ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Προστασία του Συστήματος Ελέγχου Φορτίου-Συχνότητας από Κυβερνοεπιθέσεις με Χρήση Βαθιάς Ενισχυτικής Μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Βασίλειος Δ. Δημητρόπουλος

Επιβλέπων : Νικόλαος Χατζηαργυρίου

Ομότιμος Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2025

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Προστασία του Συστήματος Ελέγχου Φορτίου-Συχνότητας από Κυβερνοεπιθέσεις με Χρήση Βαθιάς Ενισχυτικής Μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Βασίλειος Δ. Δημητρόπουλος

Επιβλέπων : Νικόλαος Χατζηαργυρίου
Ομότιμος Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 25η Φεβρουαρίου 2025.

……..	……..	……..…..……..
Νικόλαος Χατζηαργυρίου	Γεώργιος Κορρές	Παύλος Γεωργιλάκης
Ομότιμος Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2025

.....

Βασίλειος Δ. Δημητρόπουλος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Περίληψη

Τα συστήματα ηλεκτρικής ενέργειας αποτελούν πολύπλοκα δίκτυα κρίσιμης σημασίας για τη σύγχρονη κοινωνία. Για τη διασφάλιση ομαλής και ευσταθούς λειτουργίας αυτών των συστημάτων, απαιτείται η εφαρμογή ποικίλων τεχνικών προστασίας και ελέγχου. Μεταξύ αυτών, ο έλεγχος φορτίου-συχνότητας (LFC) διαδραματίζει κεντρικό ρόλο στη διατήρηση της ευστάθειας, αποτελώντας μια πολύπλοκη διαδικασία που απαιτεί τη συνεργασία μηχανικών στοιχείων με συστήματα πληροφορίας και επικοινωνιών. Η ολοένα αυξανόμενη ενοποίηση των τεχνολογιών έξυπνων δικτύων έχει κάνει τα συστήματα LFC ευάλωτα σε κυβερνοεπιθέσεις, όπως επιθέσεις έγχυσης ψευδών δεδομένων (FDIAs), οι οποίες διαστρεβλώνουν μετρήσεις αισθητήρων και αποσταθεροποιούν τη συχνότητα του δικτύου. Πολλές τεχνικές έχουν προταθεί για την ενίσχυση της ασφάλειας του LFC, με τις σύγχρονες τάσεις να εστιάζουν στην Ενισχυτική Μάθηση (RL), λόγω της ικανότητάς της να προσαρμόζεται σε δυναμικά και απρόβλεπτα περιβάλλοντα. Η RL έχει εφαρμοστεί τόσο ως ελεγκτής όσο και για το σχεδιασμό επιθέσεων σε συστήματα LFC. Ωστόσο, η χρήση της ως ελεγκτής υπό συνθήκες κυβερνοεπιθέσεων παραμένει ανεξερεύνητη, με υπάρχουσες μελέτες να αγνοούν τη διαφοροποίηση μεταξύ κανονικών διαταραχών και εσκεμμένων επιθέσεων κατά τη λειτουργία του ελέγχου. Αυτό το ερευνητικό κενό επιχειρεί να καλύψει η παρούσα εργασία, με την εισαγωγή του DRL^2FC, ενός πράκτορα RL, ως ελεγκτή για συστήματα LFC, που εκπαιδεύεται και αξιολογείται ρητά υπό συνθήκες κυβερνοεπιθέσεων. Ο πράκτορας υιοθετεί την αρχιτεκτονική Double Deep Q-Network, η οποία διαχωρίζει τις διαδικασίες επιλογής και αξιολόγησης ενεργειών προς εύρεση της βέλτιστης πολιτικής ελέγχου. Αυτή η πολιτική καθιστά τον ελεγκτή ικανό να διατηρεί ευστάθεια ακόμη και υπό συνθήκες FDI επιθέσεων που διαστρεβλώνουν μετρήσεις αισθητήρων. Το εκπαιδευτικό περιβάλλον του DRL^2FC προσομοιώνει απότομες διαταραχές φορτίου και ταυτόχρονες FDIAs, επιτρέποντας στον πράκτορα να μάθει να διακρίνει ανάμεσα σε κανονικές διακυμάνσεις και εχθρικές παρεμβολές. Η πειραματική αξιολόγηση σε δίκτυο LFC δύο περιοχών επιβεβαίωσε την ανωτερότητα του DRL^2FC. Ο πράκτορας διαχειρίστηκε επιτυχώς αλλαγές φορτίου ενώ επέφερε σταθερότητα μετά από κυβερνοεπιθέσεις. Επιπλέον, η αξιολόγηση επεκτάθηκε τόσο σε περιβάλλοντα με θόρυβο, όσο και σε σύγκριση με άλλες μεθόδους ανίχνευσης και αντιμετώπισης επιθέσεων, με εξίσου ικανοποιητικά αποτελέσματα. Τα αποτελέσματα υπογραμμίζουν τη δυνατότητα της RL να συνδυάζει έλεγχο και ασφάλεια, καθιστώντας την αποτελεσματική λύση για την προστασία κρίσιμων υποδομών έναντι δυναμικών κυβερνοαπειλών.

Λέξεις Κλειδιά

Έξυπνο Δίκτυο, Ευστάθεια Συστήματος Ηλεκτρικής Ενέργειας, Έλεγχος Φορτίου-Συχνότητας, Αυτόματος Έλεγχος Παραγωγής, Κυβερνοασφάλεια, Κυβερνοεπιθέσεις, Επιθέσεις Έγχυσης Ψευδών Δεδομένων, Βαθιά Ενισχυτική Μάθηση, Κυβερνο-Ανθεκτικός Έλεγχος

Abstract

Modern power systems constitute complex networks of critical importance to contemporary society. Ensuring the stable and reliable operation of these systems requires the implementation of diverse protection and control mechanisms. Among these, Load Frequency Control (LFC) plays a central role in maintaining grid stability, representing a sophisticated process that necessitates the collaboration of mechanical components with information and communication systems. The growing integration of smart grid technologies has rendered LFC systems vulnerable to cyberattacks, such as False Data Injection Attacks (FDIAs), which distort sensor measurements and destabilize grid frequency. Numerous techniques have been proposed to enhance LFC security, with modern trends emphasizing Reinforcement Learning (RL) due to its adaptability to dynamic and unpredictable environments. RL has been applied both as a controller and for simulating attacks in LFC systems. However, its use as a controller under cyberattack conditions remains unexplored, with existing studies neglecting the distinction between benign disturbances and malicious attacks during control operations. This research gap is addressed by the present work through the introduction of DRL^2FC, an RL-based controller for LFC systems, explicitly trained and evaluated under cyberattack scenarios. The agent adopts the Double Deep Q-Network architecture, which decouples action selection and evaluation processes to derive an optimal control policy. This policy enables the controller to maintain stability even under FDIA conditions that corrupt sensor measurements. The DRL^2FC training environment simulates abrupt load disturbances and simultaneous FDIAs, allowing the agent to learn to differentiate between normal fluctuations and adversarial perturbations. Experimental validation on a two-area LFC network confirmed the superiority of DRL^2FC. The agent successfully managed load variations while restoring stability post-cyberattack. Furthermore, evaluations extended to noisy environments and comparisons with other attack detection/mitigation methods yielded equally promising results. These findings highlight RL's potential to integrate control and security, enhancing its suitability for safeguarding critical infrastructure against evolving cyber threats.

Key Words

Smart Grid, Power System Stability, Load-Frequency Control (LFC), Automatic Generation Control (AGC), Cybersecurity, Cyberattacks, False Data Injection Attacks (FDIAs), Deep Reinforcement Learning (DRL), Cyber-Resilient Control

Ευχαριστίες

Με την εκπόνηση της παρούσας διπλωματικής ολοκληρώνεται μια μακρά πορεία. Μια πορεία που έμοιαζε ομόρροπη με αυτή του Γκρέγκορ. Μια πορεία που τελικά την ανέστρεψε το κίνητρο που εκπέμπουν τα μάτια της Ίννας. Μπορώ ξανά να βγω από το δωμάτιο.

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα Καθηγητή μου, κ. Νικόλαο Χατζηαργυρίου για την εμπιστοσύνη που μου έδειξε και την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον αντικείμενο.

Εξίσου θερμά ευχαριστώ τον συνεπιβλέποντα της εργασίας, Δρ. Ανδρέα-Δωρόθεο Συρμακέση, για την έμπνευση και την καθοδήγηση. Με χρήση της μαιευτικής μεθόδου κατέστησε την ερευνητική διαδικασία μια απολαυστική και ανταποδοτική εμπειρία.

Η εργασία αυτή θεμελιώθηκε πάνω στην, άνευ όρων και διαρκή, αγάπη και στήριξη από τους γονείς μου, Νίκη και Δημήτρη. Είναι σκοπός μου, σε κάθε βήμα, να τους φωνάζω την εκτίμηση και την ευγνωμοσύνη μου.

Αφιερώνεται στο Δημήτρη, τον νέο μου Φίλο, που βλέπουμε μαζί μπάλα.

Βασίλης Δημητρόπουλος,

Φεβρουάριος 2025

Περιεχόμενα

Ευρετήριο Σχημάτων

1

ΕΙΣΑΓΩΓΗ

1.1. Εισαγωγή στα Έξυπνα Δίκτυα Ηλεκτρικής Ενέργειας

Ο όρος Έξυπνο Δίκτυο (Smart Grid) αναφέρεται στο σύνολο της τεχνολογίας που χρησιμοποιείται για την αναβάθμιση των συστημάτων ηλεκτρικής ενέργειας, μέσω υπολογιστικών συστημάτων απομακρυσμένου ελέγχου και αυτοματοποίησης. Αυτά τα συστήματα καθίστανται δυνατά χάρη στην τεχνολογία επικοινωνίας διπλής κατεύθυνσης και την υπολογιστική επεξεργασία, που χρησιμοποιούνται εδώ και δεκαετίες σε άλλες βιομηχανίες [1],[2]. Πλέον εφαρμόζονται στα δίκτυα ηλεκτρικής ενέργειας, από τα εργοστάσια παραγωγής και τις ανανεώσιμες πηγές ενέργειας, έως τους καταναλωτές. Προσφέρουν πολλά οφέλη στις εταιρείες παροχής και στους καταναλωτές μέσω σημαντικών βελτιώσεων στην ενεργειακή απόδοση, τόσο στο δίκτυο ηλεκτρικής ενέργειας, όσο και τοπικά στις οικίες και τους επαγγελματικούς χώρους των χρηστών ενέργειας.

Για δεκαετίες, οι εταιρείες παροχής ηλεκτρικής ενέργειας έπρεπε να στέλνουν εργαζόμενους για να συλλέγουν πολλά από τα δεδομένα που απαιτούνται για την παροχή ηλεκτρισμού. Οι εργαζόμενοι διάβαζαν τις ενδείξεις των μετρητών και αναζητούσαν ελαττωματικό εξοπλισμό. Ακόμα και σήμερα, αρκετά υποσυστήματα των ηλεκτρικών δικτύων δεν έχουν ακόμη αυτοματοποιηθεί σε ικανοποιητικό βαθμό.

Ένα βασικό χαρακτηριστικό του έξυπνου δικτύου είναι οι αυτοματισμοί που επιτρέπουν στους διαχειριστές να ρυθμίζουν και να ελέγχουν χιλιάδες συσκευές από απομακρυσμένη τοποθεσία. Κάθε συσκευή στο δίκτυο εξοπλίζεται με αισθητήρες για τη συλλογή δεδομένων (μετρητές ισχύος, αισθητήρες τάσης, ανιχνευτές βλαβών, κ.λπ.), καθώς και με δυνατότητα ψηφιακής επικοινωνίας δύο κατευθύνσεων μεταξύ της συσκευής στο πεδίο και του κέντρου λειτουργίας του δικτύου της εταιρείας παροχής.

Ο αριθμός των εφαρμογών στο έξυπνο δίκτυο αυξάνεται ραγδαία, όπως και το ενδιαφέρον των ερευνητών και των επενδυτών. Η ενίσχυση της κυβερνοασφάλειας, η διαχείριση ανανεώσιμων πηγών ηλεκτρικής ενέργειας, η ενσωμάτωση ηλεκτρικών οχημάτων στο δίκτυο και η έξυπνη τιμολόγηση, αποτελούν τις σημαντικότερες προκλήσεις για την πλήρη προσαρμογή στις επιταγές της σύγχρονης εποχής. Συνεπώς, στις εταιρείες που δραστηριοποιούνται στην τεχνολογία του δικτύου περιλαμβάνονται πλέον εταιρείες τηλεπικοινωνιών, καθώς και καινοτόμες εταιρείες τεχνολογίας.

1.2. Χαρακτηριστικά των Έξυπνων Δικτύων

Σύμφωνα με το υπουργείο ενέργειας των Ηνωμένων Πολιτειών (U.S. DEPARTMENT OF ENERGY – DOE) τα βασικά χαρακτηριστικά του έξυπνου δικτύου συνοψίζονται στα εξής [16]:

Ενεργή συμμετοχή των καταναλωτών: Οι καταναλωτές βοηθούν στην εξισορρόπηση προσφοράς - ζήτησης και διασφαλίζουν την αξιοπιστία, μεταβάλλοντας τον τρόπο που χρησιμοποιούν και αγοράζουν ηλεκτρική ενέργεια. Οι καταναλωτές έχουν πλέον επιλογές που αφορούν νέες τεχνολογίες, νέες πληροφορίες σχετικά με τη χρήση της ηλεκτρικής ενέργειας, καθώς και νέες μορφές τιμολόγησης.

Φιλοξενεί όλες τις επιλογές παραγωγής και αποθήκευσης: Ένα έξυπνο δίκτυο περιέχει όχι μόνο μεγάλα, κεντρικά εργοστάσια παραγωγής, αλλά και όλο το φάσμα των διανεμημένων πηγών ενέργειας. Η ενσωμάτωση των ανανεώσιμων πηγών ενέργειας, των μικρής κλίμακας συστημάτων συνδυασμένης παραγωγής θερμότητας και ηλεκτρικής ενέργειας, καθώς και των μονάδων αποθήκευσης ενέργειας αυξάνεται ταχύτατα με συμμετοχή τόσο των προμηθευτών όσο και των εμπόρων, ακόμα και των ίδιων των ιδιωτών καταναλωτών.

Επιτρέπει νέα προϊόντα, υπηρεσίες και αγορές: Οι ορθά σχεδιασμένες και λειτουργικές αγορές ενέργειας επιτρέπουν στους καταναλωτές να επιλέγουν ανάμεσα σε ανταγωνιστικές υπηρεσίες. Οι ρυθμιστικές αρχές, οι ιδιοκτήτες/διαχειριστές και οι καταναλωτές χρειάζονται την ευελιξία να τροποποιούν τους κανόνες της επιχειρηματικής δραστηριότητας ώστε να ανταποκρίνονται στις συνθήκες λειτουργίας και της αγοράς. Μερικές από τις ανεξάρτητες μεταβλητές του δικτύου που πρέπει να διαχειρίζονται ρητά είναι: η ενέργεια, η χωρητικότητα, η τοποθεσία, ο χρόνος, ο ρυθμός μεταβολής και η ποιότητα.

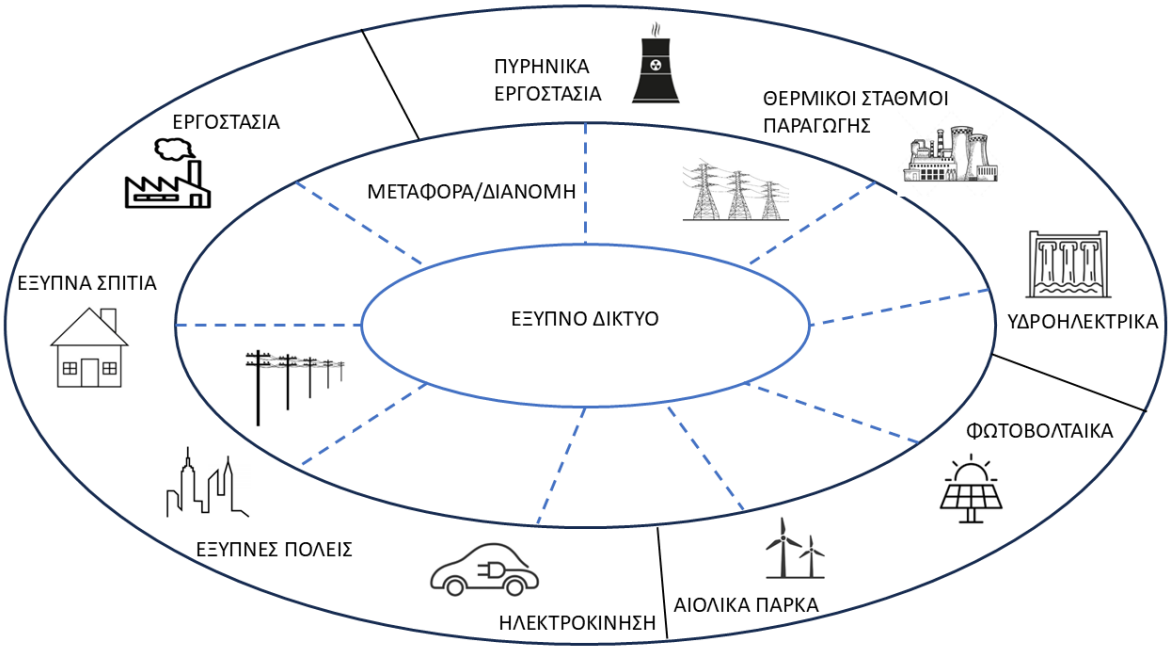
Παρέχει ποιοτική ισχύ σε όλο το φάσμα καταναλωτών: Δεν χρειάζονται όλοι οι καταναλωτές την ίδια ποσότητα ισχύος. Ένα έξυπνο δίκτυο παρέχει διαφορετικές κατηγορίες συμβάσεων ηλεκτρικής ενέργειας. Το κόστος των χαρακτηριστικών υψηλής ποιότητας ισχύος μπορεί να συμπεριληφθεί στη σύμβαση παροχής ηλεκτρικής ενέργειας. Προηγμένες μέθοδοι ελέγχου παρακολουθούν βασικά στοιχεία του συστήματος, επιτρέποντας τη γρήγορη διάγνωση και την αντιμετώπιση συμβάντων που επηρεάζουν την ποιότητα ισχύος, όπως κεραυνοί, υπερτάσεις από διακοπές και επαναφορές, βλάβες γραμμών και αρμονικές πηγές.

Λειτουργεί αποδοτικά και βελτιστοποιεί τη χρήση των πόρων: Ένα έξυπνο δίκτυο εφαρμόζει τις πιο σύγχρονες τεχνολογίες για τη βέλτιστη αξιοποίηση των πόρων του. Για παράδειγμα, η βελτιστοποιημένη χωρητικότητα μπορεί να επιτευχθεί μέσω δυναμικών ονομαστικών τιμών, οι οποίες επιτρέπουν τη χρήση των πόρων σε υψηλότερα φορτία, παρακολουθώντας συνεχώς και αξιολογώντας τις δυνατότητές τους. Η αποδοτικότητα της συντήρησης μπορεί να βελτιστοποιηθεί με τη συντήρηση βάσει κατάστασης, η οποία ειδοποιεί για την ανάγκη συντήρησης του εξοπλισμού ακριβώς τη σωστή στιγμή. Οι συσκευές ελέγχου του συστήματος μπορούν να ρυθμιστούν ώστε να μειώσουν τις απώλειες και να εξαλείψουν τη συμφόρηση. Η αποδοτικότητα της λειτουργίας αυξάνεται όταν επιλέγεται το χαμηλότερου κόστους διαθέσιμο σύστημα παροχής ενέργειας μέσω αυτών των τύπων συσκευών ελέγχου συστήματος.

Είναι ανθεκτικό σε διαταραχές, επιθέσεις και φυσικές καταστροφές: Η ανθεκτικότητα αναφέρεται στην ικανότητα ενός συστήματος να αντιδρά σε απρόβλεπτα γεγονότα, απομονώνοντας τα προβληματικά στοιχεία, ενώ το υπόλοιπο σύστημα επανέρχεται στη φυσιολογική του λειτουργία. Αυτές οι ενέργειες μειώνουν τις διακοπές παροχής στους καταναλωτές και βοηθούν τους παρόχους να διαχειρίζονται καλύτερα την υποδομή διανομής.

1.3. Τεχνολογικοί Τομείς Έξυπνων Δικτύων

Το έξυπνο δίκτυο αφορά τεχνολογίες με εφαρμογές σε ολόκληρο το ηλεκτρικό δίκτυο, από την παραγωγή έως τη μεταφορά και διανομή ηλεκτρικής ενέργειας. Ορισμένες από αυτές τις εφαρμογές έχουν ήδη αναπτυχθεί και εφαρμόζονται ευρέως, ενώ άλλες βρίσκονται ακόμα σε στάδιο ανάπτυξης και δοκιμών. Ένα πλήρως βελτιστοποιημένο σύστημα ηλεκτρικής ενέργειας θα πρέπει να ενσωματώσει όλες τους τεχνολογικούς τομείς που παρουσιάζονται στη συνέχεια.



Σχήμα 1: Σχηματική αναπαράσταση Έξυπνου Συστήματος Ηλεκτρικής Ενέργειας

1.3.1. Παρακολούθηση και έλεγχος ευρείας περιοχής

Η παρακολούθηση και η προβολή σε πραγματικό χρόνο των στοιχείων και της απόδοσης του ηλεκτρικού συστήματος σε διασυνδέσεις και μεγάλες γεωγραφικές περιοχές βοηθούν τους διαχειριστές να κατανοούν και να βελτιστοποιούν τη συμπεριφορά και την απόδοση του δικτύου. Προηγμένα συστήματα συμβάλλουν στην αποφυγή μπλακάουτ και διευκολύνουν την ενσωμάτωση ανανεώσιμων πηγών ενέργειας. Συγκεκριμένα, συστήματα Επίγνωσης Κατάστασης Ευρείας Περιοχής (Wide Area Situational Awareness - WASA) [17], συστήματα

παρακολούθησης ευρείας περιοχής (Wide Area Monitoring Systems - WAMS) [18] και της Προσαρμοστικής Προστασίας, Ελέγχου και Αυτοματοποίησης Ευρείας Περιοχής (Wide Area Adaptive Protection, Control and Automation - WAAPCA) [19] παράγουν δεδομένα για τη λήψη αποφάσεων, τον μετριασμό διαταραχών μεγάλης κλίμακας και τη βελτίωση της ικανότητας μεταφοράς και της αξιοπιστίας του συστήματος.

1.3.2. Ενσωμάτωση τεχνολογίας πληροφοριών και επικοινωνιών (Information and Communication Technology - ICT)

Η υποδομή επικοινωνίας χρησιμοποιεί ιδιωτικά δίκτυα επικοινωνίας εταιρειών κοινής ωφέλειας (ραδιοφωνικά δίκτυα, δίκτυα μετρητών) αλλά και δημόσιους παρόχους και δίκτυα (διαδίκτυο, κινητή τηλεφωνία, καλωδιακή τηλεόραση ή τηλέφωνο). Μέσω αυτής επιτυγχάνεται η μετάδοση δεδομένων ακόμα και κατά τη διάρκεια διακοπών. Μαζί με τις συσκευές επικοινωνίας, λογισμικό ελέγχου συστημάτων και λογισμικό προγραμματισμού επιχειρησιακών πόρων υποστηρίζουν την αμφίδρομη ανταλλαγή πληροφοριών μεταξύ των εμπλεκομένων μερών, επιτρέποντας την πιο αποδοτική χρήση και διαχείριση του δικτύου [20].

1.3.3. Ενσωμάτωση ανανεώσιμων και διανεμημένων πηγών παραγωγής

Η ενσωμάτωση ανανεώσιμων και διανεμημένων πηγών ενέργειας μπορεί να δημιουργήσει προκλήσεις για τον έλεγχο αυτών των πόρων και για την αποδοτική λειτουργία του συστήματος ηλεκτρικής ενέργειας. Τα συστήματα αποθήκευσης ενέργειας μπορούν να επιλύσουν τέτοια προβλήματα αποσυνδέοντας την παραγωγή και την ζήτηση ενέργειας. Τα έξυπνα δίκτυα μπορούν να βοηθήσουν μέσω της αυτοματοποίησης του ελέγχου παραγωγής να εξασφαλίσουν την εξισορρόπηση της προσφοράς και της ζήτησης [21].

1.3.4. Εφαρμογές ενίσχυσης της μεταφοράς

Το έξυπνο δίκτυο εκσυγχρονίζει τα συστήματα μεταφοράς με πλήθος εφαρμογών.

Τα Ευέλικτα Συστήματα Μεταφοράς Εναλλασσόμενου Ρεύματος (Flexible AC Transmission Systems - FACTS) χρησιμοποιούνται για την αποδοτική ροή ισχύος και τη στήριξη της τάσης. Η εφαρμογή αυτής της τεχνολογίας σε υπάρχουσες γραμμές μπορεί να βελτιώσει την αποδοτικότητα και να καθυστερήσει την ανάγκη για πρόσθετες επενδύσεις [22].

Οι τεχνολογίες Υψηλής Τάσης Συνεχούς Ρεύματος (High Voltage Direct Current - HVDC) χρησιμοποιούνται για τη σύνδεση ενεργειακών πηγών μακριά από τα κέντρα, με μειωμένες απώλειες και ενισχυμένο έλεγχο του συστήματος [23].

Η Δυναμική Αξιολόγηση Γραμμής (Dynamic Line Rating - DLR), η οποία χρησιμοποιεί αισθητήρες για την αναγνώριση της ικανότητας μεταφοράς ρεύματος ενός τμήματος του δικτύου σε πραγματικό χρόνο, μπορεί να βελτιστοποιήσει την αξιοποίηση των υπαρχόντων πόρων μεταφοράς χωρίς τον κίνδυνο υπερφόρτωσης [24].

Οι Υπεραγωγοί Υψηλής Θερμοκρασίας (High Temperature Superconductors - HTS) μπορούν επίσης να μειώσουν σημαντικά τις απώλειες μεταφοράς και να περιορίσουν τις βλάβες [25].

1.3.5. Διαχείριση δικτύου διανομής

Η παρακολούθηση και αυτοματοποίηση του δικτύου διανομής και των υποσταθμών μπορεί να μειώσει τον χρόνο διακοπής και επισκευής, να διατηρήσει το επίπεδο τάσης και να βελτιώσει τη διαχείριση των πόρων. Οι αυτοματισμοί της διανομής επεξεργάζονται πληροφορίες από αισθητήρες και μετρητές σε πραγματικό χρόνο για τον εντοπισμό βλαβών, την αυτόματη αναδιαμόρφωση των τροφοδοτών, τη βελτιστοποίηση της τάσης και της ροής άεργου ισχύος, καθώς και για τον έλεγχο της διανεμημένης παραγωγής.

Η τεχνολογία αισθητήρων επιτρέπει τη συντήρηση των συστατικών του δικτύου βάσει κατάστασης και απόδοσης, βελτιστοποιώντας την απόδοση του εξοπλισμού, μειώνοντας το κόστος συντήρησης και, συνεπώς, συμβάλλοντας στην αποτελεσματική αξιοποίηση των πόρων [26].

1.3.6. Προηγμένη υποδομή μετρητικών συστημάτων (Advanced Metering Infrastructure - AMI)

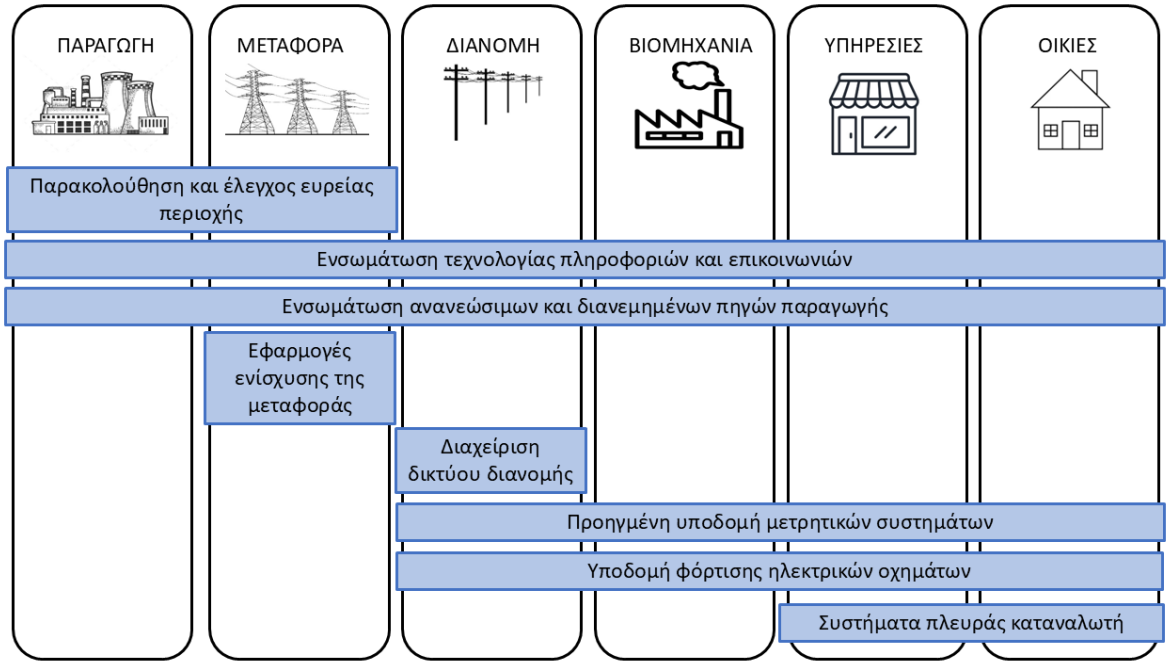
Η προηγμένη υποδομή μετρητικών συστημάτων περιλαμβάνει την ανάπτυξη μιας σειράς αυτοματισμών που επιτρέπουν τη δίοδο πληροφορίας σε αμφίδρομη κατεύθυνση, παρέχοντας στους πελάτες και τις εταιρείες κοινής ωφέλειας δεδομένα πραγματικού χρόνου για την τιμή και την κατανάλωση ηλεκτρικής ενέργειας. Τα σύγχρονα μετρητικά συστήματα διευκολύνουν τον εντοπισμό βλαβών και την αναγνώριση φαινομένων ρευματοκλοπών. Επίσης, δίνεται η δυνατότητα στους διαχειριστές να κάνουν απομακρυσμένους χειρισμούς σύνδεσης και αποσύνδεσης καταναλωτών στο δίκτυο [27].

1.3.7. Υποδομή φόρτισης ηλεκτρικών οχημάτων

Η υποδομή φόρτισης ηλεκτρικών οχημάτων διαχειρίζεται την τιμολόγηση και τον προγραμματισμό της έξυπνης φόρτισης (δικτύου προς όχημα) κατά τη διάρκεια περιόδων χαμηλής ζήτησης ενέργειας. Μακροπρόθεσμα, προβλέπεται ότι οι μεγάλες εγκαταστάσεις φόρτισης θα παρέχουν επικουρικές υπηρεσίες στο σύστημα ηλεκτρικής ενέργειας, όπως εφεδρεία ισχύος, εξομάλυνση αιχμής φόρτου και μεταφορά ενέργειας από το όχημα στο δίκτυο [28]. Αυτό θα περιλαμβάνει αλληλεπίδραση τόσο με την AMI όσο και με τα συστήματα πλευράς καταναλωτή.

1.3.8. Συστήματα πλευράς καταναλωτή

Τα συστήματα πλευράς καταναλωτή βοηθούν στη διαχείριση της κατανάλωσης ηλεκτρικής ενέργειας σε βιομηχανικό, εμπορικό και οικιακό επίπεδο. Περιλαμβάνουν συσκευές αποθήκευσης ενέργειας, έξυπνες οικιακές συσκευές και διανεμημένη παραγωγή. Σκοπός τους είναι η αποδοτική κατανάλωση ενέργειας και η μείωση της ζήτησης αιχμής [29].



Σχήμα 2: Τεχνολογίες Έξυπνων Δικτύων και πεδία εφαρμογής τους

1.4. Εξοπλισμός Έξυπνων Δικτύων

Οι υποδομές και ο εξοπλισμός των έξυπνων δικτύων χρησιμοποιούνται ήδη σε άλλες εφαρμογές, όπως η βιομηχανία και οι τηλεπικοινωνίες, και προσαρμόζονται για χρήση στις λειτουργίες του δικτύου.

1.4.1. Εξοπλισμός και υποδομές επικοινωνίας

Ο τομέας των επικοινωνιών έχει αναπτυχθεί σταδιακά και δεν είναι πλήρως ενοποιημένος. Στις περισσότερες περιπτώσεις, τα δεδομένα συλλέγονται μέσω ενός modem αντί για άμεση σύνδεση δικτύου. Τομείς για βελτίωση περιλαμβάνουν: αυτοματοποίηση υποσταθμών, απόκριση ζήτησης, αυτοματοποίηση διανομής, ασύρματα δίκτυα πλέγματος ,τεχνολογίες επικοινωνίας μέσω γραμμής μεταφοράς ρεύματος και οπτικές ίνες [30]. Οι ενοποιημένες επικοινωνίες θα επιτρέψουν τον έλεγχο σε πραγματικό χρόνο, την ανταλλαγή πληροφοριών και δεδομένων για τη βελτιστοποίηση της αξιοπιστίας του συστήματος, της αξιοποίησης των πόρων και της ασφάλειας.

1.4.2. Αισθητήρες

Οι κύριες αρμοδιότητες των αισθητήρων είναι η αξιολόγηση της συμφόρησης και της σταθερότητας του δικτύου, η παρακολούθηση της υγείας του εξοπλισμού, η πρόληψη ρευματοκλοπών και η υποστήριξη στρατηγικών ελέγχου. Ως αισθητήρες αναφέρονται τα WAMS, τα συστήματα DLR, τα συστήματα Θερμικής Αξιολόγησης σε Πραγματικό Χρόνο

(Real Time Thermal Rating - RTTR) [31], οι προηγμένοι διακόπτες και οι ψηφιακοί προστατευτικοί ηλεκτρονόμοι.

1.4.3. Έξυπνοι μετρητές

Ένα έξυπνο δίκτυο αντικαθιστά τους αναλογικούς μηχανικούς μετρητές με ψηφιακούς μετρητές που καταγράφουν τη χρήση σε πραγματικό χρόνο. Οι έξυπνοι μετρητές είναι παρόμοιοι με τους μετρητές της AMI και παρέχουν έναν δίαυλο επικοινωνίας που εκτείνεται από τα εργοστάσια παραγωγής έως τις έξυπνες ηλεκτρικές υποδοχές (smart plugs) και άλλες συσκευές που υποστηρίζουν το έξυπνο δίκτυο. Κατόπιν επιλογής του πελάτη, τέτοιες συσκευές μπορούν να απενεργοποιούνται κατά τη διάρκεια των περιόδων υψηλής ζήτησης [32].

1.4.4. Μονάδες Μέτρησης Φάσης (Phasor Measurment Units - PMUs)

Πρόκειται για υψηλής ταχύτητας αισθητήρες, κατανεμημένους σε όλο το δίκτυο, που παρακολουθούν την ποιότητας ενέργειας και ανταποκρίνονται αυτόματα σε αυτήν. Οι φάσεις είναι αναπαραστάσεις των κυματομορφών της εναλλασσόμενης τάσης. Ιδανικά, είναι ταυτόσημες σε όλο το δίκτυο και συμμορφώνονται με το σήμα αναφοράς [33].

Ήδη από τη δεκαετία του 1980, αναγνωρίστηκε ότι οι παλμοί του ρολογιού από τους δορυφόρους του Παγκόσμιου Συστήματος Εντοπισμού Θέσης (Global Positioning System - GPS) θα μπορούσαν να χρησιμοποιηθούν για πολύ ακριβείς μετρήσεις χρόνου στο δίκτυο [34]. Με μεγάλο αριθμό PMUs και τη δυνατότητα σύγκρισης των κυματομορφών από τις αναγνώσεις της εναλλασσόμενης τάσης σε όλο το δίκτυο, τα αυτοματοποιημένα συστήματα είναι σε θέση να ανταποκριθούν στις συνθήκες του συστήματος με ταχύ και δυναμικό τρόπο.

Ένα σύστημα ευρείας περιοχής μέτρησης (WAMS) είναι ένα δίκτυο PMUs που μπορεί να παρέχει παρακολούθηση σε πραγματικό χρόνο σε περιφερειακή και εθνική κλίμακα.

1.4.5. Συστήματα εποπτικού ελέγχου και απόκτησης δεδομένων (Supervisory Control and Data Acquisition - SCADA)

Ο όρος SCADA περιγράφει μια κατηγορία συστημάτων βιομηχανικού αυτομάτου ελέγχου και τηλεμετρίας. Το χαρακτηριστικό των συστημάτων SCADA είναι ότι αποτελούνται από τοπικούς ελεγκτές, που ελέγχουν επί μέρους στοιχεία και μονάδες μιας εγκατάστασης, συνδεδεμένους σε έναν Κύριο Σταθμό Εργασίας (Master Station). Ο κύριος σταθμός εργασίας μπορεί κατόπιν να επικοινωνεί τα δεδομένα που συλλέγει από την εγκατάσταση σε ένα πλήθος από σταθμούς εργασίας σε Τοπικό Δίκτυο (Local Area Network- LAN) ή και να μεταδίδει τα δεδομένα της εγκατάστασης σε μακρινά σημεία μέσω κάποιου συστήματος τηλεπικοινωνίας, όπως το ενσύρματο τηλεφωνικό δίκτυο. Επίσης είναι δυνατό ο κάθε ένας τοπικός ελεγκτής να βρίσκεται σε απομακρυσμένη τοποθεσία και να μεταδίδει τα δεδομένα προς το κύριο σταθμό εργασίας μέσω απλού καλωδίου ή μέσω ασύρματου πομποδέκτη [35].

1.4.6. Προηγμένα εξαρτήματα

Οι καινοτομίες στη υπεραγωγιμότητα, την αντοχή σε βλάβες, την αποθήκευση, την ηλεκτρονική ισχύς και τα διαγνωστικά εξαρτήματα αλλάζουν τις θεμελιώδεις ικανότητες και χαρακτηριστικά των δικτύων [36]. Οι τεχνολογίες εντός αυτών των ευρύτερων κατηγοριών περιλαμβάνουν: συσκευές ευέλικτου συστήματος μεταφοράς εναλλασσόμενου ρεύματος, HVDC, πρώτης και δεύτερης γενιάς καλώδια υπεραγωγών, καλώδια HTS, συσκευές παραγωγής και αποθήκευσης ενέργειας διανεμημένου τύπου, σύνθετοι αγωγοί και έξυπνες συσκευές.

1.4.7. Προηγμένο λογισμικό ελέγχου

Η αυτοματοποίηση του συστήματος ισχύος επιτρέπει τη γρήγορη διάγνωση και τις ακριβείς λύσεις για συγκεκριμένες διαταραχές ή διακοπές του δικτύου. Χαρακτηριστικό παράδειγμα αποτελεί το λογισμικό Παρακολούθησης και Ελέγχου Σταθερότητας Τάσης (Voltage Stability Monitoring and Control - VSMC) [37] , το οποίο χρησιμοποιεί μέθοδο ευαισθησίας με επαναλαμβανόμενο γραμμικό προγραμματισμό για να προσδιορίσει αξιόπιστα τη βέλτιστη ροή ισχύος στις γραμμές του δικτύου.

1.5. Ευστάθεια Έξυπνων Δικτύων

Η ευστάθεια είναι ένα θεμελιώδες στοιχείο της λειτουργίας των δικτύων ενέργειας, εξασφαλίζοντας ότι το σύστημα μπορεί να διατηρήσει μια κατάσταση ισορροπίας υπό κανονικές συνθήκες λειτουργίας και να επιστρέψει σε μια σταθερή κατάσταση αφού υποστεί διαταραχές. Στο πλαίσιο των έξυπνων δικτύων, η ευστάθεια γίνεται ακόμα πιο κρίσιμη λόγω της ενσωμάτωσης ανανεώσιμων πηγών ενέργειας, διανεμημένης παραγωγής και προηγμένων συστημάτων ελέγχου, τα οποία εισάγουν νέες δυναμικές και προκλήσεις. Η ευστάθεια του συστήματος ενέργειας μπορεί να κατηγοριοποιηθεί σε διάφορους τύπους, οι οποίοι αντιμετωπίζουν διαφορετικές πτυχές της συμπεριφοράς του δικτύου [38]:

1.5.1. Ευστάθεια Γωνίας Δρομέα (Rotor Angle Stability)

Η ευστάθεια γωνίας δρομέα εξασφαλίζει ότι οι σύγχρονες μηχανές παραμένουν σε φάση μετά από διαταραχές, εξισορροπώντας τη μηχανική και ηλεκτρική ροπή. Τα παραδοσιακά δίκτυα βασίζονταν στην εγγενή αδράνεια των συγχρονισμένων γεννητριών, αλλά τα έξυπνα δίκτυα ενισχύουν την ευστάθεια μέσω παρακολούθησης σε πραγματικό χρόνο (π.χ. PMU και WAMS) και προσαρμοσμένων αλγορίθμων ελέγχου. Για παράδειγμα, τα συστήματα προστασίας ταχείας δράσης μειώνουν τους κινδύνους αστάθειας από αλυσιδωτές αποτυχίες .

1.5.2. Ευστάθεια Συχνότητας (Frequency Stability)

Η ευστάθεια συχνότητας διατηρεί τη συχνότητα του δικτύου στην ονομαστική τιμή 50/60Hz κατά τις ανισορροπίες μεταξύ παραγωγής και ζήτησης. Οι αποκλίσεις συχνότητας συμβαίνουν όταν υπάρχει ανισορροπία μεταξύ της παραγωγής ισχύος και της ζήτησης φορτίου. Εάν δεν διορθωθούν αμέσως, αυτές οι αποκλίσεις μπορούν να οδηγήσουν σε αστάθεια, με αποτέλεσμα την αποκοπή φορτίων, τον αποσυγχρονισμό γεννητριών ή ακόμη και σε μαζικές διακοπές ρεύματος. Η ασταθής φύση των ανανεώσιμων πηγών ενέργειας, επιδεινώνει τις διακυμάνσεις της συχνότητας. Τα έξυπνα δίκτυα αντιμετωπίζουν αυτό το πρόβλημα μέσω της ανταπόκρισης ζήτησης (Demand Response) ,έξυπνων αντιστροφέων (Smart Inverters) και συστημάτων αποθήκευσης ενέργειας (Energy Storage Systems - ESS) που παρέχουν γρήγορη ρύθμιση συχνότητας. Μελέτες δείχνουν ότι τα συστήματα αποθήκευσης ενέργειας μπαταριών μπορούν να μειώσουν τη βύθιση συχνότητας κατά τις διακοπές ρεύματος [39].

1.5.3. Ευστάθεια Τάσης (Voltage Stability)

Η ευστάθεια τάσης εξασφαλίζει ότι οι τάσεις στους κόμβους παραμένουν εντός αποδεκτών ορίων. Οι βυθίσεις τάσης προκαλούνται κατά κύριο λόγο από ανωμαλίες που συμβαίνουν στο δίκτυο διανομής. Ειδικότερα στην Ελλάδα, είναι ένα πολύ συχνό φαινόμενο, λόγω της κακής ποιότητας δικτύου και της ελλιπούς συντήρησης. Δευτερευόντως, οι βυθίσεις τάσης προκαλούνται από σφάλματα στην εσωτερική ηλεκτρολογική εγκατάσταση βιομηχανικών καταναλωτών, καθώς και από την απότομη χρήση επιπρόσθετου φορτίου, όπως για παράδειγμα συμβαίνει με την εκκίνηση ενός κινητήρα.. Τα έξυπνα δίκτυα χρησιμοποιούν δυναμικούς ρυθμιστές τάσης (Dynamic Voltage Regulators - DVRs) [40], Στατικών Συγχρονισμένων Αντισταθμιστών (Static Synchronous Compensators – STATCOMs) [41] και διανεμημένη παραγωγή για τοπική υποστήριξη αέργου ισχύος. Για παράδειγμα, διανεμημένα συστήματα ηλιακής ενέργειας (PV) μπορούν να σταθεροποιήσουν τις τάσεις κατά τη διάρκεια αιχμών ζήτησης [42].

1.5.4. Ευστάθεια Μικρο-Σημάτων (Micro-Signal Stability)

Η ευστάθεια μικρο-σημάτων ασχολείται με τις ασταθείς ταλαντώσεις χαμηλής κλίμακας που προκαλούνται από αδύναμα χαρακτηριστικά απόσβεσης ή αλληλεπιδράσεις ελέγχου. Τα έξυπνα δίκτυα χρησιμοποιούν προηγμένους ρυθμιστές απόσβεσης και υπολογιστικά μοντέλα που εκπαιδεύονται με δεδομένα από PMUs για να προβλέψουν και να μετριάσουν τις ταλαντώσεις. Τα συστήματα παρακολούθησης ευρέος πεδίου (WAMS) έχουν βελτιώσει σημαντικά τους λόγους απόσβεσης κατά σε πιλοτικά έργα [43].

1.5.5. Ευστάθεια Μεταβατικής Κατάστασης (Transient Stability)

Η ευστάθεια μεταβατικής κατάστασης επικεντρώνεται στη διατήρηση του συγχρονισμού μετά από σοβαρές βλάβες όπως βραχυκυκλώματα. Λύσεις περιλαμβάνουν τις εικονικές σύγχρονες μηχανές (Virtual Synchronous Machines - VSMs) [44] και τους περιοριστές ρεύματος βλάβης από υπεραγωγούς (Superconducting Fault Current Limiters - SFCLs). Έχει αποδειχθεί ότι οι SFCLs μπορούν να μειώσουν τους κρίσιμους χρόνους εκκαθάρισης σε δίκτυα με υψηλή διείσδυση αιολικής ενέργειας [45].

1.5.6. Ανθεκτικότητα Κυβερνο-Φυσικών Συστημάτων (Cyber-Physical Resilience)

Η εξάρτηση των έξυπνων δικτύων από ψηφιακή επικοινωνία εισάγει κυβερνο-φυσικούς κινδύνους όπως κακόβουλη χειραγώγηση των υπολογιστικών συστημάτων ή των δικτύων επικοινωνίας. Για την περιγραφή αυτού του είδους ενεργειών έχει εισαχθεί ο όρος Κυβερνοεπιθέσεις (Cyber Attacks). Η ανθεκτικότητα απαιτεί ισχυρή κρυπτογράφηση, συστήματα ανίχνευσης εισβολών και αποκεντρωμένες αρχιτεκτονικές ελέγχου [46]. Το πρότυπο NERC CIP-014 απαιτεί φυσική ασφάλεια για κρίσιμες υποδομές [47], ενώ πλέον αναπτύσσονται εργαλεία ανίχνευσης ανωμαλιών βασισμένα σε μηχανική μάθηση.

1.5.6.1. Κίνδυνοι σχετιζόμενοι με το κυβερνο-φυσικό έξυπνο δίκτυο

Οι κίνδυνοι που απειλούν το έξυπνο δίκτυο μπορούν να συνοψιστούν ως εξής [48] :

Αυξημένη πολυπλοκότητα (Advanced Complexity): Η διασύνδεση νέων τεχνολογιών στο δίκτυο εισάγει μεγαλύτερη δυσκολία υπολογισμών.

Κίνδυνος αλυσιδωτών βλαβών (Risk of Cascading Failures): Σε ένα κυβερνο-φυσικό δίκτυο, μια τυχαία βλάβη ή μια κακόβουλη επίθεσης σε οποιονδήποτε από τους δύο τομείς μπορεί να επηρεάσει τον άλλο τομέα και να οδηγήσει σε πιθανές αλυσιδωτές βλάβες.

Αύξηση πιθανών αντιπάλων (Potential Adversaries): Καθώς αυξάνεται ο αριθμός των κόμβων του δικτύου, αυξάνονται και τα σημεία εισόδου για επιτιθέμενους, γεγονός που εισάγει πιθανούς κινδύνους. Αυτό θεωρείται ένας από τους κύριους λόγους για την αύξηση των σχετικών τύπων επιθέσεων και διεισδύσεων στο Έξυπνο Δίκτυο.

Ζητήματα ιδιωτικότητας δεδομένων (Data Privacy Issues): Σε ένα Έξυπνο Δίκτυο, η εκτεταμένη χρήση Έξυπνων Ηλεκτρονικών Συσκευών (Intelligent Electronic Devices - IEDs) έχει αυξήσει σημαντικά τη συλλογή δεδομένων και τις αμφίδρομες ροές πληροφοριών, εισάγοντας προβλήματα που σχετίζονται με την εμπιστευτικότητα των δεδομένων και την παραβίαση της ιδιωτικότητας των καταναλωτών.

1.5.6.2. Απαιτήσεις ασφαλείας έξυπνων δικτύων

Για την ασφαλή και αποδοτική λειτουργία του Έξυπνων Δικτύων, τα δεδομένα που ανταλλάσσονται μεταξύ των κόμβων αποτελούν το πιο σημαντικό στοιχείο του συστήματος. Θα πρέπει λοιπόν να πληρούν κάποιες αυστηρές προϋποθέσεις. Συγκεκριμένα:

Διαθεσιμότητα Δεδομένων (Data Availability): Έγκαιρη και αξιόπιστη πρόσβαση στη χρήση των πληροφοριών.

Ακεραιότητα Δεδομένων (Data Integrity): Η προέλευση και η ποιότητα των δεδομένων είναι γνωστές και πιστοποιημένες.

Εμπιστευτικότητα Δεδομένων (Data Confidentiality): Η προστασία της ιδιωτικότητας των τελικών χρηστών.

Είναι σημαντικό να αναφερθεί ότι ένα σωστά σχεδιασμένο πλαίσιο άμυνας κατά των κυβερνοεπιθέσεων θα πρέπει να λαμβάνει υπόψιν όχι μόνο τις στοχευμένες κυβερνοεπιθέσεις, αλλά και οι ακούσιες ανωμαλίες που σχετίζονται με ICT, όπως σφάλματα ανθρώπινων

χειρισμών, σφάλματα λογισμικού, βλάβες εξοπλισμού και, προφανώς, προβλήματα που σχετίζονται με φυσικές καταστροφές [49].

Στη διαδικασία εξέλιξης του ηλεκτρικού δικτύου σε ένα πιο έξυπνο σύστημα, εισάγεται περισσότερος αυτοματοποιημένος και αποκεντρωμένος έλεγχος. Ο κίνδυνος κυβερνοεπιθέσεων θα αυξηθεί καθώς το δίκτυο γίνεται πιο αυτοματοποιημένο και ο αριθμός των κόμβων πληθαίνει.

1.5.6.3. Ιστορικά Γεγονότα

Τις τελευταίες δεκαετίες, αρκετές επιθέσεις κυβερνοασφάλειας σε δίκτυα ενέργειας και υποδομές έξυπνων δικτύων έχουν αποδείξει τη δυνατότητα των κακόβουλων δρώντων να διαταράξουν τα ενεργειακά συστήματα, να προκαλέσουν εκτεταμένες διακοπές ρεύματος και να θέσουν σε κίνδυνο κρίσιμες υποδομές. Αυτά τα περιστατικά υπογραμμίζουν τη σημασία των ισχυρών μέτρων κυβερνοασφάλειας στο σχεδιασμό και τη λειτουργία των έξυπνων δικτύων. Παρακάτω παρατίθενται μερικά αξιόλογα ιστορικά γεγονότα:

Ο Ιός Stuxnet (2010): Αν και δεν στόχευε άμεσα σε δίκτυο ενέργειας, ο ιός Stuxnet αποτελεί ένα σημαντικό γεγονός στην ιστορία των κυβερνοεπιθέσεων σε κρίσιμες υποδομές. Ανακαλύφθηκε το 2010 και ήταν ένα εξαιρετικά εξελιγμένο κακόβουλο λογισμικό, σχεδιασμένο για να σαμποτάρει το πυρηνικό πρόγραμμα του Ιράν, στοχεύοντας σε Προγραμματιζόμενους Λογικούς Ελεγκτές (Programmable Logic Controller - PLCs) στις φυγόκεντρες μηχανές (centrifuges). Αν και ο κύριος στόχος του ήταν βιομηχανικός εξοπλισμός, ο ιός Stuxnet ανέδειξε τη δυνατότητα χρήσης κυβερνοεπιθέσεων για τη διατάραξη φυσικών συστημάτων, εγείροντας ανησυχίες για παρόμοιες απειλές στα δίκτυα ενέργειας [64].

Η Εκστρατεία Dragonfly (2011–2014): Η εκστρατεία Dragonfly, γνωστή και ως Energetic Bear, στόχευσε ενεργειακές εταιρείες στην Ευρώπη και τη Βόρεια Αμερική. Οι επιτιθέμενοι χρησιμοποίησαν ηλεκτρονικό ψάρεμα (phishing) και κακόβουλο λογισμικό για να διεισδύσουν στα δίκτυα των διαχειριστών δικτύων. Παρά το γεγονός ότι η εκστρατεία δεν οδήγησε σε φυσικές διακοπές, ανέδειξε την ικανότητα των επιτιθέμενων να συλλέγουν πληροφορίες και ενδεχομένως να σαμποτάρουν τις λειτουργίες του δικτύου [65].

Οι Κυβερνοεισβολές (Cyber Intrusions) στο Δίκτυο Ενέργειας των ΗΠΑ (2013–2014): Από το 2013 έως το 2014, το Υπουργείο Εσωτερικής Ασφάλειας των ΗΠΑ (Department of Homeland Security - DHS) ανέφερε πολλές κυβερνοεισβολές στο δίκτυο ενέργειας της χώρας. Οι επιτιθέμενοι, που πιστεύεται ότι ήταν κρατικοί δρώντες, απέκτησαν πρόσβαση στα συστήματα ελέγχου των υπηρεσιών κοινής ωφελείας, χωρίς όμως να προκαλέσουν φυσικές ζημιές. Αυτά τα περιστατικά ανέδειξαν τη δυνατότητα των αντιπάλων να πραγματοποιούν αναγνωριστικές ενέργειες και να εγκαθιστούν θέση στις υποδομές του δικτύου, ανοίγοντας το δρόμο για μελλοντικές επιθέσεις [66].

Οι Επιθέσεις στο Ουκρανικό Δίκτυο Ενέργειας (2015, 2016): Μια από τις πιο σημαντικές επιθέσεις κυβερνοασφάλειας σε δίκτυο ενέργειας συνέβη τον Δεκέμβριο του 2015, όταν οι επιτιθέμενοι κατάφεραν να διαταράξουν την παροχή ηλεκτρικής ενέργειας σε περίπου 225.000 πελάτες στην Ουκρανία. Η επίθεση, η οποία αποδόθηκε στην ρωσική ομάδα χάκερ Sandworm, περιλάμβανε τη χρήση κακόβουλου λογισμικού (BlackEnergy) για να διεισδύσουν στα συστήματα ελέγχου του δικτύου. Οι επιτιθέμενοι απομακρυσμένα αποσύνδεσαν τους

διακόπτες κυκλώματος, αχρηστεύοντας τους υποσταθμούς διανομής και προκαλώντας εκτεταμένες διακοπές ρεύματος. Αυτό το περιστατικό αποτέλεσε την πρώτη γνωστή κυβερνοεπίθεση που οδήγησε σε διακοπή ρεύματος και ανέδειξε την ευπάθεια των Συστημάτων Βιομηχανικού Ελέγχου (Industrial Control Systems - ICS) σε προηγμένα απειλητικά σενάρια [67].

Τον Δεκέμβριο του 2016, η Ουκρανία υπέστη μια ακόμα κυβερνοεπίθεση στο δίκτυο ενέργειας, η οποία επηρέασε περιοχές του Κιέβου. Αυτή η επίθεση χρησιμοποίησε μια πιο εξελιγμένη παραλλαγή κακόβουλου λογισμικού με την ονομασία CrashOverride (ή Industroyer), η οποία είχε σχεδιαστεί ειδικά για να στοχεύει στις υποδομές των δικτύων ενέργειας. Το CrashOverride αυτοματοποίησε τη διαδικασία διατάραξης των λειτουργιών του δικτύου, επιτρέποντας στους επιτιθέμενους να προκαλέσουν διακοπές ρεύματος με ελάχιστη ανθρώπινη παρέμβαση. Το περιστατικό ανέδειξε τη δυνατότητα του κακόβουλου λογισμικού να χρησιμοποιηθεί ως όπλο κατά των κρίσιμων υποδομών ενέργειας [68].

Η Επίθεση Κακόβουλου Λογισμικού και Απαγωγής Δεδομένων (Ransomware) στον Αγωγό Colonial (2021): Αν και κυρίως στόχευσε έναν αγωγό καυσίμων, η επίθεση ransomware στον αγωγό Colonial τον Μάιο του 2021 είχε σημαντικές συνέπειες για την κυβερνοασφάλεια κρίσιμων υποδομών. Η επίθεση, που πραγματοποιήθηκε από την ομάδα DarkSide, ανάγκασε τον διαχειριστή του αγωγού να κλείσει τις επιχειρήσεις του, προκαλώντας ελλείψεις καυσίμων σε όλη την ανατολική ακτή των ΗΠΑ. Το περιστατικό αυτό ανέδειξε την αυξανόμενη απειλή του ransomware για τις ενεργειακές υποδομές και την ανάγκη για ισχυρές πρακτικές κυβερνοασφάλειας [69].

Οι Επιθέσεις στον Ενεργειακό Τομέα της Αυστραλίας (2021): Το 2021, ο ενεργειακός τομέας της Αυστραλίας υπέστη μια σειρά κυβερνοεπιθέσεων που στόχευαν στους ηλεκτροπαραγωγούς σταθμούς και τους διαχειριστές δικτύου. Αν και οι λεπτομέρειες παραμένουν απόρρητες, τα περιστατικά προκάλεσαν την αυστραλιανή κυβέρνηση να εκδώσει προειδοποιήσεις σχετικά με την αυξανόμενη πολυπλοκότητα των κυβερνοαπειλών για τα ενεργειακά συστήματα. Αυτές οι επιθέσεις ανέδειξαν τη παγκόσμια διάσταση της πρόκλησης της κυβερνοασφάλειας και την ανάγκη για διεθνή συνεργασία [70].

Η Διπλή Επίθεση στο Δίκτυο Ενέργειας της Ουκρανίας (2022): Τον Οκτώβριο του 2022, η ενεργειακή υποδομή της Ουκρανίας υπέστη μια μεγάλης κλίμακας κυβερνοεπίθεση, η οποία αποδόθηκε σε ρωσικές κρατικές ομάδες χάκερ. Αυτή η επίθεση αποτέλεσε μέρος μιας συντονισμένης εκστρατείας που συνδύασε κυβερνοεπιθέσεις με φυσικές επιθέσεις με πυραύλους σε ενεργειακές εγκαταστάσεις, με στόχο την αποσταθεροποίηση του δικτύου ενέργειας της Ουκρανίας και την υπονόμευση του ηθικού του πληθυσμού καθώς πλησίαζε ο χειμώνας. Η επίθεση ανέδειξε την αυξανόμενη πολυπλοκότητα του κυβερνοπολέμου και την ενσωμάτωσή του σε σύγχρονες στρατιωτικές επιχειρήσεις [3],[71].

2

ΚΥΒΕΡΝΟΕΠΙΘΕΣΕΙΣ ΕΝΑΝΤΙΑ ΣΤΟ ΣΥΣΤΗΜΑ ΕΛΕΓΧΟΥ ΦΟΡΤΙΟΥ-ΣΥΧΝΟΤΗΤΑΣ

2.1. Εισαγωγή στον Έλεγχο Φορτίου-Συχνότητας

Σήμερα, η βιομηχανία ηλεκτρικής ενέργειας έχει μεταβεί από το σενάριο κάθετα ολοκληρωμένων επιχειρήσεων, όπου ένας μόνο φορέας κατείχε και λειτουργούσε τα συστήματα παραγωγής, μεταφοράς και διανομής παρέχοντας ενέργεια, προς το σενάριο απελευθερωμένης αγοράς, όπου ανταγωνιστικές εταιρείες πωλούν αποδεσμευμένη ισχύ σε χαμηλότερες τιμές. Επιπλέον, διάφορες συσκευές με μεγάλη ισχύ και ταχεία κατανάλωση, όπως εργοστάσια πυρηνικής σύντηξης και χαλυβουργίες, αυξάνονται σημαντικά. Όταν αυτά τα φορτία συγκεντρώνονται στα συστήματα ισχύος, μπορεί να προκαλέσουν σοβαρά προβλήματα ταλαντώσεων συχνότητας

Ο Έλεγχος Φορτίου-Συχνότητας (Load-Frequency Control - LFC) είναι ένας θεμελιώδης μηχανισμός στα συστήματα ηλεκτρικής ενέργειας που εξασφαλίζει την ισορροπία μεταξύ παραγωγής και ζήτησης ισχύος, διατηρώντας παράλληλα τη συχνότητα του συστήματος εντός αποδεκτών ορίων. Οι αποκλίσεις συχνότητας μπορεί να προκύψουν λόγω αιφνίδιων αλλαγών στο φορτίο ή στην παραγωγή, οι οποίες, αν δεν ρυθμιστούν, μπορεί να οδηγήσουν σε αστάθεια, ζημιές στον εξοπλισμό ή ακόμα και σε γενικευμένες διακοπές (Blackouts).

Οι κύριοι στόχοι του συστήματος LFC είναι:

- Η διασφάλιση μηδενικού στατικού σφάλματος για αποκλίσεις συχνότητας.
- Η ελαχιστοποίηση των μη προγραμματισμένων ροών ισχύος στις γραμμές διασύνδεσης μεταξύ γειτονικών περιοχών ελέγχου.
- Η καλή παρακολούθηση της ζήτησης φορτίου και των διαταραχών.
- Η διατήρηση αποδεκτών υπερβάσεων και χρόνου αποκατάστασης για τις αποκλίσεις συχνότητας και ισχύος στις γραμμές διασύνδεσης.

Το σύστημα LFC υλοποιείται σε πολλαπλούς βρόχους ελέγχου για την επίτευξη της σταθερότητας της συχνότητας και της ορθής κατανομής φορτίου μεταξύ των μονάδων παραγωγής [50].

Ο **πρωτεύων έλεγχος** συχνότητας είναι ο πρώτος μηχανισμός απόκρισης σε αιφνίδιες αποκλίσεις της συχνότητας. Υλοποιείται μέσω του συστήματος ρύθμισης στροφών του στροβίλου (Governor), το οποίο προσαρμόζει αυτόματα την ισχύ εξόδου των μονάδων παραγωγής με βάση τις μεταβολές της συχνότητας. Ο πρωτεύων έλεγχος βασίζεται στον Έλεγχο Πτώσης (Droop Control), όπου μια αύξηση ή μείωση του φορτίου οδηγεί σε αναλογική απόκριση της ισχύος εξόδου της γεννήτριας.

Αυτός ο μηχανισμός ελέγχου είναι αποκεντρωμένος, πράγμα που σημαίνει ότι κάθε γεννήτρια προσαρμόζει ανεξάρτητα την ισχύ της με βάση τις τοπικές μετρήσεις συχνότητας. Ο χρόνος απόκρισης του πρωτεύοντος ελέγχου είναι της τάξης των δευτερολέπτων, εξασφαλίζοντας άμεση αντίδραση στις διαταραχές της συχνότητας. Ωστόσο, ο πρωτεύων έλεγχος δεν αποκαθιστά τη συχνότητα στην ονομαστική της τιμή, αλλά απλώς μετριάζει την αρχική απόκλιση.

Ο **δευτερεύων έλεγχος** συχνότητας είναι υπεύθυνος για την αποκατάσταση της συχνότητας του συστήματος στην ονομαστική της τιμή 50/60Hz και τη διατήρηση των προγραμματισμένων ανταλλαγών ισχύος μεταξύ διασυνδεδεμένων περιοχών. Αυτός ο βρόχος ελέγχου αποτελεί μέρος του Αυτόματου Ελέγχου Παραγωγής (Automatic Generation Control - AGC) και λειτουργεί κεντρικά, συνήθως στο κέντρο ελέγχου κάθε περιοχής.

Ο μηχανισμός του δευτερεύοντος ελέγχου χρησιμοποιεί το **Σφάλμα Ελέγχου Περιοχής (Area Control Error - ACE)** [51], ο οποίος αποτελεί γραμμικό συνδυασμό της απόκλισης της συχνότητας και των αποκλίσεων ισχύος στις διασυνδετικές γραμμές. Τα σημεία λειτουργίας των γεννητριών προσαρμόζονται από Ελεγκτές (Controllers) [52], για να επαναφέρουν τη συχνότητα στην ονομαστική της τιμή.

Σε αντίθεση με τον πρωτεύοντα έλεγχο, ο οποίος δρα άμεσα, ο δευτερεύων έλεγχος λειτουργεί σε μεγαλύτερη χρονική κλίμακα της τάξης των δεκάδων δευτερόλεπτων έως λεπτού, για να εξασφαλίσει μια σταδιακή και σταθερή αποκατάσταση των αποκλίσεων συχνότητας.

Η αποτελεσματικότητα του ελέγχου συχνότητας στα ηλεκτρικά συστήματα εξαρτάται από τον συντονισμό μεταξύ των πρωτευόντων και δευτερευόντων βρόχων ελέγχου. Αρχικά, ο πρωτεύων έλεγχος παρέχει άμεση απόκριση για την αντιμετώπιση των αποκλίσεων συχνότητας, αποτρέποντας την αστάθεια. Στη συνέχεια, ο δευτερεύων έλεγχος αναλαμβάνει σταδιακά, ρυθμίζοντας το σύστημα για την αποκατάσταση της συχνότητας και τη διασφάλιση της ορθής κατανομής ισχύος μεταξύ των ελεγχόμενων περιοχών [53].

Αξίζει να αναφερθεί, αν και δεν λαμβάνεται υπόψιν για την παρούσα εργασία, ο **τριτεύων** βρόχος ελέγχου. Ο έλεγχος επιτυγχάνεται με αυτόματη ή χειροκίνητη αλλαγή στα σημεία λειτουργίας των γεννητριών με σκοπό το βέλτιστη δυνατή Οικονομική Διανομή (Economic Dispatch) της παραγόμενης ισχύος και την εγγύηση επαρκούς εφεδρείας ισχύος ανά πάσα στιγμή. Το χρονικό διάστημα δράσης του είναι της τάξης των αρκετών λεπτών [54].

Τα συστήματα LFC είναι απαραίτητα για τη διασφάλιση της ευστάθειας και αξιοπιστίας των δικτύων ηλεκτρικής ενέργειας. Καθώς τα δίκτυα ηλεκτρικής ενέργειας γίνονται ολοένα και πιο πολύπλοκα, αναπτύσσονται προηγμένες στρατηγικές LFC, συμπεριλαμβανομένων προσαρμοστικών και έξυπνων μεθόδων ελέγχου, για την ενίσχυση της ανθεκτικότητας έναντι δυναμικών διαταραχών και κυβερνοαπειλών.

2.2. Ελεγκτές για συστήματα LFC

Η σχεδίαση αποτελεσματικών ελεγκτών για συστήματα LFC αποτελεί κρίσιμο στοιχείο για τη διατήρηση της συχνότητας του δικτύου εντός αποδεκτών ορίων και την αποφυγή ζημιών σε εξοπλισμό ή μαζικών διακοπών. Στη βιβλιογραφία έχουν προταθεί ποικίλες μέθοδοι ελέγχου, οι οποίες διαφέρουν ως προς την πολυπλοκότητα, την ανθεκτικότητα και την ικανότητα αντιμετώπισης μη γραμμικών φαινομένων.

Αναλογικός-Ολοκληρωτικός-Διαφορικός (PID) Ελεγκτής: Ο PID είναι ο πιο διαδεδομένος ελεγκτής στη βιομηχανία λόγω της απλότητάς του και της ευελιξίας στη ρύθμιση των παραμέτρων K_P, K_I και K_D. Στα συστήματα LFC, ο αναλογικός όρος (K_P) καθορίζει την ένταση της αντίδρασης του ελεγκτή σε μια απόκλιση της συχνότητας , ο ολοκληρωτικός όρος (K_I) χρησιμοποιείται για την εξάλειψη του μόνιμου σφάλματος συχνότητας (steady-state error), ενώ ο διαφορικός (K_D) αντιδρά στον ρυθμό μεταβολής του σφάλματος, μειώνοντας την υπερύψωση και βελτιώνοντας την ευστάθεια. Ωστόσο, ο PID εμφανίζει περιορισμούς σε συστήματα υψηλής αδράνειας ή σε παρουσία σημαντικών χρονικών καθυστερήσεων επικοινωνίας [55].

Βέλτιστος Έλεγχος (Optimal Control): Ο βέλτιστος έλεγχος επιδιώκει να ελαχιστοποιήσει ένα δείκτη κόστους που συνδυάζει απόκλιση συχνότητας, ενεργειακή κατανάλωση και φθορά εξοπλισμού [56]. Ο Γραμμικός Τετραγωνικός Ρυθμιστής (LQR), μια ειδική περίπτωση βέλτιστου ελέγχου, βασίζεται σε γραμμικά μοντέλα κατάστασης και τετραγωνικές συναρτήσεις κόστους. Παρά την ακρίβειά του, η αποτελεσματικότητά του εξαρτάται από την ακριβή μοντελοποίηση του συστήματος και δεν λαμβάνει υπόψη μη γραμμικές συμπεριφορές ή εξωτερικές διαταραχές [57].

Εύρωστος Έλεγχος (Robust Control) :Σε αντίθεση με τον LQR, ο εύρωστος έλεγχος σχεδιάζεται να διατηρεί την σταθερότητα και απόδοση ακόμη και σε συνθήκες αβεβαιότητας μοντελοποίησης ή μεταβλητών λειτουργικών συνθηκών. Είναι ιδανικός για συστήματα LFC με υψηλή διείσδυση ανανεώσιμων πηγών, όπου οι διακυμάνσεις φορτίου είναι απρόβλεπτες. Ωστόσο, η υλοποίησή του απαιτεί πολύπλοκα μαθηματικά εργαλεία και συχνά οδηγεί σε συντηρητικές ρυθμίσεις [58].

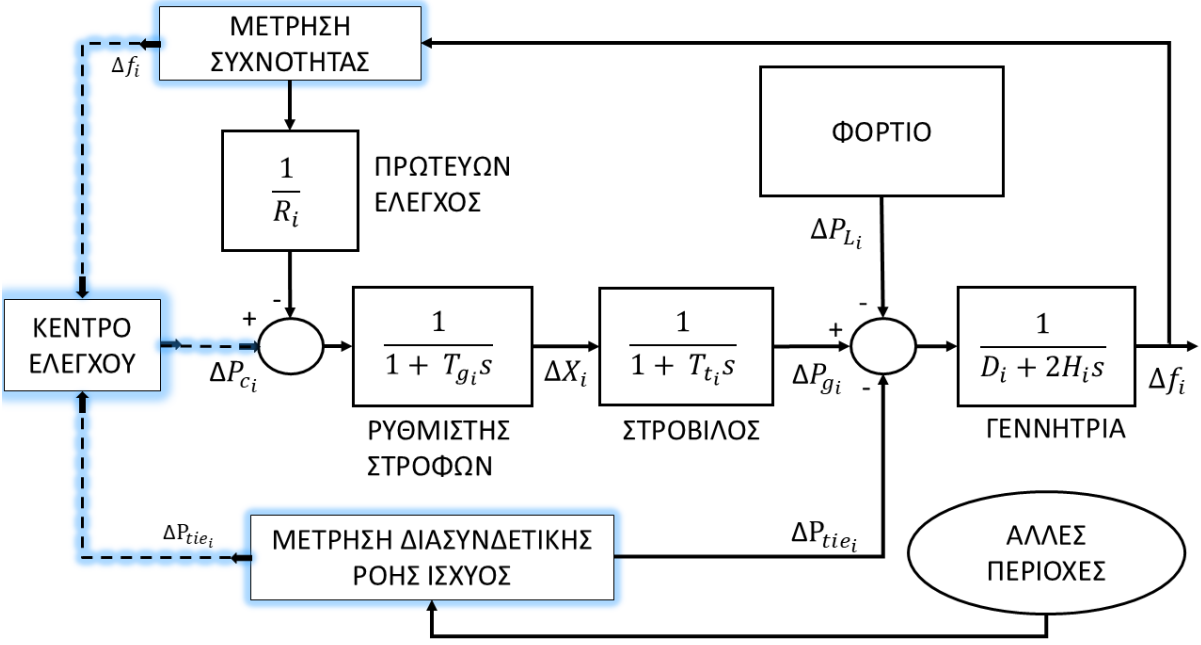
Έλεγχος με Ασαφή Λογική (Fuzzy Logic Control): Η ασαφής λογική βασίζεται σε κανόνες τύπου Αν-Τότε (π.χ. ΑΝ η συχνότητα μειώνεται γρήγορα, ΤΟΤΕ αυξάνουμε δραστικά την παραγωγή), οι οποίοι μπορούν να καλύψουν μη γραμμικές συμπεριφορές. Είναι ιδιαίτερα αποτελεσματική σε συστήματα με ασαφείς μετρήσεις ή πολλαπλούς στόχους ελέγχου. Η έλλειψη αυστηρής μαθηματικής βάσης και η δυσκολία βελτιστοποίησης των κανόνων αποτελούν τα κύρια μειονεκτήματα της μεθόδου [59].

Έλεγχος με Πρόβλεψη Μοντέλου (Model Predictive Control - MPC): Ο MPC χρησιμοποιεί ένα δυναμικό μοντέλο του συστήματος για να προβλέψει τη μελλοντική του συμπεριφορά και να βελτιστοποιήσει μια ακολουθία ενεργειών ελέγχου σε ένα πεπερασμένο ορίζοντα. Στα συστήματα LFC, αυτό επιτρέπει την αναμονή διακυμάνσεων φορτίου και την εφαρμογή προληπτικών διορθώσεων. Η υπολογιστική πολυπλοκότητα και η ανάγκη για ακριβή μοντέλα περιορίζουν την εφαρμογή του σε συστήματα υψηλών ταχυτήτων [60].

2.3. Μαθηματική Μοντελοποίηση LFC

Η γενική δομή του διαγράμματος των μπλοκ για την i-οστή ($i = 1, 2, \ldots, N$) περιοχή ελέγχου ενός συστήματος ισχύος N περιοχών με σύστημα LFC απεικονίζεται στο Σχήμα 3. Για την ανάλυση του συστήματος LFC, οι ταλαντώσεις μεταξύ των μηχανών και η απόδοση του συστήματος μεταφοράς αγνοούνται. Επίσης, για τις ανάγκες της εργασίας, αγνοείται η ύπαρξη ανανεώσιμων πηγών ενέργειας και η σύνδεση HVDC μεταξύ των περιοχών. Το μοντέλο δεν περιλαμβάνει και ορισμένες λειτουργίες υλοποίησης, όπως το φιλτράρισμα του σφάλματος ελέγχου περιοχής (ACE) [61].

Επομένως, η σύνθετη δυναμική απόδοση όλων των γεννητριών σε μια περιοχή συστήματος ισχύος μπορεί να αναπαρασταθεί από ένα ισοδύναμο μοντέλο γεννήτριας που αποτελείται από μια σύγχρονη μηχανή, έναν στρόβιλο (turbine) και έναν ρυθμιστή στροφών (governor).



Σχήμα 3: Διάγραμμα μπλοκ της i-οστής περιοχής ενός συστήματος LFC

Με βάση το Σχήμα 3, οι δυναμικές εξισώσεις του συστήματος για κάθε περιοχή i μπορούν να εξαχθούν. Οι δυναμική εξίσωση της γεννήτριας είναι η εξής:

$$\Delta \dot{f}_i = \frac{1}{2\,H_i}(\Delta P_{g_i} - D_i \Delta f_i - \Delta P_{tie_i} - \Delta P_{L_i})$$

όπου Δf_i είναι η απόκλιση συχνότητας, D_i και H_i αντιπροσωπεύουν τις σταθερές απόσβεσης και αδράνειας, αντίστοιχα. ΔP_{tie_i} και ΔP_{L_i} είναι η απόκλιση της ανταλλαγής ισχύος διασύνδεσης και η διαταραχή φορτίου αντίστοιχα.

Οι δυναμικές εξισώσεις για το ρυθμιστή στροφών και το στρόβιλο είναι:

$$\dot{\Delta P_{g_i}} = \frac{1}{T_{t_i}}(\Delta X_i - \Delta P_{g_i})$$

$$\dot{\Delta X_i} = \frac{1}{T_{g_i}}(\Delta P_{c_i} - \Delta X_i - \frac{\Delta f_i}{R_i})$$

όπου ΔP_{g_i} και ΔX_i είναι η απόκλιση ισχύος εξόδου της τουρμπίνας και η απόκλιση θέσης της βαλβίδας του ρυθμιστή στροφών, αντίστοιχα. Με T_{t_i} συμβολίζεται η σταθερά χρόνου της τουρμπίνας και T_{g_i}η σταθερά χρόνου του ρυθμιστή. R_i είναι η σταθερά πτώσης (droop), και ΔP_{c_i} είναι η απόκλιση της θέσης του ρυθμιστή ταχύτητας, που αντιπροσωπεύει το σήμα δευτερεύοντος ελέγχου.

Για ένα σύστημα ισχύος με N περιοχές, η συνολική απόκλιση ισχύος διασύνδεσης της i-οστής περιοχής δίνεται από:

$$\Delta P_{tie_i} = \Delta P_{ac_i}$$

όπου $\Delta Paci$ είναι η απόκλιση ανταλλαγής ισχύος της διασύνδεσης AC. Εάν η διασύνδεση AC δεν είναι εξοπλισμένη με έναν Μετατοπιστή Φάσης Ελεγχόμενου με Θυρίστορ (Thyristor Controlled Phase Shifter - TCPS) [62], το ΔP_{ac_i} δίνεται από:

$$\Delta P_{ac_i} = \frac{2\,\pi}{s} \sum_{j=1, j\neq i}^{N} T_{s_{ij}}\,(\Delta f_i - \Delta f_j)$$

όπου $T_{s_{ij}}$ είναι ο συντελεστής συγχρονισμού μεταξύ των περιοχών i και j, $j = 1,2,\ldots,N$, $j \neq i$.

Εάν μια γεννήτρια δεν συμμετέχει στον LFC, το σημείο ρύθμισης (setpoint) του ρυθμιστή της, ΔP_{c_i}, είναι μια σταθερή και γνωστή είσοδος στο σύστημα. Διαφορετικά, ο ρυθμιστής στροφών της γεννήτριας ρυθμίζεται από τον LFC για να παρακολουθεί τις αλλαγές του φορτίου και να ρυθμίζει την ισχύ των διασυνδετικών ροών και τη συχνότητα [71].

Ο ελεγκτής του LFC λαμβάνει ως είσοδο το Σφάλμα Ελέγχου Περιοχής (Area Control Error - ACE). Για κάθε περιοχή ελέγχου i, το ACE_i ορίζεται ως:

$$ACE_i = \Delta P_{tie_i} + \beta_i\,\Delta f_i$$

όπου βi είναι ο παράγοντας μεροληψίας (bias) της συχνότητας. Το σήμα κλειστού βρόχου βασισμένο σε ελεγκτή PID προκύπτει από:

$$\Delta P_{c_i} = -(K_{P_i} ACE_i + K_{I_i} \int ACE_i\, dt + K_{D_i} d\,\frac{dACE_i}{dt})$$

Όπου K_{P_i} το αναλογικό κέρδος (proportional gain),K_{I_i} το ολοκληρωτικό κέρδος (integral gain) και K_{D_i} το διαφορικό κέρδος (derivative gain).

Το τοπικό διάνυσμα κατάστασης που χρησιμοποιείται στην αναπαράσταση του χώρου κατάστασης της i-οστής περιοχής, δηλαδή $x_i \in R^O$, όπου O είναι η δυναμική τάξη της i-οστής περιοχής, ορίζεται ως:

$$\Delta x_i = [\Delta f_i, \quad \Delta P_{g_i}, \quad \Delta X_i, \quad ACE_i, \quad \int ACE_i, \ {}^{dACE_i}/_{dt}, \quad \Delta P_{ac_i}]^T$$

Σύμφωνα με τις προηγούμενες σχέσεις, η διαφορική-αλγεβρική συμπαγής μορφή για ολόκληρο το σύστημα που αποτελείται από N περιοχές είναι:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Ed(t) \\ y(t) = Cx(t) \end{cases}$$

όπου $x(t) = [x_1 \ x_2 \ \cdots \ x_N]^T \in R_n$ το καθολικό διάνυσμα κατάστασης ($O \cdot N = n$),

$u(t) = [\Delta P_{c_1} \ \Delta P_{c_2} \ \cdots \ \Delta P_{c_m}]^T \in R_m$ το διάνυσμα εισόδου

$d(t) = [\Delta P_{d_1} \ \Delta P_{d_2} \ \cdots \ \Delta P_{d_r}]^T \in R_r$ το διάνυσμα διαταραχών και

$y(t) \in R_p$ το διάνυσμα εξόδου.

Οι πίνακες $A \in R_{n\times n}$, $B \in R_{n\times m}$, $E \in R_{n\times r}$, $C \in R_{p\times n}$ θεωρούνται γνωστοί.

2.4. Προσομοίωση Αλλαγών Φορτίου σε Σύστημα LFC 2 Περιοχών

Το σύστημα LFC δύο περιοχών χρησιμοποιείται ευρέως για τη μελέτη της επίδρασης των επιθέσεων στον κυβερνοχώρο σε συστήματα LFC πολλαπλών περιοχών λόγω της απλότητας του. Με σκοπό την προσομοίωση της λειτουργίας ενός συμβατικού ελεγκτή για το LFC, σχεδιάστηκε σε περιβάλλον MATLAB/Simulink ένα σύστημα δύο περιοχών ονομαστικής συχνότητας 60Hz.

Τα στοιχεία του συστήματος είναι τα εξής:

Βάση Ισχύος = 1000MVA, R_3= 0.5

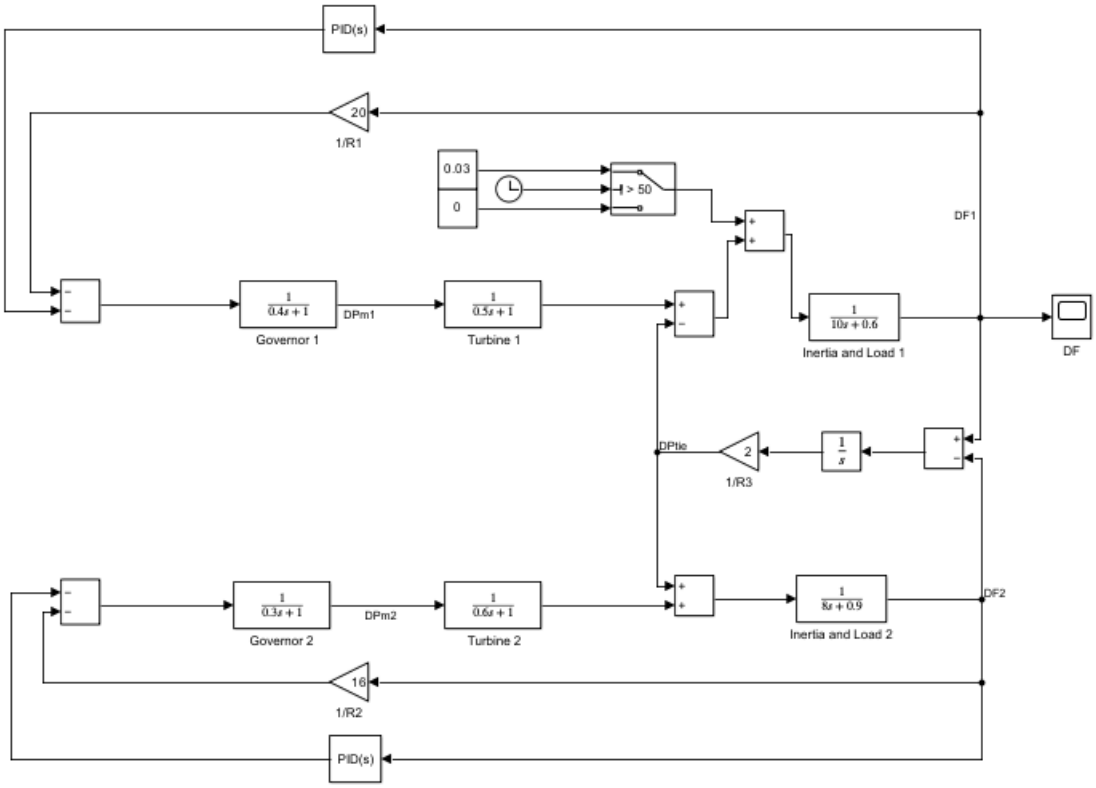
Περιοχή 1: R_1=0.05, D_1 =0.6, H_1=5, T_{g_1} = 0.2, T_{t_1}= 0.5

Περιοχή 2: R_2=0.0625, D_2=0.9, H_2=4, T_{g_2}=0.3, T_{t_2}=0,6

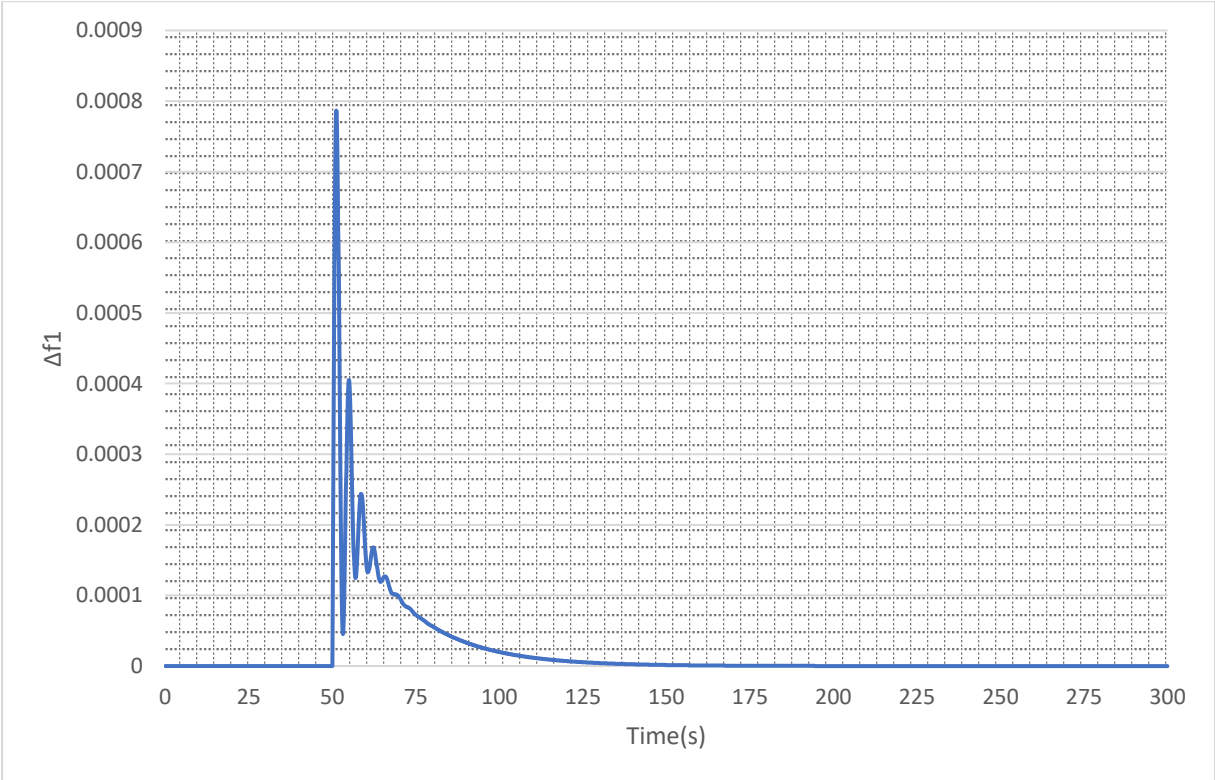
Στο Σχήμα 4 φαίνεται η διάταξη που σχεδιάστηκε.

Προσομοιώθηκαν τρεις περιπτώσεις απότομης αλλαγής φορτίου: 0.01 α.μ., 0.03 α.μ. και 0.06 α.μ.. Οι αλλαγές συμβαίνουν το 50ό δευτερόλεπτο κάθε προσομοίωσης. Συγκεκριμένα, στο ΔP_{L_i} προστέθηκε βηματική διαταραχή πλάτος αντίστοιχου με τις αλλαγές φορτίων προς πειραματισμό.

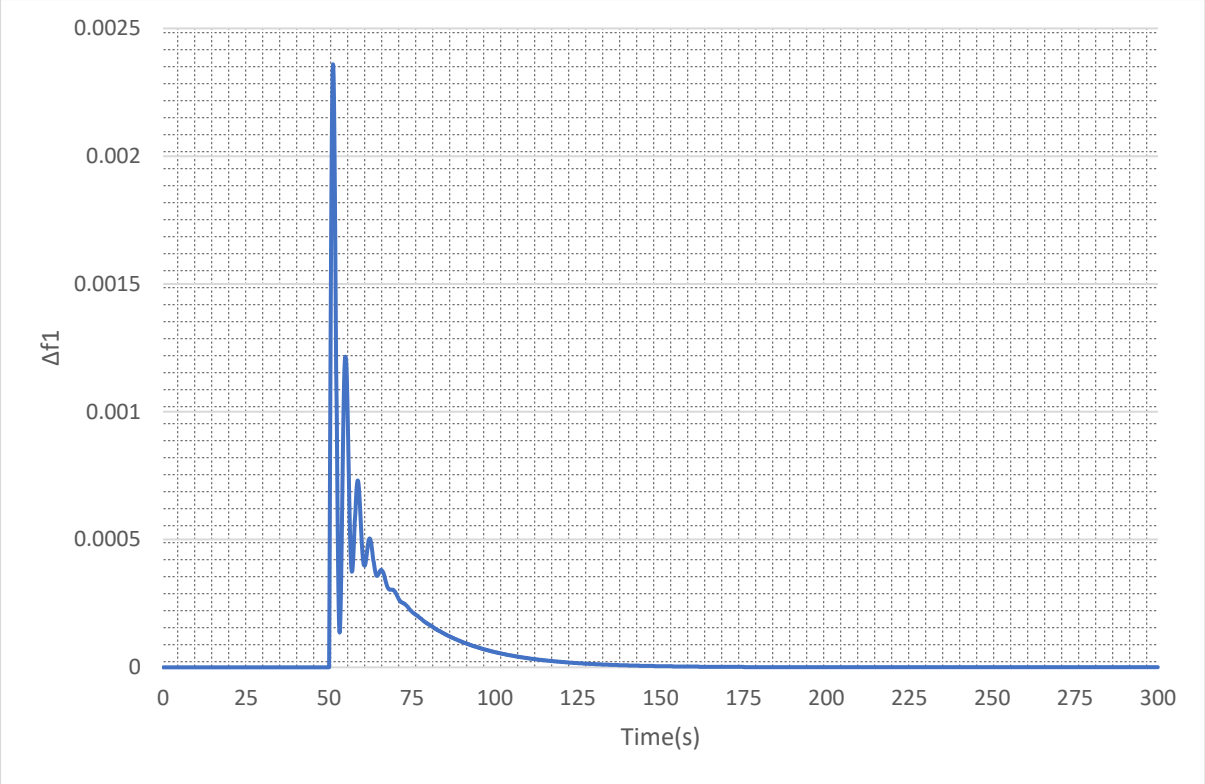
Τα αποτελέσματα των πειραμάτων φαίνονται στα Σχήματα 5-7 και θα αποτελέσουν μέτρο σύγκρισης με την προτεινόμενη μέθοδο στη συνέχεια της εργασίας.



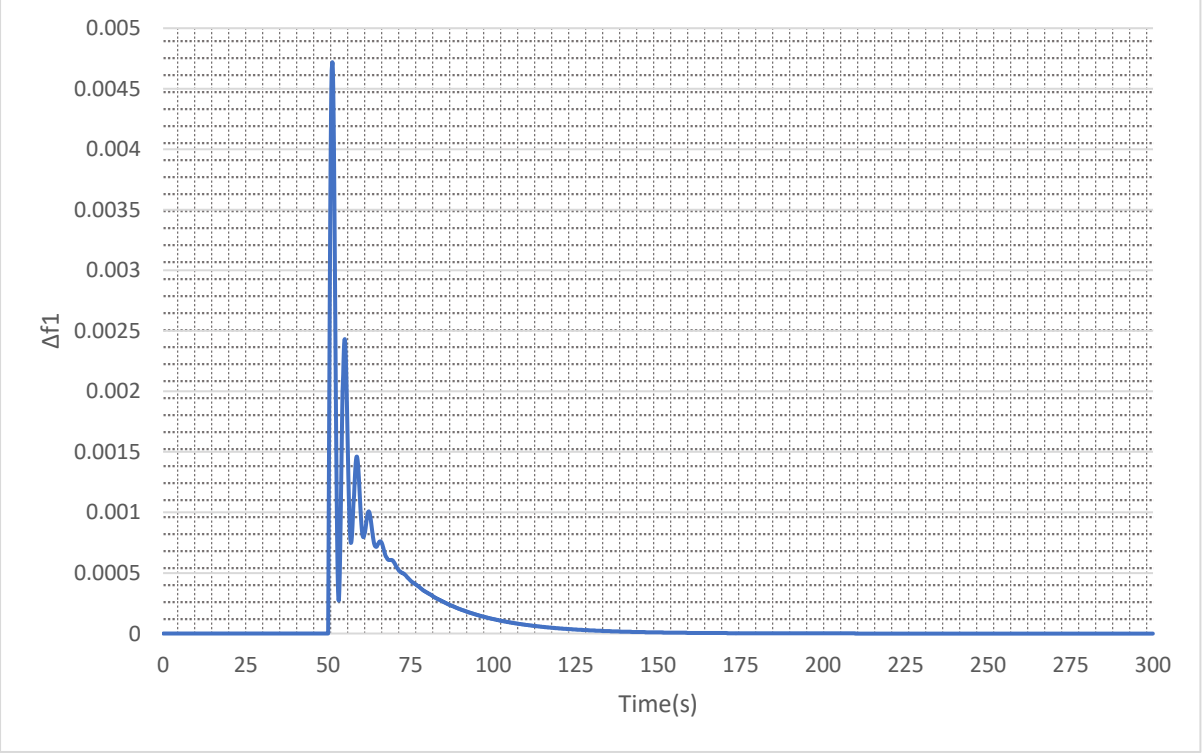
Σχήμα 4: Προσομοίωση συστήματος LFC δύο περιοχών σε περιβάλλον MATLAB/Simulink



Σχήμα 5: Απόκριση συχνότητας συστήματος LFC με συμβατικό ελεγκτή σε περίπτωση μεταβολής φορτίου 0.01 α.μ.



Σχήμα 6: Απόκριση συχνότητας συστήματος LFC με συμβατικό ελεγκτή σε περίπτωση μεταβολής φορτίου 0.03 α.μ.



Σχήμα 7: Απόκριση συχνότητας συστήματος LFC με συμβατικό ελεγκτή σε περίπτωση μεταβολής φορτίου 0.06 α.μ.

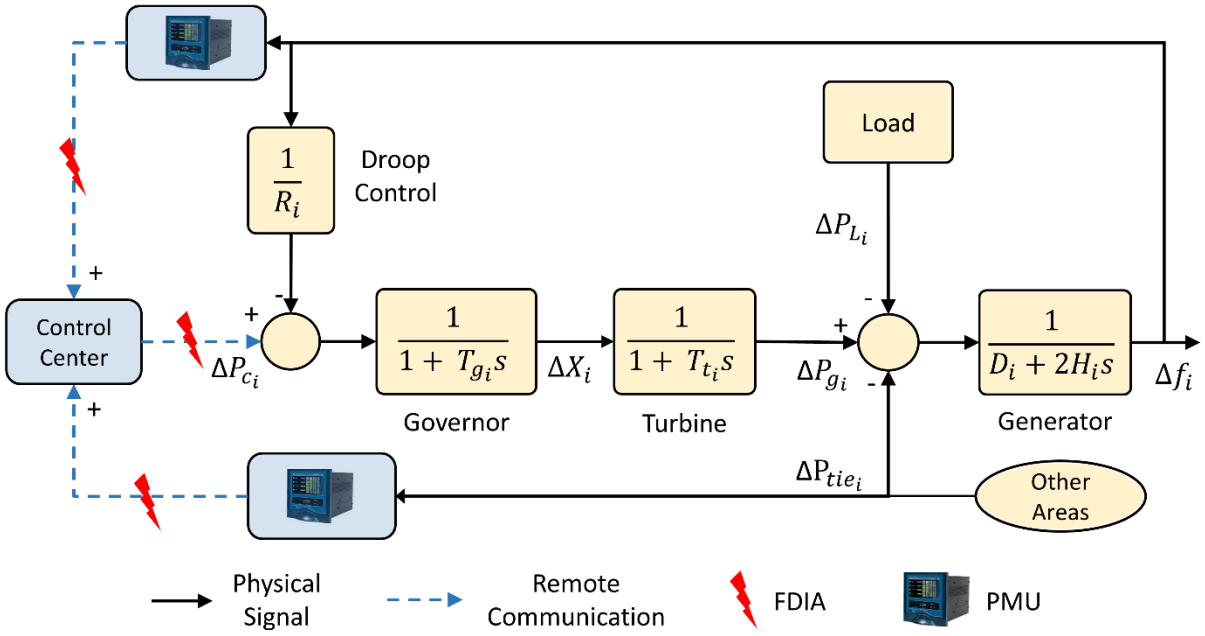
2.5. Ευπαθή Σημεία στο LFC

Η εξάρτηση από τη συλλογή και μεταφορά δεδομένων σε πραγματικό χρόνο καθιστά το σύστημα LFC ευάλωτο σε κυβερνοεπιθέσεις, οι οποίες μπορούν να διαταράξουν τη σταθερότητα του δικτύου ή ακόμη και να προκαλέσουν καθολικές διακοπές [63]. Στο Σχήμα 8 σημειώνονται τα ευπαθή σημεία της i-στης περιοχής ενός συστήματος LFC πολλών περιοχών. Συγκεκριμένα, τα κύρια σημεία εφαρμογής κακόβουλων σημάτων, όπου πιθανές επιθέσεις μπορούν να παρεμβαίνουν στη λειτουργία του συστήματος, είναι τα ακόλουθα:

Μέτρηση Συχνότητας (Κανάλι Επικοινωνίας PMU-Κέντρου Ελέγχου): Οι μονάδες μέτρησης φάσης (PMUs) αποστέλλουν δεδομένα συχνότητας από τους κόμβους του δικτύου προς το κέντρο ελέγχου. Μια επίθεση σε αυτό το κανάλι (π.χ. διακοπή σήματος, παραποίηση τιμών συχνότητας) μπορεί να οδηγήσει σε εσφαλμένες αποφάσεις ελέγχου, όπως μη αναγκαίες ρυθμίσεις παραγωγής ή υπερφόρτωση γεννητριών.

Μέτρηση Διασυνδετικής Ροής Ισχύος (Κανάλι Επικοινωνίας PMU-Κέντρου Ελέγχου): Σε συστήματα πολλών περιοχών, τα PMUs παρακολουθούν τη ροή ενέργειας μεταξύ διασυνδεδεμένων ζωνών. Κακόβουλη παρέμβαση (π.χ. εισαγωγή ψευδών δεδομένων ροής) μπορεί να διαταράξει την ισορροπία μεταξύ περιοχών, προκαλώντας διακυμάνσεις συχνότητας ή ακόμη και αποσυνδέσεις γραμμών.

Σήμα Ελέγχου (Κανάλι Επικοινωνίας Κέντρου Ελέγχου-Αθροιστή Ρυθμιστή Στροφών): Μετά την επεξεργασία των δεδομένων, το κέντρο ελέγχου στέλνει σήματα διόρθωσης στον αθροιστή που τροφοδοτεί τον ρυθμιστή στροφών. Μια επίθεση σε αυτό το στάδιο (π.χ. υπερβολική αύξηση/μείωση του σήματος ελέγχου) μπορεί να προκαλέσει υπερτάσεις συχνότητας, ζημιές σε εξοπλισμό ή αστάθεια σε ολόκληρη την περιοχή.



Σχήμα 8: Ευπαθή σημεία συστήματος LFC

2.6. Κατηγορίες Κυβερνοεπιθέσεων που απειλούν το LFC

Επιθέσεις Άρνησης Υπηρεσίας (Denial of Service - DoS): Οι επιθέσεις Άρνησης Υπηρεσίας στοχεύουν στη διακοπή της επικοινωνίας μεταξύ των στοιχείων του συστήματος LFC (π.χ. αισθητήρες, ελεγκτές και ενεργοποιητές) υπερφορτώνοντας το δίκτυο με κίνηση ή εκμεταλλευόμενες ευπάθειες στα πρωτόκολλα επικοινωνίας. Μια επιτυχής επίθεση DoS μπορεί να καθυστερήσει ή να αποκλείσει κρίσιμα σήματα ελέγχου συχνότητας, αποτρέποντας το σύστημα LFC από την αντίδραση σε ανισορροπίες σε πραγματικό χρόνο. Αυτό μπορεί να οδηγήσει σε παρατεταμένες διακυμάνσεις συχνότητας και πιθανές αλυσιδωτές αποτυχίες [72].

Επιθέσεις Έγχυσης Ψευδών Δεδομένων (False Data Injection Attacks - FDIAs): Οι επιθέσεις έγχυσης ψευδών δεδομένων περιλαμβάνουν τη χειραγώγηση των δεδομένων αισθητήρων ή των σημάτων ελέγχου που αποστέλλονται στο σύστημα LFC [4]. Οι επιτιθέμενοι εισάγουν ψευδείς μετρήσεις στο σύστημα, προκαλώντας λανθασμένες λήψεις αποφάσεων, σφάλματα στην οικονομική διανομή ενέργειας ή ακόμη και ευρείας κλίμακας αποτυχίες συστημάτων [6]. Για παράδειγμα, μια FDIA θα μπορούσε να παραπλανήσει το σύστημα LFC να υπερεκτιμήσει τη ζήτηση φορτίου, προκαλώντας περιττές αυξήσεις στην παραγωγή και ταλαντώσεις συχνότητας. Οι FDIAs είναι ιδιαίτερα επικίνδυνες γιατί μπορούν να παρακάμψουν τους παραδοσιακούς μηχανισμούς ανίχνευσης ανωμαλιών [5],[73].

Επιθέσεις Αναπαραγωγής (Replay Attacks): Σε μια επίθεση αναπαραγωγής, ο επιτιθέμενος παρεμβάλλει και αναμεταδίδει έγκυρα σήματα ελέγχου ή δεδομένα αισθητήρων στο σύστημα LFC. Για παράδειγμα, ένας επιτιθέμενος θα μπορούσε να καταγράψει μια προηγούμενη μέτρηση συχνότητας και να την αναπαράγει στον ελεγκτή, καλύπτοντας την πραγματική κατάσταση του συστήματος. Αυτό μπορεί να προκαλέσει τη λειτουργία του συστήματος LFC με παλιά ή λανθασμένα δεδομένα, οδηγώντας σε ακατάλληλες ενέργειες ελέγχου και αστάθεια συχνότητας [74].

Επιθέσεις Ενδιάμεσου Ανθρώπου (Man in the Middle - MitM) : Οι Επιθέσεις Ενδιάμεσου Ανθρώπου συμβαίνουν όταν ένας επιτιθέμενος με φυσική παρουσία παρεμβάλλει και αλλοιώνει την επικοινωνία μεταξύ των στοιχείων του συστήματος LFC. Για παράδειγμα, ένας επιτιθέμενος θα μπορούσε να τροποποιήσει τα σήματα ελέγχου που αποστέλλονται από τον ελεγκτή LFC στις γεννήτριες, προκαλώντας τις να αντιδράσουν ακατάλληλα σε διακυμάνσεις συχνότητας. Οι επιθέσεις MitM μπορούν επίσης να χρησιμοποιηθούν για την εισαγωγή κακόβουλων εντολών ή την εξαγωγή ευαίσθητων πληροφοριών για το σύστημα LFC [75].

Επιθέσεις Καθυστέρησης Χρόνου (Time-Delay Attacks): Οι επιθέσεις χρονοκαθυστέρησης εισάγουν σκόπιμες καθυστερήσεις στη μετάδοση σημάτων ελέγχου ή δεδομένων αισθητήρων στο σύστημα LFC. Ακόμα και μικρές καθυστερήσεις μπορεί να υποβαθμίσουν την απόδοση των αλγορίθμων ελέγχου συχνότητας, οι οποίοι εξαρτώνται από δεδομένα σε πραγματικό χρόνο για ακριβή λήψη αποφάσεων. Παρατεταμένες καθυστερήσεις μπορεί να οδηγήσουν σε αστάθεια, καθώς το σύστημα LFC ενδέχεται να ενεργήσει για να διορθώσει διακυμάνσεις συχνότητας που έχουν ήδη διαχειριστεί [76].

Επιθέσεις Κακόβουλου Λογισμικού και Απαγωγής Δεδομένων (Ransomware): Το κακόβουλο λογισμικό ή το ransomware μπορεί να μολύνει τα λογισμικά και τα υλικά εξαρτήματα των συστημάτων LFC, διαταράσσοντας τη λειτουργία τους ή καθιστώντας τα αδρανή. Για παράδειγμα, το ransomware θα μπορούσε να αποκλείσει τους χειριστές από τη

διεπαφή ελέγχου LFC, εμποδίζοντάς τους να ανταποκριθούν σε διακυμάνσεις συχνότητας. Το κακόβουλο λογισμικό θα μπορούσε επίσης να τροποποιήσει τους αλγορίθμους ελέγχου ή να διαφθείρει τα δεδομένα των αισθητήρων, οδηγώντας σε λανθασμένες διορθώσεις συχνότητας [77].

Εκμεταλλεύσεις Μηδενικής Ημέρας (Zero-Day Exploits): Οι εκμεταλλεύσεις zero-day στοχεύουν σε ευπάθειες στο λογισμικό ή το υλικό του συστήματος LFC, που είναι άγνωστες μέχρι εκείνη τη στιγμή. Αυτές οι επιθέσεις είναι ιδιαίτερα επικίνδυνες γιατί μπορούν να παρακάμψουν τα υπάρχοντα μέτρα ασφαλείας. Για παράδειγμα, ένας επιτιθέμενος θα μπορούσε να εκμεταλλευτεί μια ευπάθεια στο λογισμικό του ελεγκτή LFC για να αποκτήσει μη εξουσιοδοτημένη πρόσβαση και να χειριστεί τα σήματα ελέγχου συχνότητας [78].

Απειλές Εσωτερικών Χρηστών (Insider Threats): Οι απειλές εσωτερικών χρηστών περιλαμβάνουν κακόβουλες ενέργειες από άτομα με εξουσιοδοτημένη πρόσβαση στο σύστημα LFC. Για παράδειγμα, ένας δυσαρεστημένος υπάλληλος θα μπορούσε σκόπιμα να τροποποιήσει τις παραμέτρους ελέγχου ή να απενεργοποιήσει τις λειτουργίες κανονισμού συχνότητας, οδηγώντας σε αστάθεια. Οι απειλές εσωτερικών χρηστών είναι δύσκολο να ανιχνευθούν γιατί ο επιτιθέμενος έχει νόμιμη πρόσβαση και γνώση του συστήματος [79].

2.7. Κυβερνοεπιθέσεις τύπου FDI στο LFC

Οι επιθέσεις FDI εφαρμόζονται στα κανάλια μέτρησης και ελέγχου του συστήματος LFC με τη μορφή διανυσμάτων επίθεσης εισόδου που έχουν διαμορφωθεί χρησιμοποιώντας στρατηγικές διαφθοράς δεδομένων ή πρότυπα επίθεσης. Επεκτείνοντας την εξίσωση στο χώρο κατάστασης που παρουσιάστηκε στην παράγραφο 2.3, οι επιθέσεις μπορούν να εισαχθούν στο μοντέλο ως εξής:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Ed(t) \\ y(t) = Cx(t) + Da(t) \end{cases}$$

Όπου $a(t) \in R_q$ είναι το διάνυσμα επίθεσης και ο πίνακας $D \in R_{p\times q} : p \geq q + r$ θεωρείται γνωστός [80].

Στην περίπτωση των επιθέσεων τύπου FDI, το $a(t)$ είναι μια τυπική συνάρτηση επίθεσης [95]:

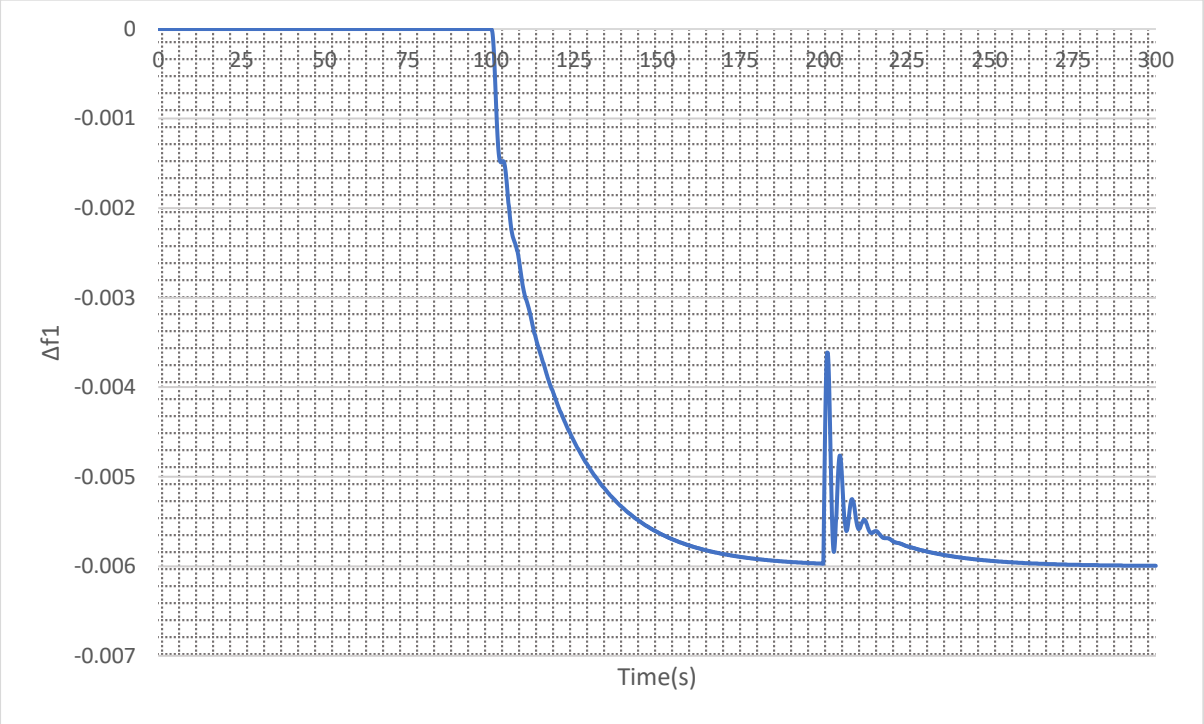
- Βηματική συνάρτηση: $a(t) = a\,u(t - t_0)$, όπου $u(t)$είναι η μοναδιαία βηματική συνάρτηση. Προκαλεί αιφνίδιες μεταβολές στη συχνότητα, μειώνοντας την ικανότητα του ελεγκτή για ομαλή προσαρμογή.
- Συνάρτηση κλίσης: $a(t) = \beta \cdot t$, που οδηγεί σε προοδευτική απόκλιση συχνότητας, δυσκολεύοντας την ανίχνευση λόγω της "αργής" εξέλιξης της επίθεσης.
- Κυματομορφές με συγκεκριμένη συχνότητα: Για ανάδυση ταλαντώσεων που ενισχύουν φαινόμενα συντονισμού.

Ο συνδυασμός του a(t) με τον πίνακα D επιτρέπει την επιλεκτική διείσδυση σε συγκεκριμένα σημεία του συστήματος.

2.8. Προσομοίωση FDIA σε Σύστημα LFC Δύο Περιοχών

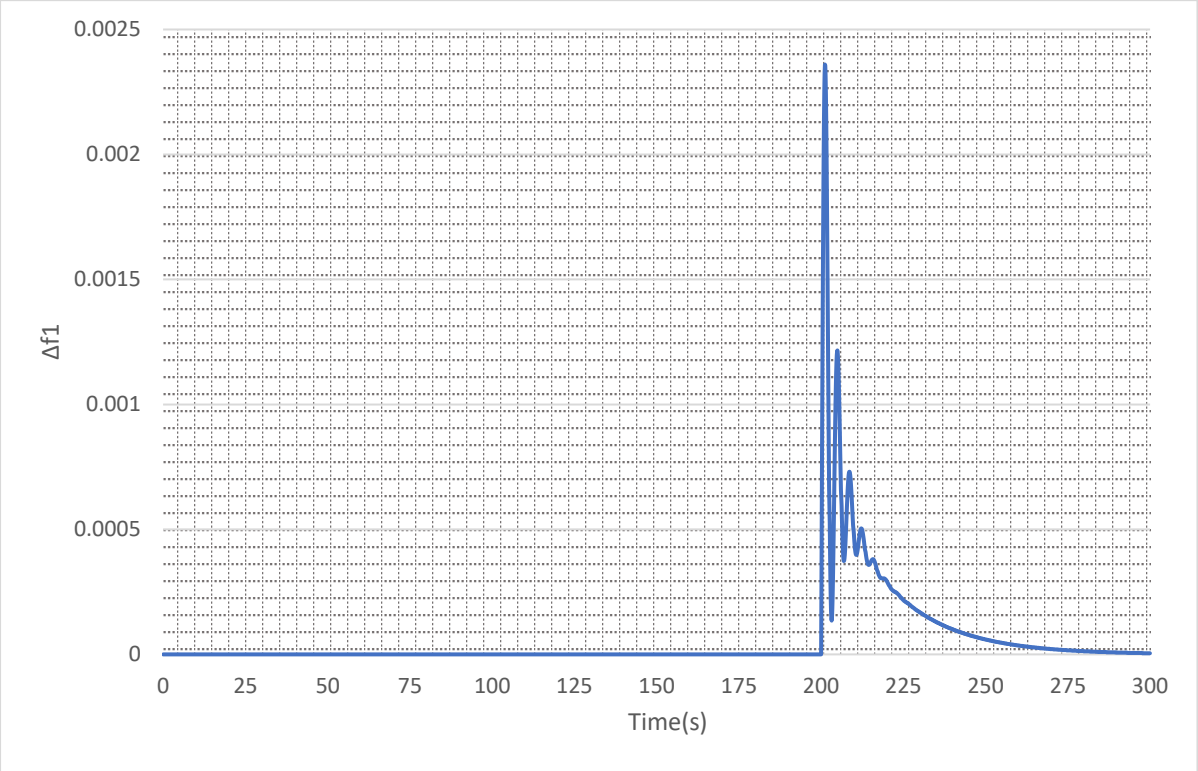
Στη συνέχεια παρουσιάζεται η επίδραση επιθέσεων ψευδούς έγχυσης δεδομένων σε ένα σύστημα ελέγχου φορτίου συχνότητας δύο περιοχών. Τα στοιχεία του συστήματος είναι τα ίδια με το σύστημα που χρησιμοποιήθηκε για την προσομοίωση αλλαγών φορτίου σε προηγούμενη παράγραφο. Ο δευτερεύων έλεγχος επιτυγχάνεται μέσω ενός PID ελεγκτή. Το σύστημα σχεδιάστηκε και προσομοιώθηκε σε περιβάλλον MATLAB/Simulink.

Αρχικά, στη μέτρηση της συχνότητας της περιοχής 1 εγχύθηκε βηματικό προσθετικό σήμα μεγέθους 0.01 α.μ. κατά το 100ό δευτερόλεπτο της προσομοίωσης . Στη δεύτερη περίπτωση, το ίδιο σήμα εφαρμόστηκε την ίδια χρονική στιγμή στη μέτρηση της διασυνδετικής ροής ισχύος. Τέλος, προσομοιώθηκε συγχρονισμένη επίθεση και στις δύο μετρήσεις με σήματα επίθεσης πλάτους 0.03αμ στη μέτρηση συχνότητας και 0.01αμ στη μέτρηση διασυνδετικής ροής ισχύος. Στα παρακάτω διαγράμματα φαίνεται η απόκριση της συχνότητας της περιοχής 1 του συστήματος στις FDIAs.

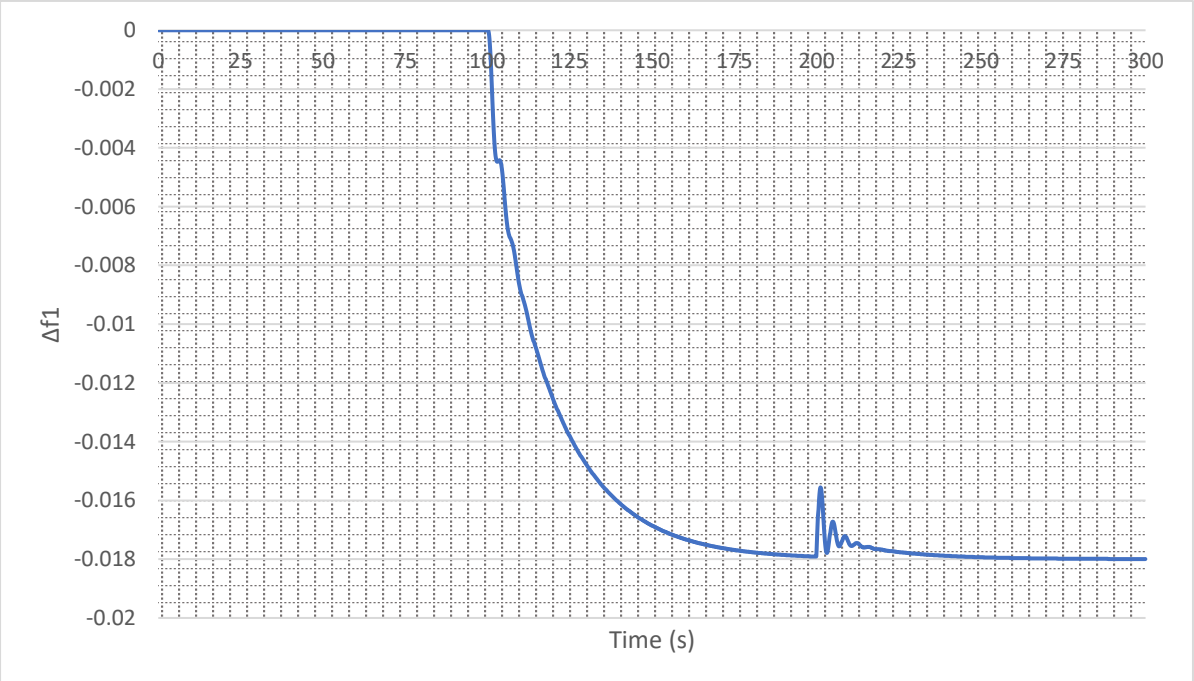


Σχήμα 9: Απόκριση συχνότητας συστήματος LFC με συμβατικό ελεγκτή σε FDIA στη μέτρηση συχνότητας

Η προσομοίωση αναδεικνύει την επικινδυνότητα των επιθέσεων έγχυσης ψευδών δεδομένων στα ευπαθή σημεία του συστήματος ελέγχου φορτίου συχνότητας. Συνεπώς, επιβεβαιώνεται και πειραματικά η ανάγκη για εντοπισμό και αντιμετώπιση αντίστοιχων επιθέσεων.



Σχήμα 10: Απόκριση συχνότητας συστήματος LFC με συμβατικό ελεγκτή σε FDIA στη μέτρηση ισχύος διασύνδεσης



Σχήμα 11: Απόκριση συχνότητας συστήματος LFC με συμβατικό ελεγκτή σε ταυτόχρονες FDIA στη μέτρηση συχνότητας και τη μέτρηση ισχύος διασύνδεσης

2.9. Μέθοδοι Άμυνας και Ανθεκτικοί Ελεγκτές Απέναντι σε Κυβερνοεπιθέσεις στο LFC

Για να μετριάσουν τους κινδύνους από τις κυβερνοεπιθέσεις στο LFC, ερευνητές και επαγγελματίες έχουν προτείνει διάφορους μηχανισμούς άμυνας και στρατηγικές ελέγχου ανθεκτικές σε επιθέσεις. Στη συνέχεια παρατίθενται μερικές από τις σημαντικότερες προσεγγίσεις για την προστασία του LFC έναντι κυβερνοεπιθέσεων.

2.9.1. Μέθοδοι Άμυνας απέναντι σε Κυβερνοεπιθέσεις στο LFC

2.9.1.1. Συστήματα Ανίχνευσης Ανωμαλιών (Anomaly Detection Systems)

Συστήματα Ανίχνευσης Εισβολών (Intrusion Detection Systems - IDS): Τα IDS μπορούν να παρακολουθούν τα δίκτυα επικοινωνίας για ασυνήθιστα μοτίβα ή αποκλίσεις από την κανονική συμπεριφορά, όπως FDIAs ή επιθέσεις άρνησης υπηρεσίας. Τα IDS που βασίζονται σε τεχνικές μηχανικής μάθησης μπορούν να προσαρμοστούν σε εξελισσόμενες στρατηγικές επιθέσεων [81].

Ανίχνευση Βασισμένη σε Στατιστική και Μοντέλα (Statistical and Model-Based Detection): Τεχνικές όπως τα φίλτρα Kalman και μέθοδοι που βασίζονται σε υπολείμματα (residual-based methods) μπορούν να ανιχνεύσουν αποκλίσεις μεταξύ των αναμενόμενων και των παρατηρούμενων καταστάσεων του συστήματος, εντοπίζοντας πιθανές επιθέσεις [82].

2.9.1.2. Ασφαλή Πρωτόκολλα Επικοινωνίας (Secure Communication Protocols)

Κρυπτογράφηση και Πιστοποίηση (Encryption and Authentication): Η εφαρμογή κρυπτογραφικών τεχνικών (π.χ. AES, RSA) και ασφαλών πρωτοκόλλων επικοινωνίας (π.χ. TLS) μπορεί να αποτρέψει την μη εξουσιοδοτημένη πρόσβαση και τον χειρισμό δεδομένων [83].

Επαλήθευση Χρονικά Στοιχειοθετημένων Δεδομένων (Time-Stamped Data Verification): Η χρήση χρονικά συγχρονισμένων δεδομένων με ψηφιακές υπογραφές διασφαλίζει την ακεραιότητα και την αυθεντικότητα των δεδομένων [84].

2.9.1.3. Εφεδρεία και Ποικιλομορφία (Redundancy and Diversity)

Επικοινωνία μέσω Πολλαπλών Καναλιών (Multi-Channel Communication): Η χρήση υπεράριθμων καναλιών επικοινωνίας μπορεί να περιορίσει την επίδραση επιθέσεων άρνησης υπηρεσίας (DoS) ή τον χειρισμό δεδομένων σε ένα μόνο κανάλι [85].

Ποικιλόμορφα Δίκτυα Αισθητήρων (Diverse Sensor Networks): Η χρήση ετερογενών αισθητήρων και πηγών δεδομένων μπορεί να παρέχει διασταυρωμένη επαλήθευση, καθιστώντας οι επιτιθέμενοι να θέσουν σε κίνδυνο ολόκληρο το σύστημα [86].

2.9.1.4 Ανθεκτικές Αρχιτεκτονικές Ελέγχου (Resilient Control Architectures)

Αποκεντρωμένος και Κατανεμημένος Έλεγχος (Decentralized and Distributed Control): Η κατανομή του ελέγχου σε πολλαπλούς κόμβους μειώνει τον κίνδυνο ενός μοναδικού σημείου αποτυχίας και περιορίζει την επίδραση τοπικών επιθέσεων [87].

Ιεραρχικές Δομές Ελέγχου (Hierarchical Control Structures): Η εφαρμογή αρχιτεκτονικών ελέγχου πολλαπλών στρωμάτων μπορεί να απομονώσει τις επιθέσεις σε συγκεκριμένα στρώματα του δικτύου, αποτρέποντάς τες από το να διαδοθούν σε ολόκληρο το σύστημα [88].

2.9.1.5. Παρακολούθηση και Απόκριση σε Πραγματικό Χρόνο (Real-Time Monitoring and Response)

Προσαρμοζόμενα Όρια Ανίχνευσης (Adaptive Thresholds): Η δυναμική προσαρμογή των ορίων ανίχνευσης με βάση τις συνθήκες του συστήματος μπορεί να βελτιώσει την ακρίβεια της ανίχνευσης επιθέσεων [89].

Σχέδια Απόκρισης σε Περιστατικά (Incident Response Plans): Προκαθορισμένα πρωτόκολλα απόκρισης μπορούν να βοηθήσουν τους χειριστές να μετριάσουν γρήγορα τις επιπτώσεις μιας επίθεσης και να αποκαταστήσουν την ευσταθή λειτουργία του συστήματος [90].

2.9.2. Ανθεκτικοί Απέναντι σε Κυβερνοεπιθέσεις Ελεγκτές (Cyberattack-Resilient Controllers)

2.9.2.1. Εύρωστες Στρατηγικές Ελέγχου (Robust Control Strategies)

Έλεγχος H–Απείρου (H-Infinity Control): Αυτή η προσέγγιση ελαχιστοποιεί την επίδραση διαταραχών και αβεβαιοτήτων, συμπεριλαμβανομένων των κυβερνοεπιθέσεων, βελτιστοποιώντας την πιο δυσμενή περίπτωση [91].

Έλεγχος Ολισθαίνουσας Κατάστασης (Sliding Mode Control - SMC): Ο SMC είναι εγγενώς ανθεκτικός στις αβεβαιότητες και μπορεί να διατηρήσει τη σταθερότητα του συστήματος ακόμη και υπό συνθήκες επίθεσης [8],[9],[92].

2.9.2.2. Ανθεκτικοί Ελεγκτές Βασισμένοι σε Παρατηρητές (Resilient Observer-Based Controllers)

Παρατηρητές Άγνωστων Εισόδων (Unknown Input Observers - UIOs): Οι UIOs μπορούν να εκτιμήσουν τις καταστάσεις του συστήματος ενώ αποζευγνύουν τις επιπτώσεις άγνωστων εισόδων, όπως οι κυβερνοεπιθέσεις [7],[93].

Παρατηρητές Ανεκτικοί σε Σφάλματα (Fault-Tolerant Observers): Οι ανεκτικοί σε σφάλματα παρατηρητές μπορούν να ανιχνεύσουν και να απομονώσουν σφάλματα ή επιθέσεις, διασφαλίζοντας ακριβή εκτίμηση της κατάστασης του συστήματος [94].

2.9.2.3. Προσεγγίσεις βασισμένες στη Θεωρία Παιγνίων (Game-Theory Approaches)

Παιχνίδια Μηδενικού Αθροίσματος (Zero-Sum Games): Η μοντελοποίηση της αλληλεπίδρασης μεταξύ επιτιθέμενων και αμυνόμενων ως ένα παιχνίδι μηδενικού αθροίσματος μπορεί να βοηθήσει στον σχεδιασμό βέλτιστων στρατηγικών άμυνας [95].

Παίγνια Stackelberg: Αυτά τα παίγνια λαμβάνουν υπόψη την ιεραρχική φύση των επιθέσεων και των αμυντικών μηχανισμών, επιτρέποντας τον σχεδιασμό προληπτικών αμυντικών στρατηγικών [96].

2.9.2.4. Υβριδικές Στρατηγικές Ελέγχου (Hybrid Control Strategies)

Έλεγχος που Ενεργοποιείται από Συμβάντα (Event-Triggered Control): Αυτή η προσέγγιση μειώνει το φόρτο επικοινωνίας και περιορίζει τις ευκαιρίες για επιθέσεις, μεταδίδοντας δεδομένα μόνο όταν είναι απαραίτητο [97].

Συστήματα Ελέγχου με Εναλλαγή (Switched Control Systems): Η εναλλαγή μεταξύ πολλαπλών τρόπων ελέγχου με βάση τις συνθήκες του συστήματος μπορεί να ενισχύσει την ανθεκτικότητα σε επιθέσεις [10], [98].

2.9.3. Σύγχρονες Τάσεις και Κατευθύνσεις

Όπως φαίνεται από το πλήθος των προτεινόμενων λύσεων στον τομέα της κυβερνοασφάλειας του LFC, είναι ένας τομέας που διαρκώς ερευνάται και εμπλουτίζεται. Παρά ταύτα, οι δυσκολίες παραμένουν. Συνεπώς, η έρευνα επικεντρώνεται σε νέες τακτικές και εργαλεία για τη διασφάλιση της ευστάθειας συχνότητας των έξυπνων δικτύων [11], [99], [100], [101], [102].

Με τη διαθεσιμότητα υπολογιστικής ισχύος σήμερα να έχει αυξηθεί θεαματικά σε σχέση με το πρόσφατο παρελθόν, καθίσταται δυνατή η αξιοποίηση της Ενισχυτικής Μάθησης (Reinforcement Learning - RL) και της Ανάλυσης Μεγάλων Δεδομένων (Big Data Analysis) για εντοπισμό κυβερνοεπιθέσεων σε πραγματικό χρόνο και την αντιμετώπισή τους [103].

2.9.3.1. Εφαρμογές Ενισχυτικής Μάθησης στο σύστημα LFC

Η Ενισχυτική Μάθηση έχει εμπνεύσει σειρά ερευνών πάνω στον έλεγχο φορτίου συχνότητας των συστημάτων ισχύος. Συγκεκριμένα, αλγόριθμοι RL έχουν εφαρμοστεί για τη ρύθμιση της ισχύος στις περιοχές ελέγχου, την αποκατάσταση της συχνότητας μετά από διαταραχές και τη βελτιστοποίηση της λειτουργίας των γεννητριών ώστε να μειώνονται οι αποκλίσεις συχνότητας και ισχύος [13],[120].

Σε συστήματα πολλών περιοχών, η χρήση πολλαπλών πρακτόρων ενισχυτικής μάθησης (MARL) επιτρέπει τη συντονισμένη δράση ανεξάρτητων ελεγκτών, βελτιώνοντας την ευστάθεια του δικτύου μέσω διανεμημένης βελτιστοποίησης και αποφυγής συγκρούσεων στόχων [121]. Επίσης, έχουν προταθεί μέθοδοι βαθιάς ενισχυτικής μάθησης με συνεχή

αναζήτηση ενεργειών, που επιτρέπει τη ρύθμιση φορτίου συχνότητας σε στοχαστικά συστήματα ισχύος με έμφαση στη προσέγγιση με βάση τα δεδομένα και την αντιμετώπιση αβεβαιοτήτων [122].

Όσον αφορά τη βελτιστοποίηση των διαδικασιών ελέγχου, έχει εισαχθεί RL-ελεγκτής για τη δυναμική προσαρμογή κανόνων ενεργοποίησης γεννητριών, εξασφαλίζοντας βέλτιστη απόκριση υπό μεταβαλλόμενο φορτίο και ελαχιστοποιώντας τη φθορά υλικού [123]. Επιπλέον ένας ελεγκτής που συνδυάζει Βαθιά RL με Ασαφή Λογική Τύπου-ΙΙ αξιοποιεί την ικανότητα της τελευταίας να διαχειρίζεται ποιοτικές πληροφορίες, προσφέροντας ανθεκτικότητα ακόμη και σε συνθήκες έλλειψης δεδομένων ή θόρυβο [124].

Στον τομέα της κυβερνοασφάλειας,, η RL έχει αξιοποιηθεί για την προσομοίωση πραγματιστικών επιθέσεων (FDIAs) σε LFC συστήματα, επιτρέποντας τη δοκιμή αμυντικών μηχανισμών υπό κρίσιμα σενάρια [14], [125]. Παράλληλα, η ενσωμάτωσή της σε υβριδικά αμυντικά πλαίσια αναδεικνύει νέες στρατηγικές για την ταυτόχρονη αντιμετώπιση φυσικών διαταραχών και κυβερνοαπειλών [126].

Στην μέχρι σήμερα δημοσιευμένη βιβλιογραφία, όμως, δεν εντοπίστηκε κάποια τεχνική ελέγχου του LFC βασισμένη στην RL που να έχει εκπαιδευτεί και δοκιμαστεί σε συνθήκες κυβερνοεπιθέσεων. Η παρούσα εργασία δοκιμάζει να καλύψει αυτό το ερευνητικό κενό με τις εξής συνεισφορές:

- Την ανάπτυξη του πρώτου κυβερνο-ανθεκτικού ελεγκτή για συστήματα LFC με τη χρήση βαθιάς ενισχυτικής μάθησης.
- Την ικανότητα του προτεινόμενου ελεγκτή να αντιμετωπίζει αποτελεσματικά FDIAs ενάντια σε κάθε ευάλωτο σε κυβερνοεπιθέσεις σημείο του συστήματος LFC.
- Την αποτελεσματικότητα της προτεινόμενης μεθόδου σε θορυβώδη περιβάλλοντα και την υπεροχή της έναντι παρόμοιων τεχνικών.

3

ΕΝΙΣΧΥΤΙΚΗ ΜΑΘΗΣΗ

"Αντί να προσπαθούμε να δημιουργήσουμε ένα πρόγραμμα που προσομοιάζει το μυαλό ενός ενήλικα, γιατί να μην προσπαθήσουμε να δημιουργήσουμε ένα που προσομοιάζει το μυαλό ενός παιδιού; Αν αυτό στη συνέχεια υποβληθεί σε μια κατάλληλη διαδικασία εκπαίδευσης, θα μπορούσαμε να αποκτήσουμε το μυαλό ενός ενήλικα" - Άλαν Τούρινγκ.

3.1. Εισαγωγή στην Ενισχυτική Μάθηση

Η Ενισχυτική Μάθηση (Reinforcement Learning - RL) είναι μια υπολογιστική προσέγγιση Μηχανικής Μάθησης (Machine Learning - ML). Διακρίνεται από άλλες προσεγγίσεις εστιάζοντας στην απόκτηση γνώσης από έναν πράκτορα μέσω της απευθείας αλληλεπίδρασής του σε ένα περιβάλλον, χωρίς να απαιτούνται εξωτερική επίβλεψη ή εκτενή σύνολα δεδομένων [12]. Η RL χρησιμοποιεί τις αρχές της Μαρκωβιανής Διαδικασίας Αποφάσεων (Markov Decision Process) για να ορίσει την αλληλεπίδραση μεταξύ ενός πράκτορα που μαθαίνει και του περιβάλλοντός του με όρους καταστάσεων, ενεργειών και ανταμοιβών. Αυτές οι μαθηματικές αρχές αποτελούν έναν απλό και πλήρη τρόπο αποτύπωσης του προβλήματος της Τεχνητής Νοημοσύνης (Artificial Intelligence, AI).

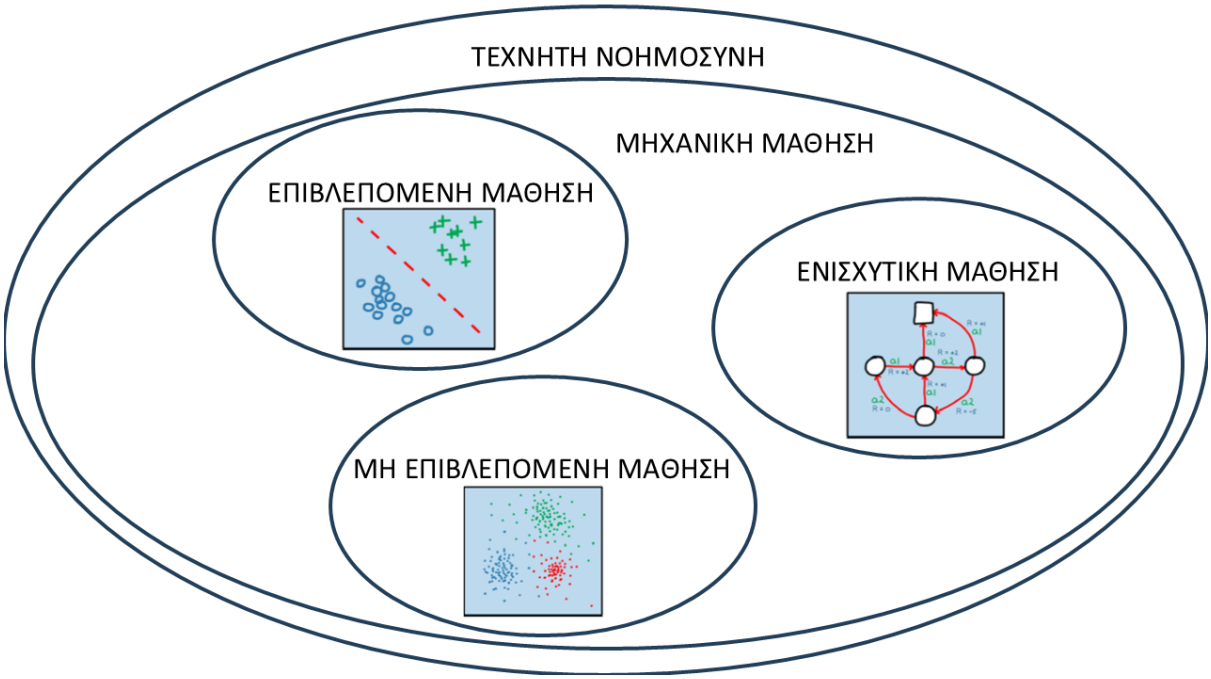
Το μέλλον RL έχει τεράστιες προοπτικές, με συνεχιζόμενες εξελίξεις που αναμένεται να φέρουν επανάσταση σε πολλούς τομείς, από τη ρομποτική και την υγειονομική περίθαλψη μέχρι τα αυτόνομα συστήματα και τα παιχνίδια. Καθώς η υπολογιστική ισχύς αυξάνεται και αναπτύσσονται πιο εξελιγμένοι αλγόριθμοι, η RL έχει τη δυνατότητα να λύσει όλο και πιο σύνθετα προβλήματα που κάποτε θεωρούνταν αδύνατα να επιλυθούν. Ο συνδυασμός της Βαθιάς Μάθησης (Deep Learning - DL) με την RL δημιουργεί την Βαθιά Ενισχυτική Μάθηση (Deep Reinforcement Learning - DRL) και έχει οδηγήσει ήδη σε σημαντικά επιτεύγματα σε τομείς όπως η Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing - NLP) και η λήψη αποφάσεων σε πραγματικό χρόνο σε δυναμικά περιβάλλοντα. Επιπλέον, η πρόοδος στα πολυ-πρακτορικά συστήματα, όπου πολλοί πράκτορες αλληλεπιδρούν μεταξύ τους, επιτρέπει την ανάπτυξη αποδοτικών και κλιμακούμενων εφαρμογών για εργασίες όπως ο έλεγχος κυκλοφορίας, τα κατανεμημένα ενεργειακά συστήματα και η διαχείριση πολύπλοκων δικτύων. Ωστόσο, παραμένουν προκλήσεις, όπως η βελτίωση της αποδοτικότητας απόδειξης, η αντιμετώπιση ηθικών ανησυχιών στη διαδικασία λήψης αποφάσεων και η ανάπτυξη αλγορίθμων RL που να γενικεύονται σε διαφορετικά περιβάλλοντα [15]. Καθώς αυτές οι προκλήσεις επιλύονται, το μέλλον της RL αναμένεται να επεκταθεί σε ένα ευρύτερο φάσμα βιομηχανιών, παρέχοντας καινοτόμες λύσεις και διαμορφώνοντας την επόμενη γενιά έξυπνων συστημάτων.

3.2. Μηχανική Μάθηση (Machine Learning - ML)

Η Εποπτευόμενη Μάθηση (Supervised Learning), η Μη Εποπτευόμενη Μάθηση (Unsupervised Learning) και η RL αποτελούν τις τρεις ευρύτερες κατηγορίες Μηχανικής Μάθησης [104].

Η Μη Εποπτευόμενη Μάθηση χρησιμοποιείται για τον εντοπισμό προτύπων ή κρυμμένων δομών σε σύνολα δεδομένων που δεν έχουν κατηγοριοποιηθεί.

Η Εποπτευόμενη Μάθηση στηρίζεται στην αντιστοίχιση της δεδομένης εισόδου σε σαφώς ορισμένες κατηγορίες δεδομένων.



Σχήμα 12: Τομείς της Τεχνητής Νοημοσύνης

Η Ενισχυτική Μάθηση σε αντίθεση με τις άλλες μεθόδους, αξιοποιεί δεδομένα που λαμβάνει από δυναμικά περιβάλλοντα. Σκοπός της είναι η εύρεση της κατάλληλης αλληλουχίας ενεργειών για την επίτευξη του βέλτιστου αποτελέσματος. Ο τρόπος για την επίλυση αυτού του προβλήματος είναι η αλληλεπίδραση μιας οντότητας λογισμικού, γνωστού ως πράκτορα (agent), με το περιβάλλον και η αξιολόγηση των ενεργειών του.

Μέσα στον πράκτορα υπάρχει μια μοναδική συνάρτηση που δέχεται σαν είσοδο την κατάσταση του περιβάλλοντος και την αντιστοιχεί σε μια ενέργεια. Η συνάρτηση αυτή ονομάζεται πολιτική και αντικαθιστά το σύνολο των στοιχείων ενός παραδοσιακού συστήματος ελέγχου [105].

3.3. Εφαρμογές Ενισχυτικής Μάθησης

Η RL έχει ευρύ φάσμα εφαρμογών σε πολλούς τομείς. Στη ρομποτική, χρησιμοποιείται για την εκπαίδευση ρομπότ να εκτελούν περίπλοκες εργασίες όπως η πλοήγηση και η συναρμολόγηση αντικειμένων [111]. Στον τομέα των αυτόνομων οχημάτων, η RL βοηθά στην εκμάθηση της λήψης αποφάσεων για την ασφαλή πλοήγηση σε κυκλοφοριακά περιβάλλοντα [112]. Στην υγειονομική περίθαλψη, εφαρμόζεται για την εξατομικευμένη θεραπεία και τη διαχείριση ασθενών [113], ενώ στον τομέα των χρηματοοικονομικών χρησιμοποιείται για τη βελτιστοποίηση χαρτοφυλακίων και την ανίχνευση απάτης[114]. Επίσης, στη βιομηχανία και την ενέργεια, η RL συμβάλλει στην αποδοτική διαχείριση πόρων και στην πρόβλεψη ζήτησης [115][116][117] καθώς και στην ανίχνευση βλαβών [118] [119].

3.4. Βασικές Αρχές της Ενισχυτικής Μάθησης

Πράκτορας (Agent) : Ο λήπτης των αποφάσεων. Λογισμικό που αλληλεπιδρά με το περιβάλλον βάσει μιας συγκεκριμένης πολιτικής που διαμορφώνει.

Ενέργειες (Actions) : Το σύνολο A των δυνατών ενεργειών. Μπορεί να είναι συνεχές, διακριτό ή συνδυασμός.

Καταστάσεις (States) : Το διάνυσμα S των μεταβλητών του περιβάλλοντος προς παρατήρηση από τον πράκτορα. Αντιπροσωπεύει το σύνολο όλων των πιθανών διαμορφώσεων του περιβάλλοντος.

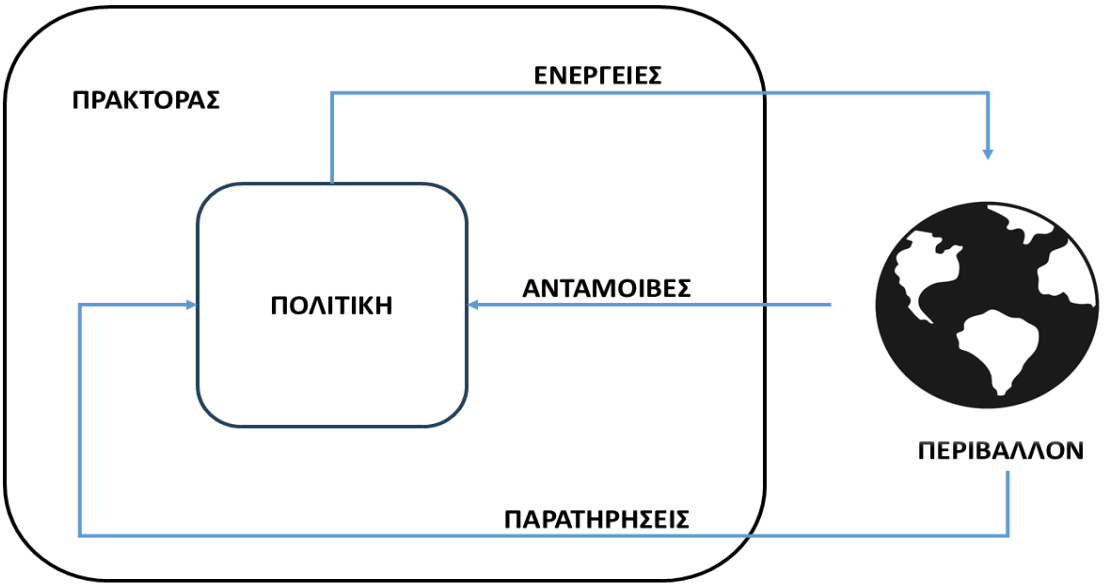
Πολιτική (Policy) : Μια συνάρτηση π(s) που καθορίζει ποια ενέργεια θα πραγματοποιήσει ο πράκτορας δεδομένης μιας κατάστασης s.

Περιβάλλον (Environment) : Αναπαράσταση ενός συστήματος με μηχανισμούς που επιτρέπουν στον πράκτορα να παρατηρεί και να ενεργεί.

Βρόχος Ανατροφοδότησης (Feedback Loop) : Ο μηχανισμός μέσω του οποίου ο πράκτορας λαμβάνει ανταμοιβές και παρατηρεί τις συνέπειες των ενεργειών του.

Ανταμοιβές (Rewards): Η τιμή που επιστρέφει από το περιβάλλον στον πράκτορα ως συνέπεια των ενεργειών του.

Επεισόδιο (Episode): Μια αλληλουχία αλληλεπιδράσεων μεταξύ του πράκτορα και του περιβάλλοντος.



Σχήμα 13: Αλληλεπίδραση πράκτορα Ενισχυτής Μάθησης με το περιβάλλον

3.4.1. Μαρκωβιανή Διαδικασία Αποφάσεων (Markov Decision Process - MDP)

Το μαθηματικό υπόβαθρο στο οποίο στηρίζεται η RL είναι η MDP [106] η οποία ορίζεται από:

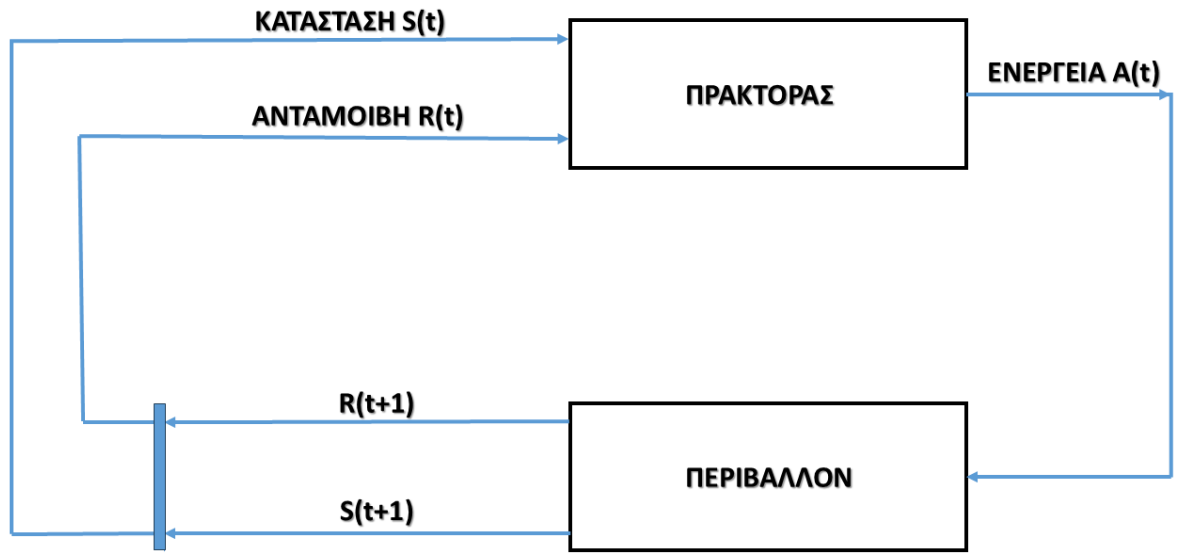
Καταστάσεις (States - S): Το σύνολο όλων των πιθανών διαμορφώσεων του περιβάλλοντος.

Ενέργειες (Actions - A): Το σύνολο των πιθανών ενεργειών που είναι διαθέσιμες στον πράκτορα.

Πιθανότητα Μετάβασης (Transition Probability - P): Η πιθανότητα μετάβασης από μία κατάσταση σε μια άλλη δεδομένης μιας ενέργειας.

Συνάρτηση Ανταμοιβής (Reward Function - R): Η άμεση ανάδραση που λαμβάνεται για την εκτέλεση μιας συγκεκριμένης ενέργειας σε μια δεδομένη κατάσταση.

Παράγοντας Απόσβεσης (Discount Factor - γ): Μια παράμετρος ($0 \leq \gamma \leq 1$) που εξισορροπεί τη σημασία των άμεσων έναντι των μελλοντικών ανταμοιβών.



Σχήμα 14 Μαρκωβιανή Διαδικασία Αποφάσεων

Οι MDPs παρέχουν έναν τυπικό τρόπο μοντελοποίησης της διαδικασίας λήψης αποφάσεων του πράκτορα, υποθέτοντας την ιδιότητα Markov, όπου η μελλοντική κατάσταση εξαρτάται μόνο από την τρέχουσα κατάσταση και την ενέργεια, και όχι από την ιστορία των καταστάσεων.

$$P[S_{t+1}|S_t] = P[S_{t+1}|S_1, S_2, \dots, S_t]$$

Η πολιτική (Π) καθορίζει τη βέλτιστης δράση του πράκτορα δεδομένης της τρέχουσας κατάστασης, ώστε να αποκομίσει τη μέγιστη ανταμοιβή. Με απλά λόγια, συνδέει τις δράσεις με τις καταστάσεις.

$$\Pi: S \to A$$

Για τον καθορισμό της βέλτιστης πολιτικής, είναι απαραίτητο να οριστούν οι αποδόσεις των ανταμοιβών του πράκτορα σε κάθε κατάσταση. Για αυτό το σκοπό, εισάγεται μια μεταβλητή που ονομάζεται Παράγοντας Απόσβεσης (γ). Αρχικά το γ λαμβάνει τιμές κοντά στο μηδέν και δίνεται προτεραιότητα στις άμεσες ανταμοιβές. Στην πορεία, το γ λαμβάνει τιμές κοντά στη μονάδα οπότε η έμφαση μετατοπίζεται στις μακροπρόθεσμες ανταμοιβές. Επομένως, η μέθοδος του αποσβεσμένου ορίζοντα στο άπειρο είναι καθοριστική για την αποκάλυψη της βέλτιστης πολιτικής.

3.4.2. Συνάρτηση Αξίας

Η **Συνάρτηση Αξίας V(s) (Value Function)** προσδιορίζει την επιστροφή ανταμοιβής σε κάθε συγκεκριμένη κατάσταση. Ο τύπος της χαρακτηρίζεται από το αναμενόμενο άθροισμα των αποσβεσμένων μελλοντικών ανταμοιβών:

$$V(s) = E\left[\sum_{t=0}^{\infty} \gamma^t r_t \,\middle|\, s_t\right.$$

Η συνάρτηση αξίας μπορεί να διαχωριστεί σε δύο συνιστώσες: την ανταμοιβή της τρέχουσας κατάστασης και την αποσβεσμένη τιμή ανταμοιβής της επόμενης κατάστασης. Αυτός ο διαχωρισμός προκύπτει από την εξίσωση του Bellman [107], όπως φαίνεται παρακάτω:

$$V^*(S) = max_a[R(s,a) + \gamma \sum_{s' \in S} P(s'|s,a)V^*(s')]$$

Αξίζει να σημειωθεί ότι οι δράσεις και οι ανταμοιβές του πράκτορα μεταβάλλονται ανάλογα με την πολιτική. Αυτό υποδηλώνει ότι η συνάρτηση αξίας είναι συγκεκριμένη για μια πολιτική.

Η βέλτιστη συνάρτηση αξίας μπορεί να προκύψει με τη βοήθεια επαναληπτικών μεθόδων, όπως οι αξιολογήσεις Monte-Carlo, ο δυναμικός προγραμματισμός (Dynamic Programming) ή η εκμάθηση με χρονική διαφορά (Temporal-Difference Learning). Η πολιτική που επιλέγει τη μέγιστη αξία λαμβάνοντας υπόψη την παρούσα κατάσταση αναφέρεται ως η βέλτιστη πολιτική. Μαθηματικά, αναπαρίσταται από την ακόλουθη εξίσωση:

$$\Pi^*(s) = argmax_a[R(s,a) + \gamma \sum_{s' \in S} P_{ss'}^a, V(s')]$$

Έτσι, η πολιτική σε κάθε χρονικό βήμα αξιολογείται με βάση τις πληροφορίες της τρέχουσας κατάστασης. Η αξιολόγηση πραγματοποιείται μέσω διαφόρων μεθόδων, όπως η Προσέγγιση Πολιτικής (Policy Iteration), το Q-learning, η Προσέγγιση Αξίας (Value Iteration) και ο Γραμμικός Προγραμματισμός.

3.4.3. Εξερεύνηση και Εκμετάλλευση

Ένα κρίσιμο στοιχείο της RL είναι η εξισορρόπηση μεταξύ εξερεύνησης και εκμετάλλευσης καθώς ο πράκτορας αλληλεπιδρά με το περιβάλλον. Ο λόγος που προκύπτει αυτό το δίλημμα είναι ότι η μάθηση γίνεται σε πραγματικό χρόνο. Αντί να βασίζεται σε ένα στατικό σύνολο δεδομένων, οι δράσεις του πράκτορα καθορίζουν ποια δεδομένα επιστρέφονται από το περιβάλλον. Οι επιλογές που κάνει ο πράκτορας καθορίζουν τις πληροφορίες που λαμβάνει και, κατά συνέπεια, τις πληροφορίες από τις οποίες μπορεί να μάθει [108].

Εάν ο πράκτορας εκμεταλλεύεται συνεχώς αυτό που θεωρεί ως την καλύτερη ενέργεια σε κάθε δεδομένη στιγμή, μπορεί να μην λάβει ποτέ πρόσθετες πληροφορίες για τις καταστάσεις που υπάρχουν πέρα από μια ενέργεια χαμηλής ανταμοιβής. Η καθαρή εκμετάλλευση μπορεί να αυξήσει τον χρόνο που απαιτείται για να βρεθεί η βέλτιστη πολιτική ή να οδηγήσει τον αλγόριθμο μάθησης να συγκλίνει σε μια μη βέλτιστη πολιτική, καθώς ολόκληρες περιοχές του χώρου καταστάσεων μπορεί να παραμείνουν ανεξερεύνητες.

Η καθαρή εξερεύνηση δεν είναι αποδοτικός τρόπος μάθησης, καθώς ο πράκτορας πιθανότατα θα ξοδέψει χρόνο καλύπτοντας μεγαλύτερο μέρος του χώρου καταστάσεων. Αν και αυτό είναι ωφέλιμο για την εύρεση μιας γενικής λύσης, η υπερβολική εξερεύνηση μπορεί να μειώσει τόσο πολύ το ρυθμό μάθησης ώστε να μην βρεθεί μια επαρκής λύση σε εύλογο χρόνο. Επομένως, οι καλύτεροι αλγόριθμοι μάθησης επιτυγχάνουν μια ισορροπία μεταξύ εξερεύνησης και εκμετάλλευσης του περιβάλλοντος.

Αν και οι αλγόριθμοι RL παρέχουν έναν απλό τρόπο ισορροπίας μεταξύ εξερεύνησης και εκμετάλλευσης, μπορεί να μην είναι προφανές πού να ρυθμιστεί αυτή η ισορροπία κατά τη διάρκεια της διαδικασίας μάθησης, ώστε ο πράκτορας να καταλήξει σε μια επαρκή πολιτική εντός του διαθέσιμου χρόνου μάθησης. Γενικά, όμως, ένας πράκτορας εξερευνά περισσότερο στην αρχή της μάθησης και σταδιακά μεταβαίνει σε έναν πιο εκμεταλλευτικό ρόλο προς το τέλος.

3.4.4. Η έννοια της Αξίας

Ένα δεύτερο κρίσιμο στοιχείο της RL είναι η έννοια της αξίας. Η αξιολόγηση της αξίας μιας κατάστασης ή μιας ενέργειας, αντί για την ανταμοιβή, βοηθά τον πράκτορα να επιλέξει την ενέργεια που θα αποφέρει τις περισσότερες ανταμοιβές με την πάροδο του χρόνου, αντί για ένα βραχυπρόθεσμο όφελος.

Ανταμοιβή: Το άμεσο όφελος από την παραμονή σε μια κατάσταση και την εκτέλεση μιας συγκεκριμένης ενέργειας.

Αξία: Οι συνολικές ανταμοιβές που αναμένει να λάβει ο πράκτορας από μια κατάσταση και στο εξής προς το μέλλον.

Φυσικά, η εκτίμηση της λήψης υψηλής ανταμοιβής μετά από πολλές διαδοχικές ενέργειες δεν σημαίνει απαραίτητα ότι η αρχική δράση είναι η καλύτερη.

Είναι αποδοτικό η εκτίμηση της αξίας να είναι σχετικά βραχυπρόθεσμη. Η ρύθμιση πόσο βραχυπρόθεσμος είναι ο πράκτορας στους υπολογισμούς του, επιτυγχάνεται αποσβένοντας τις ανταμοιβές σε μεγαλύτερο βαθμό όσο πιο μακριά βρίσκονται στο μέλλον. Αυτό γίνεται ρυθμίζοντας τον παράγοντα απόσβεσης, γ, μεταξύ 0 και 1.

$$total\ discounted\ reward = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots = \sum_{i=1}^{T} \gamma^{i-1} r_i$$

3.5. Αναπαριστώντας μια Πολιτική

Μια πολιτική είναι μια συνάρτηση που δέχεται παρατηρήσεις καταστάσεων και επιστρέφει ενέργειες. Επομένως, οποιαδήποτε συνάρτηση με την ακόλουθη σχέση εισόδου και εξόδου μπορεί να λειτουργήσει ως αναπαράσταση μιας πολιτικής :

$$\Pi : A = f(S)$$

3.5.1. Αναπαράσταση με πίνακα

Για απλά προβλήματα, η πολιτική μπορεί να αναπαρασταθεί από έναν πίνακα με την αξία κάθε διαθέσιμης ενέργειας δεδομένης της τρέχουσας κατάστασης.

Σχήμα 15: Αναπαράσταση πολιτικής με πίνακα

Η αναπαράσταση των παραμέτρων της πολιτικής σε έναν πίνακα δεν είναι εφικτή όταν ο αριθμός των ζευγών καταστάσεων - ενεργειών γίνεται μεγάλος ή άπειρος.

Η συνεχής φύση ενός προβλήματος μπορεί να αναπαρασταθεί με μια συνεχή συνάρτηση που δέχεται καταστάσεις και επιστρέφει ενέργειες. Ωστόσο, πριν ξεκινήσει η εκμάθηση των σωστών παραμέτρων σε αυτή τη συνάρτηση, θα πρέπει να οριστεί η λογική της δομή. Η λογική δομή μπορεί να είναι δύσκολο να οριστεί για συστήματα με υψηλό βαθμό ελευθερίας ή μη γραμμικά συστήματα.

Άρα, χρειάζεται ένας τρόπος αναπαράστασης μιας συνάρτησης που να μπορεί να χειριστεί συνεχείς καταστάσεις και ενέργειες, ο οποίος δεν απαιτεί μια δύσκολη προς κατασκευή λογική δομή για κάθε περιβάλλον. Ο τρόπος αυτός είναι η χρήση των Νευρωνικών Δικτύων (Neural Networks - NNs) [109].

3.5.2. Αναπαράσταση Με Νευρωνικά Δίκτυα

Ένα νευρωνικό δίκτυο είναι μια ομάδα κόμβων ή τεχνητών νευρώνων που είναι συνδεδεμένοι με τέτοιο τρόπο ώστε να επιτρέπουν στο δίκτυο να είναι ένας Καθολικός Προσεγγιστής (Universal Approximator) συναρτήσεων. Αυτό σημαίνει ότι, με την κατάλληλα συνδυασμένη διάταξη κόμβων και συνδέσεων, το δίκτυο μπορεί να ρυθμιστεί ώστε να αποτυπώνει οποιαδήποτε σχέση εισόδου και εξόδου. Παρά το γεγονός ότι η συνάρτηση μπορεί να είναι εξαιρετικά περίπλοκη, η καθολική φύση των νευρωνικών δικτύων εξασφαλίζει ότι υπάρχει ένα νευρωνικό δίκτυο κάποιας μορφής που μπορεί να το πετύχει.

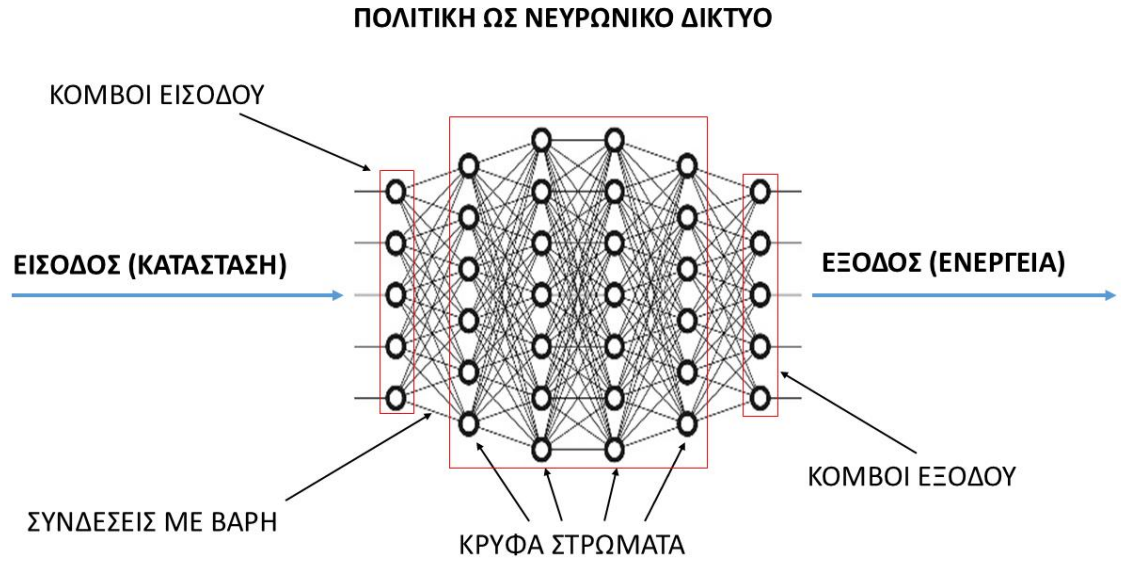
Έτσι, αντί για την εύρεση μη γραμμικής δομής συνάρτησης που λειτουργεί σε ένα συγκεκριμένο περιβάλλον, ένα νευρωνικό δίκτυο μπορεί να χρησιμοποιήσει την ίδια συνδυασμένη διάταξη κόμβων και συνδέσεων σε πολλά διαφορετικά περιβάλλοντα. Η μόνη

διαφορά είναι στις ίδιες τις παραμέτρους. Η διαδικασία μάθησης συνίσταται στη συστηματική προσαρμογή των παραμέτρων για να βρεθεί η βέλτιστη σχέση εισόδου - εξόδου.

Είναι σημαντικό να επισημανθούν ορισμένα σημεία, για να βοηθήσουν στην εξήγηση των αποφάσεων αργότερα κατά την ρύθμιση των πολιτικών. Στο παρακάτω σχήμα φαίνεται η τυπική διάρθρωση ενός νευρωνικού δικτύου. Στα αριστερά είναι οι κόμβοι εισόδου, ένας για κάθε είσοδο στη συνάρτηση, και στα δεξιά είναι οι κόμβοι εξόδου. Ανάμεσά τους βρίσκονται οι στήλες των κόμβων που ονομάζονται Κρυφά Στρώματα (Hidden Layers).

Σε ένα πλήρως συνδεδεμένο δίκτυο, υπάρχει μια σύνδεση με Βάρος (Weight) από κάθε κόμβο εισόδου σε κάθε κόμβο του επόμενου στρώματος. Ομοίως και στη συνέχεια από αυτούς τους κόμβους στο στρώμα που ακολουθεί. Τέλος με αντίστοιχο τρόπο τα δεδομένα μεταδίδονται ξανά μέχρι τους κόμβους εξόδου.

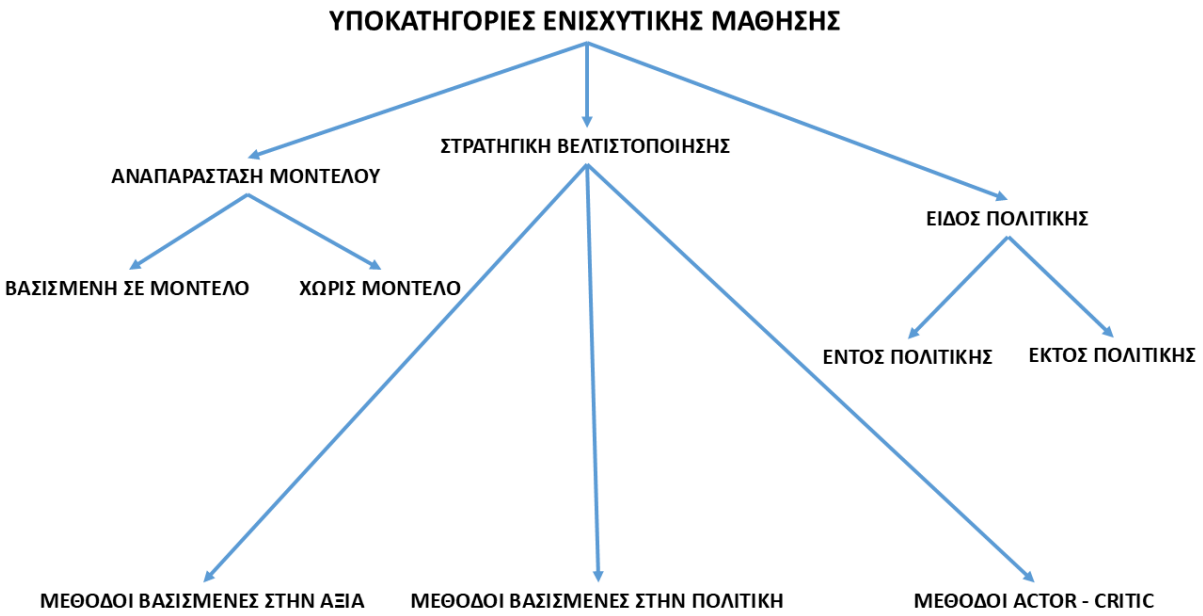
Η τιμή οποιουδήποτε δεδομένου κόμβου είναι ίση με το άθροισμα κάθε κόμβου που του παρέχει δεδομένα, πολλαπλασιασμένο με τον αντίστοιχο παράγοντα βάρους του, συν μια Μεροληψία (Bias). Αυτό το σύνολο των λειτουργιών μετασχηματίζει ουσιαστικά τις αριθμητικές τιμές των κόμβων σε ένα στρώμα στις τιμές των κόμβων του επόμενου στρώματος.



Σχήμα 16: Αναπαράσταση πολιτικής με νευρωνικό δίκτυο

Υπάρχει ένα βήμα που είναι πιθανώς από τα πιο σημαντικά στοιχεία ενός νευρωνικού δικτύου. Μετά τον υπολογισμό της τιμής ενός κόμβου, εφαρμόζεται μια Συνάρτηση Ενεργοποίησης (Activation Function) που αλλάζει την τιμή του κόμβου πριν αυτός τροφοδοτηθεί ως είσοδος στην επόμενη στρώση. Υπάρχουν πολλές διαφορετικές συναρτήσεις ενεργοποίησης. Οι συναρτήσεις αυτές είναι μη γραμμικές, κάτι που είναι κρίσιμο για τη δημιουργία ενός δικτύου που μπορεί να προσεγγίσει οποιαδήποτε συνάρτηση. Πολλές μη γραμμικές συναρτήσεις μπορούν να αναλυθούν σε έναν κατάλληλα ζυγισμένο συνδυασμό των εξόδων των συναρτήσεων ενεργοποίησης.

3.6. Κατηγοριοποίηση Αλγορίθμων Ενισχυτικής Μάθησης



Σχήμα 17: Κατηγορίες αλγορίθμων Ενισχυτικής Μάθησης

3.6.1. Βασισμένη σε Μοντέλο Ενισχυτική Μάθηση (Model-Based RL)

Ο πράκτορας δημιουργεί ένα μοντέλο του περιβάλλοντος (προσπαθεί να καταλάβει τη μηχανική διαδοχής των καταστάσεων ή της κατανομής των ανταμοιβών) και το χρησιμοποιεί για να σχεδιάζει και να λαμβάνει αποφάσεις.

3.6.1.1. Μέθοδοι RL Βασισμένες σε Μοντέλο

Προσέγγιση Αξίας (Value Iteration) : Ανανεώνει επαναλαμβανόμενα τη συνάρτηση αξίας σε κάθε κατάσταση, βρίσκοντας την μέγιστη αναμενόμενη αξία από όλες τις πιθανές ενέργειες. Η πολιτική προκύπτει από συνάρτηση αξίας, όταν μετά από επαναλήψεις αυτή συγκλίνει.

Προσέγγιση Πολιτικής (Policy Iteration) : Εναλλάσσεται μεταξύ αξιολόγησης της πολιτικής (υπολογισμός της αξίας μιας συγκεκριμένης πολιτικής) και βελτίωσης της πολιτικής (ανανεώνοντας την πολιτική σύμφωνα με την τρέχουσα τιμή της συνάρτησης αξίας) μέχρι αυτές οι δύο να συγκλίνουν.

3.6.2. Ενισχυτική Μάθηση Χωρίς Μοντέλο (Model-Free RL)

Ο πράκτορας μαθαίνει αμιγώς από την αλληλεπίδραση με το περιβάλλον μέσα από διαδικασία δοκιμής – σφάλματος (Trial and Error) και δεν χτίζει μοντέλο του περιβάλλοντος.

3.6.2.1. Διάκριση με Βάση τη Στρατηγική Βελτιστοποίησης

Μέθοδοι Βασισμένες στην Αξία (Value – Based Methods): Ο πράκτορας μαθαίνει μια συνάρτηση αξίας για να καθοδηγεί τις ενέργειές του. Η πολιτική εξάγεται έμμεσα από τη συνάρτηση αξίας.

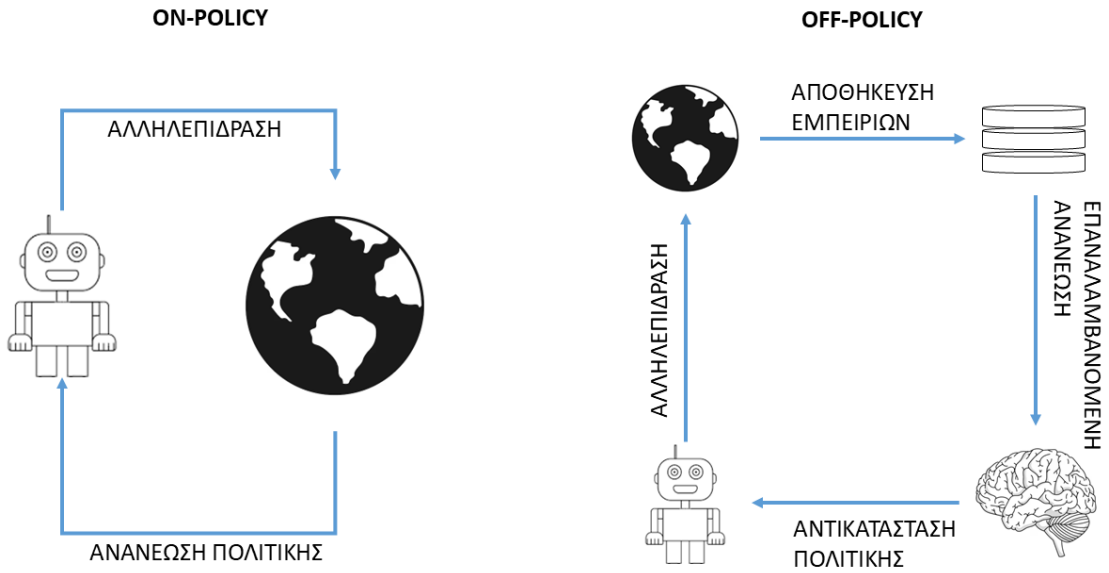
Μέθοδοι Βασισμένες στην Πολιτική (Policy – Based Methods): Ο πράκτορας μαθαίνει την πολιτική απευθείας μέσω της βελτιστοποίησης της, χωρίς τον υπολογισμό συνάρτησης αξίας. Επικεντρώνεται στην εύρεση βέλτιστης πολιτικής για τη μεγιστοποίηση των ανταμοιβών.

Μέθοδοι Παράγοντα – Κριτή (Actor – Critic Methods): Ένα υβρίδιο των άλλων μεθόδων. Ο παράγοντας ανανεώνει την πολιτική (σαν μέθοδος βασισμένη στην πολιτική) και ο κριτής αξιολογεί πόσο καλή ήταν η απόφαση που έλαβε ο παράγοντας (σαν μέθοδος βασισμένη στην αξία).

3.6.2.2. Διάκριση με Βάση τον Τύπο Πολιτικής

Εντός Πολιτικής (On-Policy): Ο πράκτορας μαθαίνει και βελτιώνει την πολιτική που χρησιμοποιεί. Αναβαθμίζει τη γνώση του μέσω ενεργειών που αποφάσισε βασισμένος στην πολιτική του.

Εκτός Πολιτικής (Off-Policy): Ο πράκτορας μαθαίνει μια πολιτική που είναι διαφορετική από αυτή που χρησιμοποιεί για να αλληλεπιδρά με το περιβάλλον. Αυτό του επιτρέπει να διαθέτει μεγαλύτερη ευελιξία, καθώς προηγούμενες εμπειρίες αποθηκεύονται σε μια Ενδιάμεση Μνήμη (Buffer).



Σχήμα 18: Σύγκριση μεθόδων RL εντός και εκτός πολιτικής

3.6.2.3. Παραδείγματα Αλγορίθμων Μη Βασισμένων σε Μοντέλο

3.6.2.3.1. On-Policy

Monte Carlo : Μια μέθοδος value-based που προσεγγίζει συναρτήσεις αξίας, με βάση το μέσο όρο των ανταμοιβών πολλών διαδοχικών επεισοδίων. Απαιτεί πλήρη εκτέλεση κάθε επεισοδίου για τον υπολογισμό της συνάρτησης αξίας. Είναι αυστηρά on-policy, μαθαίνει απευθείας από την πολιτική που ακολουθεί.

SARSA: Μια μέθοδος value-based που στηρίζεται στην χρονική διαφορά. Ανανεώνει τη συνάρτηση αξίας σύμφωνα με τα δεδομένα που προκύπτουν από την ενέργεια που επιλέγει η τρέχουσα πολιτική. Η ανανέωση της πολιτικής γίνεται μετά από κάθε ενέργεια και όχι μετά από ένα πλήρες επεισόδιο.

N-Step SARSA: Επέκταση της μεθόδου SARSA που ανανεώνει την πολιτική λαμβάνοντας υπόψη τις αθροιστικές ανταμοιβές από μια αλληλουχία N βημάτων. Η μέθοδος ισορροπεί μεταξύ της μεροληψίας του αλγορίθμου SARSA και της διακύμανσης του Monte Carlo.

Expected SARSA : Παρόμοια μέθοδος με τη SARSA, η οποία ανανεώνει την πολιτική βασιζόμενη στην αναμενόμενη αξία της επόμενης ενέργειας, ανηγμένη στο μέσο όρο της αξίας του συνόλου των ενεργειών της υπάρχουσας πολιτικής.

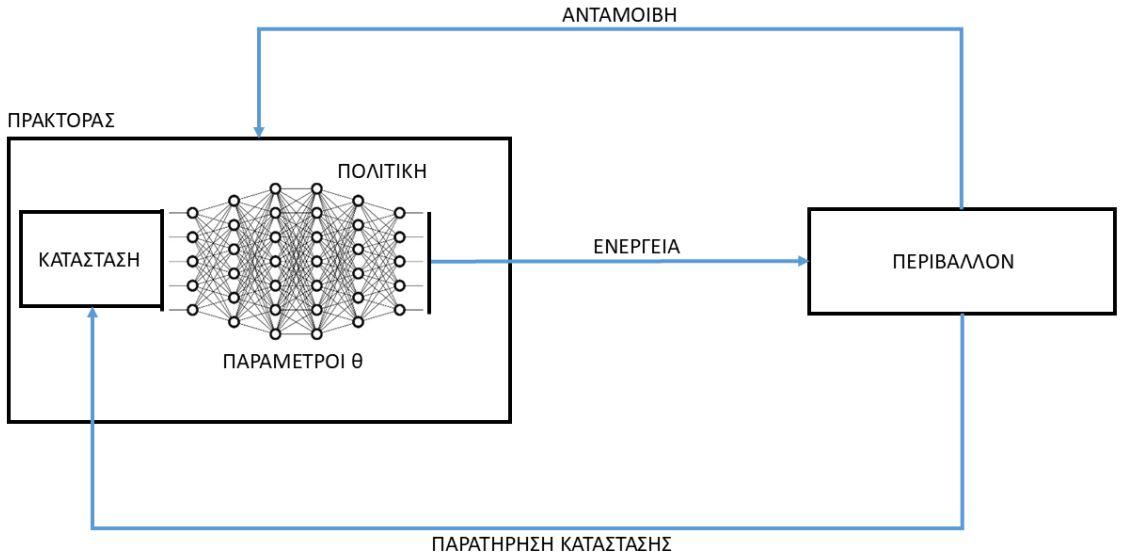
3.6.2.3.2. Off-Policy

Q–Learning : Ένας αλγόριθμος value-based βασισμένος στην χρονική διαφορά. Στόχος του η μάθηση της βέλτιστης συνάρτησης αξίας, ανανεώνοντάς την σύμφωνα με την αναμενόμενη μέγιστη μελλοντική αξία, ανεξάρτητα από την πολιτική που ακολουθεί.

Σχήμα 19: Πίνακας Q-Value

Double Q-Learning: Παραλλαγή της Q-learning που χρησιμοποιεί δύο συναρτήσεις αξίας, μια για την επιλογή ενεργειών και μια για την αξιολόγησή τους, που ανανεώνονται ασύγχρονα μεταξύ τους. Με αυτό τον τρόπο αντιμετωπίζεται πιθανή υπερεκτίμηση των βαρών της συνάρτησης αξίας της απλής μεθόδου.

Deep Q-Learning (DQN): Εισάγει τη χρήση βαθέων νευρωνικών δικτύων στη μέθοδο Q-Learning. Ιδανική για τη διαχείριση πολύπλοκων συστημάτων με πολυδιάστατο σύνολο καταστάσεων. Χρησιμοποιεί μνήμη για την αποθήκευση των εμπειριών και ένα δεύτερο δίκτυο για την ομαλή και σταθερή μάθηση.



Σχήμα 20: Διαδικασία εκπαίδευσης πράκτορα Deep Q-Learning

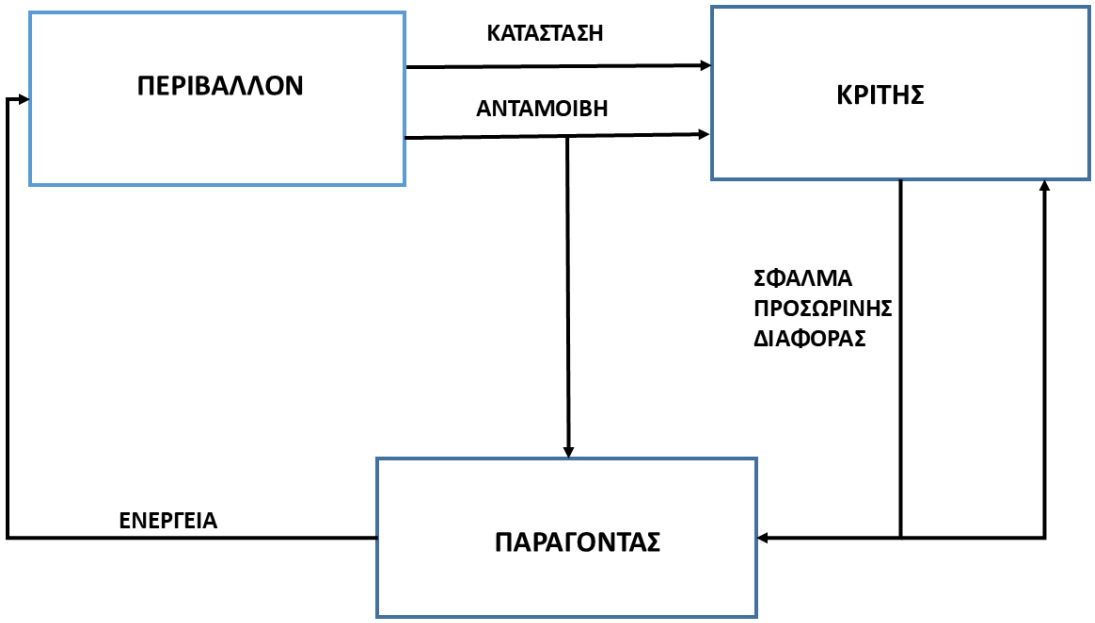
3.6.2.3.3. Αλγόριθμοι Actor-Critic

Ο παράγοντας ανανεώνει την πολιτική του βασισμένος στην αξιολόγηση του κριτή, που υπολογίζει τη συνάρτηση αξίας. Με αυτόν τον τρόπο επιτυγχάνεται πιο σταθερή μάθηση. Σε αυτή την κατηγορία αλγορίθμων συνδυάζονται χαρακτηριστικά μεθόδων βασισμένων στην αξία και μεθόδων βασισμένων στην πολιτική.

A2C (Advantage Actor-Critic): Μέθοδος εντός πολιτικής όπου πολλαπλοί «εργάτες» κάνουν ταυτόχρονη δειγματοληψία συνόλων ενεργειών-ανταμοιβών χρησιμοποιώντας την υπάρχουσα πολιτική. Ο παράγοντας ανανεώνει την πολιτική βασισμένος στην συνάρτηση πλεονεκτήματος, που προκύπτει από το μέσο όρο των συνόλων εμπειριών που έχουν ληφθεί.

A3C (Asynchronous Advantage Actor-Critic): Παρόμοια μέθοδος με την A2C, στην οποία οι εργάτες αλληλεπιδρούν με το περιβάλλον ασύγχρονα χρησιμοποιώντας την ίδια πολιτική. Με αυτόν τον τρόπο η μάθηση επισπεύδεται και περιττεύει η χρήση μνήμης εμπειριών.

PPO (Proximal Policy Optimization): Βελτιώνει την απλή μέθοδο actor-critic χρησιμοποιώντας μια περιορισμένη αντικειμενική συνάρτηση για να εξασφαλίσει τη σταθερή μάθηση.



Σχήμα 21: Διαδικασία εκπαίδευσης πράκτορα Actor-Critic

TRPO (Trust Region Policy Optimization): Μέθοδος εντός πολιτικής που βελτιστοποιεί την πολιτική με τρόπο που να παραμένει πάντα εντός μιας «περιοχής εμπιστοσύνης», αποφεύγοντας μεγάλες αλλαγές που προκαλούν αστάθεια.

DDPG (Deep Deterministic Policy Gradient): Μέθοδος εκτός πολιτικής κατάλληλη για συνεχή χώρο ενεργειών. Χρησιμοποιεί ντετερμινιστική πολιτική ως actor και Q-συνάρτηση ως critic, σε συνδυασμό με μνήμη εμπειριών και δευτερεύον δίκτυο.

SAC (Soft Actor-Critic): Μέθοδος εκτός πολιτικής που εισάγει τον όρο της εντροπίας στη συνάρτηση υπολογισμού ανταμοιβών, ενθαρρύνοντας έτσι τον πράκτορα να εξερευνήσει. Ιδιαίτερα αποτελεσματική σε πολυδιάστατους συνεχείς χώρους ενεργειών [110].

	ON-POLICY	OFF-POLICY	VALUE-BASED	POLICY-BASED	ACTOR-CRITIC
Monte Carlo	✖		✖		
SARSA	✖		✖		
N-Step SARSA	✖		✖		
Expected SARSA	✖		✖		
Q-Learning		✖	✖		
Double Q-Learning		✖			
Deep Q-Learning		✖			
Actor-Critic	✖	✖	✖	✖	✖
A2C	✖				✖
A3C	✖				✖
PPO	✖				✖
TRPO	✖				✖
DDPG		✖			✖
SAC		✖			✖

Σχήμα 22: Συνοπτικός πίνακας δημοφιλών αλγορίθμων Ενισχυτικής Μάθησης

3.7. Πλεονεκτήματα και Προκλήσεις της Ενισχυτικής Μάθησης

Πλεονεκτήματα:

- **Προσαρμοστικότητα:** Τα συστήματα RL μαθαίνουν από την αλληλεπίδραση, γεγονός που τα καθιστά κατάλληλα για δυναμικά ή ανεξερεύνητα περιβάλλοντα.
- **Κλιμακωσιμότητα**: Η RL μπορεί να χειριστεί μεγάλους χώρους καταστάσεων-ενεργειών με κατάλληλες προσεγγίσεις συναρτήσεων.

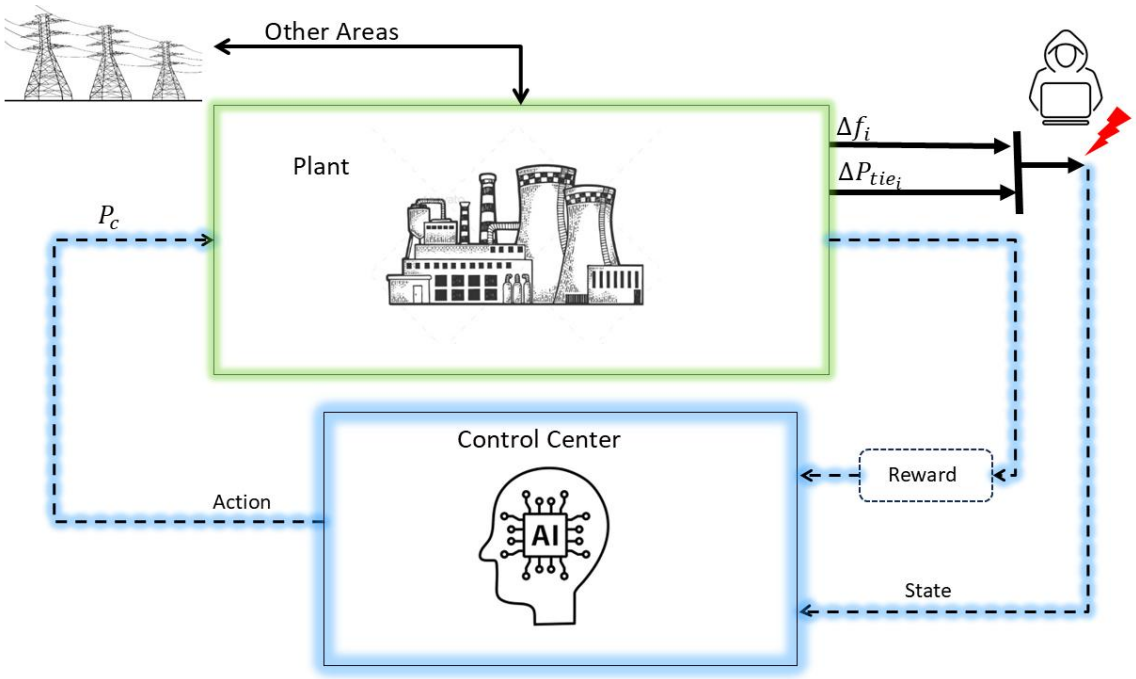
Προκλήσεις:

- **Αναποτελεσματικότητα Δειγμάτων**: Η RL συχνά απαιτεί μεγάλο αριθμό αλληλεπιδράσεων για να συγκλίνει.
- **Κίνδυνοι Εξερεύνησης**: Οι μη ασφαλείς ενέργειες κατά την εξερεύνηση μπορεί να οδηγήσουν σε ανεπιθύμητα αποτελέσματα.
- **Υπολογιστική Πολυπλοκότητα**: Η εκπαίδευση πρακτόρων RL μπορεί να είναι υπολογιστικά απαιτητική.

4

DRL^2FC : ΑΝΘΕΚΤΙΚΟΣ ΣΕ ΚΥΒΕΡΝΟΕΠΙΘΕΣΕΙΣ ΕΛΕΓΚΤΗΣ ΓΙΑ ΣΥΣΤΗΜΑΤΑ LFC

Η αναγνώριση του ερευνητικού κενού στον τομέα των ανθεκτικών σε κυβερνοεπιθέσεις ελεγκτών RL για το σύστημα LFC αποτέλεσε την έμπνευση για την παρούσα εργασία. Συγκεκριμένα, στόχος ήταν η ανάπτυξη ενός καινοτόμου ελεγκτή για το LFC, με τη μορφή ενός πράκτορα ενισχυτικής μάθησης, με σκοπό την εκπαίδευσή του στην αναγνώριση κυβερνοεπιθέσεων τύπου FDI και την αντιμετώπισή τους.



Σχήμα 23: Διάγραμμα συστήματος LFC με ελεγκτή βασισμένο σε Ενισχυτική Μάθηση

4.1. Σχεδιασμός Περιβάλλοντος LFC Κατάλληλο για Εκπαίδευση Πρακτόρων RL

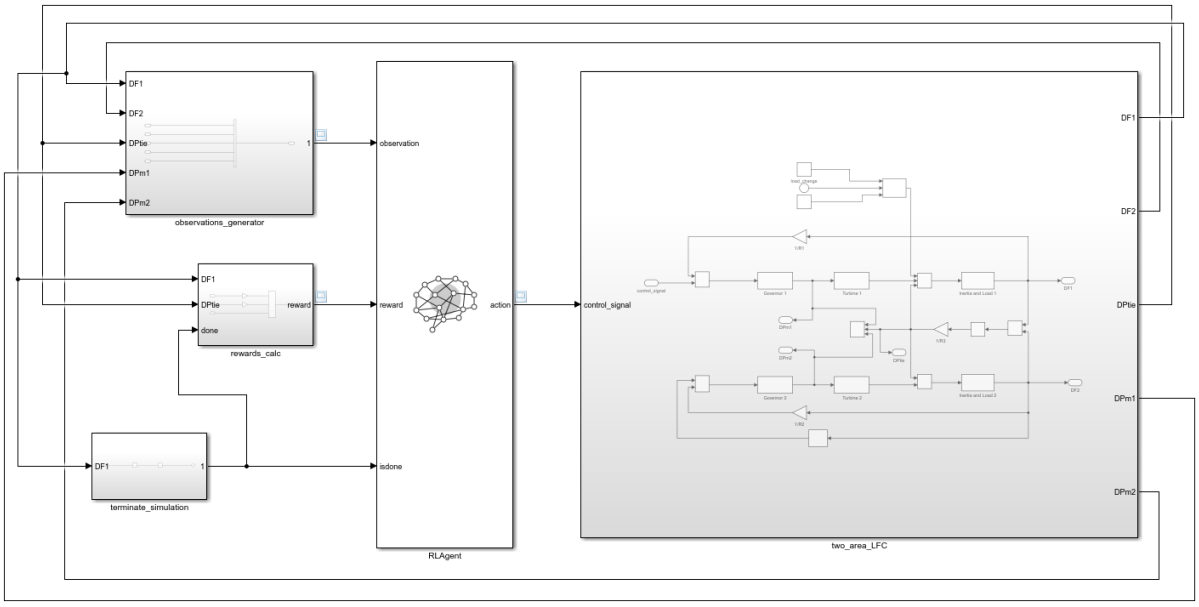
Η εκπαίδευση ενός πράκτορα να δρα ως ανθεκτικός σε κυβερνοεπιθέσεις ελεγκτής για συστήματα LFC απαιτούσε ένα ισχυρό και αποδοτικό περιβάλλον ενισχυτικής μάθησης. Το

PyTorch [127] ξεχωρίζει στην έρευνα της RL λόγω του δυναμικού γραφήματος υπολογισμών και του διαισθητικού προστακτικού προγραμματιστικού μοντέλου του, επιτρέποντας τη γρήγορη δημιουργία πρωτοτύπων αλγορίθμων RL, όπως το DQN ή το PPO. Το TensorFlow [128], με τη στατική βελτιστοποίηση γραφήματος και τις δυνατότητες διανεμημένης εκπαίδευσης, είναι ιδανικό για εφαρμογές RL ευρείας κλίμακας, με υποστήριξη από οικοσυστήματα όπως το TF-Agents. Και τα δύο πλαίσια παρέχουν ισχυρά συστήματα αυτόματης διαφόρισης (autograd) και επιτάχυνση μέσω GPU, καθιστώντας τα απαραίτητα για την ανάπτυξη και αξιολόγηση κυβερνο-ανθεκτικών RL ελεγκτών.

Αρχικά, ο στόχος ήταν να συνδυαστεί το PyTorch με το περιβάλλον LFC βασισμένο στο Simulink που χρησιμοποιήθηκε στα κεφάλαια 3 και 4, αξιοποιώντας τα πλεονεκτήματα και των δύο εργαλείων. Ωστόσο, αυτή η προσέγγιση αντιμετώπισε σημαντικές προκλήσεις στην επίτευξη απρόσκοπτης επικοινωνίας μεταξύ των δύο πλαισίων.

Για να αντιμετωπιστεί αυτό, αναπτύχθηκε μια προσαρμοσμένη μέθοδος επικοινωνίας μεταξύ Python και MATLAB/Simulink χρησιμοποιώντας το πρωτόκολλο TCP/IP . Ενώ αυτή η λύση επέτρεπε την ανταλλαγή δεδομένων, αποδείχθηκε ανεπαρκής για τις απαιτήσεις της εκπαίδευσης RL, ιδίως λόγω ζητημάτων συγχρονισμού και υπολογιστικού φόρτου.

Ως αποτέλεσμα, το MATLAB Reinforcement Learning Toolbox [129] υιοθετήθηκε ως το κύριο πλαίσιο για τον σχεδιασμό του περιβάλλοντος και την εκπαίδευση του πράκτορα. Αυτό το εργαλείο προσέφερε ολοκληρωμένη λύση, επιτρέποντας την αξιοποίηση του συστήματος που είχε ήδη διαμορφωθεί.



Σχήμα 24: Διάγραμμα συστήματος LFC κατάλληλου για εφαρμογές Ενισχυτικής Μάθησης σε περιβάλλον MATLAB/Simulink

4.1.1. Μοντέλο LFC Δύο Περιοχών

Στον πράκτορα ανατέθηκε ο δευτερεύον έλεγχος της περιοχής 1, αντικαθιστώντας τον PID ελεγκτή. Η περιοχή 2 συνέχισε να βασίζεται στον έλεγχο με PID. Σκοπός αυτής της επιλογής ήταν η απομόνωση της επίδρασης του πράκτορα στην περιοχή 1. Συνεπώς, η αξιολόγηση της επίδρασης του στην συχνότητα της περιοχής 1 παρείχε μια σαφή εικόνα για την συμπεριφορά του σε ένα πραγματικό σύστημα LFC.

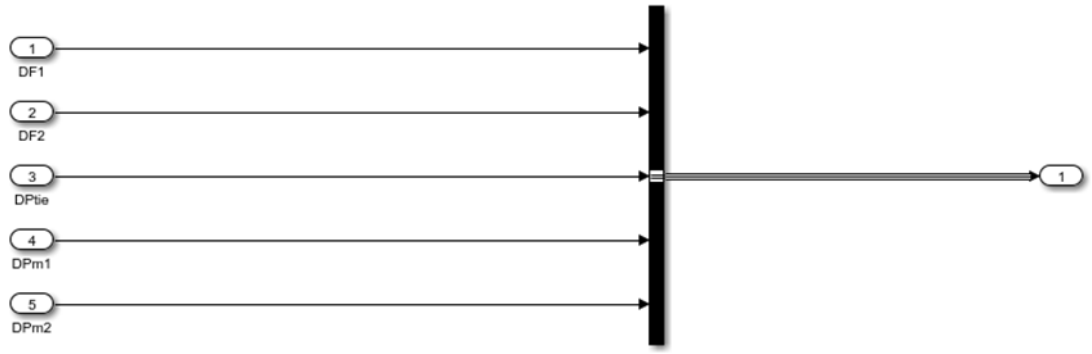
Το σύστημα LFC με τα χαρακτηριστικά που παρατέθηκαν στο κεφάλαιο 2, χρειάστηκε να τροποποιηθεί κατάλληλα για τη αλληλεπίδραση με τον πράκτορα.

Συγκεκριμένα, διαμορφώθηκε ένα διάνυσμα παρατηρήσεων (observation vector) που περιλάμβανε κρίσιμες μεταβλητές κατάστασης του συστήματος. Εκτός από τις προφανώς απαραίτητες μετρήσεις της απόκλισης συχνότητας της περιοχής 1 (Δf_1) και της διασυνδετικής ροής ισχύος (ΔP_{tie}), δόθηκε πρόσβαση στον πράκτορα και στις μετρήσεις απόκλισης συχνότητας της περιοχής 2 (Δf_2), καθώς των αποκλίσεων της μηχανικής ισχύος εξόδου των στροβίλων (ΔP_{m_1}, ΔP_{m_2}).

Το διάνυσμα παρατηρήσεων κατάστασης που διαμορφώθηκε είναι:

$$observations = [\ \Delta f_1\ \ \Delta f_2\ \ \Delta P_{tie}\ \ \Delta P_{m_1}\ \ \Delta P_{m_2}\ ]$$

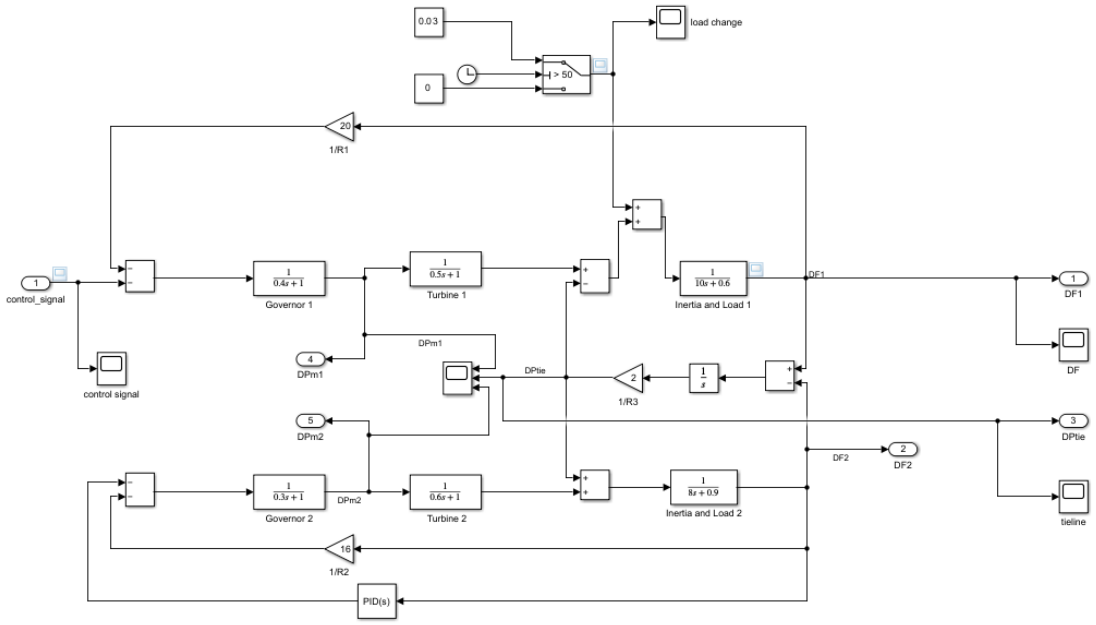
Αυτό το διάνυσμα τροφοδοτήθηκε στον πράκτορα ως είσοδος, επιτρέποντάς του να λαμβάνει αποφάσεις βάσει των τρεχουσών συνθηκών. Στο σχήμα 25 φαίνεται το υποσύστημα σύνθεσης του διανύσματος παρατηρήσεων.



Σχήμα 25: Υποσύστημα σύνθεσης διανύσματος Παρατηρήσεων

Η έξοδος του πράκτορα, η οποία αντιπροσώπευε το σήμα δευτερεύοντος ελέγχου για την περιοχή 1, τροφοδοτήθηκε πίσω στο σύστημα LFC, αντικαθιστώντας το σήμα ελέγχου από τον PID ελεγκτή.

Το ανανεωμένο σύστημα φαίνεται στο Σχήμα 26.



Σχήμα 26: Υποσύστημα προσομοίωσης συστήματος LFC δύο περιοχών

4.1.2. Συνάρτηση Ανταμοιβών

Για την αξιολόγηση των ενεργειών του πράκτορα κατά την εκπαίδευση, υλοποιήθηκε μια συνάρτηση ανταμοιβής που είχε ως στόχο την ελαχιστοποίηση των αποκλίσεων στη συχνότητα και στην διασυνδετική ροή ισχύος μεταξύ των δύο περιοχών. Η συνάρτηση αυτή σχεδιάστηκε με βάση την έννοια του Σφάλματος Ελέγχου Περιοχής το οποίο χρησιμοποιείται και σε συμβατικά συστήματα LFC για τη μέτρηση και διόρθωση των αποκλίσεων.

Η ανταμοιβή που χρησιμοποιήθηκε για την εκπαίδευση του πράκτορα δίνεται από τον τύπο:

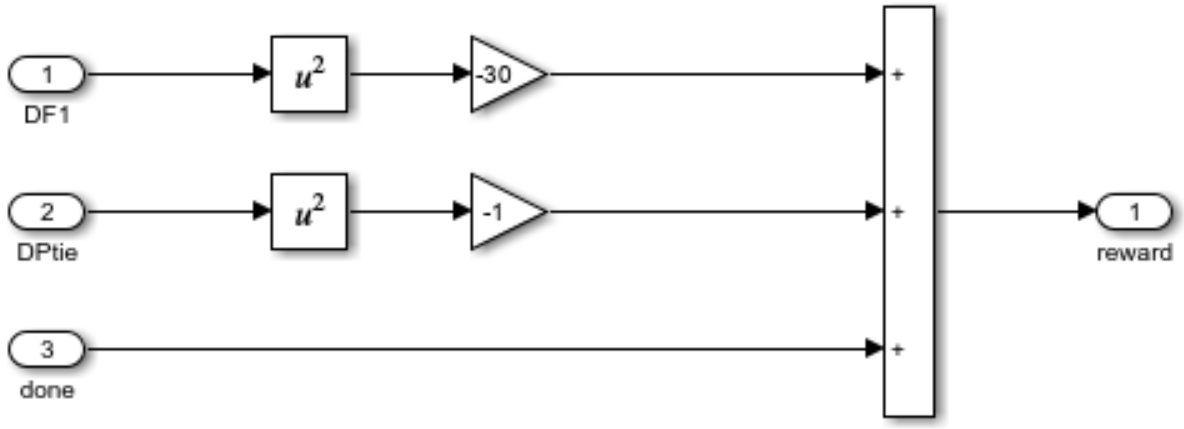
$$reward = -(b\,\Delta {f_1}^2 + \Delta {P_{tie}}^2)$$

Ο κύριος στόχος της συνάρτησης ανταμοιβής είναι να τιμωρήσει τον πράκτορα για οποιαδήποτε απόκλιση των δύο τιμών. Το αρνητικό πρόσημο στην εξίσωση εξασφαλίζει ότι οι αποκλίσεις οδηγούν σε μείωση της ανταμοιβής, ενθαρρύνοντας τον πράκτορα να διατηρεί τις αποκλίσεις όσο το δυνατόν πιο κοντά στο μηδέν.

Απώτερος στόχος της εκπαίδευσης είναι η μεγιστοποίηση των συσσωρευμένων ανταμοιβών. Ιδανικά, ο πράκτορας θα πρέπει να συσσωρεύει μια συνολική ανταμοιβή όσο το δυνατόν πιο κοντά στο μηδέν στο τέλος κάθε επεισοδίου εκπαίδευσης.

Η σταθερά b επιλέχθηκε να είναι 30 μετά από πολλές δοκιμές με διαφορετικές τιμές. Η επιλογή αυτής της τιμής βασίστηκε στην ικανότητά της να εξισορροπεί τη σημασία της απόκλισης συχνότητας σε σχέση με την απόκλιση ισχύος διασύνδεσης εξασφαλίζοντας ότι ο πράκτορας δίνει την κατάλληλη προσοχή και στις δύο παραμέτρους.

Στο Σχήμα 27 φαίνεται το υποσύστημα υπολογισμού των ανταμοιβών.



Σχήμα 27: Υποσύστημα υπολογισμού Ανταμοιβών

4.2. Σχεδιασμός Πράκτορα

Η επιλογή της κατάλληλης μεθόδου RL για το σχεδιασμό και την εκπαίδευση ενός πράκτορα είναι καίριας σημασίας για την εξαγωγή μιας αποδοτικής πολιτικής ελέγχου για το εκάστοτε πρόβλημα. Μετά από αρκετούς πειραματισμούς, επιλέχθηκε η χρήση της μεθόδου Double Deep Q-Network [130], μιας βελτιστοποιημένης παραλλαγής της Deep Q-Network (DQN), μιας επέκτασης του Q-Learning.

4.2.1. Double Deep Q-Network (Double DQN)

Ο αλγόριθμος Double DQN είναι μια επέκταση του αλγορίθμου DQN, σχεδιασμένος να αντιμετωπίζει το πρόβλημα της υπερεκτίμησης των Q-τιμών (Q-value overestimation) που συχνά εμφανίζεται στις μεθόδους Q-learning. Συνδυάζει τα πλεονεκτήματα του DQN με μια πιο ισχυρή προσέγγιση στην εκτίμηση των Q-τιμών, οδηγώντας σε βελτιωμένη σταθερότητα και απόδοση [6].

4.2.1.1. Διαμόρφωση του Πράκτορα

Στο Double DQN, ο πράκτορας αλληλεπιδρά με ένα περιβάλλον για να μάθει μια βέλτιστη πολιτική (optimal policy) εκτιμώντας τη συνάρτηση Q-τιμών (Q-value function), η οποία αντιπροσωπεύει την αναμενόμενη συσσωρευτική ανταμοιβή (cumulative reward) για την εκτέλεση μιας ενέργειας σε μια δεδομένη κατάσταση. Είναι αλγόριθμος κατάλληλος για περιβάλλοντα με συνεχή χώρο παρατηρήσεων. Ο πράκτορας αποτελείται από τα ακόλουθα βασικά στοιχεία:

Q-Network (Online Network): Το Q-network είναι ένα νευρωνικό δίκτυο με παραμέτρους θ. Προσεγγίζει τη συνάρτηση Q-τιμών , $Q(s, \alpha; \theta)$, όπου:

- s είναι η τρέχουσα κατάσταση (state)
- α είναι η ενέργεια (action) που εκτελείται

- θ είναι οι παράμετροι του νευρωνικού δικτύου.

Αυτό το δίκτυο χρησιμοποιείται για την επιλογή ενεργειών κατά τη διάρκεια της εκπαίδευσης και της αξιολόγησης.

Target Network: Το target network είναι ένα ξεχωριστό νευρωνικό δίκτυο με παραμέτρους θ^- Χρησιμοποιείται για τον υπολογισμό των target Q-τιμών κατά τη διάρκεια της εκπαίδευσης.

Το target network ενημερώνεται περιοδικά, κάθε C βήματα αντιγράφοντας τα βάρη από το online network :

$$\theta^- \leftarrow \theta$$

Η χρήση του target network βοηθά στη σταθεροποίηση της εκπαίδευσης παρέχοντας σταθερές target τιμές.

Προσωρινή Μνήμη Εμπειριών (Experience Replay Buffer): Η experience replay buffer αποθηκεύει πλειάδες (tuples) από προηγούμενες εμπειρίες $(s_t, \alpha_t, r_t, s_{t+1}, done)$, όπου:

- s_t είναι η κατάσταση (state) τη χρονική στιγμή t
- α_t είναι η ενέργεια (action) που εκτελέστηκε τη χρονική στιγμή t
- r_t είναι η ανταμοιβή (reward) που ελήφθη
- s_{t+1} είναι η επόμενη κατάσταση (next state)
- $done$ είναι μια σημαία (flag) που υποδεικνύει αν το επεισόδιο έχει ολοκληρωθεί.

Κατά τη διάρκεια της εκπαίδευσης, δείγματα (mini-batches) από εμπειρίες επιλέγονται τυχαία από το replay buffer για να σπάσει η συσχέτιση μεταξύ διαδοχικών δειγμάτων και να βελτιωθεί η αποδοτικότητα της μάθησης.

Επιλογή Ενέργειας (Action Selection): Ο πράκτορας χρησιμοποιεί μια πολιτική ε-greedy (ε-greedy policy) για να ισορροπήσει την εξερεύνηση και την εκμετάλλευση:

- Με πιθανότητα ϵ, ο πράκτορας επιλέγει μια τυχαία ενέργεια (exploration).
- Με πιθανότητα 1−ϵ, ο πράκτορας επιλέγει την ενέργεια με την υψηλότερη Q-τιμή σύμφωνα με το online network (exploitation).

4.2.1.2. Αλγόριθμος Εκπαίδευσης

Βήμα 1: Αρχικοποίηση Δικτύων και Replay Buffer

- Αρχικοποίηση του online network $Q(s, \alpha\,; \theta)$ με τυχαία βάρη θ
- Αρχικοποίηση του target network $Q(s, \alpha\,;\, \theta^{-})$ με τυχαία βάρη θ^{-}
- Αρχικοποίηση άδειου Replay Buffer.

Βήμα 2: Αλληλεπίδραση με το Περιβάλλον

Για κάθε χρονική στιγμή t:

1. Παρατήρηση της τρέχουσας κατάστασης s_t
2. Επιλογή ενέργειας α_t χρησιμοποιώντας την πολιτική ε-greedy:

$$\alpha_t = \begin{cases} \text{τυχαία ενέργεια, με πιθανότητα } \varepsilon \\ argmax_{\alpha} Q(s, \alpha; \theta)\text{ , με πιθανότητα } 1-\varepsilon \end{cases}$$

3. Εκτέλεση της ενέργειας α_t στο περιβάλλον και παρατήρηση της ανταμοιβής r_t, της επόμενης κατάστασης s_{t+1} και του $done$
4. Αποθήκευση της εμπειρίας $(s_t, \alpha_t, r_t, s_{t+1}, done)$ στο Replay Buffer.

Βήμα 3: Δειγματοληψία Mini-Batch από την προσωρινή μνήμη εμπειριών

Τυχαία επιλογή ενός mini-batch M στοιχείων από εμπειρίες $(s_t, \alpha_t, r_t, s_{t+1}, done)$ από το Replay Buffer.

Βήμα 4: Υπολογισμός Target Q-Τιμών

Για κάθε εμπειρία στο mini-batch:

- Αν $done_i$ είναι true (το επεισόδιο ολοκληρώθηκε), η target Q-τιμή ορίζεται ως η άμεση ανταμοιβή:

$$y_i = r_i$$

- Αν $done_i$ είναι false, χρησιμοποιείται η προσέγγιση Double DQN για τον υπολογισμό της target Q-τιμής:
    1. Χρησιμοποιείται το online network για επιλογή της καλύτερης ενέργειας για την επόμενη κατάσταση:

$$\alpha_{i+1}^{*} = \; argmax_a Q(s_{i+1}, a; \theta)$$

    2. Χρησιμοποιείται το target network για την αξιολόγηση της Q-τιμής για την επιλεγμένη ενέργεια:

$$y_i = r_i + \gamma Q(s_{i+1}, a^*_{i+1}; \theta^-)$$

όπου γ είναι ο παράγοντας έκπτωσης (Discount Factor), $0 < \gamma < 1$

Βήμα 5: Ενημέρωση του Online Network

1. Υπολογισμός της απώλειας (loss) ως Μέσο Τετραγωνικό Σφαλμα (Mean Squared Error) μεταξύ των προβλεπόμενων Q-τιμών και των target Q-τιμών:

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - Q(s_{i+1}, a; \theta))^2$$

όπου M είναι το μέγεθος του mini-batch.

2. Στη συνέχεια εκτέλεση μεθόδου Διαβάθμισης Κλίσης (Gradient Descent) για βελτιστοποίηση των βαρών του δικτύου.

Βήμα 6: Ενημέρωση του Target Network

Το target network ενημερώνεται περιοδικά αντιγράφοντας τα βάρη από το online network:

$$\theta^- \leftarrow \theta$$

Αυτή η ενημέρωση μπορεί να γίνεται κάθε C βήματα ή στο τέλος κάθε επεισοδίου.

Βήμα 7: Επανάληψη

Τα Βήματα 2-6 επαναλαμβάνονται μέχρι ο πράκτορας να συγκλίνει σε μια βέλτιστη πολιτική ή να φτάσει σε έναν προκαθορισμένο αριθμό επεισοδίων.

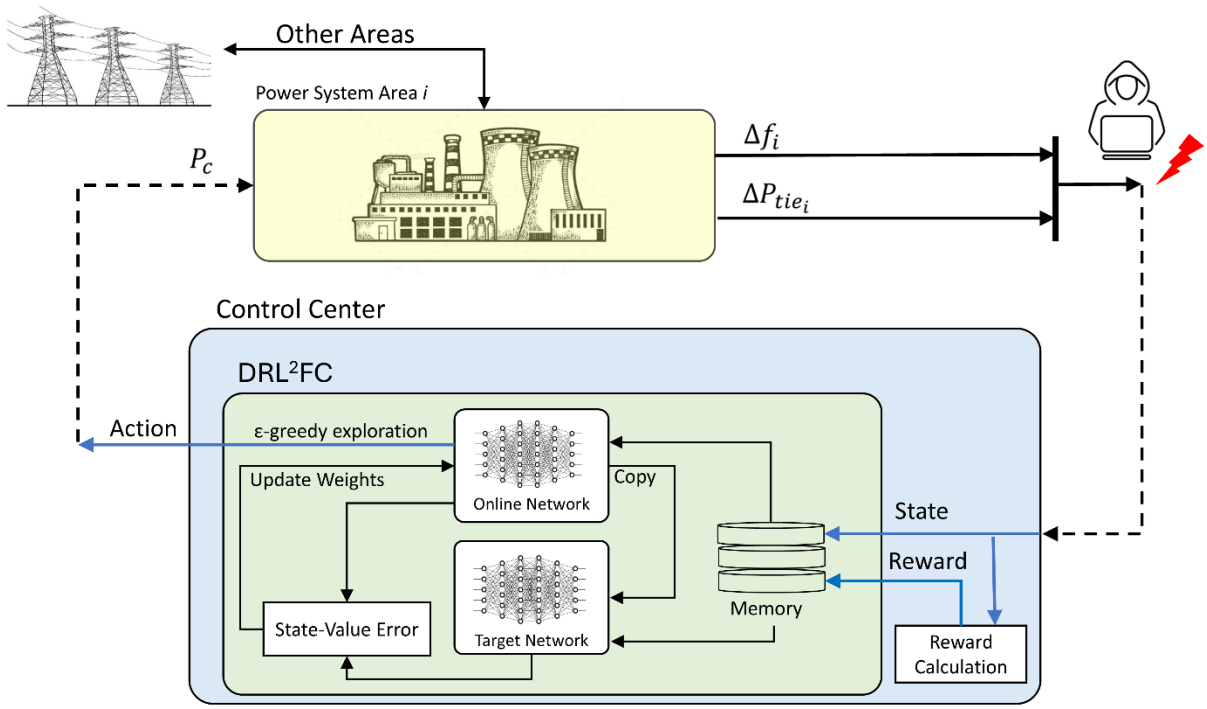
Η διαδικασία εκπαίδευσης παράγει ως έξοδο μια πολιτική ελέγχου $\pi(\alpha|s)$, κωδικοποιημένη στα βάρη θ του online δικτύου $Q(s, \alpha; \theta)$. Αυτή η πολιτική αντιστοιχίζει την τρέχουσα κατάσταση του συστήματος στη βέλτιστη ενέργεια ελέγχου , επιτρέποντας στον πράκτορα να ανταποκρίνεται αποτελεσματικά σε δυναμικές συνθήκες και διαταραχές σε πραγματικό χρόνο.

4.2.2. Υλοποίηση του Πράκτορα και Εκπαίδευση

Κάθε ένα από τα νευρωνικά δίκτυα του πράκτορα αποτελούνταν από δύο κρυφά στρώματα των 128 κόμβων, με συνάρτηση ενεργοποίησης Rectified Linear Unit (ReLU). Για λόγους υπολογιστικής απλούστευσης, ο χώρος ενεργειών διακριτοποιήθηκε στο εύρος [-0.06, 0.06] με βήμα 0.01. Όσον αφορά την επιλογή ενεργειών, ο παράγοντας ε εκκίνησε από τη μονάδα και μειωνόταν με βήμα 0.001.

Η εκπαίδευση του πράκτορα πραγματοποιήθηκε σε δύο στάδια. Αρχικά έμαθε να λειτουργεί σαν απλός ελεγκτής LFC και να διαχειρίζεται απότομες αλλαγές στο φορτίο της περιοχής 1. Στη συνέχεια, στο περιβάλλον εκπαίδευσης προτέθηκαν FDIAs, με σκοπό να μάθει να τις διακρίνει και να τις αντιμετωπίζει αποτελεσματικά. Το νευρωνικό δίκτυο που προέκυψε ονομάστηκε $\boldsymbol{DRL^2FC}$.

Το online δίκτυο ενημερωνόταν σε κάθε βήμα, ενώ το target δίκτυο ενημερωνόταν στο τέλος κάθε επεισοδίου χρησιμοποιώντας τον αλγόριθμο SGDM με ρυθμό μάθησης (learning rate) 0.0001 και παράγοντα έκπτωσης γ=0.98. Η εκπαίδευση αξιοποίησε μια μνήμη επανάληψης χωρητικότητας 10000 εμπειριών, εξάγοντας από αυτή mini-batches των 256 εμπειριών σε κάθε βήμα.



Σχήμα 28: Διαδικασία εκπαίδευσης πράκτορα Double DQN σε σύστημα LFC

4.2.2.1. Εκπαίδευση στη Διαχείριση Αλλαγών Φορτίου

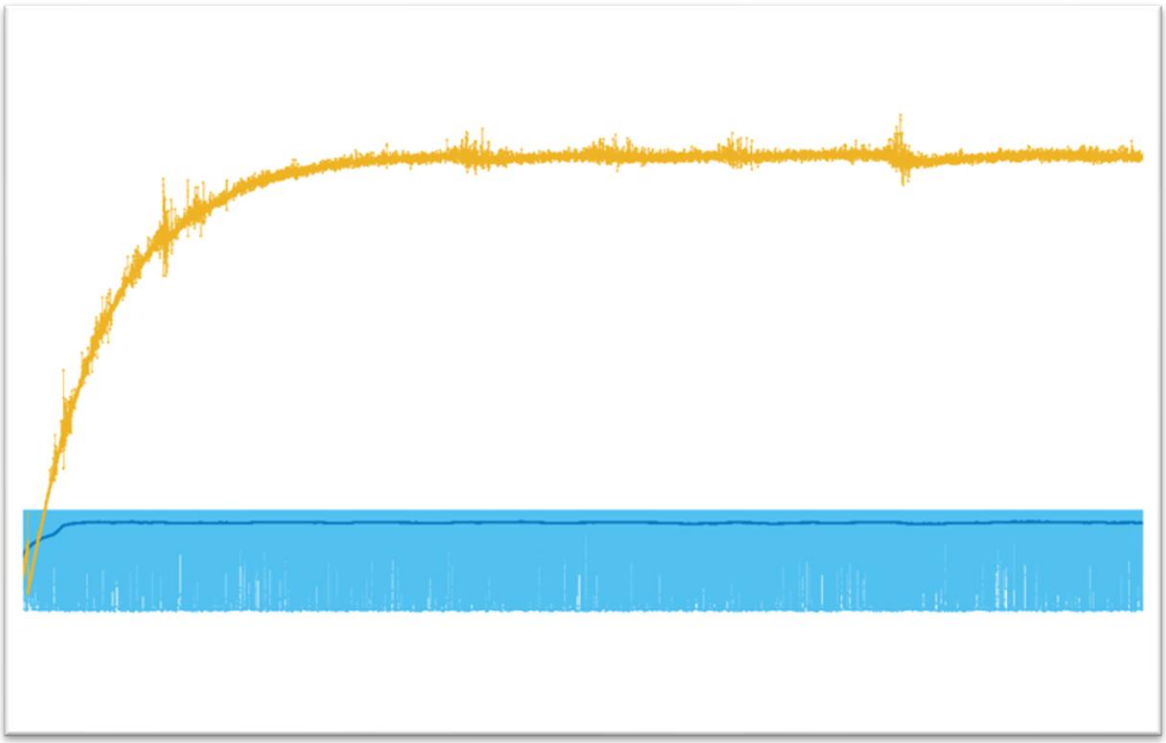
Η διαδικασία εκπαίδευσης του πράκτορα στο σύστημα LFC δύο περιοχών διήρκεσε 10000 επεισόδια, με κάθε επεισόδιο να αποτελείται από 3000 βήματα με χρονικό βήμα (timestep) 0.1 δευτερόλεπτα.

Κατά τη διάρκεια κάθε επεισοδίου, κάθε 30 δευτερόλεπτα επενεργούσε στο σήμα P_{L_i} βηματική διαταραχή τυχαίας τιμής στην περιοχή [-0.05, 0.05].

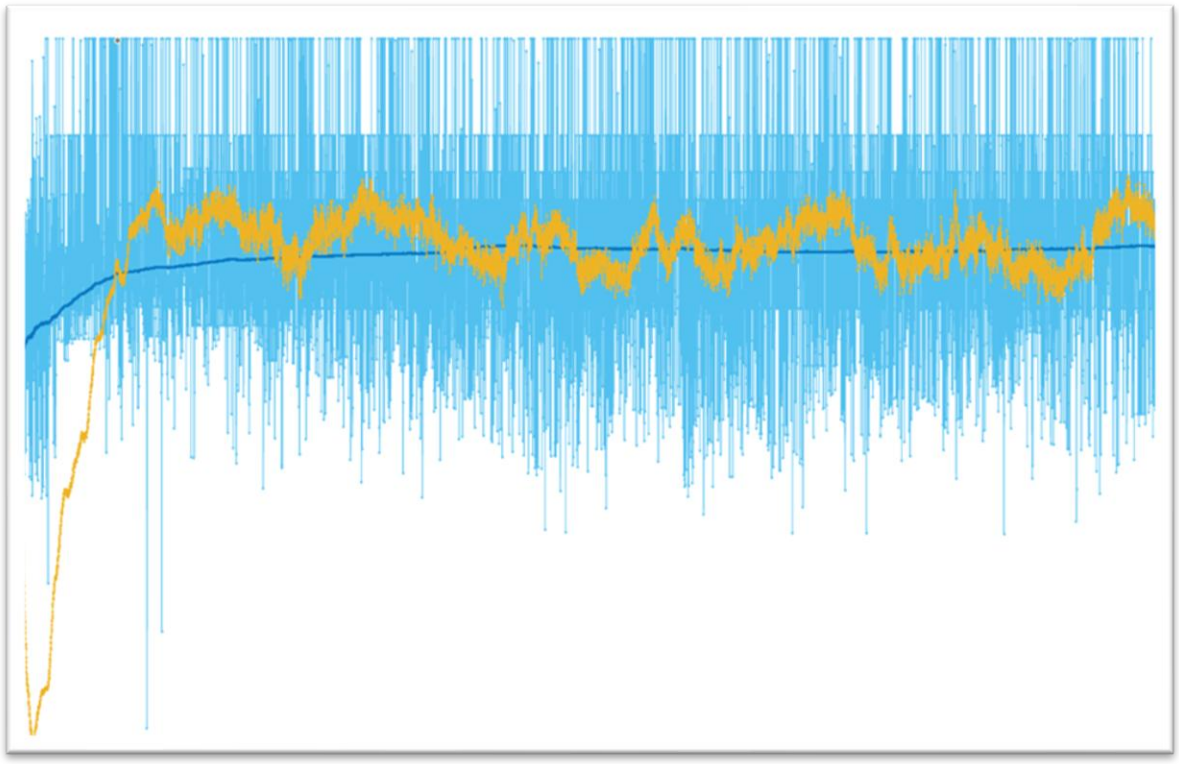
Η πρόοδος του DRL^2FC φαίνεται στο Σχήμα 29 που έχει εξαχθεί από το περιβάλλον εκπαίδευσης του MATLAB Reinforcement Learning Toolbox. Η κίτρινη γραμμή αντιπροσωπεύει τις αναμενόμενες ανταμοιβές του πράκτορα, οι γαλάζιες κουκίδες την απόδοση κάθε επεισοδίου, ενώ η μπλε σκούρα γραμμή αναπαριστά τον κινητό μέσο όρο της απόδοσης σε 200 επεισόδια.

4.2.2.2. Εκπαίδευση υπό Κυβερνοεπιθέσεις

Για να ενισχυθεί η ανθεκτικότητα έναντι κυβερνοεπιθέσεων, το DRL^2FC εκπαιδεύτηκε εκ νέου για 2000 επεισόδια χρησιμοποιώντας την ίδια διαμόρφωση, με τυχαίες βηματικές διαταραχές στην περιοχή $[-0.05,0.05]$ να εισάγονται στα κανάλια μεταφοράς δεδομένων για να προσομοιωθούν FDIAs. Η πρόοδος του πράκτορα φαίνεται στο Σχήμα 30 που έχει εξαχθεί από το περιβάλλον εκπαίδευσης του MATLAB Reinforcement Learning Toolbox



Σχήμα 29: Πρόοδος εκπαίδευσης πράκτορα σε κανονικές συνθήκες λειτουργίας

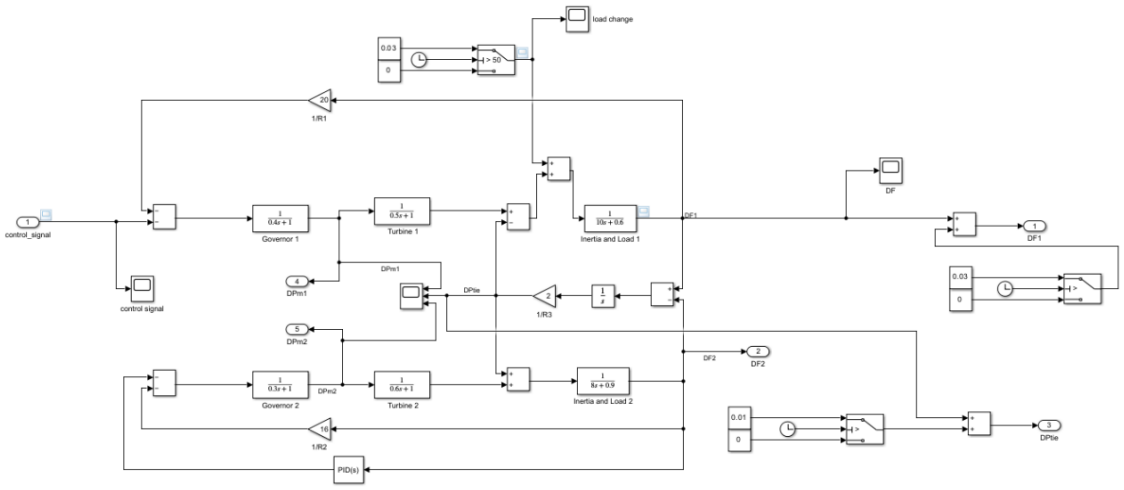


Σχήμα 30: Πρόοδος εκπαίδευσης πράκτορα σε περιβάλλον υπό FDIAs

5

ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Για την πειραματική αξιολόγηση του εκπαιδευμένου πράκτορα DRL^2FC, διαμορφώθηκε σε περιβάλλον MATLAB/Simulink ένα σύστημα ελέγχου φορτίου συχνότητας δύο περιοχών, με χαρακτηριστικά όμοια με του συστήματος στο οποίο έγινε η εκπαίδευση.

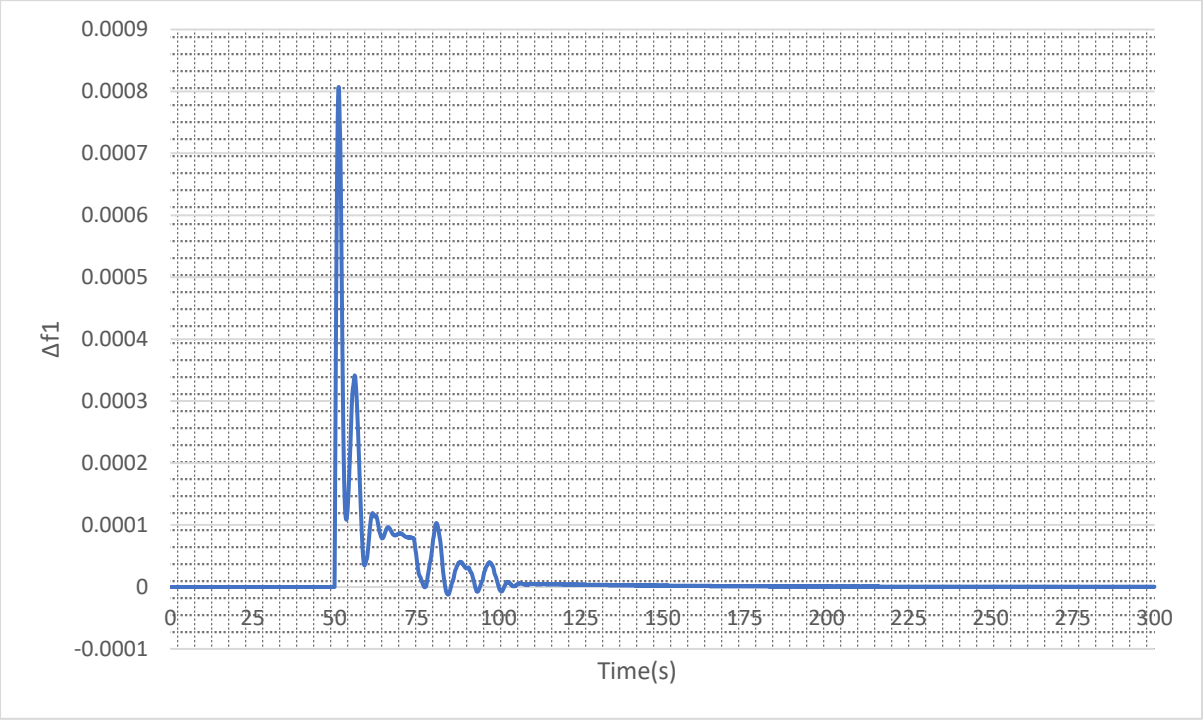


Σχήμα 31: Σύστημα Δοκιμών

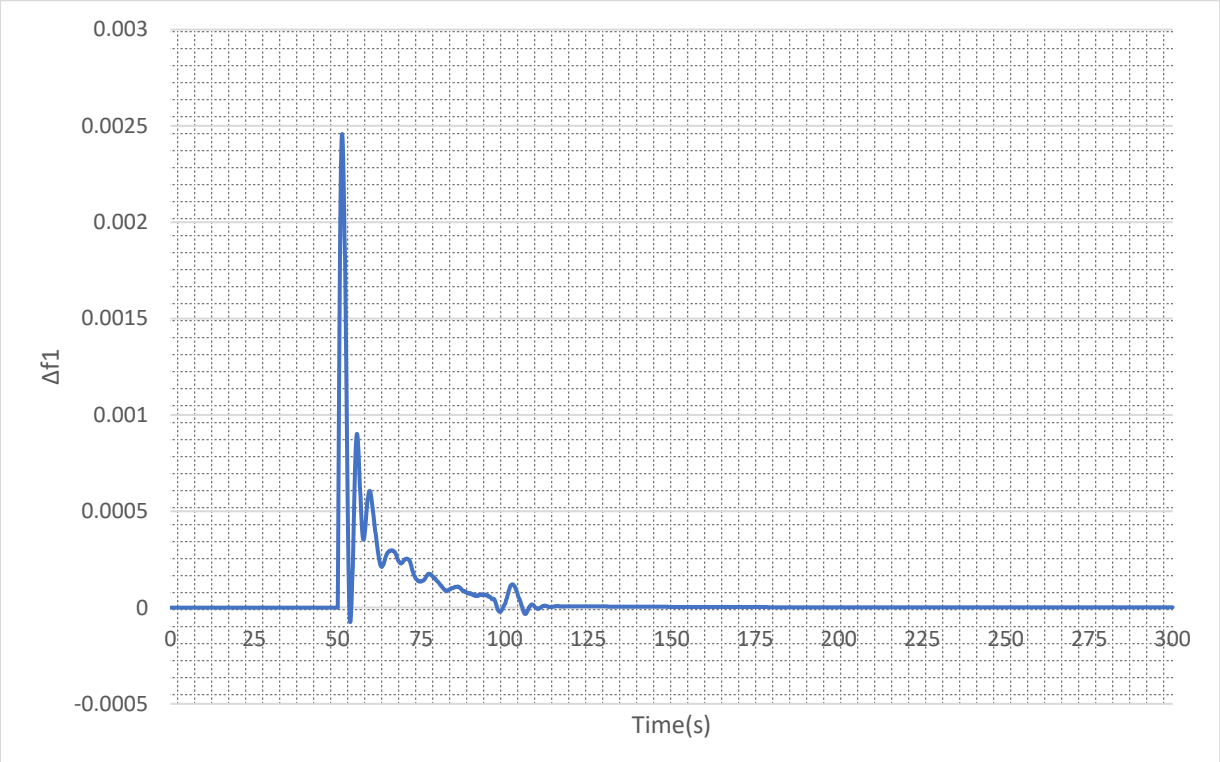
5.1. Ικανότητα Διαχείρισης Αλλαγών Φορτίου

Αρχικά ο πράκτορας δοκιμάστηκε σε απλές μεταβολές του φορτίου του συστήματος πλάτους 0.01 α.μ. και 0.03 α.μ

Και στις δύο περιπτώσεις, η διακύμανση της συχνότητας του συστήματος επανέρχεται σε μηδενική τιμή, με τρόπο και σε χρόνο αντίστοιχο με ενός συμβατικού ελεγκτή PID.

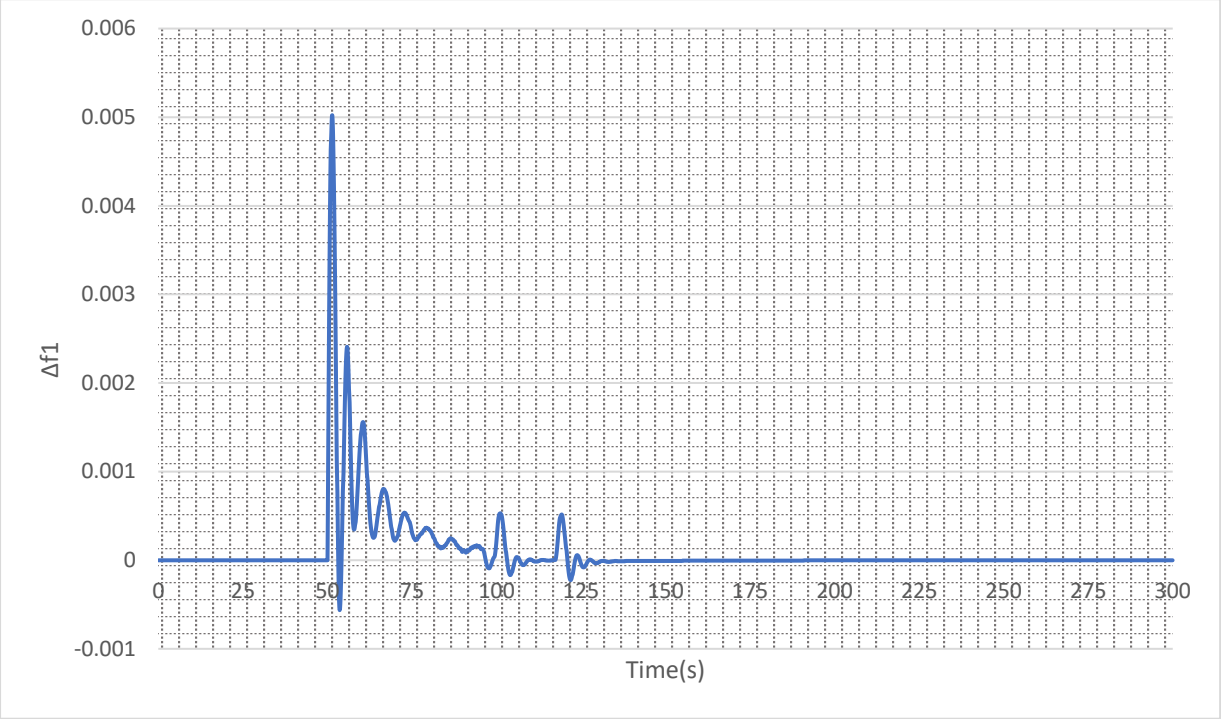


Σχήμα 32: Απόκριση συχνότητας συστήματος με DRL^2FC κατά την απότομη αλλαγή φορτίου πλάτους 0.01 α.μ.



Σχήμα 33: Απόκριση συχνότητας συστήματος με DRL^2FC,κατά την απότομη αλλαγή φορτίου πλάτους 0.03 α.μ.

Για την περαιτέρω αξιολόγηση του DRL^2FC, δοκιμάστηκε και σε μεταβολή φορτίου 0.06 α.μ.. Η μεταβολή αυτή ήταν εκτός του εύρους τιμών στο οποίο έγινε η εκπαίδευση.



Σχήμα 34: Απόκριση συχνότητας συστήματος με DRL^2FC κατά την απότομη αλλαγή φορτίου πλάτους 0.06 α.μ.

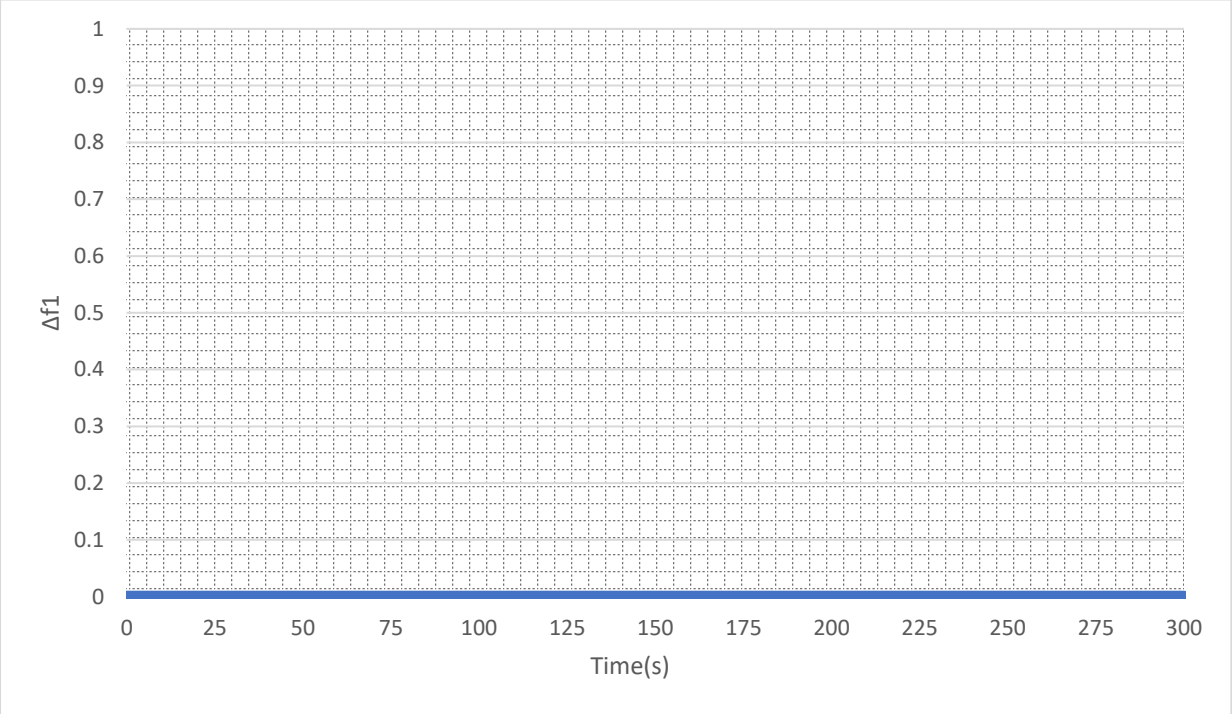
Όπως φαίνεται στο Σχήμα 30, η συχνότητα και σε αυτή την περίπτωση επέστρεψε στην ονομαστική της τιμή. Ο χρόνος που χρειάστηκε το DRL^2FC για να το επιτύχει είναι μεγαλύτερος, περίπου κατά 10%, σε σχέση με τις περιπτώσεις στις οποίες είχε εκπαιδευτεί.

Συγκρίνοντας τα αποτελέσματα αυτά με την απόδοση του PID ελεγκτή, όπως παρουσιάστηκε στο κεφάλαιο 2, ο πράκτορας σημειώνει ικανοποιητικές επιδόσεις χωρίς να υστερεί σε κάποιο από τα σενάρια.

5.2. Επίθεση Τύπου FDI Στη Μέτρηση Συχνότητας

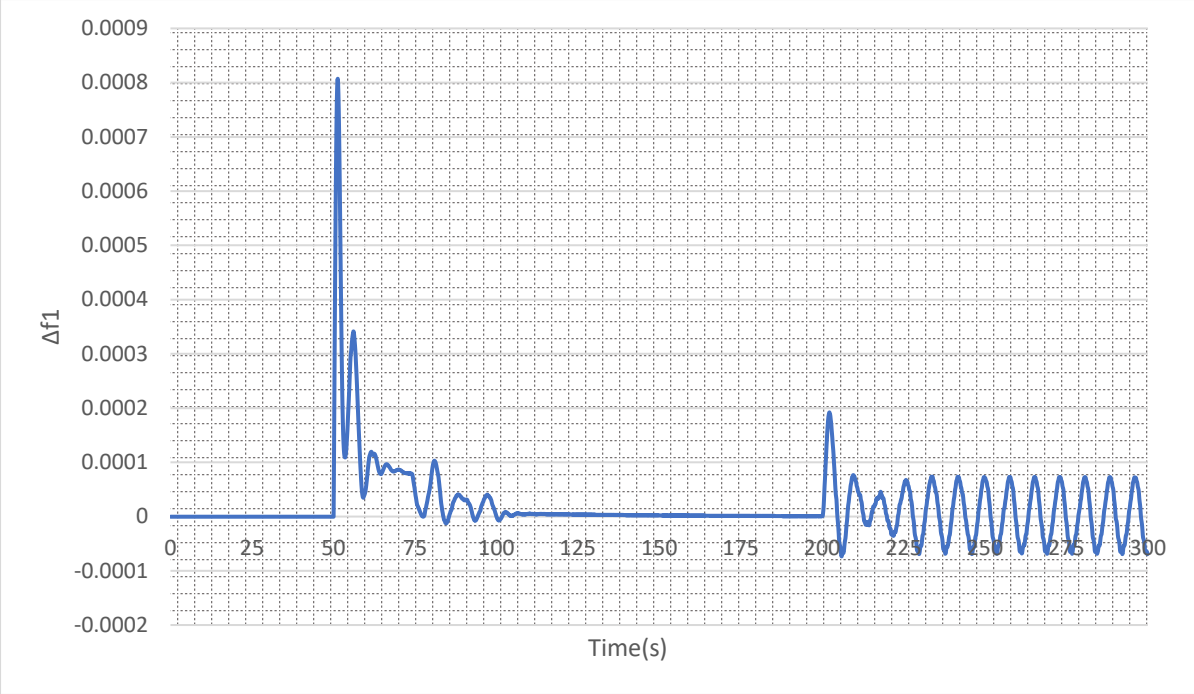
Εφαρμόστηκε στη μέτρηση συχνότητας της περιοχής 1 το σήμα επίθεσης που παρουσιάστηκε στο κεφάλαιο 3 (0.01 α.μ.). Η επίθεση πραγματοποιείται το 100ό δευτερόλεπτο της προσομοίωσης.

Ο πράκτορας αγνοεί απολύτως το σήμα επίθεσης διατηρώντας την τιμή της συχνότητας στην ονομαστική της τιμή. Η επίθεση αποτυγχάνει.

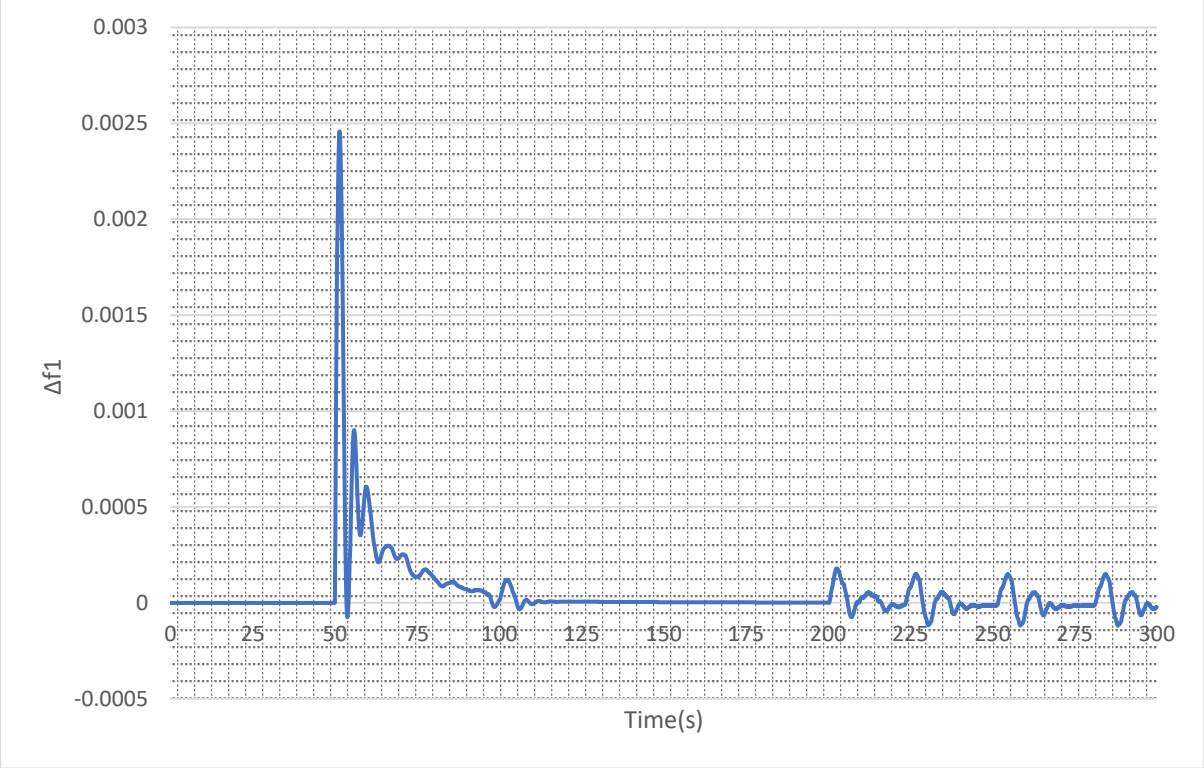


Σχήμα 35: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας υπό ηρεμία

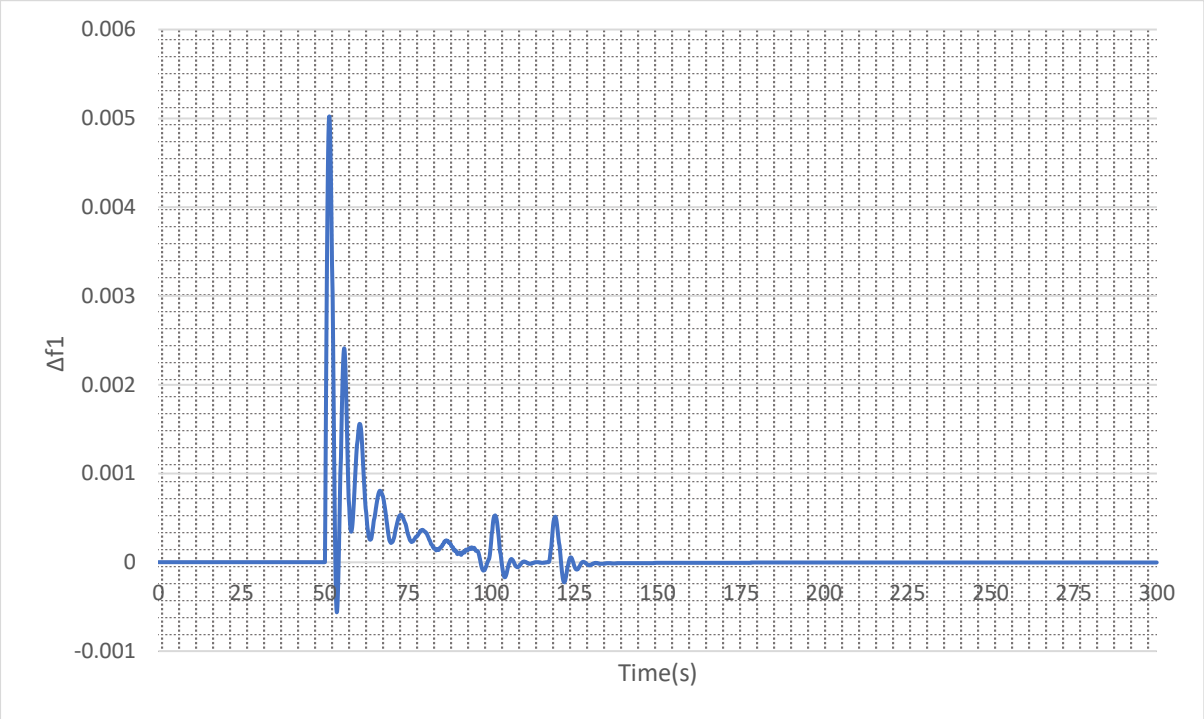
Στη συνέχεια παρουσιάζεται η συμπεριφορά του DRL^2FC σε περιπτώσεις όπου έχει ήδη διαχειριστεί αλλαγή του φορτίου της περιοχής 1. Συγκεκριμένα, τη χρονική στιγμή 50s συμβαίνει η αλλαγή φορτίου, ενώ το σήμα κυβερνοεπίθεσης εφαρμόζεται στο 200ό δευτερόλεπτο.



Σχήμα 36: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας υπό διαχείριση μεταβολής φορτίου 0.01 α.μ.



Σχήμα 37: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας υπό διαχείριση μεταβολής φορτίου 0.03 α.μ.



Σχήμα 38: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας υπό διαχείριση μεταβολής φορτίου 0.06 α.μ.

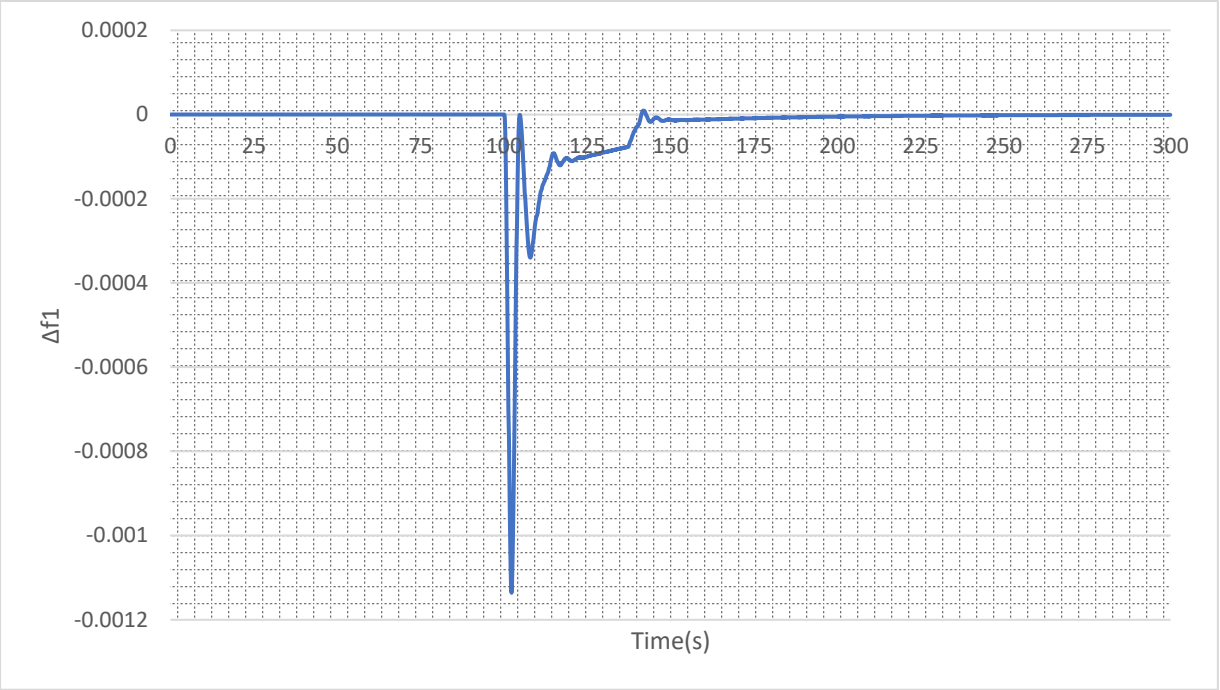
Η επίδραση της επίθεσης είναι εμφανής, σε καμία περίπτωση όμως η συχνότητα δεν αποκλίνει επικίνδυνα από την ονομαστική της τιμή.

Αξίζει να σημειωθεί ότι όσο μεγαλύτερη αλλαγή φορτίου αντιμετωπίζει ο πράκτορας, τόσο μικρότερη είναι η επίδραση του σήματος επίθεσης στο σύστημα, έως ανυπόστατη στην περίπτωση της αλλαγής φορτίου 0.06 α.μ. .

Η απόδοση του DRL^2FC είναι σαφώς ανώτερη σε σχέση με την απόδοση του PID ελεγκτή.

5.3. Επίθεση Τύπου FDI στη Μέτρηση Διασυνδετικής Ροής Ισχύος

Παρόμοιες δοκιμές έγιναν και για την επίθεση στη μέτρηση διασυνδετικής ροής ισχύος. Συγκεκριμένα, σε κατάσταση ηρεμίας το σήμα επίθεσης εφαρμόστηκε στο 100ό δευτερόλεπτο.

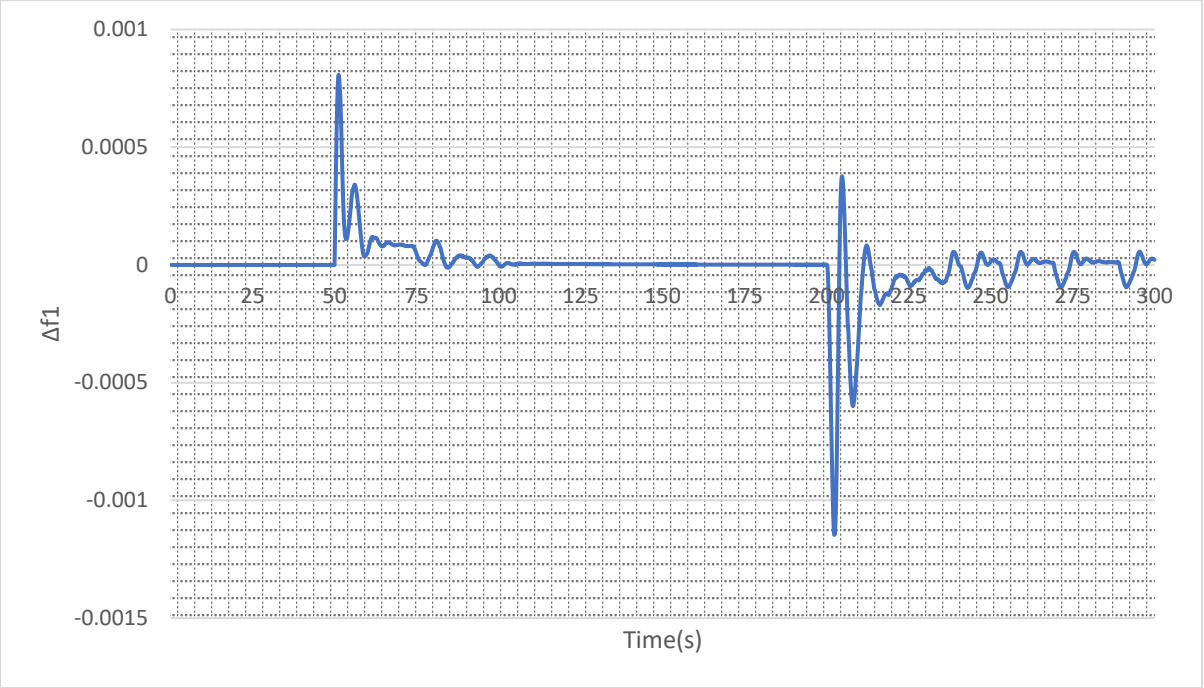


Σχήμα 39: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση διασυνδετικής ισχύος υπό ηρεμία

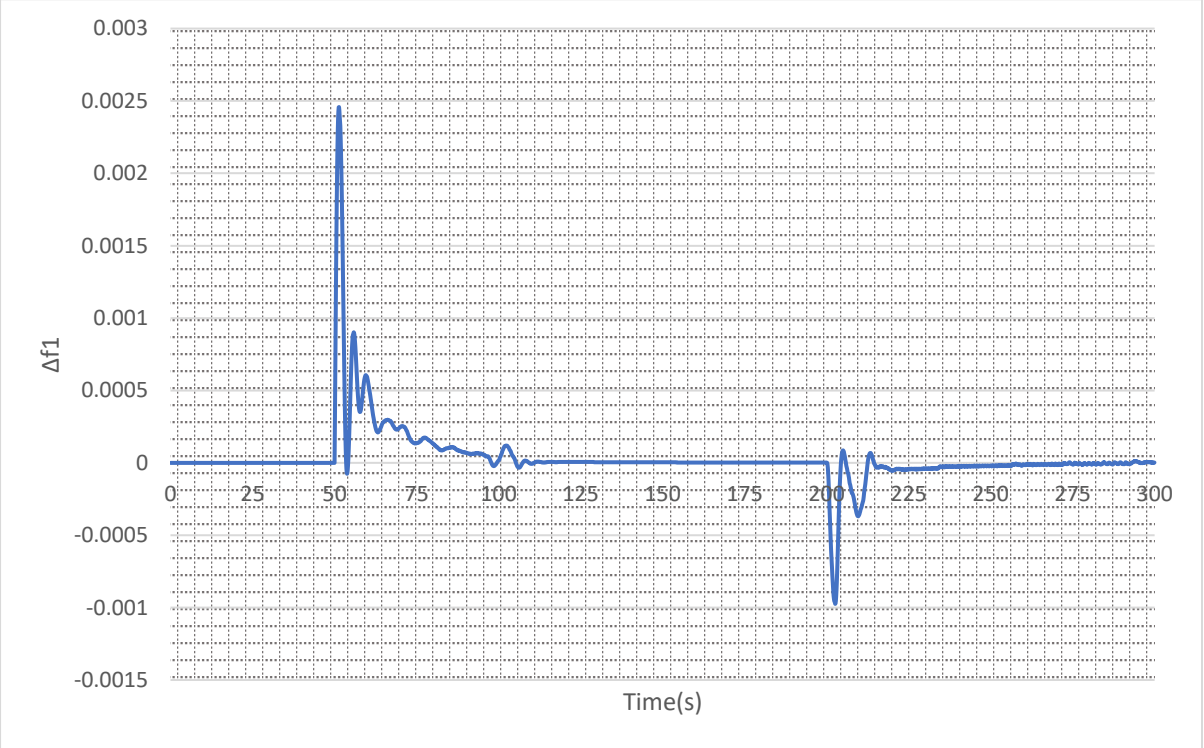
Η συχνότητα του συστήματος αποκλίνει, όμως το DRL^2FC εφαρμόζοντας κατάλληλο σήμα ελέγχου καταφέρνει να εξαλείψει την επίδρασή της σε ικανοποιητικό χρονικό διάστημα.

Συγκρίνοντας το Σχήμα 35 με το αντίστοιχο γράφημα του κεφαλαίου 3 για τον PID ελεγκτή παρατηρείται σημαντικά μικρότερη απόκλιση της συχνότητας καθώς και γρηγορότερη αποκατάστασή της στην ονομαστική τιμή.

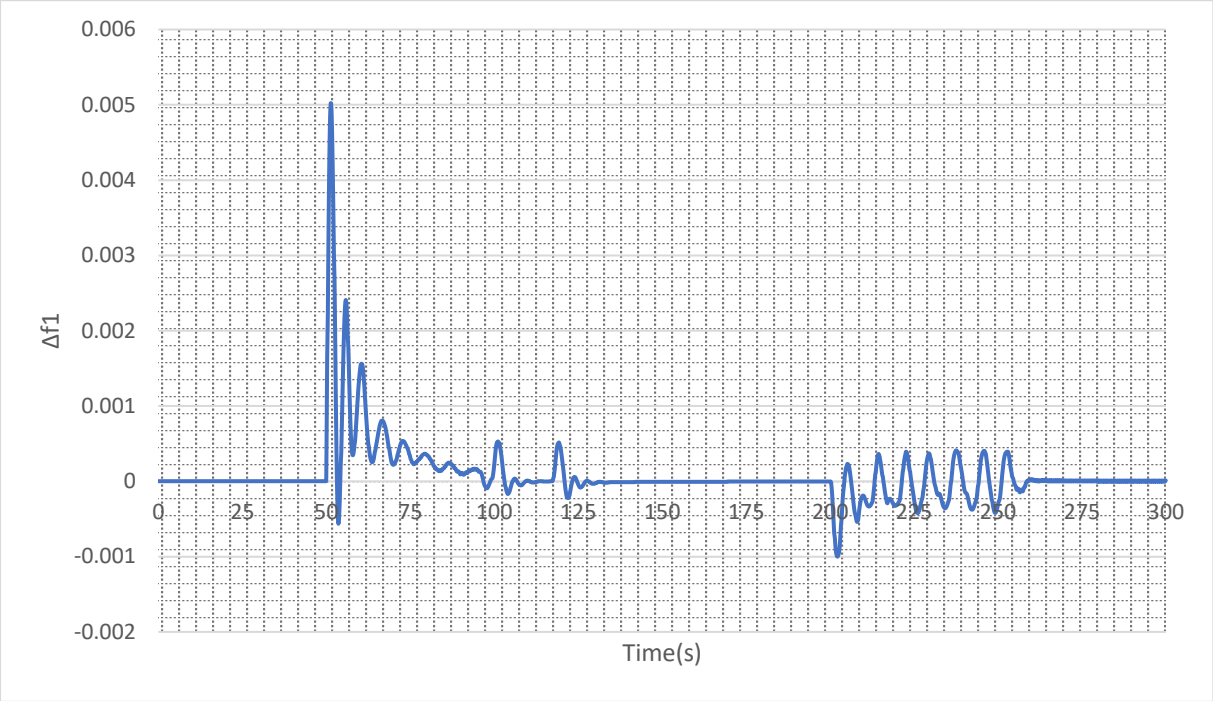
Και σε αυτή τη σειρά δοκιμών εξετάστηκε η συμπεριφορά του πράκτορα ενώ έχει ήδη εφαρμόσει σήμα ελέγχου για τη διαχείριση αλλαγής φορτίου. Προσομοιώθηκαν τρία σενάρια αλλαγής φορτίου στο 50ό δευτερόλεπτο, ακολουθούμενης από σήμα επίθεσης τη χρονική στιγμή t=50s.



Σχήμα 40: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση διασυνδετικής ισχύος υπό διαχείριση μεταβολής φορτίου 0.01 α.μ.



Σχήμα 41: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση διασυνδετικής ισχύος υπό διαχείριση μεταβολής φορτίου 0.03 α.μ.



Σχήμα 42: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση διασυνδετικής ισχύος υπό διαχείριση μεταβολής φορτίου 0.06 α.μ.

Σε κάθε σενάριο, ο πράκτορας αντιλαμβάνεται την επίθεση και επιχειρεί να την αντιμετωπίσει.

Στην πρώτη περίπτωση μετριάζει σημαντικά την επίδρασή της χωρίς να καταφέρει να σταθεροποιήσει το σύστημα στην ονομαστική συχνότητα. Η ταλάντωση της συχνότητας περιορίζεται σε πολύ μικρές τιμές, ενδεχομένως ανεκτές από το σύστημα για την ευσταθή του λειτουργία.

Στις επόμενες περιπτώσεις, όπου έχει διαχειριστεί μεγαλύτερες σε πλάτος μεταβολές, επιτυγχάνεται η σταθεροποίηση της συχνότητας στην ονομαστική της τιμή. Διαπιστώνεται λοιπόν άμεση συσχέτιση του σήματος χειρισμού με την απόδοση του DRL^2FC κατά την επίθεση.

5.4. Ταυτόχρονη Επίθεση Τύπου FDI στη Μέτρηση Συχνότητας και στη Μέτρηση Διασυνδετικής Ροής Ισχύος

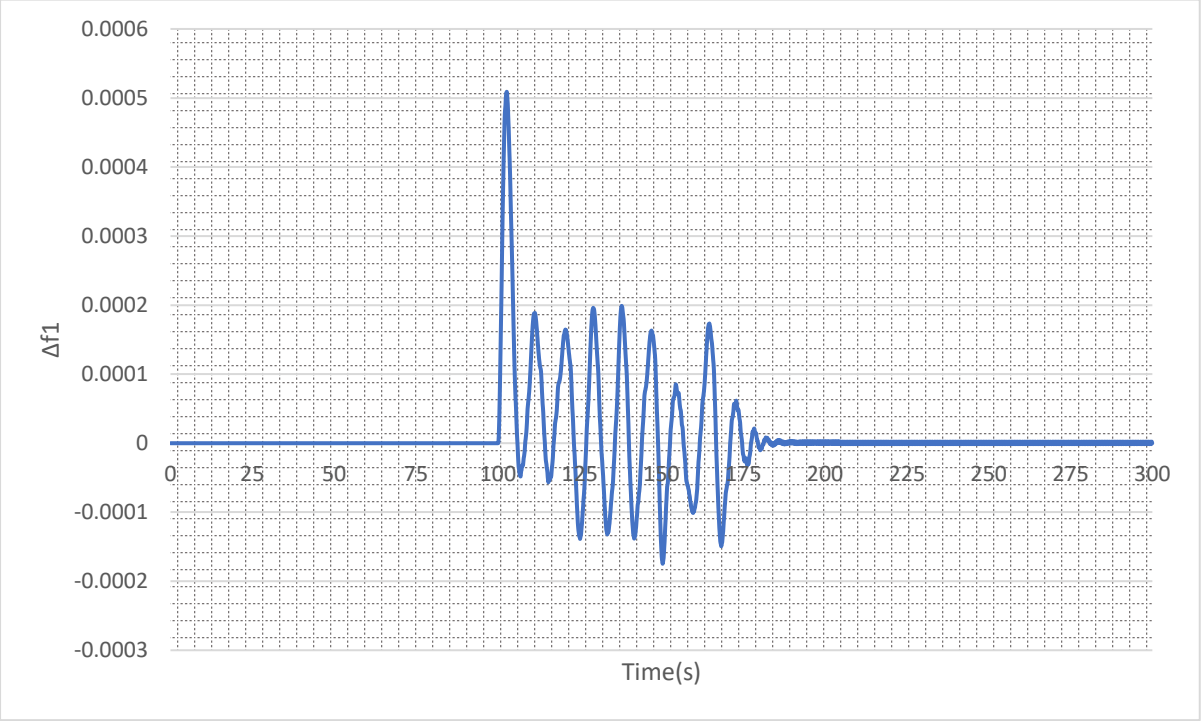
Εφόσον ο επιτιθέμενος έχει τη δυνατότητα έγχυσης ψευδών δεδομένων και στις δύο βασικές μετρήσεις που φτάνουν στον πράκτορα, μπορεί να μεγιστοποιήσει την επίδραση που θα έχει η επίθεση στο σύστημα, όπως φάνηκε στο γράφημα του κεφαλαίου 3.

Ο πράκτορας δοκιμάστηκε στο ίδιο πρότυπο κυβερνοεπίθεσης ενώ βρισκόταν σε κατάσταση ηρεμίας, με την επίθεση να εκδηλώνεται κατά το 100ό δευτερόλεπτο.

Σε αντίθεση με τον ελεγκτή PID, το DRL^2FC αναγνωρίζει και αντιμετωπίζει επιτυχώς την έγχυση σημάτων πλάτους 0.03 α.μ. στη μέτρηση συχνότητας και 0.01 α.μ. στη μέτρηση διασυνδετικής ροής ισχύος, όπως φαίνεται Σχήμα 39.

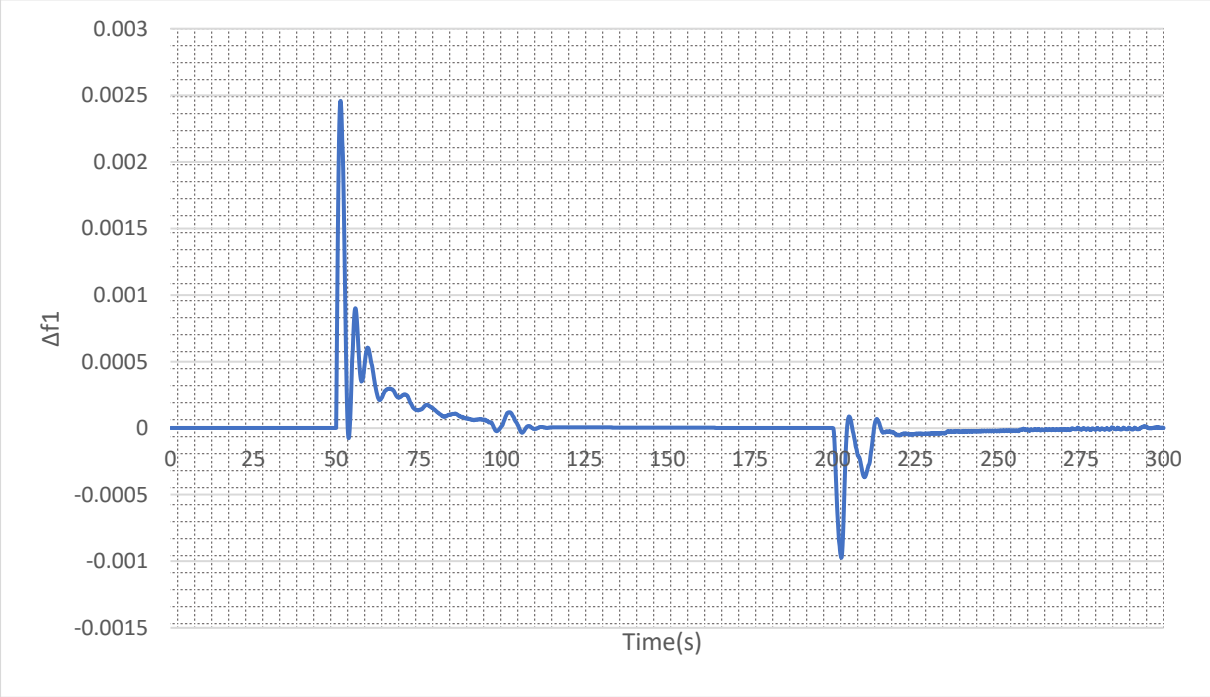
Όμοιο πρότυπο επίθεσης εφαρμόστηκε και για τις τρεις περιπτώσεις κατά της οποίες ο πράκτορας έχει προβεί σε χειρισμό για να επαναφέρει τη συχνότητα μετά από αλλαγή φορτίου. Η συμπεριφορά του πράκτορα είναι παρόμοια με την περίπτωση που παρουσιάστηκε στην παράγραφο 6.3 .

Συνεπώς, η παραποίηση της μέτρησης διασυνδετικής ροής ισχύος επισκιάζει την αντίστοιχη της μέτρησης συχνότητας, παρά τη σημαντική διαφορά πλάτους των σημάτων επίθεσης.

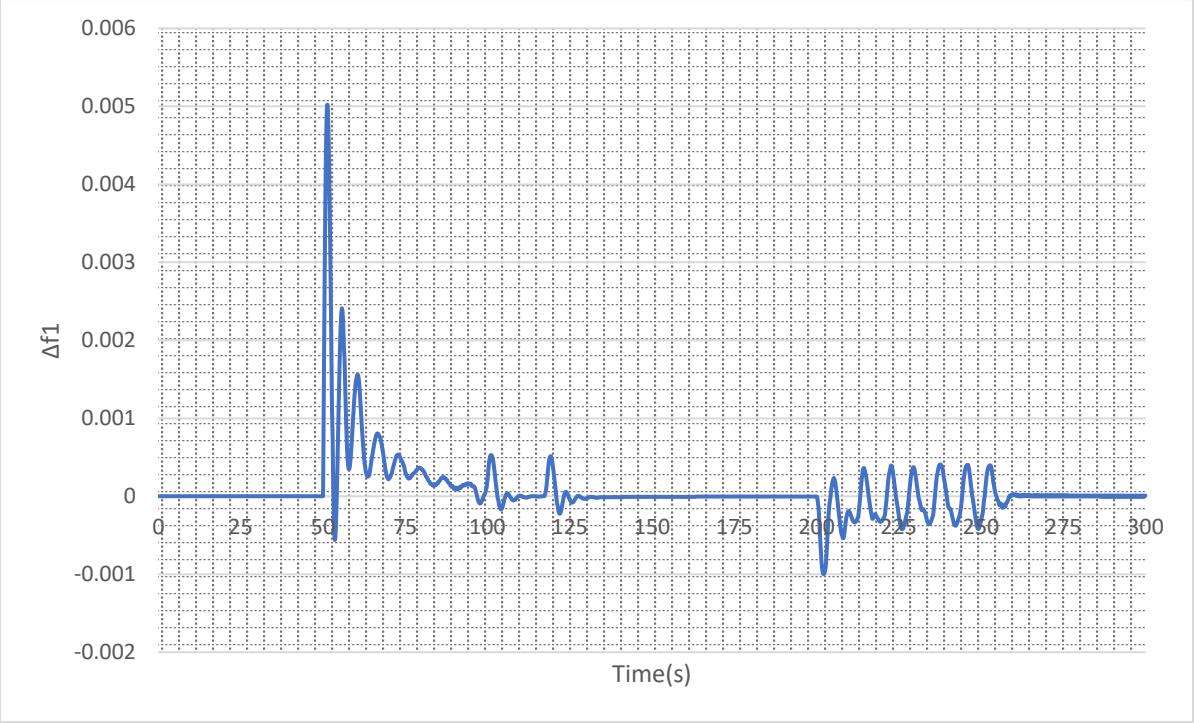


Σχήμα 43: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας και στη μέτρηση διασυνδετικής ισχύος υπό ηρεμία

Σχήμα 44: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας και στη μέτρηση διασυνδετικής ισχύος υπό διαχείριση μεταβολής φορτίου 0.01 α.μ.



Σχήμα 45: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας και στη μέτρηση διασυνδετικής ισχύος υπό διαχείριση μεταβολής φορτίου 0.03 α.μ.



Σχήμα 46: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας και τη σμέτρηση διασυνδετικής ισχύος υπό διαχείριση μεταβολής φορτίου 0.06 α.μ.

5.5. Απόδοση σε Περιβάλλον με Θόρυβο

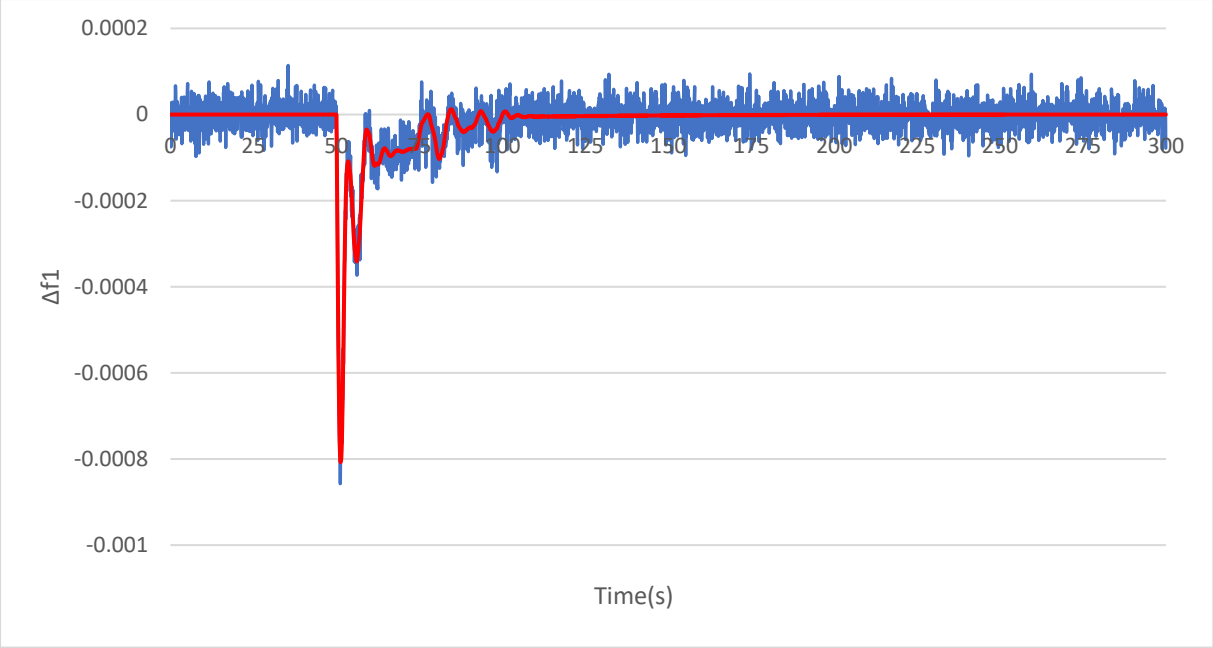
Ενώ η ανάλυση απόδοσης του πράκτορα που προηγήθηκε παρουσιάζει ενθαρρυντικά αποτελέσματα, δεν μπορεί να εγγυηθεί τη βιωσιμότητά του σε περιβάλλοντα υψηλής αβεβαιότητας. Για να διερευνηθεί η ανθεκτικότητα του DRL^2FC υπό τέτοιες συνθήκες, πραγματοποιήθηκε ανάλυση αβεβαιότητας σε θορυβώδη περιβάλλοντα.

Το θορυβώδες περιβάλλον προσομοιώνεται με την έγχυση Γκαουσιανού θορύβου στις μετρήσεις του συστήματος LFC στην τρίτη περίπτωση της παραγράφου 7.4. Τυπικά, ο εγχυμένος θόρυβος τροποποιεί τον χώρο δράσεων σε συνάρτηση με τον χρόνο t ως εξής:

$$s_n(t) = s(t) + w(t)$$

όπου το s_n συμβολίζει το θορυβώδη χώρο ενεργειών και σ τον εγχυμένο θόρυβο, δηλαδή $w(t) \sim N(0, \sigma^2)$.

Στο Σχήμα 47 παρουσιάζεται η απόκριση συχνότητας του συστήματος LFC χωρίς θόρυβο και με θόρυβο, υπό την FDIA που παρουσιάστηκε στο τρίτο σενάριο της προηγούμενης παραγράφου.



Σχήμα 47: Απόκριση συχνότητας συστήματος με DRL^2FC σε FDIA στη μέτρηση συχνότητας και τη μέτρηση διασυνδετικής ισχύος σε περιβάλλον με και χωρίς θόρυβο.

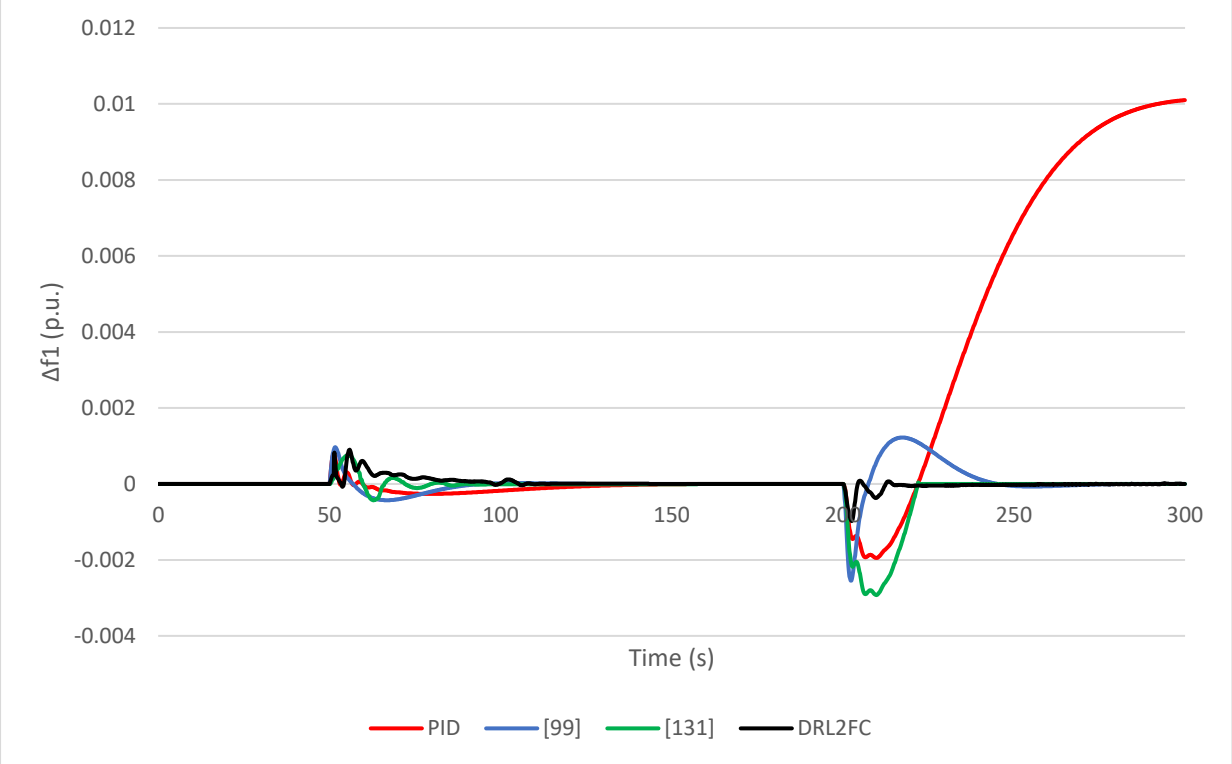
Ο πράκτορας κατάφερε να χειριστεί την αλλαγή φορτίου αποτελεσματικά και να ελαχιστοποιήσει τον αντίκτυπο της επίθεσης στο σύστημα, καθώς συμπεριφέρεται παρόμοια με τις ιδανικές συνθήκες. Τα αποτελέσματα επιβεβαιώνουν την ανθεκτικότητα της προτεινόμενης μεθοδολογίας έναντι θορυβωδών μετρήσεων, παρά τις προκλήσεις που προκαλούν τέτοια φαινόμενα.

5.6. Σύγκριση Απόδοσης σε Σχέση με Άλλους Ελεγκτές Ανθεκτικούς σε Κυβερνοεπιθέσεις

Για να αναδειχθεί η επιστημονική συνεισφορά αυτής της μεθοδολογίας, πραγματοποιήθηκε συγκριτική ανάλυση μεταξύ του πράκτορα και άλλων υφιστάμενων τεχνικών ελέγχου LFC ανθεκτικών σε επιθέσεις. Προς αυτόν τον σκοπό, στο DRL^2FC εφαρμόζεται αλλαγή φορτίου και κυβερνοεπίθεση, μαζί με έναν βελτιστοποιημένο PID ελεγκτή και τις μεθοδολογίες που περιγράφονται στα [131] και [99]. Οι αντίστοιχες αποκρίσεις συχνότητας των συγκρινόμενων συστημάτων απεικονίζονται στο Σχήμα 48.

Αν και όλοι οι εξεταζόμενοι ελεγκτές μπορούν να σταθεροποιήσουν το LFC μετά την αλλαγή φορτίου στα 50s, η συμπεριφορά τους κατά τη διάρκεια κυβερνοεπιθέσεων διαφέρει σημαντικά. Συγκεκριμένα, το σύστημα LFC που καθοδηγείται από τον βελτιστοποιημένο PID απορυθμίζεται πλήρως όταν αντιμετωπίζει την FDIA στα 200s. Επιπλέον, ενώ οι μέθοδοι των [131] και [99] μπορούν να αποκαταστήσουν τη συχνότητα του συστήματος στην ονομαστική της τιμή υπό την κυβερνοεπίθεση, παρουσιάζουν σημαντικά μεγαλύτερους χρόνους ανόδου και μεγαλύτερες υπερτάσεις σε σύγκριση με τον πράκτορα.

Τα αποτελέσματα αποδεικνύουν την υπεροχή του πράκτορα έναντι άλλων σχετικών τεχνικών ελέγχου, υπογραμμίζοντας περαιτέρω την αξία της προτεινόμενης μεθοδολογίας.



Σχήμα 48: Συγκριτική δοκιμή απόδοσης με άλλους ελεγκτές

6

ΕΠΙΛΟΓΟΣ

6.1. Συμπεράσματα

Σε αυτήν την εργασία, προτάθηκε μια νέα μέθοδος ελέγχου φορτίου-συχνότητας, ανθεκτική σε κυβερνοεπιθέσεις, για έξυπνα ενεργειακά συστήματα. Η μέθοδος βασίστηκε στις αρχές της Ενισχυτικής Μάθησης και υλοποιήθηκε ως ένας πράκτορας επιφορτισμένος με το δευτερεύον έλεγχο ενός συστήματος LFC δύο περιοχών. Αξιοποιώντας τον αλγόριθμο Double DQN, μια μεθοδολογία που προτείνεται για πρώτη φορά σε αυτόν τον τομέα, ο πράκτορας κατάφερε μέσα από κατάλληλη εκπαίδευση να μάθει τη δυναμική του περιβάλλοντος LFC και να προστατεύσει τα δίκτυα ηλεκτρικής ενέργειας από FDIAs.

Οι κύριες συνεισφορές αυτής της εργασίας συνοψίζονται ακολούθως:

- Η υλοποίηση του DRL^2FC , του πρώτου κυβερνο-ανθεκτικού ελεγκτή για συστήματα LFC με χρήση βαθιάς ενισχυτικής μάθησης.
- Η ικανότητα του DRL^2FC να αντιμετωπίζει αποτελεσματικά FDIAs ενάντια σε κάθε ευάλωτο σε κυβερνοεπιθέσεις σημείο του συστήματος LFC.
- Η αποτελεσματικότητα του DRL^2FC σε θορυβώδη περιβάλλοντα και η υπεροχή του έναντι παρόμοιων τεχνικών.

Τα πειραματικά αποτελέσματα επιβεβαίωσαν ότι ο πράκτορας μπορεί να μειώσει επιτυχώς τον αντίκτυπο των κυβερνοεπιθέσεων στο LFC, είναι ανθεκτικός σε θορυβώδεις μετρήσεις και υπερέχει έναντι άλλων υπαρχόντων ελεγκτών ανθεκτικών σε κυβερνοεπιθέσεις.

Η RL διαθέτει δυναμική για την ενίσχυση της κυβερνο-ανθεκτικότητας σε συστήματα LFC. Μέσω της αυτόνομης μάθησης βέλτιστων πολιτικών ελέγχου, με βάση τη συνεχή αλληλεπίδραση με το δίκτυο, η RL επιτρέπει προσαρμοστική λήψη αποφάσεων σε δυναμικές και εχθρικές συνθήκες. Σε αντίθεση με τους παραδοσιακούς στατικούς ελεγκτές, οι πράκτορες RL μπορούν να ανιχνεύουν ανωμαλίες και να προσαρμόζουν τον έλεγχο για τη μείωση των επιπτώσεών τους. Αυτή η προσαρμοστικότητα διασφαλίζει ευστάθεια συχνότητας, ακόμη και όταν το σύστημα αντιμετωπίζει απειλές. Επιπλέον, η ικανότητα της RL να βελτιστοποιεί την απόδοση υπό αβεβαιότητα, όπως διακυμάνσεις φορτίου ή παραποιημένα δεδομένα αισθητήρων, την καθιστά κρίσιμο εργαλείο για την προστασία σύγχρονων δικτύων ηλεκτροδότησης έναντι ολοένα και πιο εξελιγμένων κυβερνο-φυσικών επιθέσεων.

6.2. Δυσκολίες και Προκλήσεις

Η ανάπτυξη και η εκπαίδευση του πράκτορα ανέδειξαν μια σειρά σημαντικών προκλήσεων, οι οποίες είναι απαραίτητο να τεκμηριωθούν για μελλοντικούς ερευνητές που ασχολούνται με παρόμοιες εργασίες.

Καινοτομία της Προσέγγισης: Δεδομένου ότι η ανάπτυξη ανθεκτικού σε κυβερνοεπιθέσεις ελεγκτή για το LFC βασισμένο στην Ενισχυτική Μάθηση αποτελεί αχαρτογράφητη περιοχή, δεν υπήρχαν πλαίσια εργασίας ή μέθοδοι για άμεση εφαρμογή. Αυτό επέβαλε εκτεταμένες και επαναλαμβανόμενες προσαρμογές στον σχεδιασμό συναρτήσεων ανταμοιβής, αναπαραστάσεων χώρου καταστάσεων και ενεργειών. Η έλλειψη συγκρίσιμων μελετών δημιούργησε αβεβαιότητες στην επικύρωση ενδιάμεσων αποτελεσμάτων, απαιτώντας πειραματισμούς μέσω δοκιμών και σφαλμάτων.

Περιορισμοί Συνεργασίας Εργαλείων: Αρχικά, σχεδιάστηκε η χρήση πλαισίων ενισχυτικής μάθησης βασισμένων σε Python, όπως το PyTorch ή το TensorFlow, σε συνδυασμό με προσομοιώσεις MATLAB/Simulink για το περιβάλλον του LFC. Ωστόσο, η δυσκολία επικοινωνίας μεταξύ των βασισμένων σε Python πλαισίων και του περιβάλλοντος Simulink, λόγω έλλειψης ευέλικτων διασυνδέσεων, καθιστούσε ανέφικτη την ανταλλαγή δεδομένων σε πραγματικό χρόνο. Αυτός ο τεχνικός περιορισμός οδήγησε στην επιλογή του MATLAB Reinforcement Learning Toolbox ως εναλλακτικής λύσης, παρά το γεγονός ότι το κλειστό οικοσύστημα του MATLAB περιορίζει την πρόσβαση σε προηγμένες βιβλιοθήκες RL ανοιχτού κώδικα (π.χ. custom layers, αλγορίθμους βελτιστοποίησης) και σε βελτιστοποιήσεις από την ευρύτερη κοινότητα. Επιπλέον, η προσαρμογή της αρχιτεκτονικής του πράκτορα RL περιορίστηκε από την έλλειψη ευελιξίας στη διαχείριση νευρωνικών δικτύων σε σύγκριση με τα PyTorch/TensorFlow. Η εμπειρία αυτή υπογραμμίζει την ανάγκη ανάπτυξης υβριδικών λύσεων που συνδυάζουν τα πλεονεκτήματα του Simulink στην προσομοίωση δυναμικών συστημάτων με την ευελιξία των βασισμένων σε Python RL πλαισίων.

Περιορισμοί Υπολογιστικών Πόρων: Η εκπαίδευση του RL agent χωρίς τη χρήση GPU επιτάχυνσης επιμήκυνε σημαντικά τη φάση της πειραματικής ανάπτυξης. Κάθε επανάληψη εκπαίδευσης απαιτούσε ώρες έως ημέρες σε υλικό με CPU, μειώνοντας την ικανότητα για βελτιστοποίηση παραμέτρων. Αυτό το μειονέκτημα οδήγησε σε απλοποιήσεις στην αρχιτεκτονική του νευρωνικού δικτύου και περιόρισε την εξερεύνηση πολυδιάστατων χώρων καταστάσεων-ενεργειών.

Οι παραπάνω προκλήσεις υπογραμμίζουν τις υποχωρήσεις που έγιναν σε επίπεδα καινοτομίας, εργαλείων και διαθεσιμότητας πόρων στο σχεδιασμό και την εκπαίδευση του πράκτορα. Παρά το γεγονός ότι προκάλεσαν καθυστερήσεις και συμβιβασμούς, η αντιμετώπισή τους προσέφερε πολύτιμες πρακτικές γνώσεις για τη χρήση RL σε πραγματικά κυβερνο-φυσικά συστήματα, όπως η ανάγκη για ευέλικτο σχεδιασμό εργαλειοθηκών και κλιμακούμενης υπολογιστικής υποδομής. Μελλοντικές εργασίες θα μπορούσαν να εξαλείψουν αυτά τα ζητήματα μέσω υβριδικών πλαισίων MATLAB-Python και χρήσης GPU σε cloud περιβάλλοντα.

6.3. Προτάσεις για μελλοντική έρευνα

Η ανάπτυξη του DRL^2FC ανοίγει πολλά ελπιδοφόρα πεδία για περαιτέρω έρευνα και βελτίωση. Οι παρακάτω προτάσεις αποσκοπούν στην αντιμετώπιση των περιορισμών της τρέχουσας προσέγγισης και στην εξερεύνηση νέων διαστάσεων των κυβερνο-ανθεκτικών συστημάτων ελέγχου:

Πράκτορας με Συνεχή Χώρο Δράσης: Η τρέχουσα υλοποίηση χρησιμοποιεί έναν διακριτό χώρο δράσης, κάτι που μπορεί να περιορίζει την ακρίβεια και την προσαρμοστικότητα του ελεγκτή. Μελλοντική έρευνα θα μπορούσε να εστιάσει στην ανάπτυξη ενός πράκτορα με συνεχή χώρο δράσης, αξιοποιώντας αλγορίθμους όπως το Deep Deterministic Policy Gradient (DDPG) ή το Proximal Policy Optimization (PPO). Αυτό θα επέτρεπε πιο λεπτομερείς ρυθμίσεις ελέγχου και πιθανώς θα βελτίωνε την απόκριση του συστήματος σε δυναμικές μεταβολές φορτίου και κυβερνοεπιθέσεις.

Εκπαίδευση Υπό Σύνθετες Κυβερνοεπιθέσεις: Ενώ ο τρέχων πράκτορας εκπαιδεύτηκε σε απλοποιημένα σενάρια αλλαγών φορτίου και επιθέσεων, μελλοντική έρευνα θα πρέπει να ενσωματώνει πιο εξελιγμένες και ποικίλες κυβερνο-επιθετικές συνθήκες. Η εκπαίδευση του πράκτορα υπό αυτές τις σύνθετες συνθήκες θα ενίσχυε την ανθεκτικότητά του και θα τον προετοίμαζε για πραγματικά εχθρικά περιβάλλοντα.

Δοκιμές με Hardware-in-the-Loop (HIL): Για την επικύρωση της πρακτικής εφαρμογής του πράκτορα, προτείνονται δοκιμές με Hardware-in-the-Loop (HIL). Αυτή η προσέγγιση θα περιλάμβανε την ενσωμάτωση του ελεγκτή με φυσικά υλικά συστήματα ενός δικτύου ηλεκτρικής ενέργειας, παρέχοντας μια πιο ρεαλιστική αξιολόγηση της απόδοσής του υπό πραγματικές συνθήκες λειτουργίας.

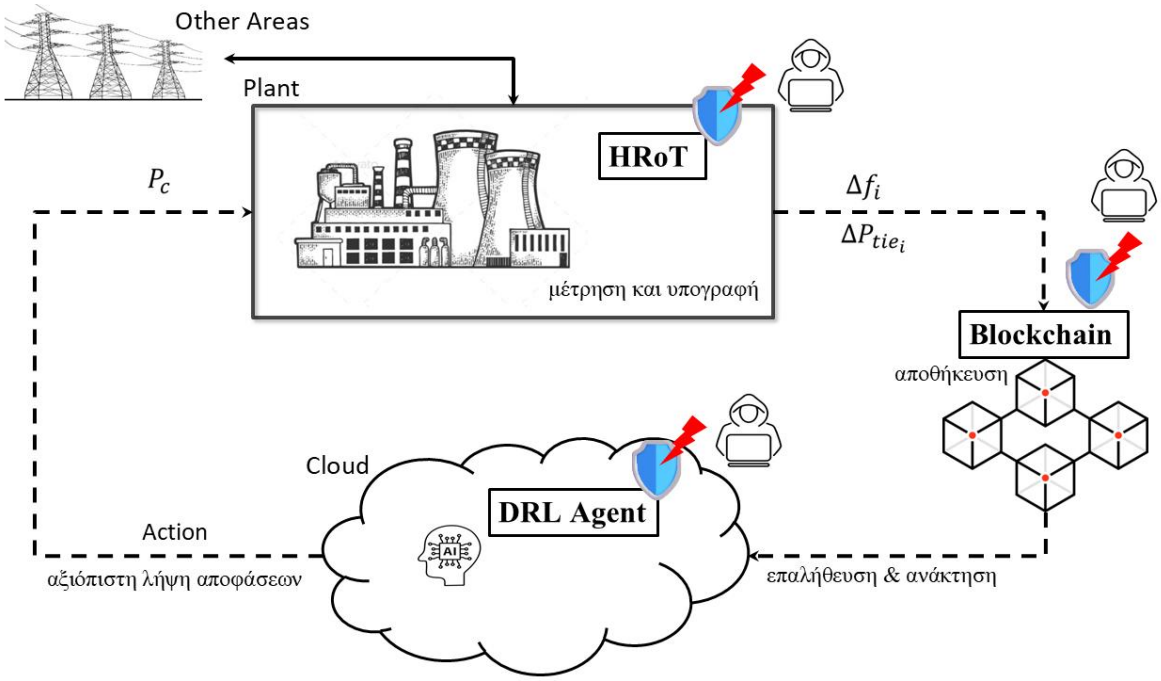
Προσέγγιση Πολλών Πρακτόρων: Ένα πλαίσιο πολυ-πρακτορικής ενισχυτικής μάθησης (Milti-Agent Reinforcement Learning - MARL) θα μπορούσε να εξερευνηθεί για την αποκέντρωση της διαδικασίας ελέγχου και τη βελτίωση της κλιμακωσιμότητας. Επιτρέποντας σε πολλούς πράκτορες να συνεργάζονται για τη διαχείριση διαφορετικών περιοχών του συστήματος ηλεκτρικής ενέργειας, αυτή η προσέγγιση θα μπορούσε να ενισχύσει την ανοχή σε σφάλματα και την ανθεκτικότητα έναντι επιθέσεων.

Πράκτορας για Συστήματα Πολλών Περιοχών: Η επέκταση του τρέχοντος μοντέλου LFC για δύο περιοχές σε ένα πολυ-περιοχικό σύστημα ηλεκτρικής ενέργειας θα αντανακλούσε καλύτερα τα πραγματικά έξυπνα δίκτυα. Ένας πράκτορας για πολυ-περιοχικά συστήματα θα μπορούσε να σχεδιαστεί για τον συντονισμό της ρύθμισης συχνότητας σε διασυνδεδεμένες περιοχές, παρέχοντας μια πιο ολοκληρωμένη λύση για τη σταθερότητα μεγάλων συστημάτων ηλεκτρικής ενέργειας.

Η υλοποίηση αυτών των ιδεών θα μπορούσε να προωθήσει σημαντικά την εξέλιξη της τεχνολογίας σε αυτόν τον τομέα. Προς ενίσχυση της συνεργατικής καινοτομίας, ο πράκτορας, το περιβάλλον δοκιμών και τα συναφή εργαλεία διατίθενται δημόσια στο αποθετήριο GitHub[1]. Για τον ίδιο σκοπό, η έρευνα που πραγματοποιήθηκε κατατέθηκε με τη μορφή επιστημονικού άρθρου στο 16ο συνέδριο IEEE PowerTech - Kiel 2025 [132].

[1] https://github.com/vassildi7/DRL2FC

Τέλος, με βάση τα παραπάνω και τις σύγχρονες τάσεις στον τομέα της κυβερνοασφάλειας, προτείνεται επέκταση της μεθόδου που αναπτύχθηκε, εμπλουτίζοντάς την με τις τεχνολογίες του Blockchain και του Hardware Root of Trust (RoT), όπως φαίνεται στο Σχήμα 49. Η κλιμακωσιμότητα, ο υπολογιστικός φόρτος και η συνεργασία με παλαιότερα συστήματα απαιτούν περαιτέρω μελέτη. Υβριδικές αρχιτεκτονικές Blockchain και ελαφροί αλγόριθμοι συναίνεσης (π.χ. Proof of Authority) θα μπορούσαν να βελτιστοποιήσουν την αποδοτικότητα. Η πρακτική επικύρωση σε πειραματικά περιβάλλοντα ή μικροδίκτυα θα γεφυρώσει τη θεωρία με την πράξη.



Σχήμα 49: DRL^2FC 2.0

Βιβλιογραφία

[1] P. Eder-Neuhauser, T. Zseby, J. Fabini, and G. Vormayr, “Cyber Attack Models for Smart Grid Environments”, Sustainable Energy, Grids and Networks, 2017.

[2] Syrmakesis, A.D., Alcaraz, C. & Hatziargyriou, N.D. “Classifying resilience approaches for protecting smart grids against cyber threats.” Int. J. Inf. Secur. 21, 1189–1210 (2022). https://doi.org/10.1007/s10207-022-00594-7

[3] A. Presekal, A. Stefanov, I. Semertzis, and P. Palensky, “Spatio-Temporal Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems using Enhanced GC-LSTM”, IEEE Transactions on Smart Grid, p. 1, 2024.

[4] Y. L. L. Cheng, “Relentless False Data Injection Attacks Against Kalman-Filter-Based Detection in Smart Grid”, IEEE Transactions on Control of Network Systems, vol. 9, p. 2022.

[5] Y. Liu, L. Cheng, and D. Ye, “Stealthy False Data Injection Attacks Against the Summation Detector in Cyber-Physical Systems”, IEEE Transactions on Industrial Cyber-Physical Systems, vol. 2, p. 391–403, 2024.

[6] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, “A Review of False Data Injection Attacks Against Modern Power Systems”, IEEE Transactions on Smart Grid, vol. 8, p. 1630 – 1638, July 2017.

[7] X. Zhao, Z. Ma, X. Shi, and S. Zou, “Attack Detection and Mitigation Scheme of Load Frequency Control Systems Against False Data Injection Attacks”, IEEE Transactions on Industrial Informatics, vol. 20, August 2024.

[8] A. D. Syrmakesis, H. H. Alhelou and N. D. Hatziargyriou, "Novel SMO-Based Detection and Isolation of False Data Injection Attacks Against Frequency Control Systems", in IEEE Transactions on Power Systems, vol. 39, no. 1, pp. 1434-1446, Jan. 2024, doi: 10.1109/TPWRS.2023.3242015.

[9] A. D. Syrmakesis, H. H. Alhelou and N. D. Hatziargyriou, "A Novel Cyber Resilience Method for Frequency Control in Power Systems Considering Nonlinearities and Practical Challenges", in IEEE Transactions on Industry Applications, vol. 60, no. 2, pp. 2176-2190, March-April 2024, doi: 10.1109/TIA.2023.3332702

[10] Z. Zhang, J. Hu, J. Lu, J. Cao, and F. E. Alsaadi, “Preventing False Data Injection Attacks in LFC System via the Attack-Detection Evolutionary Game Model and KF Algorithm”, IEEE Transactions on Network Science and Engineering, vol. 9, Nov.-Dec. 2022.

[11] Y. Li, R. Huang, and L. Ma “False Data Injection Attack and Defense Method on Load Frequency Control”, IEEE Transactions on Smart Grid, vol. 8, 15 February 2021.

[12] R. Sutton and A. Barto, “Reinforcement Learning: An Introduction”, IEEE Transactions on Neural Networks, vol. 9, September 1998.

[13] C. Chen, M. Cui, X. Fang, B. Ren, and Y. Chen, “Load Altering Attack- Tolerant Defense Strategy for Load Frequency Control System”, Applied Energy, vol. 280, 15 December 2020.

[14] A. S. Mohamed and D. Kundur, “On the Use of Reinforcement Learning for Attacking and Defending Load Frequency Control”, IEEE Transactions on Smart Grid, vol. 15, May 2024.

[15] O. Moulin, F. Vincent-Lavet, P. Elbers, and M. Hoogendoorn, «Improving Generalization to New Environments and Removing Catastrophic Forgetting in Reinforcement Learning by Using an Eco-System of Agents,» σε 2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), November 2022.

[16] U.S. Dept. of Energy and Nat. Energy Technol. Lab., "Smart Grid System Report," Rep. DOE/EE-2361, 2018.

[17] C. Alcaraz and J. Lopez, "Wide-Area Situational Awareness for Critical Infrastructure Protection," in Computer, vol. 46, no. 4, pp. 30-37, April 2013

[18] J. Guo, Y. Wang, D. Zhang, Y. Yu and X. Xie, "A data exchange method of WAMS based on IEC61970," 2011 International Conference on Advanced Power System Automation and Protection, Beijing, China, 2011, pp. 2254-2257

[19] W. Ding, B. He and H. Wang, "Discuss of wide-area relay protection based on wide-area information," 2011 International Conference on Advanced Power System Automation and Protection, Beijing, China, 2011, pp. 1862-1868

[20] S. F. Bush; S. Goel; G. Simard, "IEEE Vision for Smart Grid Communications: 2030 and Beyond Roadmap," in IEEE Vision for Smart Grid Communications: 2030 and Beyond Roadmap , vol., no., pp.1-19, 3 Sept. 2013

[21] T. Atasoy, H. E. Akınç and Ö. Erçin, "An analysis on smart grid applications and grid integration of renewable energy systems in smart cities," 2015 International Conference on Renewable Energy Research and Applications (ICRERA)

[22] T. U. Okeke and R. G. Zaher, "Flexible AC Transmission Systems (FACTS)," 2013 International Conference on New Concepts in Smart Cities: Fostering Public and Private Alliances (SmartMILE), Gijon, Spain, 2013, pp. 1-4

[23] IEEE Guide for Technical Requirements for Hybrid High-Voltage Direct Current Transmission Protection and Control Equipment," in IEEE Std 2837-2023 , vol., no., pp.1-37, 29 Sept. 2023

[24] C. J. Wallnerström, Y. Huang and L. Söder, "Impact From Dynamic Line Rating on Wind Power Integration," in IEEE Transactions on Smart Grid, vol. 6, no. 1, pp. 343-350, Jan. 2015

[25] B. W. McConnell, "Applications of high temperature superconductors to direct current electric power transmission and distribution," in IEEE Transactions on Applied Superconductivity, vol. 15, no. 2, pp. 2142-2145, June 2005

[26] J. Liu, Z. Zhao, J. Ji and M. Hu, "Research and application of wireless sensor network technology in power transmission and distribution system," in Intelligent and Converged Networks, vol. 1, no. 2, pp. 199-220, Sept. 2020

[27] D. Bian, M. Kuzlu, M. Pipattanasomporn and S. Rahman, "Analysis of communication schemes for Advanced Metering Infrastructure (AMI)," 2014 IEEE PES General Meeting | Conference & Exposition, National Harbor, MD, USA, 2014, pp. 1-5

[28] M. van der Kam and R. Bekkers, "Mobility in the Smart Grid: Roaming Protocols for EV Charging," in IEEE Transactions on Smart Grid, vol. 14, no. 1, pp. 810-822, Jan. 2023

[29] H. Albataineh, M. Nijim and D. Bollampall, "The Design of a Novel Smart Home Control System using Smart Grid Based on Edge and Cloud Computing," 2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE), Oshawa, ON, Canada, 2020, pp. 88-91

[30] IEEE Vision for Smart Grid Communications: 2030 and Beyond," in IEEE Vision for Smart Grid Communications: 2030 and Beyond , vol., no., pp.1-390, 31 May 2013

[31] M. Mahmoudian Esfahani and G. R. Yousefi, "Real Time Congestion Management in Power Systems Considering Quasi-Dynamic Thermal Rating and Congestion Clearing Time," in IEEE Transactions on Industrial Informatics, vol. 12, no. 2, pp. 745-754, April 2016

[32] M. Orlando et al., "A Smart Meter Infrastructure for Smart Grid IoT Applications," in IEEE Internet of Things Journal, vol. 9, no. 14, pp. 12529-12541, 15 July15, 2022

[33] D. Georgakopoulos and S. Quigg, "Precision Measurement System for the Calibration of Phasor Measurement Units," in IEEE Transactions on Instrumentation and Measurement, vol. 66, no. 6, pp. 1441-1445, June 2017

[34] X. Huang, J. Wang, Y. He, S. Cheng, Y. Wang and J. Yao, "Research on Intelligent Meter System Based on GPS Automatic Positioning," 2023 International Conference on Power, Electrical Engineering, Electronics and Control (PEEEC), Athens, Greece, 2023, pp. 51-55,

[35] G. M. Coates, K. M. Hopkinson, S. R. Graham and S. H. Kurkowski, "A Trust System Architecture for SCADA Network Security," in IEEE Transactions on Power Delivery, vol. 25, no. 1, pp. 158-169, Jan. 2010

[36] B. Wojszczyk, "Deployment of advanced Smart Grid solutions - Global examples & lessons learned," 2012 IEEE PES Innovative Smart Grid Technologies (ISGT), Washington, DC, USA, 2012, pp. 1-1

[37] N. Duan, C. Huang, C. -C. Sun and L. Min, "Smart Meters Enabling Voltage Monitoring and Control: The Last-Mile Voltage Stability Issue," in IEEE Transactions on Industrial Informatics, vol. 18, no. 1, pp. 677-687, Jan. 2022

[38] P. Kundur, Power System Stability and Control. New York, NY, USA: McGraw-Hill, 1994

[39] S. Ahmad, M. Shafiullah, C. B. Ahmed and M. Alowaifeer, "A Review of Microgrid Energy Management and Control Strategies," in IEEE Access, vol. 11, pp. 21729-21757, 2023

[40] Li Shang, L. Peh and N. K. Jha, "Power-efficient Interconnection Networks: Dynamic Voltage Scaling with Links," in IEEE Computer Architecture Letters, vol. 1, no. 1, pp. 6-6, Jan. 2002

[41] IEEE Draft Guide for the Functional Specifications for Transmission Static Synchronous Compensator (STATCOM) Systems," in IEEE P1052/08, September 2018 , vol., no., pp.1-130, 5 Oct. 2018

[42] A. Tbaileh, C. Mishra and K. Thomas, "PV impacts on dynamic voltage stability," SoutheastCon 2017, Concord, NC, USA, 2017, pp. 1-5

[43] H. Zhou et al., "On Small Signal Stability Early Warning Based on Measurement Data of WAMS," 2010 Asia-Pacific Power and Energy Engineering Conference, Chengdu, China, 2010, pp. 1-4

[44] O. Mo, S. D'Arco and J. A. Suul, "Evaluation of Virtual Synchronous Machines With Dynamic or Quasi-Stationary Machine Models," in IEEE Transactions on Industrial Electronics, vol. 64, no. 7, pp. 5952-5962, July 2017

[45] K. Yasuda et al., "Research & development of superconducting fault current limiter in Japan," in IEEE Transactions on Applied Superconductivity, vol. 15, no. 2, pp. 1978-1981, June 2005

[46] K. Bhat; V. Sundarraj; S. Sinha; A. Kaul, "IEEE Cyber Security for the Smart Grid," in IEEE Cyber Security for the Smart Grid , vol., no., pp.1-122, 16 Sept. 2013

[47] North American Electric Reliability Corporation (NERC), CIP-014-3 — Physical Security. Atlanta, GA, USA: NERC, 2020.

[48] I. Zografopoulos, J. Ospina, X. Liu and C. Konstantinou, "Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies," in IEEE Access, vol. 9, pp. 29775-29818, 2021

[49] K. Bhat; V. Sundarraj; S. Sinha; A. Kaul, "IEEE Cyber Security for the Smart Grid," in IEEE Cyber Security for the Smart Grid , vol., no., pp.1-122, 16 Sept. 2013

[50] ΠΑΡΑΓΩΓΗ ΗΛΕΚΤΡΙΚΗΣ ΕΝΕΡΓΕΙΑΣ ΚΑΙ ΕΛΕΓΧΟΣ ΣΥΧΝΟΤΗΤΑΣ ΚΑΙ ΤΑΣΕΩΣ, Β.Κ.ΠΑΠΑΔΙΑΣ ΚΑΙ Κ.ΒΟΥΡΝΑΣ, ΕΚΔΟΣΕΙΣ ΣΥΜΜΕΤΡΙΑ, 1991

[51] J. W. Simpson-Porco, "On Area Control Errors, Area Injection Errors, and Textbook Automatic Generation Control," in IEEE Transactions on Power Systems, vol. 36, no. 1, pp. 557-560, Jan. 2021

[52] IEEE Recommended Practice for Application of Controllers and Automation to Industrial and Commercial Power Systems," in IEEE P3001.11/D8, October 2015 , vol., no., pp.1-81, 1 Jan. 2015

[53] ΑΝΑΛΥΣΗ ΣΥΣΤΗΜΑΤΟΣ ΗΛΕΚΤΡΙΚΗΣ ΕΝΕΡΓΕΙΑΣ, Β.ΠΑΠΑΔΙΑΣ, ΕΚΔΟΣΕΙΣ ΕΜΠ, 1985

[54] X. S. Han, H. B. Gooi and D. S. Kirschen, "Dynamic economic dispatch: feasible and optimal solutions," in IEEE Transactions on Power Systems, vol. 16, no. 1, pp. 22-28, Feb 2001

[55] K. Kim, P. Rao and J. A. Burnworth, "Self-Tuning of the PID Controller for a Digital Excitation Control System," in IEEE Transactions on Industry Applications, vol. 46, no. 4, pp. 1518-1524, July-Aug. 2010

[56] A. Gurpegui, E. Tegling and A. Rantzer, "Minimax Linear Optimal Control of Positive Systems," in IEEE Control Systems Letters, vol. 7, pp. 3920-3925, 2023

[57] A. Ilka and N. Murgovski, "Novel Results on Output-Feedback LQR Design," in IEEE Transactions on Automatic Control, vol. 68, no. 9, pp. 5187-5200, Sept. 2023

[58] H. Bevrani, "On Future of Robust Control in Smart Grids," 2014 Smart Grid Conference (SGC), Tehran, Iran, 2014, pp. 1-2

[59] G. Feng, "A Survey on Analysis and Design of Model-Based Fuzzy Control Systems," in IEEE Transactions on Fuzzy Systems, vol. 14, no. 5, pp. 676-697, Oct. 2006

[60] A. N. Venkat, I. A. Hiskens, J. B. Rawlings and S. J. Wright, "Distributed MPC Strategies With Application to Power System Automatic Generation Control," in IEEE Transactions on Control Systems Technology, vol. 16, no. 6, pp. 1192-1206, Nov. 2008

[61] T. N. Pham, H. Trinh, and L. V. Hien, “Load frequency control of power systems with electric vehicles and diverse transmission links using distributed functional observers,” IEEE Trans. Smart Grid, vol. 7, no. 1, pp. 238–252, Jan. 2016.

[62] Xing Kai and G. Kusic, "Application of thyristor-controlled phase shifters to minimize real power losses and augment stability of power systems," in IEEE Transactions on Energy Conversion, vol. 3, no. 4, pp. 792-798, Dec. 1988

[63] S. K. Tripathi, V. P. Singh, R. K. Patel and A. S. Pandey, "Impact of Different types of Cyber-Physical Attack and its Prevention on Load Frequency Control," 2023 IEEE 15th International

Conference on Computational Intelligence and Communication Networks (CICN), Bangkok, Thailand, 2023, pp. 410-414

[64] D. P. Fidler, "Was Stuxnet an Act of War? Decoding a Cyberattack," in IEEE Security & Privacy, vol. 9, no. 4, pp. 56-59, July-Aug. 2011

[65] Nell Nelson, "The Impact of Dragonfly Malware on Industrial Control Systems", Global Information Assurance Certification Paper, 2016

[66] U.S. Department of Homeland Security, "The 2014 Quadrennial Homeland Security Review"

[67] ICS-CERT, “Cyber-Attack Against Ukrainian Critical Infrastructure”, (25 February 2016), United States Department of Homeland Security.

[68] Andy Greenberg, "Crash Override’: The Malware That Took Down a Power Grid", Wired, June 2017

[69] Jen Easterly, "The Attack on Colonial Pipeline: What We’ve Learned & What We’ve Done Over the Past Two Years", CISA’s Cybersecurity Advisory Committee, May 7 2023

[70] Leonard Bernardone, "Energy One systems hit by cyber attack: Australian energy sector a growing target", Information Age Australia, 2023

[71] Human Rights Watch, "Ukraine: Russian Attacks on Energy Grid Threaten Civilians", December 2022

[72] X. Chi, X. Liu, X. Jia and L. Zhang, "Networked Load Frequency Control of Multi-Rate Sampled-Data Multi-Area Interconnected Power Systems Under Denial-of-Service Attacks," 2023 China Automation Congress (CAC), Chongqing, China, 2023, pp. 9056-9060

[73] Y. Li, R. Huang and L. Ma, "False Data Injection Attack and Defense Method on Load Frequency Control," in IEEE Internet of Things Journal, vol. 8, no. 4, pp. 2910-2919, 15 Feb.15, 2021

[74] H. Guo, Z. -H. Pang, J. Sun and J. Li, "An Output-Coding-Based Detection Scheme Against Replay Attacks in Cyber-Physical Systems," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 10, pp. 3306-3310, Oct. 2021

[75] L. Hadjidemetriou et al., "Demonstration of Man in the Middle Attack on a Feeder Power Factor Correction Unit," 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), The Hague, Netherlands, 2020, pp. 126-130

[76] P. Ganesh et al., "Learning-Based Simultaneous Detection and Characterization of Time Delay Attack in Cyber-Physical Systems," in IEEE Transactions on Smart Grid, vol. 12, no. 4, pp. 3581-3593, July 2021

[77] Oleg Skulkin, “Incident Response Techniques for Ransomware Attacks: Understand modern ransomware attacks and build an incident response strategy to work through them “, Packt Publishing, 2022

[78] I. Mbona and J. H. P. Eloff, "Detecting Zero-Day Intrusion Attacks Using Semi-Supervised Machine Learning Approaches," in IEEE Access, vol. 10, pp. 69822-69838, 2022

[79] H. Bao, R. Lu, B. Li and R. Deng, "BLITHE: Behavior Rule-Based Insider Threat Detection for Smart Grid," in IEEE Internet of Things Journal, vol. 3, no. 2, pp. 190-205, April 2016

[80] Y. Li, R. Huang and L. Ma, "False Data Injection Attack and Defense Method on Load Frequency Control," in IEEE Internet of Things Journal, vol. 8, no. 4, pp. 2910-2919, 15 Feb.15, 2021

[81] S. Han, M. Xie, H. -H. Chen and Y. Ling, "Intrusion Detection in Cyber-Physical Systems: Techniques and Challenges," in IEEE Systems Journal, vol. 8, no. 4, pp. 1052-1062, Dec. 2014

[82] Z. Xu and M. Liu, "An Intrusion Detection Method based on PCA and ResNet," 2021 3rd International Conference on Applied Machine Learning (ICAML), Changsha, China, 2021, pp. 52-56,

[83] J. -L. Tsai and N. -W. Lo, "Secure Anonymous Key Distribution Scheme for Smart Grid," in IEEE Transactions on Smart Grid, vol. 7, no. 2, pp. 906-914, March 2016

[84] Y. Ding, B. Wang, Y. Wang, K. Zhang and H. Wang, "Secure Metering Data Aggregation With Batch Verification in Industrial Smart Grid," in IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6607-6616, Oct. 2020

[85] L. Chen, X. Dong, X. Kuang, B. Chen and D. Hong, "Towards Ubiquitous Power Distribution Communication: Multi-service Access and QoS Guarantees for IoT Applications in Smart Grid," 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), Chengdu, China, 2019, pp. 894-898

[86] L. Li, X. Zhao, S. Geng and Y. Zhang, "An Efficient Partially Overlapping Channels Assignment for Smart Grid IoT With Differentiated QoS," in IEEE Access, vol. 7, pp. 165207-165216, 2019

[87] Wu Lizhen, Hao Xiaohong, Li Zhengzhe, Luo Zuohao and Zhou Hu, "Decentralized control approach for voltage unbalance compensation in islanded microgrid," 2016 IEEE 8th International Power Electronics and Motion Control Conference (IPEMC-ECCE Asia), Hefei, 2016, pp. 1776-1781

[88] Y. Han, P. Shen, X. Zhao and J. M. Guerrero, "Control Strategies for Islanded Microgrid Using Enhanced Hierarchical Control Structure With Multiple Current-Loop Damping Schemes," in IEEE Transactions on Smart Grid, vol. 8, no. 3, pp. 1139-1153, May 2017

[89] A. G. Phadke and J. S. Thorp, "Real-time monitoring and control of smart power grids," Proc. IEEE, vol. 105, no. 11, pp. 1879–1905, Nov. 2017

[90] A. Ipakchi and F. Albuyeh, Smart Grid: Communication-Enabled Intelligence for the Electric Power Grid. Hoboken, NJ, USA: Wiley-IEEE Press, 2022

[91] H. Li, Y. Wang, and D. Liu, "Robust H∞ control for power systems with multiple time delays and uncertainties," IEEE Trans. Smart Grid, vol. 10, no. 2, pp. 2152–2163, Mar. 2019

[92] M. Rahmani and H. R. Karimi, "Sliding mode control for load frequency regulation in smart grids with renewable energy sources," IEEE Trans. Ind. Electron., vol. 65, no. 8, pp. 6312–6321, Aug. 2018

[93] S. Mishra and Y. Mo, "Unknown input observer-based detection of cyber-attacks in smart grids," IEEE Trans. Smart Grid, vol. 13, no. 2, pp. 1096–1107, Mar. 2022

[94] S. Das, L. Liu, and V. Venkatasubramanian, "Distributed fault-tolerant observers for resilient smart grids under sensor and actuator faults," IEEE Trans. Ind. Informat., vol. 19, no. 5, pp. 6432–6443, May 2023

[95] C. Wang, Q. Li, and H. Zhang, "A zero-sum game framework for cyber-physical security in smart grids: Attack-defense strategies," IEEE Trans. Inf. Forensics Security, vol. 17, pp. 2389–2403, 2022

[96] R. Zhang, Y. Liang, and S. H. Low, "A Stackelberg game approach to cyber-physical security in smart grids with distributed energy resources," IEEE Trans. Smart Grid, vol. 14, no. 4, pp. 2732–2745, Jul. 2023

[97] Y. Li, H. Zhang, and C. Zhao, "Event-triggered resilient control for smart grids under denial-of-service attacks: A decentralized approach," IEEE Trans. Ind. Informat., vol. 19, no. 10, pp. 10045–10056, Oct. 2023

[98] J. Wang, L. Zhang, and M. Liu, "Resilient switched control for cyber-physical smart grids under dynamic false data injection attacks," IEEE Trans. Power Syst., vol. 39, no. 2, pp. 2456–2469, Mar. 2024

[99] A. D. Syrmakesis, H. H. Alhelou and N. D. Hatziargyriou, "A Novel Cyberattack-Resilient Frequency Control Method for Interconnected Power Systems Using SMO-Based Attack Estimation," in IEEE Transactions on Power Systems, vol. 39, no. 4, pp. 5672-5686, July 2024, doi: 10.1109/TPWRS.2023.3340744

[100] Andrew D. Syrmakesis, Cristina Alcaraz, Nikos D. Hatziargyriou, “DAR-LFC: A data-driven attack recovery mechanism for Load Frequency Control, International Journal of Critical Infrastructure Protection,” Volume 45, 2024, 100678, ISSN 1874-5482, https://doi.org/10.1016/j.ijcip.2024.100678

[101] Syrmakesis Andrew D. , Hatziargyriou Nikos D., "Cyber resilience methods for smart grids against false data injection attacks: categorization, review and future directions", Frontiers in Smart Grids, vol.3, 2024, ISSN=2813-4311, DOI 10.3389/frsgr.2024.1397380

[102] Syrmakesis Andreas Dorotheos, "A Hybrid Framework for the Cyber Resilience Enhancement of Frequency Control in Smart Grids", Ph.D. dissertation, School of Electrical and Computer Engineering, National Technical University of Athens, 2025

[103] Y. Wang, C. Chen, and H. Li, "Big data analytics for cybersecurity in smart grids: A machine learning approach to anomaly detection," IEEE Trans. Smart Grid, vol. 12, no. 5, pp. 4052–4065, Sep. 2021

[104] K. P. Murphy, Machine Learning: A Probabilistic Perspective. Cambridge, MA, USA: MIT Press, 2012

[105] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018

[106] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming. Hoboken, NJ, USA: Wiley, 1994

[107] R. Bellman, Dynamic Programming. Princeton, NJ, USA: Princeton University Press, 1957

[108] L. Weng, M. Van Hasselt, T. Schaul, and Z. Wang, "A survey on exploration strategies in reinforcement learning," IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 8, pp. 4129–4148, Aug. 2022

[109] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529–533, Feb. 2015

[110] X. Wang, Y. Chen, Z. Zhang, and Y. Liu, "A survey on reinforcement learning: A taxonomy, recent advances, and challenges," IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 10, pp. 7125–7150, Oct. 2023

[111] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep neural networks for robotic manipulation and navigation using reinforcement learning," IEEE Trans. Robot., vol. 34, no. 4, pp. 887–899, Aug. 2018

[112] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning for autonomous vehicles in traffic: A survey," IEEE Trans. Intell. Transp. Syst., vol. 20, no. 6, pp. 2065–2081, Jun. 2019

[113] B. P. Ginsburg, M. L. Littman, and S. R. Kosorok, "Personalized treatment recommendation systems using reinforcement learning," IEEE J. Biomed. Health Inform., vol. 24, no. 3, pp. 789–801, Mar. 2020

[114] Y. Deng, F. Bao, and Y. Kong, "Deep reinforcement learning for financial portfolio optimization and fraud detection," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 7, pp. 2487–2500, Jul. 2020

[115] A. G. Buchholz et al., "Reinforcement learning for industrial resource management and demand forecasting," IEEE Trans. Ind. Informat., vol. 17, no. 9, pp. 6023–6033, Sep. 2021

[116] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning for energy-efficient smart grid operations," IEEE Trans. Smart Grid, vol. 12, no. 5, pp. 4052–4065, Sep. 2021

[117] M. G. Bellemare et al., "Autonomous demand-side management via reinforcement learning," IEEE Trans. Power Syst., vol. 36, no. 5, pp. 4567–4579, Sep. 2021

[118] R. S. Sutton, A. G. Barto, and H. van Hasselt, "Reinforcement learning for predictive maintenance and fault detection in energy systems," IEEE Trans. Sustain. Energy, vol. 13, no. 2, pp. 1122–1133, Apr. 2022

[119] K. Arulkumaran et al., "Deep reinforcement learning for structural health monitoring and anomaly detection," IEEE Trans. Instrum. Meas., vol. 71, pp. 1–12, 2022

[120] H. Bevrani and T. Hiyama, "Reinforcement learning-based load frequency control in multi-area power systems under stochastic renewable integration," IEEE Trans. Sustain. Energy, vol. 10, no. 4, pp. 1782–1793, Oct. 2019.

[121] D. Vrabie, K. G. Vamvoudakis, and F. L. Lewis, "Adaptive multi-agent differential games for distributed optimal frequency control in power systems," IEEE Trans. Smart Grid, vol. 11, no. 3, pp. 2439–2450, May 2020.

[122] V. Mnih et al., "Continuous control with deep reinforcement learning for stochastic power system frequency regulation," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 9, pp. 3416–3428, Sep. 2020

[123] X. Zhang, Y. Liu, and H. He, "Dynamic generator activation tuning via reinforcement learning for adaptive load frequency control," IEEE Trans. Power Syst., vol. 36, no. 5, pp. 4567–4579, Sep. 2021

[124] J. Qiang, H. Wang, and K. Zhang, "Deep reinforcement learning with interval type-2 fuzzy logic for robust frequency control under data uncertainty," IEEE Trans. Fuzzy Syst., vol. 30, no. 8, pp. 2939–2951, Aug. 2022

[125] Y. Mo, S. Mishra, and B. Sinopoli, "Reinforcement learning-based false data injection attack simulation and defense in LFC systems," IEEE Trans. Smart Grid, vol. 14, no. 1, pp. 789–802, Jan. 2023

[126] A. Sanjab, W. Saad, and T. Başar, "Hybrid cyber-physical resilience framework for power systems: Integrating reinforcement learning and game theory," IEEE Trans. Control Netw. Syst., vol. 9, no. 2, pp. 621–634, Jun. 2022

[127] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in Adv. Neural Inf. Process. Syst. (NeurIPS), 2019, pp. 8026–8037

[128] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in Proc. USENIX Symp. Oper. Syst. Des. Implement. (OSDI), 2016, pp. 265–283

[129] MathWorks, MATLAB Reinforcement Learning Toolbox. Natick, MA, USA: MathWorks, 2023

[130] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in Proc. AAAI Conf. Artif. Intell., 2016, pp. 2094–2100

[131] H. H. Alhelou and P. Cuffe, "A Dynamic-State-Estimator-Based Tolerance Control Method Against Cyberattack and Erroneous Measured Data for Power Systems," IEEE Transactions on Industrial Informatics, vol. 18, no. 7, pp. 4990‑4999, 2022

[132] V Dimitropoulos, AD Syrmakesis, N Hatziargyriou, "DRL2FC: An Attack-Resilient Controller for Automatic Generation Control Based on Deep Reinforcement Learning", submitted to IEEE PowerTech 2025, arXiv preprint arXiv:2404.16974, 2024