



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

Διαδίκτυο των Πραγμάτων

Μελέτη Αναδρομικών Νευρωνικών Δικτύων ως Φυσικές μη
Κλωνοποιήσιμες Συναρτήσεις για εφαρμογές Κυβερνο-
Ασφάλειας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΕΩΡΓΙΟΥ ΚΟΚΚΑΛΗ ΚΟΣΜΑ

Επιβλέπων : Χάρης Μεσσαριτάκης

Μέλη εξεταστικής επιτροπής: Σκιάνης Χαράλαμπος , Χάρης Μεσσαριτάκης , Κωστούλας
Θεόδωρος

Σάμος, [Ιούλιος 2024]

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πρόλογος και ευχαριστίες

Αυτή η διπλωματική εργασία δημιουργήθηκε κάτω από μεγάλη προσωπική πίεση σε συνδυασμό με μία απαιτητική δουλειά. Οπότε θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν ψυχολογικά με τις κουβέντες τους και τις συζητήσεις, τους φίλους, τους γονείς και την οικογένειά μου. Αν και μου αρέσει να θεωρώ τους φίλους μου σαν ένα κομμάτι της οικογένειάς μου, είτε αυτοί είναι συνάδελφοι, είτε βρίσκονται στο εξωτερικό, είτε κάνουν καποέρα, είτε απλά μιλάμε και κάνουμε βόλτες ή παίζουμε επιτραπέζια.

Ακόμη θα ήθελα να ευχαριστήσω τον κ. Μεσαριτάκη που δέχτηκε να με αναλάβει στην επίβλεψη της διπλωματικής καθώς επίσης και τον υποψήφιο Διδάκτορα Δημήτρη Δερμάνη που παρά το φορτωμένο του πρόγραμμα έβρισκε χρόνο για να με καθοδηγήσει.

Η διπλωματική αυτή δεν θα μπορούσε να γίνει πραγματικότητα χωρίς όλους εσάς και για αυτό και σας την αφιερώνω.

© [2024]

του/της

[ΓΕΩΡΓΙΟΣ ΚΟΚΚΑΛΗΣ ΚΟΣΜΑΣ]

Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πίνακας περιεχομένων

1	ΦΥΣΙΚΕΣ ΜΗ ΚΛΩΝΟΠΟΙΗΣΙΜΕΣ ΣΥΝΑΡΤΗΣΕΙΣ (Physical Unclonable Functions - PUF).....	2
1.1	Προδιαγραφές	3
1.2	Βασικές έννοιες και βασικά χαρακτηριστικά	3
1.3	Είδη ΦΜΚΣ	4
1.3.1	Μη ηλεκτρονικές ΦΜΚΣ.....	5
1.3.2	Ηλεκτρονικές ΦΜΚΣ	5
1.3.3	ΦΜΚΣ Ποιτίου.....	5
1.4	Εσωτερικές και μη Εσωτερικές ΦΜΚΣ.....	6
1.4.1	<u>Εξωτερικές έναντι Εσωτερικών Αξιολογήσεων</u>	6
1.4.2	Εμφανείς έναντι Σιωπηλές Τυχαίες Διακυμάνσεις.....	7
1.5	Εγγενείς, μη Εγγενείς ΦΜΚΣ	8
1.5.1	Μη εγγενείς ΦΜΚΣ	8
1.5.2	Εγγενείς ΦΜΚΣ.....	10
1.6	Ασθενείς και ισχυρές ΦΜΚΣ.....	11
1.6.1	Χαρακτηριστικά των Αδύναμων ΦΜΚΣ.....	11
1.6.2	Χαρακτηριστικά των ισχυρών ΦΜΚΣ	12
1.7	Εγγενής ΦΜΚΣ (Intrinsic PUF).....	13
1.7.1	ΦΜΚΣ Συγκριτής	13
1.7.2	ΦΜΚΣ Κυκλικού Ταλαντωτή (Ring Oscillator Puf).....	14
1.7.3	ΦΜΚΣ Σφάλματος (Glitch Puf)	15
1.7.4	SRAM.....	16
1.7.5	Νευρομορφική ΦΜΚΣ.....	17
1.8	Σύνοψη.....	18
2	ΜΗΤΡΙΚΕΣ ΚΑΙ ΕΡΓΑΛΕΙΑ ΤΩΝ ΜΗ ΚΛΩΝΟΠΟΙΗΣΙΜΩΝ ΣΥΝΑΡΤΗΣΕΩΝ (Physical Unclonable Functions - PUF).....	19
2.1	Τι εννοούμε με τον όρο μη κλωνοποιήσιμα.....	19
2.2	Πειραματική Διαδικασία.....	19
2.3	Αποκρίσεις ΦΜΚΣ	20
2.3.1	<u>Intra-distance</u>	20
2.3.2	<u>Inter-distance</u>	21
2.4	NIST TEST	22
2.4.1	Τύποι γεννητριών τυχαίων αριθμών	23

2.4.2	Έλεγχος (testing)	24
2.4.3	Χρήση σε ΦΜΚΣ	24
3	ΑΝΑΔΡΟΜΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	26
3.1	ΑΝΑΔΡΟΜΙΚΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ (RECURSIVE NEURAL NETWORKS)	26
3.2	Εκπαίδευση ΑΝΔ.....	27
3.3	Περιγραφή ΑΝΔ	27
3.4	ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΕΙΣΑΓΩΓΗ ΣΤΑ ESN	30
3.4.1	ΕΠΙΣΤΡΟΦΗ ΣΦΑΛΜΑΤΟΣ (BACK PROPAGATION).....	30
3.4.2	ESN (ECHO STATE NETWORK).....	34
3.5	Πρακτική Προσέγγιση για την Παραγωγή reservoir	41
3.5.1	<u>Προτεραιότητα Παραμέτρων</u>	41
3.5.2	<u>Η διάταξη για την επιλογή παραμέτρων</u>	41
3.5.3	Επιλογή Παραμέτρων με Χειροκίνητο Τρόπο	42
3.5.4	Αυτοματοποιημένη Επιλογή Παραμέτρων	43
4	ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ.....	46
4.1	Εισαγωγή.....	46
4.1.1	ESN ως ΦΜΚΣ	46
4.2	Υλοποίηση	46
4.2.1	Inter distance	50
4.2.2	Inter challenge	51
4.2.3	Intra reservoir.....	52
	Παράρτημα Ι [NARMA]	54
	Παράρτημα ΙΙ [ΒΑΣΙΚΟΣ ΚΩΔΙΚΑΣ ESN]	59
5	Bibliography	64

Λίστα Εικόνων

Εικόνα 1 (Α) Ένα τυπικό σχήμα ΦΜΚΣ. Επισκόπηση: (Β) μίας συμβατικής οπτικής ΦΜΚΣ και (Γ) μίας ανθεκτικής οπτικής ΦΜΚΣ . Εδώ, η ανθεκτικότητα αναφέρεται στην επίτευξη ανοσίας έναντι αναγνώρισης και περιβαλλοντικών μεταβολών. [2].....	9
Εικόνα 2: Arbiter Puf [3]	14
<i>Εικόνα 3 Βασική δομή του RO-PUF [4]</i>	<i>14</i>
<i>Εικόνα 4. Παραδοσιακή διαμορφώσιμη δομή RO [4]</i>	<i>15</i>
Εικόνα 5 ΦΜΚΣ σφάλματος όπως προτάθηκε από τον Shimizu [5]	16
Εικόνα 6 Ένα κελί SRAM [6]	17
Εικόνα 7 Σχηματική περίληψη του Neuromorphic PUF. Τα Xin και Yout είναι οι παράμετροι μίας χρονοσειράς και είναι δημόσια, τα Wint είναι σύνθετα βάρη που συνδέονται με το υλικό και τα Wout θεωρούνται ιδιωτικά. Τα ψηφιακά βάρη γραμμικής παλινδρόμησης (linear regrestion) που προκύπτουν μέσω της εκπαίδευσης θεωρούνται η απόκριση της ΦΜΚΣ. [7].....	18
Εικόνα 8 Τυπική δομή ενός feedforward δικτύου τροφοδότησης (αριστερά) και ενός αναδρομικού δικτύου δεξιά [11].....	26
<i>Εικόνα 9 Διακεκομμένα βέλη υποδεικνύουν συνδέσεις που είναι δυνατές αλλά δεν απαιτούνται.</i>	<i>28</i>
<i>Εικόνα 10 Πάνω είσοδος αρχικό σήμα Κατω είσοδος διάρκεια Έξωδος Ορθογώνιο σήμα επιθυμητής διάρκειας [13]</i>	<i>29</i>
<i>Εικόνα 11 Επίδοση ενός ΑΝΔ εκπαιδευμένου στην διεργασία του χρονομέτρου. Στην τελευταία γραφική παράσταση, η συνεχής γραμμή αντιπροσωπεύει την επιθυμητή έξοδο (teacher data). Η γραμμή με τις τελείες : έξοδος του δικτύου.</i>	<i>29</i>
<i>Εικόνα 12 Εικόνα ενός ESN</i>	<i>35</i>
Εικόνα 13 Generative and Predictive Mode	46
Εικόνα 14 Κομμάτι του κώδικα για την ρύθμιση των δεδομένων εκπαίδευσης	47
Εικόνα 15 Μέγεθος Reservoir και ρυθμού Διαρροής.....	47
Εικόνα 16 Random Seed	47
Εικόνα 17 Φασματική ακτίνα.....	47
Εικόνα 18 MSE Mean Square Error	48
Εικόνα 19 Μπλε χρώμα σήμα στην έξοδο , πράσινο χρώμα επιθυμητή έξοδος.	48
Εικόνα 20 Ενεργοποίηση βαρών εξόδου.....	49
Εικόνα 21 Ιστόγραμμα Τιμών Βαρών	49
Εικόνα 22 Έξοδος NSGAI	49
Εικόνα 23 MSE ανά μήκος εκπαίδευσης	50
Εικόνα 24 Ραβδόγραμμα Βαρών πριν και μετά την ομογενοποίηση	51
Εικόνα 25 Hamming distance.....	51
Εικόνα 26 Mean Hamming.....	52
Εικόνα 27 Τυχαία κατανομή	52

Εικόνα 28 Mean Hamming.....	52
Εικόνα 29 Γράφημα Hamming Distance κατανομών.....	53
Εικόνα 30 Κυματομορφή NARMA10.....	57

Ακρωνύμια

Instance	Παράδειγμα ενός Puf
Teacher data	Δεδομένα εκπαίδευσης
Puf	Φυσικές μη κλωνοποιήσιμες συναρτήσεις
RNN	Παλινδρομικά Νευρωνικά Δίκτυα
Challenge	Πρόκληση

Περίληψη

Η διπλωματική εργασία ασχολείται με τη χρήση και προσομοίωση Φυσικών μη κλωνοποιήσιμων συναρτήσεων (Physical Unclonable Functions - PUFs) από (Echo State Networks – ESNs που είναι ένας τύπος Αναδρομικών Νευρωνικών Δικτύων.

Οι ΦΜΚΣ είναι μηχανισμοί που εκμεταλλεύονται τις μικρές διαφορές στις φυσικές ιδιότητες υλικών για τη δημιουργία μοναδικών και δύσκολα αντιγράψιμων ψηφιακών αποτυπωμάτων. Χρησιμοποιούνται ευρέως σε εφαρμογές ασφαλείας, όπως στην ταυτοποίηση συσκευών, την προστασία από παραποίηση και την κρυπτογραφία.

Το ESN είναι ένα είδος αναδρομικού νευρωνικού δικτύου που μπορεί να χρησιμοποιηθεί για την πρόβλεψη χρονοσειρών.

Στην εργασία αυτή, μελετάτε ο τρόπος με τον οποίο οι ΦΜΚΣ μπορούν να χρησιμοποιηθούν ως μέσα ασφαλείας. Συγκεκριμένα, εξετάζεται η δυνατότητα ύπαρξης ενός νευρομορφικού PUF με την χρήση ενός ESN ως κλειδί ασφαλείας, καθώς και η χρήση μεθόδων που επιβεβαιώνουν αυτή την δυνατότητα.

Θα εξετάσουμε τις προδιαγραφές που συνήθως πρέπει να πληροί ένα σύστημα φυσικών μη κλωνοποιήσιμων συναρτήσεων για να είναι αποτελεσματικό για διάφορες εφαρμογές. Ενδεικτικά πρέπει να είναι **αξιόπιστο, μοναδικό με τυχαιότητα και εύκολη ταυτοποίηση**.

Επειδή οι ΦΜΚΣ δεν ορίζονται απλά και μόνο ως μια συνάρτηση αλλά και βασίζονται σε μια ποικιλία τεχνολογιών και υλικών χωρίζονται σε διάφορα είδη.

Με ηλεκτρονικά και μη ηλεκτρονικά στοιχεία, ΦΜΚΣ πυριτίου, εσωτερικές και εξωτερικές, εγγενείς μη εγγενείς, ασθενείς και ισχυρές, εμφανείς σιωπηλές.

Θα δούμε ΦΜΚΣ με λειτουργία, συγκριτή, κυκλικού ταλαντωτή, σφάλματος, Slam, νευρομορφικές.

Θα μελετήσουμε τα χαρακτηριστικά τους, τα πλεονεκτήματα και τα μειονεκτήματά τους.

Θα εξετάσουμε τις αποκρίσεις εσωτερικών μη εσωτερικών ΦΜΚΣ, τύπους γεννητριών τυχαίων αριθμών, δοκιμές και χρήσεις τους σε ΦΜΚΣ.

Στη συνέχεια θα δούμε τα νευρωνικά δίκτυα, βασικά μοντέλα και χαρακτηριστικά τους. Οι βασικοί αλγοριθμικοί τύποι εκπαίδευσης. Πως φτιάχνεται ένα ESN, λειτουργία του, γενικές παράμετροι, μέγεθος, αραιότητα, φασματική ακτίνα, ρυθμός διαρροής. Προτεραιότητες παραμέτρων, τρόποι για την επιλογή τους είτε χειροκίνητα είτε αυτοματοποιημένα. Αλγόριθμοι και αξιολογήσεις τους.

Και τέλος το πειραματικό μέρος που ασχολείται με την προσομοίωση μίας ΦΜΚΣ με την χρήση ενός echo state network ESN σε python και τις διαδικασίες ελέγχου της ΦΜΚΣ.

Λέξεις Κλειδιά: Ασφάλεια Υλικού, Φυσικές μη κλωνοποιήσιμες συναρτήσεις, Νευρωνικά Δίκτυα,

Abstract

This thesis deals with the use and simulation of Physical Unclonable Functions (PUFs) by Echo State Networks (ESNs), a type of Recurrent Neural Networks. PUFs are mechanisms that exploit minute differences in the physical properties of materials to create unique and hard-to-replicate digital fingerprints. They are widely used in security applications, such as device identification, tamper protection, and cryptography.

The ESN is a kind of recurrent neural network that can be used for time series prediction. In this work, we study how PUFs can be used as security measures. Specifically, we examine the possibility of a neuromorphic PUF using an ESN, as well as methods that confirm this potential.

We will explore the specifications that a physical unclonable function system typically needs to meet to be effective for various applications. Indicatively, it should be reliable, unique with randomness, and easily identifiable. Since PUFs are not defined solely as a function but also based on a variety of technologies and materials, they are classified into various types.

PUFs can include electronic and non-electronic components, silicon PUFs, internal and external, intrinsic and extrinsic, weak and strong, visible and silent. We will examine PUFs with functions such as comparator, ring oscillator, error, SLAM, and neuromorphic. We will study their characteristics, advantages, and disadvantages.

We will examine the responses of internal and non-internal PUFs, types of random number generators, tests, and their uses in PUFs. Then, we will look at neural networks, basic models, and their characteristics. The main algorithmic types of training, how to create an ESN, its operation, general parameters, size, sparsity, spectral radius, leakage rate, parameter priorities, ways to choose them either manually or automatically, algorithms, and their evaluations.

Finally, we will delve into the experimental part, which deals with the simulation of a PUF using an echo state network ESN in Python and the procedures for verifying the PUF.

Keywords: *Physical Unlonable Functions, Hardware security, Neural Networks, Echo State Networks*

1

ΦΥΣΙΚΕΣ ΜΗ

ΚΛΩΝΟΠΟΙΗΣΙΜΕΣ

ΣΥΝΑΡΤΗΣΕΙΣ (Physical Unclonable

Functions - PUF)

Οι Φυσικές μη κλωνοποιήσιμες συναρτήσεις (PUF) αναφέρονται σε ένα είδος κρυπτογραφικής τεχνολογίας που βασίζεται στις ξεχωριστές φυσικές διαφορές που εμφανίζονται στα φυσικά στοιχεία ενός συστήματος ή μιας συσκευής το οποίο μπορεί να θεωρηθεί και σαν δαχτυλικό αποτύπωμα της συσκευής. Οι φυσικές μη κλωνοποιήσιμες συναρτήσεις χρησιμοποιούν αυτές τις μοναδικές φυσικές ιδιαιτερότητες για να παράγουν μοναδικά κλειδιά, αριθμούς ή αναγνωριστικά που μπορούν να χρησιμοποιηθούν για διάφορους σκοπούς, όπως την πιστοποίηση και αυθεντικοποίηση συσκευών ή χρηστών. [1] Το αυξημένο ενδιαφέρον αυτών των συναρτήσεων για εφαρμογές ασφάλειας υλικού και πληροφοριών έχει διευρυνθεί γοργά σε σχεδόν όλους τους τομείς της φυσικής επιστήμης. [2]

Οι φυσικές μη κλωνοποιήσιμες συναρτήσεις εκμεταλλεύονται φυσικά φαινόμενα και διαδικασίες που είναι δύσκολο να αναπαραχθούν, καθιστώντας τα κλειδιά ή τις ταυτότητες που παράγονται από αυτά πολύ δύσκολο να αντιγραφούν. Ένα από τα πιο συνηθισμένα φυσικά φαινόμενα που χρησιμοποιούνται σε φυσικές μη κλωνοποιήσιμες συναρτήσεις είναι οι μικρο-διαφοροποιήσεις στις ηλεκτρικές ιδιότητες των στοιχείων, όπως τα τρανζίστορ ή οι αντιστάσεις, που παρατηρούνται λόγω μικροσκοπικών διαφορών ασυνέχειας υλικού ή μικρών αστοχιών κατά την κατασκευή. [2]

Αυτό προκαλεί μία ενδιαφέρουσα προσέγγιση για ασφάλεια με χαμηλό κόστος, καθώς εκμεταλλεύονται τις ήδη υπάρχουσες φυσικές ιδιότητες των συστημάτων πράγμα που τα καθιστά ιδανικά για IoT συσκευές προκειμένου ώστε να εισάγουμε μία ασφάλεια σε αυτά τα συστήματα μικρού κόστους και χαμηλής επεξεργαστικής ισχύος [3]

Για συντομία σε αυτή την διπλωματική οι φυσικές μη κλωνοποιήσιμες συναρτήσεις θα αναφέρονται ως ΦΜΚΣ

1.1 Προδιαγραφές

Ένα σύστημα φυσικών μη κλωνοποιήσιμων συναρτήσεων συνήθως πρέπει να πληροί τις παρακάτω προδιαγραφές για να είναι αποτελεσματικό για διάφορες εφαρμογές:

1. **Αξιοπιστο:** Οι ΦΜΚΣ πρέπει να παρέχει συνεπώς την ίδια απόκριση όταν παρουσιάζεται με την ίδια πρόκληση υπό τις ίδιες συνθήκες. Πρέπει να αντέχει στις διακυμάνσεις που προκαλούνται από περιβαλλοντικούς παράγοντες ή θόρυβο.
2. **Μοναδικό:** Κάθε παράδειγμα ΦΜΚΣ πρέπει να έχει μια μοναδική απόκριση που το ξεχωρίζει από άλλα παραδείγματα της ίδιας κατηγορίας ΦΜΚΣ. Αυτή η μοναδικότητα συμβάλλει στην ασφάλεια και διασφαλίζει ότι δύο 'ίδιες' ΦΜΚΣ δεν παράγουν ίδιες αποκρίσεις.
3. **Τυχαιότητα:** Οι αποκρίσεις της ΦΜΚΣ πρέπει να φαίνονται τυχαίες για έναν παρατηρητή που δεν διαθέτει το μυστικό κλειδί ή γνώση για τους εσωτερικούς μηχανισμούς του ΦΜΚΣ. Αυτή η τυχαιότητα ενισχύει την ασφάλεια και την απροβλεψιμότητα.
4. **Εύκολη Ταυτοποίηση:** Η διαδικασία λήψης και επαλήθευσης της απόκρισης του ΦΜΚΣ πρέπει να είναι απλή και αποτελεσματική. Δεν πρέπει να απαιτεί περίπλοκους υπολογισμούς ή εκτενή πρωτόκολλα, καθιστώντας την κατάλληλη για πρακτικά σενάρια ταυτοποίησης.

Η συνδυασμένη εκπλήρωση αυτών των προδιαγραφών ενισχύει την αξιοπιστία, την ασφάλεια και τη χρησιμότητα της ΦΜΚΣ για σκοπούς ταυτοποίησης. Πρέπει να ληφθεί υπόψη ότι η πρακτική υλοποίηση και αποτελεσματικότητα αυτών των προδιαγραφών μπορεί να διαφέρει ανάλογα με τη συγκεκριμένη τεχνολογία PUF και το πλαίσιο εφαρμογής. [16]

1.2 Βασικές έννοιες και βασικά χαρακτηριστικά

Τα βασικά χαρακτηριστικά των ΦΜΚΣ χωρίζονται σε

- Κατηγορία ΦΜΚΣ (Class Puf)

Η έννοια που περιγράφεται αναφέρεται στον τύπο ή την κατηγορία μιας κατασκευής η οποία συμβολίζεται με τον όρο P και αντιστοιχεί στην πλήρη περιγραφή ενός συγκεκριμένου τύπου κατασκευής. Μια κατηγορία ΦΜΚΣ παρέχει μια διαδικασία δημιουργίας "P.Create", που χρησιμοποιείται για να φτιάξει νέα παραδείγματα (instances) του P. Για να φτιάξουμε εισόδους για το P PUF χρησιμοποιούμε την $P.Create(r^c)$ όπου το r^c πιθανή τυχαιότητα στη δημιουργία των εισόδων. Άρα με την P.create αναπαριστούμε τη λεπτομερή φυσική διαδικασία παραγωγής για να υλοποιηθεί ο σχεδιασμός της ΦΜΚΣ.

- ΦΜΚΣ στιγμιότυπο (PUF Instance)

Μία ΦΜΚΣ (puf) Instance αποτελεί μια ξεχωριστή υλοποίηση ενός στιγμιότυπου μίας φυσικής μη κλωνοποιήσιμης συνάρτησης (P).

$$P \equiv \{puf_i \leftarrow P.Create(r_i^c) : \forall i, r_i^c \leftarrow \{0, 1\}^*\}$$

Η περίπτωση μίας φυσικής μη κλωνοποιήσιμης συνάρτησης (puf) θεωρείται ως μια συγκεκριμένη, σταθεροποιημένη κατάσταση της κατασκευής που περιγράφεται από την κλάση της μη κλωνοποιήσιμης συνάρτησής της. Η κατασκευή πολλών κλάσεων ΦΜΚΣ επιτρέπει τη δυνατότητα να διαμορφώνεται μέρος της κατάστασης της περίπτωσης P , δηλαδή δεν είναι απαραίτητο να μεταβάλλεται από την διαδικασία δημιουργίας αλλά μπορεί να οριστεί και να αλλάξει εύκολα με τη χρήση εξωτερικής εισόδου πχ διαφορά θερμοκρασίας. Όταν απαιτείται, ορίζουμε το μεταβλητό μέρος της κατάστασης μιας περίπτωσης ΦΜΚΣ ως "πρόκληση" (challenge), η οποία εφαρμόζεται στην περίπτωση ΦΜΚΣ. Το σύνολο όλων των δυνατών προκλήσεων x που μπορούν να εφαρμοστούν σε μια περίπτωση της κλάσης PUF P συμβολίζεται ως X_P . Πέρα όμως από την αλλαγή θερμοκρασίας μία πρόκληση μπορεί να είναι να βάλουμε στην είσοδο του συστήματος της ΦΜΚΣ ένα σήμα όπως έναν ηλεκτρικό παλμό ή μία σειρά δυαδικών bits μέχρι και μία ακτίνα lazer [6]

- Αξιολόγηση ΦΜΚΣ (PUF Evaluation)

Κάθε στιγμιότυπο μίας φυσικής μη κλωνοποιήσιμης συνάρτησης παρέχει μία διαδικασία αξιολόγησης puf.Eval η οποία παράγει ένα αποτέλεσμα που αντιπροσωπεύει μία μέτρηση της περίπτωσης ΦΜΚΣ. Όταν ένα στιγμιότυπο ΦΜΚΣ μπορεί να δεχτεί μία "πρόκληση" γράφουμε $\text{puff}(x).\text{Eval}$. Το $\text{puf}(x).\text{Eval}$ είναι επίσης μια πιθανοτική διαδικασία, την οποία καθορίζουμε όταν είναι απαραίτητο ως $\text{puf}(x).\text{Eval}(r^E \leftarrow \{0,1\}^*)$

Άρα προκύπτει ότι το αποτέλεσμα της αξιολόγησης (evaluation) είναι το αποτέλεσμα ενός φυσικού πειράματος, το οποίο παράγει μια συγκεκριμένη μέτρηση της φυσικής κατάστασης του στιγμιότυπου μίας φυσικής μη κλωνοποιήσιμης συνάρτησης εκείνη την χρονική στιγμή αυτή η μέτρηση μπορεί να ονομαστεί και απόκριση της περίπτωσης ΦΜΚΣ. Η κατηγορία όλων των πιθανών τιμών απόκρισης που μια περίπτωση PUF μιας κλάσης P μπορεί να παράγει σημειώνεται ως y_P . Το αποτέλεσμα μίας αξιολόγησης ενός στιγμιότυπου ΦΜΚΣ μπορεί να επηρεαστεί από εξωτερικούς παράγοντες όπως για παράδειγμα την θερμοκρασία περιβάλλοντος την πραγματική τάση της παροχής κ.α.. Άρα όταν κάνουμε μία αξιολόγηση πρέπει να προσδιορίζουμε και τις συνθήκες στις οποίες έγινε. Όταν παραλείπετε η αναφορά σε αυτές τότε λέμε ότι η αξιολόγηση έγινε στις ονομαστικές συνθήκες λειτουργίας [16]

1.3 Είδη ΦΜΚΣ

Με μία μαθηματική ματιά, μία ΦΜΚΣ δεν είναι αυστηρά μια συνάρτηση, διότι μια μόνο είσοδος (πρόκληση) μπορεί να συσχετίζεται με περισσότερες από μία εξόδους (απόκριση) λόγω των μη ελέγξιμων επιδράσεων του φυσικού περιβάλλοντος και του τυχαίου θορύβου στην δημιουργία της απόκρισης. Άρα όπως προέκυψε και από τα προηγούμενα μια ΦΜΚΣ ορίζεται ως πιθανοτική συνάρτηση, δηλαδή μια συνάρτηση για την οποία μέρος της εισόδου αποτελεί μια μη ελέγξιμη τυχαία μεταβλητή. Αυτό καμιά φορά γίνεται προφανές παρέχοντας μια επιπλέον είσοδο r σε μια συνάρτηση, που αντιπροσωπεύει ένα (μη προκαθορισμένο) αριθμό δίκαιων ρίψεων κέρματος. Εναλλακτικά, μπορούμε να θεωρήσουμε το αποτέλεσμα μιας πιθανοτικής συνάρτησης ως μια τυχαία μεταβλητή με μια συγκεκριμένη κατανομή πιθανοτήτων, αντί για μια προκαθορισμένη τιμή.

Οι φυσικές κατασκευές των ΦΜΚΣ και αυτών που μοιάζουν με ΦΜΚΣ βασίζονται σε μία ποικιλία τεχνολογιών και υλικών, όπως γυαλί, πλαστικό, χαρτί, ηλεκτρονικά εξαρτήματα και

ολοκληρωμένα κυκλώματα πυριτίου (ICs). Μια πιθανή κατηγοριοποίηση των κατασκευών ΦΜΚΣ βασίζεται στο αν αποτελούνται ή όχι από ηλεκτρονικά στοιχεία. Άρα οι ΦΜΚΣ μπορούν να χωριστούν σε ηλεκτρονικά και μη.

1.3.1 Μη ηλεκτρονικές ΦΜΚΣ

Οι μη-ηλεκτρονικές ΦΜΚΣ είναι αυτές που βασίζονται σε μη-ηλεκτρονικές τεχνολογίες. Αναφερόμαστε ειδικά στη βασικό σχεδιασμό της ΦΜΚΣ με μη-ηλεκτρονικά στοιχεία,. [6] Για παράδειγμα, η τυχαία δομή των ινών ενός φύλλου χαρτιού ή η τυχαία ανάκλαση των χαρακτηριστικών διάθλασης ενός οπτικού μέσου. Εδώ να σημειωθεί ότι η ονομασία "μη-ηλεκτρονικό" αντανakλά την προέλευση της συμπεριφοράς της ΦΜΚΣ και όχι τον τρόπο με τον οποίο επεξεργάζονται ή αποθηκεύονται οι αποκρίσεις της ΦΜΚΣ, καθώς αυτά είναι δευτερεύουσες λειτουργίες του μπορούν να εφαρμοστούν με τη χρήση ηλεκτρονικών και ψηφιακών υποσυστημάτων.

1.3.2 Ηλεκτρονικές ΦΜΚΣ

Αντίθετα, οι ηλεκτρονικές ΦΜΚΣ χρησιμοποιούν ηλεκτρονικά εξαρτήματα για τη βασική τους λειτουργία ,δηλαδή κατασκευές που περιλαμβάνουν τυχαίες διακυμάνσεις στα ηλεκτρονικά τους χαρακτηριστικά, όπως η αντίσταση, η χωρητικότητα κ.λπ. Φυσικά αυτού του τύπου ΦΜΚΣ είναι πιο εύκολο να συνδεθούν με άλλα ηλεκτρονικά εξαρτήματα και είναι πιο εύκολα στην επεξεργασία λόγω των ψηφιακών αποκρίσεων τους. [6]Είναι σημαντικό να σημειώσουμε ότι δεν υπάρχει πάντα αυστηρός διαχωρισμός μεταξύ ηλεκτρονικών και μη-ηλεκτρονικών ΦΜΚΣ.

Κάτι άλλο που αξίζει να σημειωθεί είναι ότι οι διακυμάνσεις στην ηλεκτρονική συμπεριφορά είναι σχεδόν πάντα αποτέλεσμα των διακυμάνσεων σε μη-ηλεκτρονικές παραμέτρους σε ένα χαμηλότερο αφαιρετικό φυσικό επίπεδο, όπως για παράδειγμα το μήκος και η συνοχή υλικού ενός μεταλλικού καλωδίου καθορίζουν την αντίστασή του, η επιφάνει² α και η απόσταση μεταξύ δύο αγωγών καθορίζουν τη χωρητικότητά του, κ.λπ.[17]

1.3.3 ΦΜΚΣ Πυριτίου

Οι ΦΜΚΣ πυριτίου είναι τα πιο γνωστά και διαδεδομένα στην κατηγορία των ηλεκτρονικών ΦΜΚΣ. Αυτού του τύπου ΦΜΚΣ είναι ολοκληρωμένα ηλεκτρονικά κυκλώματα που εμφανίζουν συμπεριφορά ΦΜΚΣ και ενσωματώνονται σε ένα κύκλωμα πυριτίου (silicon chip) Αυτές οι φυσικές διακυμάνσεις είναι μοναδικές για κάθε σύνθεση πυριτίου, καθιστώντας τις ΦΜΚΣ πυριτίου κατάλληλες για αναγνώριση σε συστήματα ασφαλείας, αναγνώρισης και πιστοποίησης. Επειδή οι ΦΜΚΣ πυριτίου μπορούν να συνδεθούν απευθείας σε τυπικά ψηφιακά κυκλώματα που ενσωματώνονται στον ίδιο το chip, μπορούν να χρησιμοποιηθούν αμέσως στην κρυπτογραφική υλοποίηση. Για αυτόν τον λόγο, είναι προφανές ότι οι ΦΜΚΣ πυριτίου είναι ιδιαίτερα ενδιαφέροντα για λύσεις ασφαλείας

1.4 Εσωτερικές και μη Εσωτερικές ΦΜΚΣ

Εσωτερικές ΦΜΚΣ (I-PUFs): Οι εσωτερικές ΦΜΚΣ παίρνουν τη μοναδικότητά τους και την απροβλεψιμότητά τους από τα ενσωματωμένα φυσικά χαρακτηριστικά ή ιδιότητες του hardware ή του υλικού που χρησιμοποιείται για την υλοποίησή τους. Αυτά τα χαρακτηριστικά είναι μοναδικά για κάθε περίπτωση ΦΜΚΣ. [6] Για παράδειγμα, η συγκεκριμένη διάταξη των τρανζίστορ σε ένα ολοκληρωμένο κύκλωμα πυριτίου ή η τυχαία διακύμανση της αντίστασης των καλωδίων μπορούν να χρησιμοποιηθούν για τη δημιουργία εσωτερικών ΦΜΚΣ. Οι εσωτερικές ΦΜΚΣ θεωρούνται γενικά πιο ασφαλείς, καθώς η μοναδικότητά τους είναι βαθιά ριζωμένη σε φυσικά χαρακτηριστικά που είναι δύσκολο να αντιγραφούν. Αυτή η μοναδικότητα μπορεί να θεωρηθεί και σαν δαχτυλικό αποτύπωμα της πλακέτας [9]

Μη-εσωτερικές ΦΜΚΣ (NI-PUFs): Αντίθετα, οι μη-εσωτερικές ΦΜΚΣ δεν βασίζονται σε ενσωματωμένα φυσικά χαρακτηριστικά, αλλά χρησιμοποιούν πρόσθετα εξαρτήματα ή χαρακτηριστικά για να δημιουργήσουν αποκρίσεις. Αυτές οι αποκρίσεις μπορεί ακόμα να είναι μοναδικές και απρόβλεπτες, αλλά δεν σχετίζονται με τα εσωτερικά φυσικά χαρακτηριστικά του υλικού. Παραδείγματα μη-εσωτερικών ΦΜΚΣ περιλαμβάνουν ΦΜΚΣ που χρησιμοποιούν λογισμικούς αλγόριθμους ή κρυπτογραφικές λειτουργίες για να δημιουργήσουν αποκρίσεις. Οι μη-εσωτερικές ΦΜΚΣ μπορεί να είναι περισσότερο ευάλωτες σε ορισμένους τύπους επιθέσεων, καθώς βασίζονται σε επιπρόσθετα εξαρτήματα ή αλγόριθμους.

Η επιλογή μεταξύ εσωτερικών και μη-εσωτερικών ΦΜΚΣ εξαρτάται από τη συγκεκριμένη εφαρμογή και τις απαιτήσεις ασφαλείας. Οι εσωτερικές ΦΜΚΣ συχνά προτιμώνται για εφαρμογές όπου υψηλότερο επίπεδο ασφάλειας και μοναδικότητας είναι ουσιώδες, ενώ οι μη-εσωτερικές ΦΜΚΣ μπορεί να είναι πιο κατάλληλες για εφαρμογές όπου προτιμάται η ευελιξία και η ευκολία υλοποίησης.

1.4.1 Εξωτερικές έναντι Εσωτερικών Αξιολογήσεων

Όπως έχει αναφερθεί μια αξιολόγηση της ΦΜΚΣ είναι μια φυσική μέτρηση που μπορεί να παρουσιάζεται με πολλούς τρόπους. Υπάρχουν δύο τρόποι που μπορεί να γίνει μία αξιολόγηση εξωτερικά και εσωτερικά. Μια εξωτερική αξιολόγηση είναι μια μέτρηση χαρακτηριστικών που είναι εξωτερικά παρατηρήσιμα ή πραγματοποιείται χρησιμοποιώντας εξοπλισμό που είναι εξωτερικός του φυσικό στοιχείου που περιέχει τα χαρακτηριστικά. Όπως για παράδειγμα όταν μετράμε ένα transistor σε ένα εργαστήριο ηλεκτρονικής. Από την άλλη πλευρά, υπάρχουν οι εσωτερικές αξιολογήσεις, δηλαδή μετρήσεις εσωτερικών χαρακτηριστικών που πραγματοποιούνται από εξοπλισμό που είναι πλήρως ενσωματωμένος στην ίδια την περίπτωση. Αυτό μπορεί να φαίνεται ως μια αρκετά τυχαία διάκριση, αλλά υπάρχουν ορισμένα σημαντικά πλεονεκτήματα για τις εσωτερικές αξιολογήσεις:

Το πρώτο πλεονέκτημα είναι αποκλειστικά πρακτικό. Έχοντας τον εξοπλισμό μέτρησης ενσωματωμένο σε κάθε περίπτωση (instance) της ΦΜΚΣ σημαίνει ότι κάθε περίπτωση ΦΜΚΣ μπορεί να αξιολογήσει τον εαυτό της χωρίς καμία εξωτερική επιρροή ή περιορισμό. Οι εσωτερικές αξιολογήσεις είναι συνήθως πιο ακριβείς, καθώς υπάρχει λιγότερος χώρος για εξωτερικές επιρροές και σφάλματα μέτρησης.

Δεύτερον, υπάρχει ένα αρκετά σημαντικό πλεονέκτημα σε ό,τι αφορά την ασφάλεια των εσωτερικών αξιολογήσεων, καθώς οι αποκρίσεις της ΦΜΚΣ προέρχονται από μέσα από το αντικείμενο ενσωμάτωσης. Αυτό σημαίνει ότι, όσο διατηρεί μια περίπτωση την απόκρισή της

κρυμμένη από τον έξω κόσμο, μπορεί να θεωρηθεί ως εσωτερικό μυστικό. Αυτό είναι ιδιαίτερα σημαντικό όταν το αντικείμενο ενσωμάτωσης είναι ένα ψηφιακό κύκλωμα, αφού σε αυτήν την περίπτωση η απόκριση της ΦΜΚΣ μπορεί να χρησιμοποιηθεί αμέσως ως μυστική είσοδος για την ενσωματωμένη υλοποίηση ενός κρυπτογραφικού αλγορίθμου.

Ένα πιθανό μειονέκτημα της εσωτερικής αξιολόγησης είναι ότι πρέπει να εμπιστευόμαστε τον ενσωματωμένο εξοπλισμό μέτρησης, καθώς είναι αδύνατο να επαληθευτεί εξωτερικά αν οι μετρήσεις πραγματοποιούνται όπως αναμένεται. Αντίθετα, στην εξωτερική αξιολόγηση είναι συχνά εφικτό να παρακολουθήσετε την ίδια τη διαδικασία μέτρησης κατά τη διάρκεια της εκτέλεσής της, προσφέροντας μεγαλύτερη διαφάνεια και επαληθευσιμότητα.

Η επιλογή μεταξύ εξωτερικών και εσωτερικών αξιολογήσεων εξαρτάται από τις απαιτήσεις της εφαρμογής και την ασφάλεια που απαιτείται. Κάθε μέθοδος έχει τα πλεονεκτήματά της και τα μειονεκτήματά της και μπορεί να επιλεγεί ανάλογα με την περίπτωση .

1.4.2 Εμφανείς έναντι Σιωπηλές Τυχαίες Διακυμάνσεις

Μια δεύτερη διάκριση βασισμένη στην κατασκευή λαμβάνει υπόψη την πηγή της τυχαιότητας των αξιολογούμενων χαρακτηριστικών σε μία ΦΜΚΣ. Ο διαδικασία κατασκευής των περιπτώσεων ΦΜΚΣ μπορεί να περιλαμβάνει μια ρητή διαδικασία τυχαιοποίησης με σκοπό την εισαγωγή τυχαίων χαρακτηριστικών που αργότερα θα μετρηθούν κατά την αξιολόγηση της ΦΜΚΣ. Υπάρχουν ορισμένα λεπτά, αλλά ενδιαφέροντα πλεονεκτήματα στις κατασκευές ΦΜΚΣ που βασίζονται σε ασαφείς τυχαίες διακυμάνσεις:

- Από πρακτικής άποψης, οι ασαφείς τυχαίες διακυμάνσεις συνήθως δεν επιφέρουν επιπλέον κόστος, επειδή υπάρχουν φυσικά ως μέρος της διαδικασίας παραγωγής. Αυτές οι διακυμάνσεις είναι εγγενείς και αναπόφευκτες και προκύπτουν κατά την κατασκευή προϊόντων.
- Αντίθετα, η εισαγωγή ρητών διαδικασιών τυχαιοποίησης μπορεί να είναι δαπανηρή, ειδικά όταν λαμβάνονται υπόψη οι περιορισμοί χρονισμού στη ροή κατασκευής υψηλού όγκου προϊόντων. Η ρητή τυχαιοποίηση περιλαμβάνει την εισαγωγή επιτηδευμένης τυχαιότητας σε ένα προϊόν, πράγμα που μπορεί να απαιτήσει επιπρόσθετους πόρους, σχεδιασμό και πιο προσεκτική υλοποίηση.
- Οι ασαφείς τυχαίες διακυμάνσεις στην κατασκευή, γενικά δεν είναι επιθυμητές, καθώς μπορούν να επηρεάσουν αρνητικά την απόδοση παραγωγής. Οι κατασκευαστές λαμβάνουν συνήθως εκτεταμένα μέτρα για την ελαχιστοποίηση των διακυμάνσεων της διαδικασίας όσο το δυνατόν περισσότερο. Ωστόσο, είναι τεχνικά αδύνατο να εξαλείψουμε εντελώς όλες τις τυχαίες επιδράσεις κατά τη διάρκεια της κατασκευής. Οι διακυμάνσεις διαδικασίας είναι αναπόφευκτες στη διαδικασία κατασκευής και θα υπάρχουν πάντα σε κάποιο βαθμό.
- Ένα κύριο πλεονέκτημα ασφαλείας των κατασκευών βασισμένων σε ασαφείς τυχαίες διακυμάνσεις ΦΜΚΣ είναι ότι ακόμη και οι ίδιοι οι κατασκευαστές, που έχουν τον πλήρη έλεγχο της διαδικασίας παραγωγής και πρόσβαση σε όλες τις λεπτομέρειες, δεν μπορούν να αφαιρέσουν ή να ελέγξουν τα τυχαία χαρακτηριστικά που αποτελούν τον πυρήνα της λειτουργίας της ΦΜΚΣ. Αυτή η εγγενής τυχαιότητα προσθέτει ένα επιπρόσθετο επίπεδο ασφάλειας, διότι δεν μπορεί να υποστεί κάποια επεξεργασία ή παρεμβολή, ακόμη και από όσους είναι εξοικειωμένοι με τη διαδικασία κατασκευής.

Συνοψίζοντας για αυτή την ενότητα, οι ασαφείς τυχαίες διακυμάνσεις είναι φυσικές και παρέχουν ένα πλεονέκτημα ασφάλειας στις ΦΜΚΣ, διότι δυσκολεύουν ακόμη και τον κατασκευαστή να ελέγξει ή να αφαιρέσει τα τυχαία χαρακτηριστικά που συμβάλλουν στη μοναδική συμπεριφορά της ΦΜΚΣ.

1.5 Εγγενείς, μη Εγγενείς ΦΜΚΣ

1.5.1 Μη εγγενείς ΦΜΚΣ

Υπάρχουν διάφορες κατασκευές ΦΜΚΣ που έχουν προταθεί και τις κατηγοριοποιούμε ως μη-εγγενείς σύμφωνα με αυτή την κατηγοριοποίηση, π.χ.:

Η οπτική ΦΜΚΣ και η ΦΜΚΣ επίστρωσης είναι δύο παραδείγματα μη-εγγενών ΦΜΚΣ που βασίζονται σε διαφορετικές αρχές λειτουργίας:

1. **Οπτική ΦΜΚΣ:** Όπως προτάθηκε από τον Pappu και τους συνεργάτες, η οπτική ΦΜΚΣ βασίζεται στο μοναδικό σχήμα speckle που παράγεται όταν λαμβάνετε μια λείζερ ακτινοβολία και αυτή διασχίζει ένα διαφανές υλικό με τυχαία διασκορπισμένα σωματίδια διασκορπισμού. Η οπτική ΦΜΚΣ εξάγει μοναδικά χαρακτηριστικά από το μοτίβο των speckles και τις τυχαίες ιδιότητες τους.
2. **ΦΜΚΣ επίστρωσης:** Όπως προτάθηκε από τον Tuyls και συνεργάτες, η ΦΜΚΣ επίστρωσης βασίζεται στη μοναδική χωρητική αντίσταση μεταξύ δύο μεταλλικών γραμμών σε ένα ολοκληρωμένο κύκλωμα πυριτίου, το οποίο έχει επικαλυφθεί με ένα υλικό που περιέχει τυχαία διασκορπισμένα διηλεκτρικά σωματίδια. Η ΦΜΚΣ επίστρωσης εξάγει μοναδικά χαρακτηριστικά από την χωρητική αντίσταση μεταξύ αυτών των μεταλλικών γραμμών και τις τυχαίες ιδιότητές της.

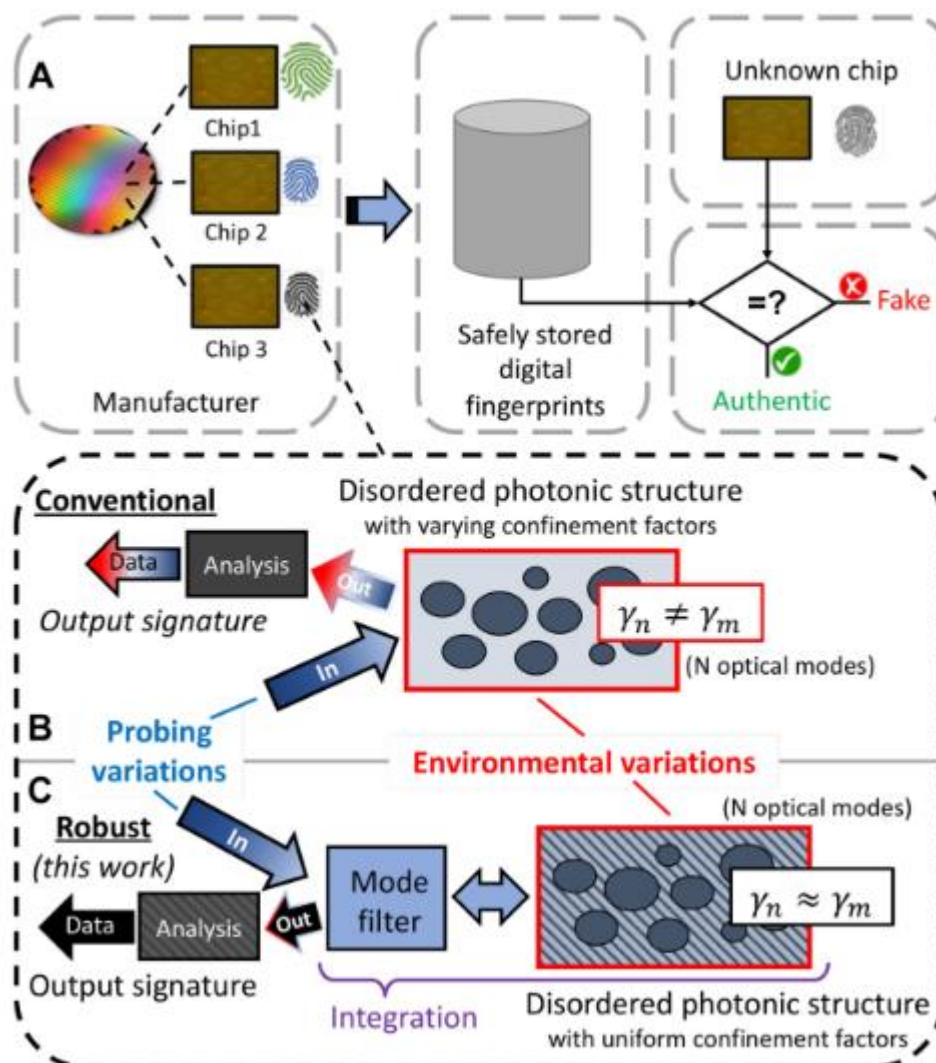
Και οι δύο αυτές ΦΜΚΣ θεωρούνται μη-εγγενείς επειδή αξιολογούνται εξωτερικά, δηλαδή μέσω της παρατήρησης των χαρακτηριστικών τους από έξω, και επειδή οι τυχαίες ιδιότητές τους εισάγονται ρητά μέσω της κατανομής των διασκορπισμένων σωματιδίων διασκορπισμού στο υλικό τους.[17]

1.5.1.1 Οπτική ΦΜΚΣ

Με ελκυστικό τρόπο, η τεράστια δυνατότητα αποθήκευσης πληροφοριών και η πλούσια φυσική των φωτονικών συστημάτων προσφέρουν τη δυνατότητα δημιουργίας συσκευών ασφαλείας, τόσο παθητικών όσο και ενεργών, που λειτουργούν τόσο σε κλασικό όσο και σε κβαντικό επίπεδο. Επιπλέον, οπτικά συστήματα χαρακτηριζόμενα από περίπλοκα και διανεμημένα πολλαπλά φαινόμενα σκέδασης με υψηλή ευαισθησία κατά την κατασκευή μπορούν να αποδειχθούν δύσκολα στη μοντελοποίηση ή στο ρυθμιστικό επίπεδο. Έτσι, παραμένει ανοικτό το ερώτημα σχετικά με το εάν η αναδυόμενη τεχνολογία θα μπορούσε να επιτρέψει επιτυχείς φυσικές ή μηχανικής μάθησης επιθέσεις σε εξελισσόμενες οπτικές ΦΜΚΣ.

Στο οπτικό πεδίο, τα σχήματα speckle είναι γνωστά για την υψηλή τους πολυπλοκότητα και μοναδικότητα που προκύπτουν από την υποκείμενη φωτονική αταξία, καθιστώντας τα κατάλληλα για τη δημιουργία μεγάλης 'intra-chip' ποικιλίας. Ωστόσο, όπως φαίνεται στο σχήμα Β παρακάτω, τα φαινόμενα speckle από συμβατικά ατακτοφωτικά συστήματα είναι αρκετά ευαίσθητα στο probing και τις περιβαλλοντικές διακυμάνσεις. Έτσι, οι οπτικές ΦΜΚΣ που προκύπτουν από μοτίβα laser speckle έχουν πολύ μεγάλη inter-chip διακύμανση και γενικά δε θεωρούνται ως ανθεκτική τεχνολογία, καθώς απαιτούν ακριβή οπτική ευθυγράμμιση, κλίσης, πόλωσης, θερμοκρασίας και σταθερή 2D χωρική ανάλυση οπτικής απεικόνισης για μέτρηση και επαλήθευση.

Για να υλοποιηθεί μια ανθεκτική και κλιμακούμενη τεχνολογία ΦΜΚΣ, είναι επιτακτική η ενίσχυση της 'inter-chip' διακύμανσης ενώ ταυτόχρονα πρέπει να ελαχιστοποιηθεί η 'intra-chip' διακύμανση. Επομένως, η έμφαση πρέπει να τεθεί στην απόκτηση υπογραφών που είναι ορατές, εύκολες στην μέτρησης, υπολογιστικά αποδοτικές για ανάλυση και ανθεκτικές σε μια



σειρά από συνθήκες μέτρησης και/ή περιβαλλοντικές συνθήκες.

Εικόνα 1 (Α) Ένα τυπικό σχήμα ΦΜΚΣ. Επισκόπηση: (Β) μίας συμβατικής οπτικής ΦΜΚΣ και (Γ) μίας ανθεκτικής οπτικής ΦΜΚΣ. Εδώ, η ανθεκτικότητα αναφέρεται στην επίτευξη ανοσίας έναντι αναγνώρισης και περιβαλλοντικών μεταβολών. [2]

Περιορίζοντας το φως σε ένα κυματοδηγό στην επιφάνεια ενός τσιπ είναι μια εναλλακτική μέθοδος για τον έλεγχο της χωρικής τυχειότητας και θα μπορούσε να πραγματοποιηθεί με έναν υψηλά ενσωματωμένο τρόπο χωρίς εικονική ή διακριτική δειγματοληψία σε σειριακές συσκευές, ενώ αποθηκεύει μεγάλη ποσότητα πληροφοριών στη συχνότητα ή την κρουστική απόκριση. Παράλληλα, γίνονται έρευνες σε ΦΜΚΣ σε φωτονικά τσιπ χρησιμοποιώντας χαοτικούς οπτικούς μηχανισμούς. Το οπτικό χάος βασίζεται στην ευαισθησία στις αρχικές συνθήκες, που αλλοιώνει δραματικά τα προσεγγιστικά πεδία speckle που προκύπτουν από ένα μεγάλο αριθμό χωρικών τρόπων με μεταβαλλόμενο Q/V . Ωστόσο, αυτοί οι τρόποι επίσης εκδηλώνουν γενικά μεταβλητούς παράγοντες συγκέντρωσης τύπου γ_n (διάδραση ύλης φωτός).

$$\gamma_n = \frac{\int_{\text{activate}} \varepsilon(\vec{r}) |E_n(\vec{r})|^2 dV}{\int \varepsilon(\vec{r}) |E_n(\vec{r})|^2 dV}$$

Έτσι, τα διάφορα modes μπορεί να επηρεαστούν διαφορετικά σε σχέση με την συχνότητα, ανάλογα με την $\frac{\Delta\omega_n}{\omega_n} = -(\Delta n/n_g)\gamma_n$.

Οπου Δn είναι ο δείκτης επιρροής της συχνότητας και n_g είναι ο δείκτης επιρροής της συχνότητας για την ομάδα που ανήκει το μέσο.

Αυτό καθιστά τις συσκευές αυτού του είδους εξαιρετικά ευαίσθητες σε όλες τις συνθήκες, συμπεριλαμβανομένων τόσο στον κύκλο κατασκευής τους όσο και των περιβαλλοντικών διακυμάνσεων, όπως οι μεταβολές της θερμοκρασίας ή οι μη γραμμικές επιδράσεις. Από την άλλη το Rayleigh backscatter από single-mode Fiber (είναι ένας τύπος ίνας που επιτρέπει ένα διάνυσμα φωτός) [10] κυματοδηγό παρέχει σταθερές και αξιόπιστες τυχαίες υπογραφές. Υποστηρίζοντας μόνο έναν εγκάρσιο τρόπο με ομοιόμορφο γ_n , η τυχαία συχνοτική απόκριση τέτοιων συσκευών δεν "εξελίσσεται" αλλά μετακινείται κατά προσδιορισμένο τρόπο ως απόκριση σε περιβαλλοντικούς παράγοντες, όπως η θερμοκρασία, η ένταση ή η γήρανση.

Αυτές οι υπογραφές των ανακλάσεων που έρχονται πίσω, ωστόσο, είναι κατ' ουσίαν πολύ ασθενείς (χαμηλή ορατότητα), απαιτώντας την ανίχνευσή αυτών των ανακλάσεων στο χρονικό ή συχνοτικό πεδίο, χωρίς να είναι δυνατό να χρησιμοποιηθούν στη μετάδοση. Η αντιμετώπιση όλων αυτών των ζητημάτων είναι κρίσιμη για την αξιοποίηση της τεχνολογίας ΦΜΚΣ και την προώθηση και των δύο κλασικών και κβαντικών τεχνολογιών ασφαλείας υλικού.

1.5.2 Εγγενείς ΦΜΚΣ

Μια ενδογενής ΦΜΚΣ είναι ένας τύπος ΦΜΚΣ που αντλεί τα μοναδικά αναγνωριστικά ή τις αποκρίσεις της από εγγενείς φυσικές παραλλαγές που συμβαίνουν φυσικά κατά τη διαδικασία κατασκευής μιας συσκευής ή ενός τσιπ.

Τα βασικά χαρακτηριστικά και οι αρχές των ενδογενή ΦΜΚΣ (Intrinsic PUFs):

Παραλλαγές κατασκευής: Οι Εγγενείς ΦΜΚΣ εκμεταλλεύονται μικρές, ανεξέλεγκτες και συχνά τυχαίες παραλλαγές που προκύπτουν φυσικά κατά την κατασκευή συσκευών ή τσιπ

ημιαγωγών. Αυτές οι παραλλαγές μπορεί να περιλαμβάνουν παραλλαγές στις ιδιότητες των τρανζίστορ, στο πάχος οξειδίου ή σε άλλα φυσικά χαρακτηριστικά.

Χωρίς πρόσθετα στοιχεία: Οι εγγενείς ΦΜΚΣ δεν απαιτούν πρόσθετα εξαρτήματα υλικού ή εξωτερικούς αισθητήρες. Βασίζονται αποκλειστικά στις εγγενείς παραλλαγές εντός της συσκευής.

Εγγενής ασφάλεια: Επειδή οι παραλλαγές είναι εγγενείς στη διαδικασία κατασκευής και δεν μπορούν εύκολα να ελεγχθούν ή να αναπαραχθούν, οι εγγενείς ΦΜΚΣ θεωρούνται ιδιαίτερα ασφαλείς έναντι προσπαθειών κλωνοποίησης ή παραποίησης.

Ελάχιστη επιβάρυνση: Οι Εγγενής ΦΜΚΣ δεν προσθέτουν σημαντική επιβάρυνση στη διαδικασία κατασκευής. Χρησιμοποιούν τις υπάρχουσες παραλλαγές χωρίς την ανάγκη για σκόπιμες διαδικασίες τυχαιοποίησης.

Αξιοπιστία: Οι εγγενείς ΦΜΚΣ μπορούν να είναι αξιόπιστοι και ανθεκτικοί καθ' όλη τη διάρκεια ζωής της συσκευής, καθώς οι υποκείμενες παραλλαγές κατασκευής είναι σταθερές.

Στην πράξη, οι Εγγενής ΦΜΚΣ μπορούν να υλοποιηθούν με διάφορες τεχνικές, όπως η μέτρηση των μεταβολών καθυστέρησης στα κυκλώματα, οι διακυμάνσεις τάσης ή τα αναλογικά χαρακτηριστικά των τρανζίστορ. Οι εξαγόμενες αποκρίσεις μπορούν στη συνέχεια να χρησιμοποιηθούν για ασφαλείς σκοπούς που έχουν αναφερθεί και προηγουμένως παραγωγή κρυπτογραφικών κλειδιών κ.α.. Οι εγγενείς ΦΜΚΣ ευνοούνται σε εφαρμογές όπου η ισχυρή ασφάλεια και η μοναδικότητα είναι απαραίτητες και είναι ιδιαίτερα πολύτιμες σε σενάρια όπου πρόσθετα εξαρτήματα υλικού ή αισθητήρες δεν είναι εφικτά ή επιθυμητά.

Αξίζει να σημειωθεί ότι οι Εγγενής ΦΜΚΣ αποτελούν μια κατηγορία ΦΜΚΣ, ενώ υπάρχουν και άλλοι τύποι ΦΜΚΣ, όπως οι ισχυρά ΦΜΚΣ, οι ασθενείς ΦΜΚΣ κ.ο.κ., ο καθένας με τα δικά του χαρακτηριστικά και τις δικές του περιπτώσεις χρήσης.

1.6 Ασθενείς και ισχυρές ΦΜΚΣ

Οι δύο πιο σημαντικοί υποτύποι των ΦΜΚΣ, είναι τα λεγόμενα "Αδύναμα ΦΜΚΣ" και "Ισχυρά ΦΜΚΣ".

1.6.1 Χαρακτηριστικά των Αδύναμων ΦΜΚΣ

Οι Αδύναμες ΦΜΚΣ αποτελούν ουσιαστικά μια νέα μορφή αποθήκευσης μυστικών κλειδιών σε εύαλωτο hardware, προσφέροντας μια εναλλακτική στην ROM, Flash ή άλλου τύπου μνήμες (NVMs). Όπως όλες οι ΦΜΚΣ, οι Αδύναμες ΦΜΚΣ εμφανίζουν κάποια εσωτερική φυσική ασυνέχεια, αδύνατο στο να αντιγραφεί και διαθέτουν κάποιου είδους μηχανισμό πρόκλησης-απόκρισης που εκμεταλλεύεται αυτήν την ασυνέχεια. Πέραν από αυτό, τα χαρακτηριστικά τους χαρακτηρίζονται ως εξής:

1. Λίγες προκλήσεις: Μία Αδύναμη ΦΜΚΣ έχει πολύ λίγες, σταθερές προκλήσεις, συνήθως μόνο μία πρόκληση ανά παράδειγμα ή αλλιώς περίπτωση ΦΜΚΣ.
2. Περιορισμένης πρόσβασης αποκρίσεις: Σε όλες τις περιπτώσεις εκτός από ελάχιστες, η διεπαφή πρόκλησης-απόκρισης μίας Αδύναμης ΦΜΚΣ πρέπει να περιορίζεται στην

πρόσβαση. Υποθέτετε ότι κάποιος μη εξουσιοδοτημένος δεν μπορεί να έχει πρόσβαση στις απαντήσεις της Αδύναμης ΦΜΚΣ, ακόμη και αν έχει στην κατοχή του το hardware που φέρει η ΦΜΚΣ.

1.6.2 Χαρακτηριστικά των ισχυρών ΦΜΚΣ

Οι ισχυρές ΦΜΚΣ, σε αντίθεση με τις αδύναμες ΦΜΚΣ, παρουσιάζουν διάφορα χαρακτηριστικά που τα καθιστούν μοναδικά και ασφαλή. Τα χαρακτηριστικά αυτά περιγράφονται παρακάτω:

1. Πολλές προκλήσεις: Οι ισχυρές ΦΜΚΣ προσφέρουν έναν τεράστιο αριθμό πιθανών προκλήσεων, ιδανικά εκθετικά πολλές, σε κάποια παράμετρο του συστήματος. Αυτός ο εκτεταμένος χώρος προκλήσεων καθιστά ανέφικτο για έναν αντίπαλο να εκτελέσει πλήρη ανάγνωση όλων των ζευγών προκλήσεων-απαντήσεων (CRPs), ακόμη και αν έχει στη φυσική του κατοχή τη ΦΜΚΣ για παρατεταμένο χρονικό διάστημα.
2. Μη προβλεψιμότητα: Ακόμη και αν ένας αντίπαλος έχει γνώση ενός μεγάλου υποσυνόλου CRPs, δεν μπορεί να προεκτείνει ή να προβλέψει τα υπόλοιπα, άγνωστα CRPs. Οι ισχυρές ΦΜΚΣ παρουσιάζουν υψηλή απροβλεψιμότητα στις αποκρίσεις τους.
3. Μη προστατευμένη διεπαφή πρόκλησης-απάντησης: Στις περισσότερες εφαρμογές των ισχυρών ΦΜΚΣ, θεωρείται ότι διαθέτουν μια ελεύθερα πρόσβαση στη διεπαφή ή μηχανισμό πρόκλησης-απάντησης. Αυτό σημαίνει ότι οποιοσδήποτε έχει στη κατοχή τη ΦΜΚΣ ή το υλικό που το περιέχει μπορεί να εφαρμόσει αυθαίρετες προκλήσεις μία ισχυρή ΦΜΚΣ και να διαβάσει τις αντίστοιχες απαντήσεις.

Αυτά τα χαρακτηριστικά διαφοροποιούν τις ισχυρές ΦΜΚΣ από τις αδύναμες ΦΜΚΣ και έχουν σημαντικές επιπτώσεις στην ασφάλειά τους:

Περιορισμοί πρόσβασης: Σε αντίθεση με τις ασθενείς ΦΜΚΣ, τα οποία συχνά βασίζονται σε περιορισμούς πρόσβασης για την προστασία των απαντήσεών τους, οι ισχυρά ΦΜΚΣ υποθέτουν μια μη προστατευμένη διεπαφή πρόκλησης-απάντησης. Ως αποτέλεσμα, δεν υπάρχουν περιορισμοί πρόσβασης στις απαντήσεις των ΦΜΚΣ.

Επεμβατικές επιθέσεις: Οι επεμβατικές επιθέσεις στις αποκρίσεις ΦΜΚΣ καθίστανται ως επί το πλείστον παρωχημένες για τις ισχυρές ΦΜΚΣ λόγω της ελεύθερα προσβάσιμης διεπαφής πρόκλησης-απάντησης. Η απουσία περιορισμών πρόσβασης σημαίνει ότι ένας αντίπαλος μπορεί εύκολα να αλληλοεπιδράσει με τη ΦΜΚΣ.

Ωστόσο, αξίζει να σημειωθεί ότι η ελεύθερα προσβάσιμη διεπαφή πρόκλησης-απάντησης των ισχυρών ΦΜΚΣ έχει και μειονεκτήματα. Για να παραμείνουν ασφαλείς, οι ισχυρές ΦΜΚΣ πρέπει να έχουν μεγάλο αριθμό CRPs και οι απαντήσεις τους μπορεί να υπόκεινται σε επιθέσεις μοντελοποίησης, καθώς οι αντίπαλοι μπορούν εύκολα να συλλέγουν απαντήσεις. Παρά τις προκλήσεις αυτές, η πολυπλοκότητα και το απρόβλεπτο των ισχυρών ΦΜΚΣ τα καθιστούν πολύτιμο εργαλείο για την ενίσχυση της ασφάλειας σε διάφορες εφαρμογές.

Αυτά τα χαρακτηριστικά είναι κρίσιμα για την κατανόηση των ιδιοτήτων ασφαλείας και των πρακτικών επιπτώσεων των ισχυρών ΦΜΚΣ στην ασφάλεια υλικού και την κρυπτογραφία. [1]

1.7 Εγγενής ΦΜΚΣ (*Intrinsic PUF*)

Όλες οι γνωστές κατασκευές εσωτερικών ΦΜΚΣ είναι ΦΜΚΣ από πυρίτιο που βασίζονται σε τυχαίες διακυμάνσεις της διαδικασίας που συμβαίνουν κατά τη διάρκεια της διαδικασίας κατασκευής της πλακέτας .

1.7.1 ΦΜΚΣ Συγκριτής

Μία ΦΜΚΣ Συγκριτής (Arbiter PUF) είναι μία είδους εγγενούς ΦΜΚΣ καθυστέρησης που βασίζεται στις διαφορές καθυστέρησης εντός ψηφιακών κυκλωμάτων για να δημιουργήσει μια μοναδική και απρόβλεπτη απόκριση, συνήθως σε μορφή δυαδικής σειράς.

Τυπική λειτουργία μίας ΦΜΚΣ Συγκριτής:

Βασική Αρχή: Μία ΦΜΚΣ Συγκριτής εκμεταλλεύεται τις διακυμάνσεις στα μονοπάτια καθυστέρησης εντός ενός ψηφιακού κυκλώματος. Αποτελείται από δύο μονοπάτια, που συχνά αναφέρονται ως "Α" και "Β," τα μονοπάτια καθυστέρησης σχεδιάζονται να είναι απόλυτα συμμετρικά, δηλαδή οποιαδήποτε διαφορά στην καθυστέρηση προκαλείται αποκλειστικά από τις διακυμάνσεις της διαδικασίας. Αυτά τα μονοπάτια χρησιμοποιούνται για να συγκρίνουν τη σχετική καθυστέρηση των σημάτων που διέρχονται μέσα από αυτά.

Διάδοση Σήματος: Όταν μία πρόκληση εφαρμόζεται στα μονοπάτια "Α" και "Β," το σήμα θα διαδοθεί μέσα από κάθε μονοπάτι με ελαφρώς διαφορετικές καθυστερήσεις λόγω των διακυμάνσεων που προκύπτουν από την κατασκευή, τη θερμοκρασία, την τάση και άλλους παράγοντες περιβάλλοντος.

Σύγκριση : Τα σήματα από τα μονοπάτια "Α" και "Β" στη συνέχεια συγκρίνονται χρησιμοποιώντας έναν συγκριτή (arbiter), που είναι ένας ψηφιακός συγκριτής. Ο arbiter ανιχνεύει ποιο σήμα από τα μονοπάτια "Α" και "Β" φτάνει πρώτο και δημιουργεί μια δυαδική απόκριση βάσει αυτής της σύγκρισης.

Δημιουργία Απόκρισης: Η έξοδος του arbiter αποτελεί την απόκριση της ΦΜΚΣ. Αυτή η απόκριση είναι μοναδική για κάθε συγκεκριμένο αντίγραφο της ΦΜΚΣ λόγω των ενσωματωμένων διακυμάνσεων στα μονοπάτια καθυστέρησης. Μπορεί να αναπαρασταθεί ως δυαδική συμβολοσειρά, όπου τα bits αντιπροσωπεύουν τα αποτελέσματα πολλαπλών κύκλων προκλήσεων-αποκρίσεων.

Ζευγάρι Προκλήσεων-Αποκρίσεων: Για να ανακτήσετε την απόκριση της ΦΜΚΣ, εφαρμόζετε μια συγκεκριμένη πρόκληση (είσοδος) στο Arbiter PUF, και μετράτε την απόκριση. Αυτή η διαδικασία μπορεί να επαναληφθεί πολλές φορές για να ληφθεί μια σταθερή και αξιόπιστη απόκριση της ΦΜΚΣ.

Οι ΦΜΚΣ συγκριτές προσφέρουν αρκετά πλεονεκτήματα στις εφαρμογές ασφάλειας:

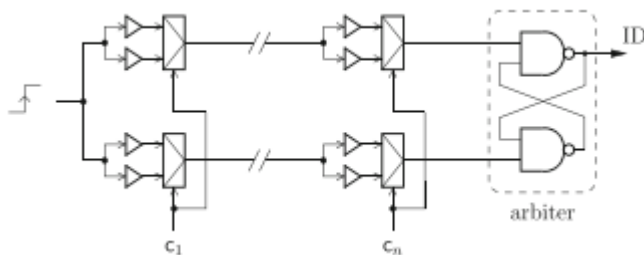
Απρόβλεπτοτητα: Οι ενσωματωμένες διακυμάνσεις στα μονοπάτια καθυστέρησης καθιστούν τις αποκρίσεις των ΦΜΚΣ συγκριτών υψηλά απρόβλεπτες και δύσκολες να αντιγραφούν ή να αναπαραχθούν.

Συμπαγής Σχεδίαση: Οι ΦΜΚΣ συγκριτές μπορούν να υλοποιηθούν σε σχετικά μικρές περιοχές των ενσωματωμένων κυκλωμάτων, καθιστώντας τα κατάλληλα για την ασφάλεια υλικού σε περιορισμένα περιβάλλοντα.

Χαμηλή Κατανάλωση Ενέργειας: Συνήθως, είναι συστατικά χαμηλής κατανάλωσης ενέργειας, καθιστώντας τα κατάλληλα για συσκευές που λειτουργούν με μπαταρία. [2]

Ωστόσο, οι ΦΜΚΣ συγκριτές έχουν επίσης ορισμένα μειονεκτήματα, συμπεριλαμβανομένης της ευαισθησίας σε επιθέσεις μοντελοποίησης και της ανάγκης για μηχανισμούς διόρθωσης σφαλμάτων για τον υπολογισμό του θορύβου και των μεταβολών στις αποκρίσεις με την πάροδο του χρόνου. [3]

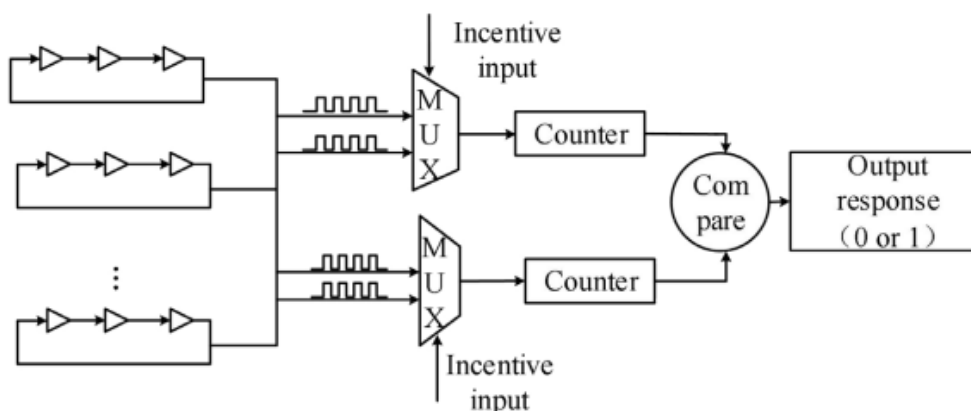
Συνολικά, οι ΦΜΚΣ συγκριτές αποτελούν ένα επιτυχημένο εργαλείο στην ασφάλεια υλικού, παρέχοντας έναν τρόπο για τη δημιουργία μοναδικών και απρόβλεπτων αναγνωριστικών ή κρυπτογραφικών κλειδιών βασισμένων στα ενσωματωμένα φυσικά χαρακτηριστικά των ψηφιακών κυκλωμάτων. [2]



Εικόνα 2: Arbiter Puf [3]

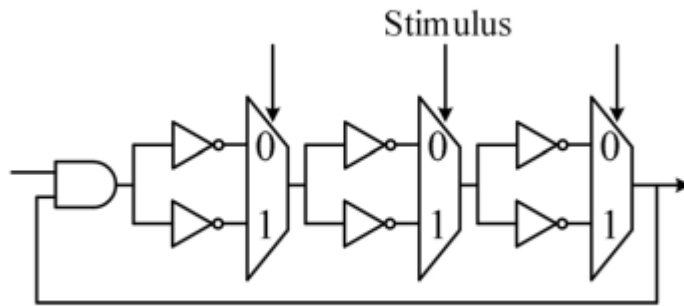
1.7.2 ΦΜΚΣ Κυκλικού Ταλαντωτή (Ring Oscillator Puf)

Το RO-PUF ανήκει και αυτό στην κατηγορία των ΦΜΚΣ καθυστέρησης και αποτελείται κυρίως από κύκλωμα κυκλικών ταλαντωτών, πολυπλέκτες, και μετρητές. Η βασική δομή του παρουσιάζεται στην εικόνα 2. Ο επιλογέας δεδομένων MUX λαμβάνει τα εξωτερικά σήματα εξόδου των δύο ταλαντωτών RO στα N κυκλώματα ταλάντωσης και το εξωτερικό σήμα εισόδου του συστήματος. Οι δύο μετρητές μετρούν αντίστοιχα τις συχνότητες ταλάντωσης των δύο κυκλικών ταλαντωτών RO κατά την ίδια χρονική περίοδο και τελικά εξάγονται μέσω του συγκριτή. Το αποτέλεσμα της συχνότητας ταλάντωσης, είναι λογικό 1 ή 0.



Εικόνα 3 Βασική δομή του RO-PUF [4]

Η δομή του παραδοσιακού παραμετροποιήσιμου κυκλώματος κυκλικής ταλάντωσης παρουσιάζεται στο εικόνα 3. Αυτή η μέθοδος παραμετροποίησης αποκτά μια συγκεκριμένη συχνότητα περιστροφής αλλάζοντας το εισερχόμενο σήμα εξάρτησης, ώστε το σήμα συχνότητας εξόδου να έχει υψηλότερη ασφάλεια. Ωστόσο, η δομή του παραμετροποιήσιμου RO-PUF που υλοποιείται με την παραδοσιακή μέθοδο είναι απλή, παράγει λίγα ζεύγη πρόκλησης-απάντησης (CRPs), η αξιοπιστία και η μοναδικότητα του σχεδιασμένου κυκλώματος με αυτό δεν είναι υψηλές, και το εισερχόμενο σήμα δεν είναι τυχαίο. Ένας επιτιθέμενος μπορεί να δημιουργήσει ένα μοντέλο επίθεσης συλλέγοντας CRPs. [4]



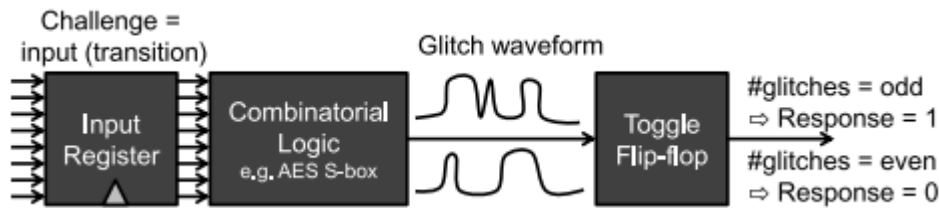
Εικόνα 4. Παραδοσιακή διαμορφώσιμη δομή RO [4]

1.7.3 ΦΜΚΣ Σφάλματος (Glitch Puf)

Ένας τρίτος τύπος κατασκευής ΦΜΚΣ βασίζεται στη συμπεριφορά σφάλματος (glitch) των λογικών κυκλωμάτων. Ένα κύκλωμα που δεν έχει εσωτερική κατάσταση, σημαίνει ότι η έξοδος του σε κατάσταση ισορροπίας καθορίζεται πλήρως από τα σήματα εισόδου του. Ωστόσο, όταν η λογική τιμή της εισόδου αλλάξει, μπορούν να συμβούν μεταβατικά φαινόμενα, δηλαδή μπορεί να χρειαστεί κάποιο χρονικό διάστημα προτού η έξοδος υιοθετήσει την τελική της τιμή. Όταν μετريέται με ακρίβεια, η συμπεριφορά σφάλματος (glitch) ενός τέτοιου κυκλώματος μπορεί να χρησιμοποιηθεί ως απόκριση ΦΜΚΣ.

Η αρχική τους πρόταση αποτελείται από μια περίπλοκη αρχιτεκτονική που πραγματοποιεί μία δειγματοληψία με υψηλή συχνότητα στο ολοκληρωμένο κύκλωμα καταγράφοντας την κηματομορφή σφάλματος και ένα κύκλωμα κβαντοποίησης των δειγμάτων που δημιουργεί ένα bit απόκρισης βασισμένο στα δεδομένα που δειγματολυντήθηκαν. Συνδέοντας απλώς μια συνδυαστική έξοδο σε μια εναλλαγή flip-flop, η τιμή της εναλλαγής flip-flop μετά τη μεταβατική φάση θα ισούται με την ισοτιμία του αριθμού των glitch που συνέβησαν

Για να βελτιωθεί η αξιοπιστία των αποκρίσεων ΦΜΚΣ, χρησιμοποιείται μια τεχνική bit-masking. Αυτή η μάσκα bit εφαρμόζεται στη συνέχεια στην απόκριση ΦΜΚΣ κατά τις επόμενες διαδικασίες ελέγχου ταυτοποίησης ή επαλήθευσης. Η μάσκα bit βοηθά στην προσθήκη ενός επιπλέον επιπέδου ασφάλειας και απρόβλεπτης ικανότητας στις αποκρίσεις ΦΜΚΣ.



Εικόνα 5 ΦΜΚΣ σφάλματος όπως προτάθηκε από τον Shimizu [5]

1.7.4 SRAM

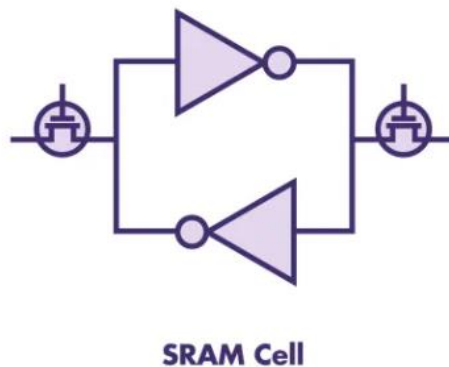
Στατική μνήμη τυχαίας προσπέλασης ή SRAM το κάθε κελί μνήμης κατασκευάζεται με έξι τρανζίστορ (MOSFET). Η λειτουργικότητα λογικής μνήμης ενός κελιού προέρχεται από δύο σταυρωτά συνδεδεμένους αντιστροφείς στον πυρήνα του, καθένας από τους οποίους κατασκευάζεται από δύο MOSFET, ένα p-MOS και ένα n-MOS. Από ηλεκτρονική άποψη, αυτό το κύκλωμα περιέχει έναν βρόχο θετικής ανάδρασης που ενισχύει την τρέχουσα κατάσταση του.

Η αρχή λειτουργίας ενός SRAM ΦΜΚΣ βασίζεται στη μεταβατική συμπεριφορά ενός κελιού μνήμης όταν ενεργοποιείται. Το προτιμώμενο αρχικό σημείο λειτουργίας ενός κελιού SRAM καθορίζεται από τη διαφορά "ισχύος" των MOSFET στο κύκλωμα του αντιστροφέα με σταυρωτή σύζευξη.

Για λόγους αποδοτικότητας και απόδοσης, τα τυπικά κελιά SRAM σχεδιάζονται ώστε να έχουν τέλεια προσαρμοσμένους αντιστροφείς. Η πραγματική διαφορά ισχύος μεταξύ των δύο αντιστροφέων,

Κυψέλες με μικρή αναντιστοιχία εξακολουθούν να έχουν μια προτιμώμενη αρχική σταθερή κατάσταση η οποία καθορίζεται από το πρόσημο της αναντιστοιχίας, αλλά λόγω του θορύβου τάσης υπάρχει μια μη αμελητέα πιθανότητα να ενεργοποιηθούν στη μη προτιμώμενη κατάστασή τους. Η τελική τους σταθερή κατάσταση θα είναι σε μεγάλο βαθμό τυχαία για κάθε ενεργοποίηση.

Το μέγεθος της επίδρασης των μεταβολών της διεργασίας στην τυχαία αναντιστοιχία των συσκευών στις κυψέλες SRAM έχει ως αποτέλεσμα οι περισσότερες κυψέλες να έχουν μια προτιμώμενη αλλά ξεχωριστή για το κάθε κελί αρχική κατάσταση, και μόνο λίγα κελιά έχουν ασθενή προτίμηση ή καθόλου προτίμηση. Αυτό σημαίνει ότι η κατάσταση ενεργοποίησης ενός τυπικού κελιού SRAM παρουσιάζει ισχυρή συμπεριφορά PUF. Μεγάλες συστοιχίες κελιών SRAM είναι σε θέση να παρέχουν χιλιάδες έως εκατομμύρια bits απόκρισης για αυτό το SRAM PUF. Η διεύθυνση ενός συγκεκριμένου κελιού σε μια τέτοια συστοιχία μπορεί να θεωρηθεί η πρόκληση του SRAM PUF.



Εικόνα 6 Ένα κελί SRAM [6]

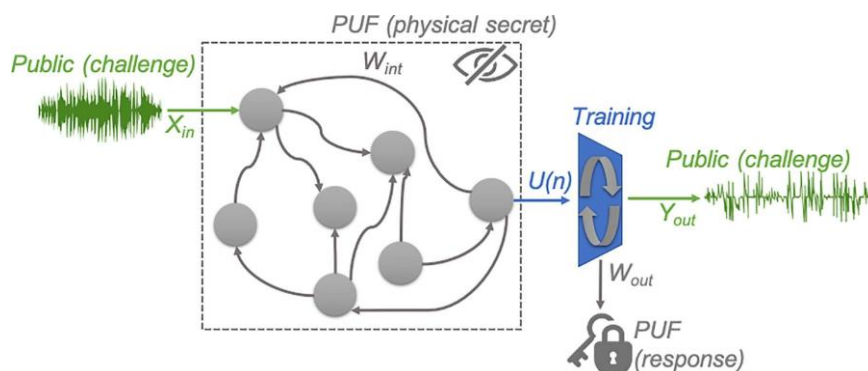
1.7.5 Νευρομορφική ΦΜΚΣ

Η νευρομορφική ΦΜΚΣ μπορεί να πραγματοποιηθεί χρησιμοποιώντας οποιοδήποτε τύπο νευρωνικού δικτύου, αν και τα RNN και τα RC (το πως λειτουργεί το RNN και το RC περιγράφεται στο μεθεπόμενο κεφάλαιο) είναι πιο κατάλληλα για την εργασία, λόγω της σύνθετης δυναμικής σχέσης μεταξύ της εξόδου τους και των κρυφών επιπέδων. Ακολουθώντας την αρχιτεκτονική RC, ένα γνωστό ζεύγος χρονικών σειρών $\{X_{in}, Y_{out}\}$ που συνδέονται μαθηματικά το X_{in} είναι η είσοδος που αντλείται από μια τυχαία κατανομή και το Y_{out} είναι η έξοδος.

Χρησιμοποιώντας τη συμβατική εκπαίδευση νευρωνικού δικτύου τα βάρη εξόδου του RC (W_{out}) προσαρμόζονται έτσι ώστε να μιμηθούν την ακολουθία της χρονοσειράς εξόδου

$$Y_{out}(i) = W_{out} U(i).$$

Επομένως, είναι εύκολο να συμπεράνουμε ότι το W_{out} ρυθμίζεται επίσης από τις ακριβείς τιμές του W_{int} . Ακολουθώντας αυτές τις παρατηρήσεις, εάν το W_{int} διαφέρει σε σύγκριση με το W_{ideal} , η διαδικασία εκπαίδευσης θα επιτρέψει ακόμα στο RC να συγκλίνει στη σωστή ακολουθία Y_{out} , αλλά σε κάθε περίπτωση θα χρησιμοποιήσει ένα διαφορετικό σύνολο W'_{out} για να το πετύχει, καθώς το RC είναι ένα σύνθετο δυναμικό σύστημα του οποίου καταστάσεις εξαρτώνται από την εσωτερική του συνδεσιμότητα. Με βάση αυτό, το W'_{out} μπορεί να θεωρηθεί ως η απόκριση της ΦΜΚΣ για μια γνωστή είσοδο. Το μη αντιγράψιμο χαρακτηριστικό της ΦΜΚΣ είναι τα σύνθετα βάρη του RC (W_{int}), ενώ η απόκριση της ΦΜΚΣ είναι τα ψηφιακά βάρη εξόδου στη γραμμική παλινδρόμηση (W_{out}) που συνδέονται άμεσα και με τα (Y_{out} και W_{int}). Επομένως, ανάλογα με τον αριθμό των ζευγών πρόκλησης που χρησιμοποιούνται για την εκπαίδευση RC, η προτεινόμενη έννοια μπορεί να χρησιμοποιηθεί είτε ως ασθενές είτε ως ισχυρό PUF. [7]



Εικόνα 7 Σχηματική περίληψη του Neuromorphic PUF. Τα X_{in} και Y_{out} είναι οι παράμετροι μίας χρονοσειράς και είναι δημόσια, τα W_{int} είναι σύνθετα βάρη που συνδέονται με το υλικό και τα W_{out} θεωρούνται ιδιωτικά. Τα ψηφιακά βάρη γραμμικής παλινδρόμησης (linear regression) που προκύπτουν μέσω της εκπαίδευσης θεωρούνται η απόκριση της ΦΜΚΣ. [7]

1.8 Σύνοψη

Συνοψίζοντας για το δεύτερο κεφάλαιο:

Η συντόμευση "PUF" αναφέρεται στο "Physical Unclonable Function". Οι ΦΜΚΣ είναι συστήματα ηλεκτρονικής ασφάλειας που εκμεταλλεύονται μικρές διαφορές στη φυσική δομή ή στη συμπεριφορά των ηλεκτρονικών στοιχείων για τη δημιουργία μοναδικών και δύσκολα αντιγράψιμων ψηφιακών υπογραφών. Οι ΦΜΚΣ χρησιμοποιούνται ευρέως σε εφαρμογές ασφαλείας, όπως στην ταυτοποίηση συσκευών, την προστασία από παραποίηση και την κρυπτογραφία. Ένα από τα κύρια πλεονεκτήματα των ΦΜΚΣ είναι η μοναδικότητα και η απεικόνιση της φυσικής δομής, κάνοντάς τα δύσκολα να αναπαραχθούν από κακόβουλους επιτιθέμενους.

2

ΜΗΤΡΙΚΕΣ ΚΑΙ

ΕΡΓΑΛΕΙΑ ΤΩΝ ΜΗ

ΚΛΩΝΟΠΟΙΗΣΙΜΩΝ

ΣΥΝΑΡΤΗΣΕΩΝ (Physical

Unclonable Functions - PUF)

2.1 Τι εννοούμε με τον όρο μη κλωνοποιήσιμα

Δεν είναι σαφής ο ορισμός του μη κλωνοποιήσιμου γιατί πρέπει να οριστεί, τι ορίζουμε ως κλώνο. Σε πολλά σενάρια, παρατηρεί κανείς μόνο τη συμπεριφορά στην είσοδο και την έξοδο μίας ΦΜΚΣ και σπάνια τα πραγματικά φυσικά στοιχεία που αποτελείται. Σε αυτήν την έννοια, οποιαδήποτε κατασκευή που μιμείται με ακρίβεια τη συμπεριφορά μίας συγκεκριμένης ΦΜΚΣ θα μπορούσε να θεωρηθεί ως αντίγραφο. Από την άλλη πλευρά, αν λάβουμε υπόψη τα φυσικά στοιχεία, το αντίγραφο αποκτά διαφορετική σημασία. Για παράδειγμα, θα μπορούσαμε να απαιτήσουμε από ένα αντίγραφο να έχει την ίδια συμπεριφορά και την ίδια φυσική εμφάνιση με το πρωτότυπο, ή ακόμα πιο αυστηρά, να απαιτήσουμε από ένα αντίγραφο να παραχθεί με την ίδια διαδικασία παραγωγής όπως το πρωτότυπο. Ο διαχωρισμός μεταξύ της πρώτης και της δεύτερης ερμηνείας του αντιγράφου γίνεται συχνά σαφής ως μαθηματικού έναντι φυσικού αντιγράφου. [7]

2.2 Πειραματική Διαδικασία

Ένα πείραμα (N_{pruf} , N_{chal} , N_{meas}) σε μια κατηγορία ΦΜΚΣ P αποτελεί έναν πίνακα τιμών απόκρισης της ΦΜΚΣ που έχουν προκύψει από παρατηρήσεις. Ένα πείραμα (N_{pruf} , N_{chal} , N_{meas}) περιέχει $N_{\text{pruf}} \times N_{\text{chal}} \times N_{\text{meas}}$ τιμές. (Όπου N_{pruf} ο αριθμός των ΦΜΚΣ περιπτώσεων N_{chal} ο αριθμός των προκλήσεων και N_{meas} οι διακριτές αξιολογήσεις απόκρισης για κάθε πρόκληση σε κάθε περίπτωση ΦΜΚΣ.)

Τα αποτελέσματα του πειράματος είναι οι αποκρίσεις μιας ΦΜΚΣ που θεωρούνται συνήθως μια μεταβλητή τυχαία.

Οι τιμές απόκρισης της ΦΜΚΣ που παρατηρούνται κατά τη διάρκεια ενός πειράματος ταξινομούνται με διάφορους τρόπους. Τρία σημαντικά "πεδία" στη σειρά των παρατηρημένων αποκρίσεων της ΦΜΚΣ είναι προφανή: (i) αποκρίσεις από διαφορετικές περιπτώσεις της ΦΜΚΣ, (ii) αποκρίσεις από την ίδια περίπτωση της ΦΜΚΣ αλλά σε διαφορετικές προκλήσεις και (iii) αποκρίσεις από την ίδια περίπτωση της ΦΜΚΣ στην ίδια πρόκληση αλλά από διακριτές αξιολογήσεις.

Για να αξιολογηθεί η εφαρμοσιμότητα μιας κατηγορίας ΦΜΚΣ, απαιτείται πληροφορία σχετικά με την κατανομή των αποκρίσεων της ΦΜΚΣ. N_{puf} (τυχαίες) περιπτώσεις ΦΜΚΣ από την εξεταζόμενη κατηγορία ΦΜΚΣ P , προκαλούμενες από το ίδιο σύνολο N_{chal} (τυχαίων) προκλήσεων, με N_{meas} διακριτές αξιολογήσεις απόκρισης για κάθε πρόκληση σε κάθε περίπτωση ΦΜΚΣ.

2.3 Αποκρίσεις ΦΜΚΣ

Υπάρχουν δύο είδη αποκρίσεων που εξετάζουμε στις ΦΜΚΣ η intra-distance και η inter-distance

2.3.1 Intra-distance

Το "intra-distance" στις ΦΜΚΣ αναφέρεται στην απόσταση μεταξύ των αποκρίσεων που παράγονται από την ίδια ΦΜΚΣ στιγμιότυπο υπό διαφορετικές προκλήσεις. [8]

2.3.1.1 Intra-distance Statistics

Για τον σχέδιο δημιουργίας όλων των εφαρμογών που βασίζονται σε ΦΜΚΣ, η γνώση σχετικά με τη κατανομή του D_P^{intra} έχει μεγάλη σημασία για τη ανάλυση του κατά πόσο μπορούμε να αναπαράγουμε μια κατηγορία ΦΜΚΣ. Στατιστικά αυτής της κατανομής παρουσιάζονται συνήθως ως ένα από τα πιο σημαντικά στοιχεία μέτρησης ποιότητας για μια κατηγορία ΦΜΚΣ.

Με βάση ένα πείραμα $ExperimentP(N_{puf}, N_{chal}, N_{meas})$, μπορεί να υπολογιστεί ένας νέος πίνακας $D_{Exp}(P)_{intra}$

$$D_{Exp}(P)_{intra} = \left[\text{dist} \left[y_i^{(j_1)}(x_k); y_i^{(j_2)}(x_k) \right] \right]_{\forall 1 \leq i \leq N_{puf}; \forall 1 \leq k \leq N_{chal}; \forall 1 \leq j_1 \neq j_2 \leq N_{meas}}.$$

Όπου dist είναι η απόσταση μεταξύ των μετρήσεων πάνω στο σύνολο αποκρίσεων Y . Συνήθως οι αποκρίσεις θεωρούνται ως δυαδικοί διανυσματικοί πίνακες και η μετρική απόστασης που χρησιμοποιείται είναι η απόσταση Hamming ή η κλασματική απόσταση Hamming.

Η απόκριση μιας ΦΜΚΣ θεωρείται μια τυχαία μεταβλητή. Για να αξιολογηθεί η χρησιμότητα μιας κλάσης ΦΜΚΣ, απαιτείται πληροφορία σχετικά με τις αποκρίσεις της ΦΜΚΣ.

Μια εκτίμηση του μέσου της κατανομής του D_P^{intra} ή $E[D_P^{intra}]$ προκύπτει από το μέσο του δειγματικού συνόλου του $D_{Exp}(P)_{intra}$:

$$\mu_P^{intra} = \overline{D_{Exp}(P)_{intra}} = \frac{2}{N_{puf} \cdot N_{chal} \cdot N_{meas} \cdot (N_{meas} - 1)} \cdot \sum D_{Exp}(P)_{intra}$$

Μία εκτίμηση της σταθερής απόκλισης της κατανομής της $D_{\mathcal{P}}^{\text{intra}}$ ή $\Sigma[D_{\mathcal{P}}^{\text{intra}}]$ βρίσκεται από την σχέση:

$$\sigma_{\mathcal{P}}^{\text{intra}} = \sqrt{\frac{2}{N_{\text{puf}} \cdot N_{\text{chal}} \cdot N_{\text{meas}} \cdot (N_{\text{meas}} - 1) - 2} \cdot \sum (\mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{intra}} - \mu_{\mathcal{P}}^{\text{intra}})^2}$$

Όπως αναφέρθηκε και πιο πριν οι ΦΜΚΣ επηρεάζονται από εξωτερικούς παράγοντες οπότε έχει σαν αποτέλεσμα να επηρεάζεται intra απόσταση των αποκρίσεων της ΦΜΚΣ. Η απόσταση μεταξύ δύο αποκρίσεων της ΦΜΚΣ που αξιολογούνται για την ίδια ΦΜΚΣ στιγμιότυπο και για την ίδια πρόκληση, αλλά κάτω από διαφορετικές συνθήκες $\alpha 1$ και $\alpha 2$, είναι συνήθως μεγαλύτερη από την εσωτερική απόσταση μεταξύ των ίδιων αποκρίσεων που αξιολογούνται κάτω από μία σταθερή συνθήκη. Για εφαρμογές που βασίζονται σε ΦΜΚΣ, είναι σημαντικό να γνωρίζουμε τη χειρότερη περίπτωση, δηλαδή τη μεγαλύτερη εσωτερική απόσταση που μπορεί να εμφανιστεί για ένα δεδομένο εύρος συνθηκών αξιολόγησης με βάση μια συγκεκριμένη αναφορική συνθήκη α_{ref} .

Για να βρούμε την χειρότερη εσωτερική απόσταση σε ένα δεδομένο εύρος συνθηκών, συνήθως αξιολογούμε την εσωτερική απόσταση κάτω από τις ακραίες τιμές αυτού του εύρους. Αυτό έχει νόημα υπό τη λογική υπόθεση ότι η εσωτερική απόσταση ως προς την αναφορική συνθήκη αυξάνεται όταν μετακινούμαστε μακριά από το σημείο αναφοράς. Τα άκρα ενός εύρους συνθηκών ονομάζονται "περιπτώσεις γωνιών" (corner cases). Για παράδειγμα, ας υποθέσουμε ότι χρειάζεται να βρούμε τη συμπεριφορά της χειρότερης εσωτερικής απόστασης σε ένα εύρος θερμοκρασίας περιβάλλοντος $T_{env} = -40 \text{ }^{\circ}\text{C} \dots 85 \text{ }^{\circ}\text{C}$. Η αναφορική συνθήκη έχει οριστεί στη θερμοκρασία δωματίου, $\alpha_{ref} = (T_{env} = 25 \text{ }^{\circ}\text{C})$, και εξετάζονται οι εσωτερικές αποστάσεις στις περιπτώσεις γωνιών $\alpha 1 = (T_{env} = -40 \text{ }^{\circ}\text{C})$ και $\alpha 2 = (T_{env} = 85 \text{ }^{\circ}\text{C})$.

2.3.2 Inter-distance

Το "inter-distance" στις ΦΜΚΣ αναφέρεται στην απόσταση μεταξύ των αποκρίσεων που παράγονται από διαφορετικές ΦΜΚΣ στιγμιότυπα υπό τις ίδιες προκλήσεις. [9]

2.3.2.1 Inter-distance Statistics

Η κατανομή της τυχαίας μεταβλητής $\mathbf{D}_{\mathcal{P}}^{\text{inter}}$ είναι ένα σημαντικό μέτρο για μια κατηγορία ΦΜΚΣ, ώστε την περίπτωση να χαρακτηρίσει την μοναδικότητα των στιγμιότυπων της. Βασισμένοι στις παρατηρούμενες αποκρίσεις $\mathbf{Y}_{\text{Exp}(\mathcal{P})}$ ενός πειράματος ΦΜΚΣ $\text{ExperimentP}(N_{\text{puf}}, N_{\text{chal}}, N_{\text{meas}})$, μπορεί να υπολογιστεί ένας νέος πίνακας $\mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{inter}}$ από παρατηρούμενες inter αποστάσεις μεταξύ των αποκρίσεων.

$$\mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{inter}} = \left[\text{dist} \left[y_{i_1}^{(j)}(x_k); y_{i_2}^{(j)}(x_k) \right] \right]_{\forall 1 \leq i_1 \neq i_2 \leq N_{\text{put}}; \forall 1 \leq k \leq N_{\text{chal}}; \forall 1 \leq j \leq N_{\text{mess}}}$$

Από αυτόν τον δειγματικό πίνακα αποστάσεων μεταξύ των αποκρίσεων, μπορούν να υπολογιστούν οι παρακάτω στατιστικές τιμές για να παρέχουν μια εκτίμηση των παραμέτρων της κατανομής $\mathbf{D}_{\mathcal{P}}^{\text{inter}}$:

Ο δειγματικός μέσος της $\mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{inter}}$ είναι μία εκτίμηση της $E[D_{\mathcal{P}}^{\text{inter}}]$:

$$\mu_{\mathcal{P}}^{\text{inter}} = \overline{\mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{inter}}} = \frac{2}{N_{\text{put}} \cdot (N_{\text{put}} - 1) \cdot N_{\text{chal}} \cdot N_{\text{meas}}} \cdot \sum \mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{inter}}.$$

Μία εκτίμηση της σταθερής απόκλισης της κατανομής της $\Sigma[D_{\mathcal{P}}^{\text{inter}}]$ αποκτιέται από:

$$\sigma_{\mathcal{P}}^{\text{inter}} = \sqrt{\frac{2}{N_{\text{put}} \cdot (N_{\text{put}} - 1) \cdot N_{\text{chal}} \cdot N_{\text{meas}} - 2} \cdot \sum (\mathbf{D}_{\text{Exp}(\mathcal{P})}^{\text{inter}} - \mu_{\mathcal{P}}^{\text{inter}})^2}$$

Η απόσταση μεταξύ των inter αποκρίσεων των ΦΜΚΣ είναι επίσης ευάλωτη σε διάφορες συνθήκες αξιολόγησης. Για την δημιουργία τυπικών εφαρμογών που βασίζονται σε ΦΜΚΣ γίνονται σε ένα ασφαλές περιβάλλον υπό ελεγχόμενες συνθήκες και η αξιολόγηση της inter απόστασης πραγματοποιείται σε συνθήκες αναφοράς α_{ref} που είναι οι πιο σημαντικές. Επομένως, σε αυτήν την περίπτωση δεν απαιτείται καμία επέκταση του ορισμού της απόστασης μεταξύ των αποκρίσεων σε μεταβαλλόμενες συνθήκες. Σε εφαρμογές όπου ο αρχική αξιολόγηση και οι μετρήσεις που παίρνονται γίνονται στο πεδίο, πρέπει να λυφθούν υπόψιν οι συνθήκες που γίνεται αξιολόγηση γιατί επηρεάζεται η ευαισθησία της inter απόστασης μεταξύ των αποκρίσεων.

2.4 NIST TEST

Η ανάγκη για τυχαίους και ψευδοτυχαίους αριθμούς προκύπτει σε πολλές κρυπτογραφικές εφαρμογές. Για παράδειγμα, συνηθισμένα κρυπτοσυστήματα χρησιμοποιούν κλειδιά που πρέπει να δημιουργούνται με τυχαίο τρόπο. Πολλά κρυπτογραφικά πρωτόκολλα απαιτούν επίσης τυχαίες ή ψευδοτυχαίες εισόδους σε διάφορα σημεία, για παράδειγμα, για βοηθητικές ποσότητες που χρησιμοποιούνται στη δημιουργία ψηφιακών υπογραφών ή για τη δημιουργία προκλήσεων σε πρωτόκολλα αυθεντικοποίησης. [10]

Το Εθνικό Ινστιτούτο Προτύπων και Τεχνολογίας (NIST) πιστεύει ότι αυτές οι διαδικασίες είναι χρήσιμες για τον εντοπισμό αποκλίσεων μιας δυαδικής ακολουθίας από την τυχαιότητα. Ωστόσο, ένας tester πρέπει να σημειώσει ότι οι εμφανείς αποκλίσεις από την τυχαιότητα μπορεί να

οφείλονται είτε σε μία κακά σχεδιασμένη γεννήτρια είτε σε ανωμαλίες που εμφανίζονται στη δυναδική ακολουθία που ελέγχεται. [10]

2.4.1 Τύποι γεννητριών τυχαίων αριθμών

Υπάρχουν δύο βασικοί τύποι γεννητριών που χρησιμοποιούνται για την παραγωγή τυχαίων ακολουθιών: οι γεννήτριες τυχαίων αριθμών και οι γεννήτριες ψευδοτυχαίων αριθμών. Για κρυπτογραφικές εφαρμογές, και οι δύο αυτοί τύποι γεννητριών παράγουν μία σειρά από μηδενικά και άσσους που μπορεί να χωριστεί σε υποσειρές ή μπλοκ τυχαίων αριθμών. Οι τυχαίοι και ψευδοτυχαίοι αριθμοί που παράγονται για κρυπτογραφικές εφαρμογές πρέπει να είναι μη προβλέψιμοι. Στην περίπτωση των ψευδοτυχαίων αριθμογεννητών (PRNGs), το seed πρέπει να είναι άγνωστο, ο επόμενος αριθμός εξόδου στην ακολουθία πρέπει να είναι μη προβλέψιμος, παρά τυχόν γνώσης προηγούμενων τυχαίων αριθμών στην ακολουθία. Αυτή η ιδιότητα είναι γνωστή ως προς τα εμπρός απροβλεψιμότητα.

Επίσης, δεν πρέπει να είναι δυνατό να προσδιοριστεί το seed από τη γνώση οποιωνδήποτε παραγόμενων τιμών (δηλαδή, απαιτείται επίσης οπισθοδρομική απροβλεψιμότητα). Δεν πρέπει να εμφανίζεται καμία συσχέτιση μεταξύ ενός seed και οποιασδήποτε τιμής που παράγεται από αυτό το seed· κάθε στοιχείο της ακολουθίας θα πρέπει να φαίνεται ότι είναι το αποτέλεσμα ενός ανεξάρτητου τυχαίου γεγονότος με πιθανότητα 1/2. [10]

2.4.1.1 Γεννήτρια τυχαίων αριθμών

Ο πρώτος τύπος γεννήτριας ακολουθιών είναι μία γεννήτρια τυχαίων αριθμών (RNG). Μία RNG χρησιμοποιεί μια μη-προσδιορίσιμη πηγή (δηλαδή η πηγή εντροπίας), μαζί με κάποια συνάρτηση επεξεργασίας (δηλαδή, διαδικασία εκχύλισης εντροπίας) για να παράγει τυχαιότητα. Η χρήση μιας διαδικασίας εκχύλισης είναι απαραίτητη για να ξεπεραστούν οι αδυναμίες στην πηγή εντροπίας που οδηγούν στην παραγωγή μη τυχαίων αριθμών (π.χ., η εμφάνιση μεγάλων ακολουθιών μηδενικών ή άσσων).

Τα αποτελέσματα μίας RNG μπορούν να χρησιμοποιηθούν απευθείας ως τυχαίοι αριθμοί ή μπορεί να τροφοδοτηθούν σε μια γεννήτρια ψευδοτυχαίων αριθμών (PRNG).

Για παράδειγμα, μια φυσική πηγή όπως ο ηλεκτρονικός θόρυβος μπορεί να περιέχει μια υπέρθεση κανονικών δομών, όπως κύματα ή άλλα περιοδικά φαινόμενα, τα οποία μπορεί να φαίνονται τυχαία, αλλά κρίνονται ως μη τυχαία με στατιστική ανάλυση.

Για κρυπτογραφικούς σκοπούς, το αποτέλεσμα των RNG πρέπει να είναι απρόβλεπτο. Ωστόσο, κάποιες φυσικές πηγές (π.χ., διανύσματα ημερομηνίας/ώρας) είναι αρκετά προβλέψιμες. Ωστόσο, τα αποτελέσματα που προκύπτουν από τον RNG μπορεί ακόμα να είναι ανεπαρκή όταν αξιολογούνται με στατιστικά τεστ. [10]

2.4.1.2 Γεννήτρια ψευδοτυχαίων αριθμών

Ο δεύτερος τύπος γεννήτριας είναι μια γεννήτρια ψευδοτυχαίων αριθμών (PRNG). Μια PRNG χρησιμοποιεί μία ή περισσότερους εισόδους και παράγει πολλούς "ψευδοτυχαίους" αριθμούς. Οι

είσοδοι στις PRNG ονομάζονται seeds. Σε περιβάλλοντα όπου απαιτείται απροβλεψιμότητα, το ίδιο seed πρέπει να είναι τυχαίο και απρόβλεπτο. Επομένως, από προεπιλογή, μια PRNG θα πρέπει να λαμβάνει τα seeds της από τις εξόδους ενός RNG. Δηλαδή, μια PRNG απαιτεί ένα RNG ως σύντροφο.

Τα αποτελέσματα μίας PRNG είναι τυπικά καθορισμένες συναρτήσεις του seed. Δηλαδή, όλη η πραγματική τυχειότητα περιορίζεται στη δημιουργία του seed. Η καθορισμένη φύση της διαδικασίας οδηγεί στον όρο "ψευδοτυχαίο". Δεδομένου ότι κάθε στοιχείο μιας ψευδοτυχαίας ακολουθίας μπορεί να αναπαραχθεί από το seed της, μόνο ο σπόρος χρειάζεται να αποθηκευτεί εάν απαιτείται η αναπαραγωγή ή η επαλήθευση της ψευδοτυχαίας ακολουθίας.

2.4.2 Έλεγχος (testing)

Διάφορα στατιστικά τεστ μπορούν να εφαρμοστούν σε μια ακολουθία στην προσπάθεια να συγκριθεί και να αξιολογηθεί μία ακολουθία από μία γεννήτρια σε σχέση με μια πραγματικά τυχαία ακολουθία. Η τυχειότητα είναι μια πιθανοτική ιδιότητα, δηλαδή οι ιδιότητες μιας τυχαίας ακολουθίας μπορούν να χαρακτηριστούν και να περιγραφούν με πιθανοτικούς όρους. Το πιθανό αποτέλεσμα των στατιστικών τεστ, όταν εφαρμόζονται σε μια πραγματικά τυχαία ακολουθία, είναι γνωστό εκ των προτέρων και μπορεί να περιγραφεί με πιθανοτικούς όρους. Υπάρχει ένας απεριόριστος αριθμός δυνατών στατιστικών τεστ, καθένα από τα οποία αξιολογεί την παρουσία ή την απουσία ενός "προτύπου" το οποίο, αν ανιχνευθεί, θα υποδήλωνε ότι η ακολουθία είναι μη τυχαία. Επειδή υπάρχουν τόσα πολλά τεστ για να κριθεί εάν μια ακολουθία είναι τυχαία ή όχι, κανένα συγκεκριμένο πεπερασμένο σύνολο τεστ δεν θεωρείται "πλήρες". Επιπλέον, τα αποτελέσματα των στατιστικών ελέγχων πρέπει να ερμηνευτούν με προσοχή προκειμένου να αποφευχθούν εσφαλμένα συμπεράσματα σχετικά με μία συγκεκριμένη γεννήτρια.

2.4.3 Χρήση σε ΦΜΚΣ

Η χρήση του NIST test σε μία ΦΜΚΣ θα μπορούσε να περιλαμβάνει τη δοκιμή των ψηφιακών υπογραφών που παράγονται από τη ΦΜΚΣ ως προς την τυχειότητα, την ομοιότητα, την ανθεκτικότητα σε διάφορες επιθέσεις και άλλες παραμέτρους ασφάλειας. Το NIST test θα μπορούσε να χρησιμοποιηθεί για να επιβεβαιώσει την αξιοπιστία της ΦΜΚΣ ως προς την παραγωγή μοναδικών και δύσκολο στην αναπαράσταση και αναπαραγωγή ψηφιακών υπογραφών.

Επιπλέον, η χρήση του NIST test θα μπορούσε να βοηθήσει στην αξιολόγηση της ανθεκτικότητας της ΦΜΚΣ σε διάφορες επιθέσεις, όπως η μοντελοποίηση, η αναπαραγωγή ή η επίθεση με εκφορτίσεις ηλεκτροστατικής απεικόνισης. Αυτή η εκτεταμένη αξιολόγηση μπορεί να βοηθήσει τους κατασκευαστές ΦΜΚΣ να διασφαλίσουν ότι οι σχεδιασμοί τους πληρούν τις απαιτήσεις ασφάλειας και ανθεκτικότητας που επιθυμούν.

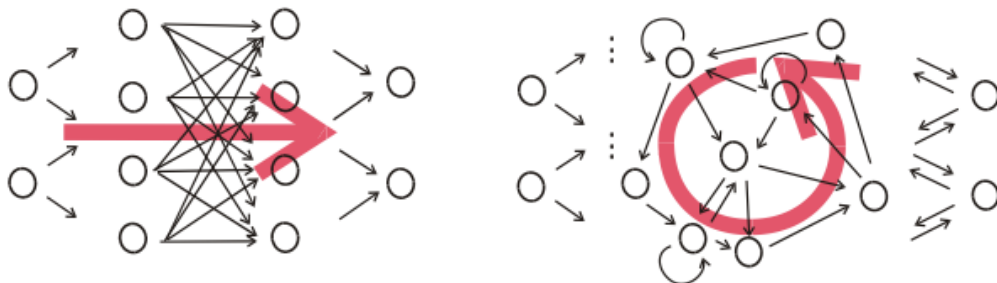
3

ΑΝΑΔΡΟΜΙΚΑ

ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

3.1 ΑΝΑΔΡΟΜΙΚΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ (RECURSIVE NEURAL NETWORKS)

Υπάρχουν δύο κύριοι τύποι νευρωνικών δικτύων, τα feedforward και τα recurrent. Στα feedforward δίκτυα, η ενεργοποίηση "διοχετεύεται" μέσω του δικτύου από τις μονάδες εισόδου στις μονάδες εξόδου. Αντίθετα, ένα αναδρομικό νευρωνικό δίκτυο (ΑΝΔ) έχει (τουλάχιστον ένα) κυκλικό μονοπάτι συναπτικών συνδέσεων δηλαδή η έξοδος ενός κόμβου του δικτύου τροφοδοτεί την είσοδο ενός άλλου κόμβου που μπορεί να τροφοδοτεί την είσοδο του προηγούμενου ή την είσοδο του εαυτού του.



Εικόνα 8 Τυπική δομή ενός feedforward δικτύου τροφοδότησης (αριστερά) και ενός αναδρομικού δικτύου δεξιά [11]

Ένα ανακυκλούμενο νευρωνικό δίκτυο (RNN) έχει τουλάχιστον μία κυκλική διαδρομή του, συναπτικές συνδέσεις.

Βασικά χαρακτηριστικά:

Από την φύση όλα τα βιολογικά νευρωνικά δίκτυα είναι επαναλαμβανόμενα.

Από μαθηματικής άποψης τα αναδρομικά νευρωνικά (ΑΝΔ) υλοποιούν δυναμικά συστήματα

Ένας θεμελιώδης νόμος στην θεωρία των Νευρωνικών Δικτύων είναι ότι όταν τα Αναδρομικά Νευρωνικά Δίκτυα (RNNs) σχεδιάζονται με τη σωστή αρχιτεκτονική και διαδικασία εκπαίδευσης, διαθέτουν μια καθολική ιδιότητα προσέγγισης, που σημαίνει ότι μπορούν θεωρητικά να προσεγγίσουν αυθαίρετα δυναμικά συστήματα με αυθαίρετη ακρίβεια, υπό ορισμένες

προϋποθέσεις και συνθήκες. Η ιδιότητα αυτή αποτελεί θεμελιώδες αποτέλεσμα στη θεωρία των νευρωνικών δικτύων.

Υπάρχουν διάφοροι τύποι αλγορίθμων εκπαίδευσης, χωρίς να υπερτερεί κάποιος

Τα ΑΝΔ προορίζονται ως μοντέλα "μαύρου κουτιού" για την επεξεργασία σήματος. [11]

Χρησιμοποιούνται πολύ για την επεξεργασία φυσικής γλώσσας. Έχοντας ένα ΑΝΔ τα βάρη εφαρμόζονται αναδρομικά σε μια δομημένη είσοδο για να λάβουμε μια δομημένη πρόβλεψη [12]

3.2 Εκπαίδευση ΑΝΔ

Τα ΑΝΔ εκπαιδεύονται με μεθόδους επίβλεψης (supervised training)

Κατά την εκπαίδευση των ΑΝΔ ξεκινάμε με δεδομένα εκπαίδευσης (teachers data):

εμπειρικά παρατηρημένα από την έξοδο του μοντέλου μας ή τεχνητά κατασκευασμένες χρονοσειρές εισροών-εκροών, οι οποίες αντιπροσωπεύουν παραδείγματα της επιθυμητής συμπεριφοράς του μοντέλου που καλείται ο ΑΝΔ να αναπαραστήσει.

Τα δεδομένα εκπαίδευσης (teacher data) χρησιμοποιούνται για την εκπαίδευση ενός ΑΝΔ έτσι ώστε να ταιριάζει με την επιθυμητή έξοδο του συστήματος ή ελπίζοντας ότι το ΑΝΔ στη συνέχεια θα γενικεύει σε νέες εισόδους. Δηλαδή, όταν το εκπαιδευμένο ΑΝΔ λαμβάνει μια ακολουθία εισόδου η οποία είναι κατά κάποιο τρόπο παρόμοια με την ακολουθία εισόδου εκπαίδευσης, θα πρέπει να παράγει μια έξοδο που μοιάζει με την έξοδο του αρχικού συστήματος που έχει κληθεί το ΑΝΔ να αναπαραστήσει.

Ένα μεγάλο πρόβλημα στην εκπαίδευση με επίβλεψη είναι το φαινόμενο της υπερειδίκευσης (overfitting) το οποίο γίνεται όταν το εκπαιδευόμενο μοντέλο δίνει ακριβείς προβλέψεις για δεδομένα εκπαίδευσης αλλά όχι για νέα δεδομένα. Ωστόσο, αυτό τείνει να μην αποτελεί πρόβλημα για την εκπαίδευση των ΑΝΔ, επειδή οι γνωστές μέθοδοι εκπαίδευσης δυσκολεύονται να προσαρμόσουν καλά τα δεδομένα εκπαίδευσης. [11]

3.3 Περιγραφή ΑΝΔ

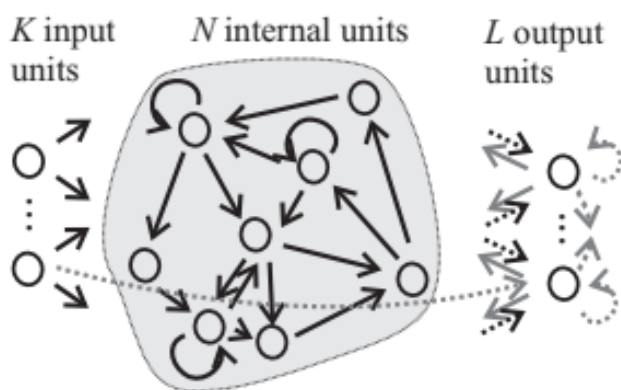
Ένα ΑΝΔ από νευρώνες που έχουν συναπτικούς συνδέσμους μεταξύ τους, των οποίων η συναπτική ισχύς κωδικοποιείται από ένα βάρος. Συνήθως διακρίνει κανείς νευρώνες εισόδου, εσωτερικούς (κρυφούς) εξόδου. Σε μια δεδομένη χρονική στιγμή, ένας νευρώνας έχει μια ενεργοποίηση. Συμβολίζουμε τις ενεργοποιήσεις των νευρώνων εισόδου με $u(n)$, των εσωτερικών νευρώνων με $x(n)$, των νευρώνων εξόδου με $y(n)$. Μερικές φορές αγνοούμε τη διάκριση εισόδου/εσωτερικής/εξόδου και τότε χρησιμοποιούμε το $x(n)$ με μετωνυμικό τρόπο.

Οι νευρώνες εισόδου για χρονικό βήμα n είναι $u(n) = (u_1(n), \dots, u_K(n))$, των εσωτερικών μονάδων είναι

$x(n) = (x_1(n), \dots, x_N(n))$, και των μονάδων εξόδου $y(n) = (y_1(n), \dots, y_L(n))$.

Οι συνδέσεις βαρών συλλέγονται σε έναν πίνακα βαρών $N \times K$, $W_{\square}^{in} = (w_{ij}^{in})$ για τα εισερχόμενα βάρη, σε έναν πίνακα $N \times N$, $W = (W_{ij}^{in})$ για τις εσωτερικές συνδέσεις, σε έναν πίνακα $L \times (K + N + L)$, $W_{\square}^{out} = (w_{ij}^{out})$ για τις συνδέσεις προς τις μονάδες εξόδου, και σε έναν πίνακα $N \times L$, $W_{\square}^{back} = (w_{ij}^{back})$ για τις συνδέσεις που προβάλλουν πίσω από την έξοδο προς τις εσωτερικές μονάδες.

Στα ΑΝΔ επιτρέπονται συνδέσεις απευθείας από τις εισόδους προς τις εξόδους και συνδέσεις μεταξύ των εξόδων. Δεν επιβάλλονται περαιτέρω περιορισμοί στην τοπολογία του δικτύου που προκαλείται από τα εσωτερικά βάρη. Συνήθως προτείνεται ότι οι εσωτερικές συνδέσεις W δημιουργούν αναδρομικούς δρόμους μεταξύ των εσωτερικών μονάδων. Στην επόμενη εικόνα παρουσιάζεται η βασική αρχιτεκτονική του δικτύου που εξετάζεται εδώ.



Εικόνα 9 Διακεκομμένα βέλη υποδεικνύουν συνδέσεις που είναι δυνατές αλλά δεν απαιτούνται.

Η ενεργοποίηση των εσωτερικών νευρώνων ενημερώνεται σύμφωνα με

$$x(n + 1) = f(W^{in}Wu(n + 1) + Wx(n) + W^{back}y(n))$$

Όπου το $u(n+1)$ αποτελεί την είσοδο που δίνεται, όπου $f = (f_1, \dots, f_N)$ είναι οι συναρτήσεις εξόδου της εισόδου των νευρώνων το f υποδηλώνει τη συνάρτηση εφαρμογής ενός νευρώνα συνάρτησης μεταφοράς, f (αλλιώς λέγεται ως συνάρτηση ενεργοποίησης, συνάρτηση εξόδου της μονάδας, ή συνάρτηση συμπίεσης). Χρησιμοποιούνται πολλοί τύποι συναρτήσεων όπως συνάρτηση sigmoid $f = \tanh$, ή όπως η $f=1$ που χρησιμοποιείται σε γραμμικά δίκτυα με $f = 1$.

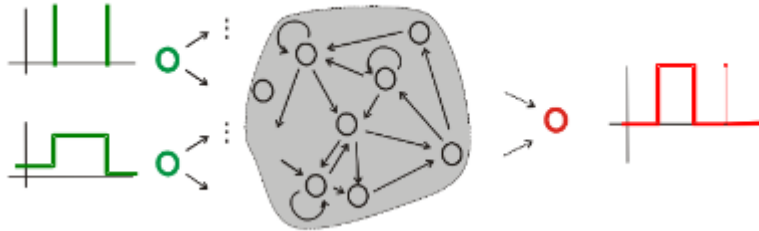
$$y(n + 1) = f^{out}(W^{out}(u(n + 1), x(n + 1), y(n))),$$

Μια μηδενική τιμή βάρους μπορεί να ερμηνευθεί ως "καμία σύνδεση". Σημειώστε ότι οι νευρώνες εξόδου μπορεί να έχουν συνδέσεις όχι μόνο από εσωτερικές νευρώνες αλλά και από νευρώνες εισόδου και πιο σπάνια μπορεί και μονάδες εξόδου.

Παράδειγμα: ένα μικρό δίκτυο χρονοδιακόπτη

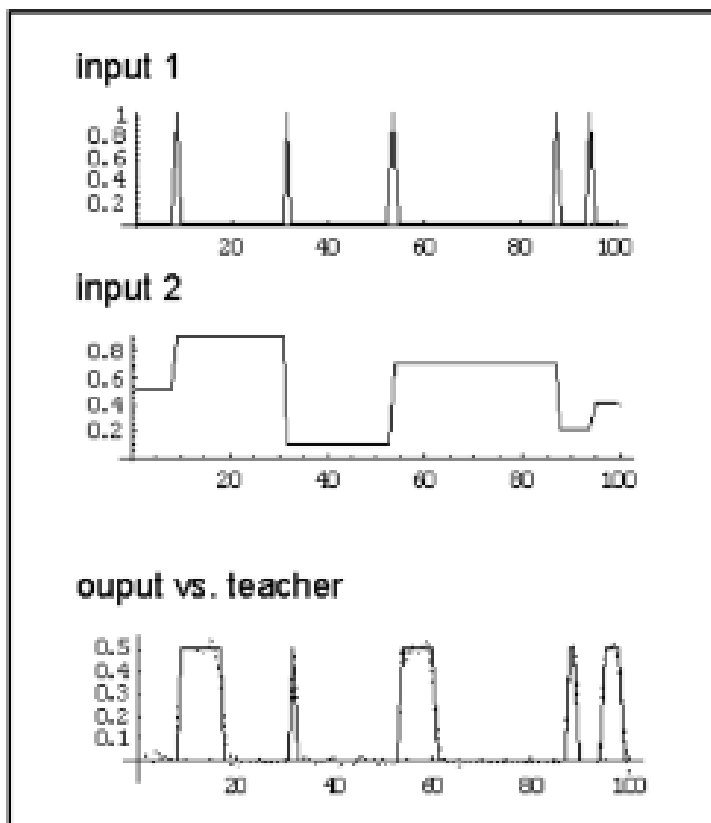
Ας εξετάσουμε ένα σύστημα εισόδου-εξόδου για χρονομέτρηση. Το σήμα εισόδου αποτελείται από δύο στοιχεία. Το πρώτο στοιχείο, $u1(n)$, είναι συνήθως 0, αλλά καμιά φορά αλλάζει σε 1. Το δεύτερο, $u2(n)$, παίρνει τιμές από 0.1 έως 1.0 με βήμα 0.1 και παίρνει μια νέα (τυχαία) τιμή από αυτές κάθε φορά που το $u1(n)$ γίνεται 1. Η επιθυμητή έξοδος είναι 0.5 για 10 επί $u2(n)$ χρονικά

βήματα μετά τη στιγμή που το $u_1(n)$ έγινε 1, αλλιώς είναι 0. Αυτό αντιστοιχεί στο να υλοποιεί έναν χρονοδιακόπτη, όπου το $u_1(n)$ είναι το "σήμα εκκίνησης" και το $u_2(n)$ είναι η επιθυμητή διάρκεια.



Εικόνα 10 Πάνω είσοδος αρχικό σήμα Κατω είσοδος διάρκεια
Έξωδος Ορθογώνιο σήμα επιθυμητής διάρκειας [13]

Το παρακάτω σχήμα δείχνει τα ίχνη της εισόδου και της εξόδου που δημιουργήθηκαν από ένα ΑΝΔ εκπαιδευμένο σε αυτήν την εργασία σύμφωνα με τη μέθοδο ESN:



Εικόνα 11 Επίδοση ενός ΑΝΔ εκπαιδευμένου στην διεργασία του χρονομέτρου. Στην τελευταία γραφική παράσταση, η συνεχής γραμμή αντιπροσωπεύει την επιθυμητή έξοδο (teacher data). Η γραμμή με τις τελείες : έξοδος του δικτύου.

Προφανώς, σε αυτή η εργασία απαιτείται από το ΑΝΔ να λειτουργεί ως μνήμη: πρέπει να διατηρεί πληροφορίες σχετικά με το "σήμα εκκίνησης" για αρκετά χρονικά βήματα. Αυτό είναι δυνατό επειδή οι εσωτερικές αναδρομικές συνδέσεις επιτρέπουν στο "σήμα εκκίνησης" να "ανακλάται" στις ενεργοποιήσεις των εσωτερικών μονάδων. Γενικά, εργασίες που απαιτούν κάποια μορφή μνήμης είναι υποψήφιες για μοντελοποίηση από ΑΝΔ.

3.4 ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΕΙΣΑΓΩΓΗ ΣΤΑ ESN

Η εκπαίδευση των Αναδρομικών Νευρωνικών Δικτύων ΑΝΔ είναι πολύ δύσκολη. Τα ΑΝΔ, αντιπροσωπεύουν ένα πολύ ισχυρό γενικό εργαλείο, συνδυάζοντας μεγάλη δυναμική μνήμη και υψηλά προσαρμόσιμες υπολογιστικές ικανότητες. Είναι το μοντέλο Μηχανικής Μάθησης (ML) που μοιάζει περισσότερο με τα βιολογικά εγκεφαλικά δίκτυα. [13]

3.4.1 ΕΠΙΣΤΡΟΦΗ ΣΦΑΛΜΑΤΟΣ (BACK PROPAGATION)

Η επιστροφή σφάλματος (BP) είναι έως και σήμερα μία από τις σημαντικότερες μεθόδους στην εκπαίδευση των τεχνητών νευρωνικών δικτύων. Έχει γίνει η τυπική μέθοδος εκπαίδευσης ειδικά για τα feed-forward νευρωνικά δίκτυα (FFNNs). Λαμβάνουμε υπόψη ένα πολυστρωματικό perceptron (MLP) με k κρυμμένα επίπεδα. Αυτά μαζί με το επίπεδο των εισόδων και το επίπεδο των εξόδων δίνει συνολικά $k+2$ επίπεδα. Τα οποία τα αριθμούμε από $0 \dots, k+1$. Ο αριθμός των μονάδων εισόδου είναι K , των μονάδων εξόδου είναι L , και οι μονάδες στο κρυφό επίπεδο m είναι N^m . Το βάρος της j -οστής μονάδας στο επίπεδο m και της i -οστής μονάδας στο επίπεδο $m+1$ εκφράζεται ως w_{ij}^m . Η ενεργοποίηση της i -οστής μονάδας στο επίπεδο m είναι x_i^m (για $m = 0$ αυτή είναι μια τιμή εισόδου, για $m = k+1$ μια τιμή εξόδου). [14]

Τα δεδομένα εκπαίδευσης για την εκπαίδευση του feedforward δικτύου αποτελούνται από T ζεύγη δεδομένων εισόδου-εξόδου.

$$u(n) = (x_1^0(n), \dots, x_K^0(n))', d(n)(d_1^{k+1}(n), \dots, d_L^{k+1}(n))'$$

όπου το n υποδηλώνει την εκπαιδευτική περίπτωση. Η ενεργοποίηση των μη μονάδων εισόδου υπολογίζεται σύμφωνα με την εξίσωση:

$$x_1^{m+1}(n) = f\left(\sum_{j=1, \dots, N^m} w_{1j}^m x_j(n)\right)$$

Παρουσιάζοντας στην είσοδο του teacher $u(t)$, η προηγούμενη ενημέρωση της εξίσωσης χρησιμοποιείται για τον υπολογισμό των ενεργοποιήσεων των μονάδων στα επόμενα κρυφά επίπεδα, μέχρι να έχουμε μια απάντηση από το δίκτυο.

$$y(n) = (x_1^{k+1}(n), \dots, x_L^{k+1}(n))'$$

Ο στόχος της εκπαίδευσης είναι να βρεθεί ένα σύνολο βαρών έτσι ώστε η συνολική τετραγωνική απόκλιση να ελαχιστοποιηθεί.

$$E = \sum_{n=1, \dots, T} \|d(n) - y(n)\|^2 = \sum_{n=1, \dots, T} E(n)$$

Βελτιστοποιούμε τα βάρη του δικτύου έτσι ώστε η συνολική τετραγωνική απόκλιση, δηλαδή η διαφορά μεταξύ της πραγματικής εξόδου του δικτύου και της επιθυμητής εξόδου, να ελαχιστοποιηθεί. Αυτό πραγματοποιείται αυξάνοντας βαθμιαία τα βάρη του δικτύου κατά μήκος της κατεύθυνσης της κλίσης του σφάλματος ως προς τα βάρη, χρησιμοποιώντας τη μέθοδο της κλίσης.

$$\frac{\partial E}{\partial w_{ij}^m} = \sum_{t=1, \dots, T} \frac{\partial E(n)}{\partial w_{ij}^m}$$

Και κάνοντας χρήση ενός μικρού ρυθμού μετάδοσης γ .

$$new\ w_{ij}^m = w_{ij}^m - \gamma \frac{\partial E}{\partial w_{ij}^m}$$

Αυτή είναι η εξίσωση που χρησιμοποιείται στη λεγόμενη μέθοδο μάθησης πακέτου (batch learning), όπου τα νέα βάρη υπολογίζονται μετά την παρουσίαση όλων των δειγμάτων εκπαίδευσης. Μια τέτοια διαδικασία που περιλαμβάνει όλα τα δείγματα ονομάζεται εποχή. Σε μια διαδικασία εποχής, το σύνολο των εκπαιδευτικών δεδομένων παρουσιάζεται στο νευρωνικό δίκτυο, και τα βάρη ενημερώνονται μετά την επεξεργασία όλων των δειγμάτων αυτού του συνόλου. Κάθε φορά που ολοκληρώνεται αυτή η διαδικασία, ολοκληρώνεται και μία "εποχή". Πριν από την πρώτη εποχή, τα βάρη αρχικοποιούνται, συνήθως με μικρούς τυχαίους αριθμούς, για να ξεκινήσει η διαδικασία εκπαίδευσης. Αντίθετα, στη σταδιακή μάθηση, τα βάρη ανανεώνονται μετά από κάθε εκπαιδευτικό δείγμα, κάνοντας τη διαδικασία εκπαίδευσης πιο ευέλικτη και ελαφριά σε υπολογιστικούς πόρους.

$$new\ w_{ij}^m = w_{ij}^m - \gamma \frac{\partial E(n)}{\partial w_{ij}^m}$$

Το κεντρικό θέμα σε αυτήν τη μέθοδο είναι ο υπολογισμός της απόκλισης του σφάλματος.

$$\frac{\partial E(n)}{\partial w_{ij}^m}$$

Ο αλγόριθμος αντίστροφης διάδοσης (backpropagation) είναι μία μέθοδος για την εκτέλεση αυτών των υπολογισμών. Εδώ είναι η διαδικασία για το πώς λειτουργεί σε μια εποχή :

Είσοδος: Τρέχον βάρη w_{ij}^m , εκπαιδευτικά δείγματα.
Έξοδος: Νέα βάρη.

Βήματα Υπολογισμού:

1. Για κάθε δείγμα n , υπολογίζονται οι ενεργοποιήσεις των εσωτερικών και εξωτερικών μονάδων .
2. Υπολογίζονται, προχωρώντας αντίστροφα μέσω $m = k+1, k, \dots, 1$, για κάθε μονάδα x_i^m , η διάδοσης σφάλματος $\delta_i^m(n)$.

$$\delta_i^{k+1}(n) = (d_i(n) - y_i(n)) \frac{\theta f(u)}{\theta u} \Big|_{u=z_i^{k+1}}$$

Για το επίπεδο εξόδου

$$\delta_j^m(n) = \sum_{i=1}^{N^{m+1}} \delta_i^{m+1} w_{ij}^m \frac{\theta f(u)}{\theta u} \Big|_{u=z_j^m}$$

Για τα κρυφά επίπεδα

$$z_i^m(n) = \sum_{j=1}^{N^{m-1}} x_j^{m-1}(n) w_{ij}^{m-1}$$

Μαθηματικά, ο όρος διάδοσης σφάλματος $\delta_j^m(n)$ αντιπροσωπεύει το κλίση του σφάλματος ως προς το δυναμικό της μονάδας x_i^m .

$$\frac{\theta E}{\theta u} \Big|_{u=z_j^m}$$

Η προσαρμογή των βαρών γίνεται με βάση

$$new w_{ij}^m = w_{ij}^m + \gamma \sum_{t=1}^T \delta_i^m(n) x_j^{m-1}(n)$$

Μετά από κάθε τέτοια εποχή, υπολογίζεται το σφάλμα. Σταματά όταν το σφάλμα πέσει κάτω από ένα προκαθορισμένο όριο, ή όταν η μεταβολή του σφάλματος πέσει κάτω από ένα άλλο προκαθορισμένο όριο, ή όταν ο αριθμός των εποχών υπερβαίνει ένα προκαθορισμένο μέγιστο αριθμό εποχών. Πολλές τέτοιες εποχές μπορεί να απαιτηθούν μέχρι να επιτευχθεί ένα επαρκώς μικρό σφάλμα.

Μια εποχή απαιτεί πολλαπλασιασμούς και προσθέσεις, με M το συνολικό πλήθος των συνδέσεων του δικτύου.

Αυτή η διαδικασία επαναλαμβάνεται συνήθως για πολλές εποχές μέχρι η απόδοση του δικτύου να συγκλίνει σε ένα ικανοποιητικό επίπεδο. Ο αλγόριθμος αντίστροφης διάδοσης είναι βασικός για

την εκπαίδευση νευρωνικών δικτύων, προσαρμόζοντας τα βάρη για να ελαχιστοποιήσει το σφάλμα μεταξύ των προβλεπόμενων και των πραγματικών εξόδων.

Η βασική προσέγγιση κατά της κλίσης (και η υλοποίησή της με τον αλγόριθμο backpropagation) είναι περίφημη για την αργή σύγκλιση, διότι ο ρυθμός μάθησης γ πρέπει τυπικά να είναι μικρός για να αποφευχθεί η αστάθεια. Υπάρχουν πολλές τεχνικές επιτάχυνσης, όπως δυναμικά σχήματα προσαρμογής της συνιστώσας μάθησης.

Όπως και με όλες τις τεχνικές κατά της κλίσης στις επιφάνειες σφάλματος, ο αλγόριθμος backpropagation βρίσκει μόνο έναν τοπικό ελάχιστο σφάλματος. Αυτό το πρόβλημα μπορεί να αντιμετωπιστεί με διάφορα μέτρα, όπως η προσθήκη θορύβου κατά τη διάρκεια της εκπαίδευσης (προσεγγιστικές μεθόδους ανάπτυξης) για να αποφευχθεί η παγίδευση σε φτωχά ελάχιστα, ή επανάληψη ολόκληρης της μάθησης από διάφορες αρχικές ρυθμίσεις βαρών, ή αρχικοποίηση των βαρών ανάλογα με προ υπάρχουσες πληροφορίες.

Ένα άλλο πρόβλημα είναι η επιλογή μιας κατάλληλης τοπολογίας δικτύου (αριθμός και μέγεθος των κρυμμένων επιπέδων). Και πάλι, μπορεί κανείς να χρησιμοποιήσει προηγούμενες πληροφορίες με αρχικές τιμές στα βάρη ή να πραγματοποιήσει συστηματική αναζήτηση

Συνολικά, ενώ το βασικό backpropagation είναι διαφανές και εύκολο στην χρήση, σημαντική είναι η εμπειρία και υπομονή για να επιτευχθούν καλά αποτελέσματα σε μη-απλές εργασίες

Οι μέθοδοι BP έχουν επίσης επεκταθεί στα ΑΝΔ αλλά με μερική επιτυχία. Ένας από τους περιορισμούς της μεθόδου BP για τα ΑΝΔ είναι ότι οι διακλαδώσεις μπορούν να καταστήσουν την εκπαίδευση μη συγκλίνουσα. Ακόμα και όταν συγκλίνουν, αυτή η σύγκλιση είναι αργή, υπολογιστικά ακριβή και μπορεί να οδηγήσει σε κακά τοπικά ελάχιστα. [14]

3.4.1.1 Feedforward backpropagation

Ο αλγόριθμος (feedforward backpropagation) δεν μπορεί να χρησιμοποιηθεί άμεσα σε ΑΝΔ, επειδή προϋποθέτει ότι οι συνδέσεις μεταξύ των μονάδων του δικτύου δημιουργούν μια ταξινόμηση χωρίς ανατροφοδότηση. Η λύση της προσέγγισης BPTT είναι να "ξεδιπλώσει" το αναδρομικό δίκτυο στον χρόνο, στοιβάζοντας αντίγραφα του ίδιου ΑΝΔ και ανακατευθύνοντας τις συνδέσεις μέσα στο δίκτυο (διακλαδώσεις) για να πάρουμε συνδέσεις μεταξύ των επόμενων αντιγράφων. Αυτό δίνει ένα δίκτυο feedforward που είναι ευαίσθητο στον αλγόριθμο backpropagation. [12]

Αυτό έχει μόνο μερική επιτυχία. Ένα από τα συναφή προβλήματα της μεθόδου BP για τα ΑΝΔ είναι ότι οι διακλαδώσεις μπορούν να καταστήσουν την εκπαίδευση να μην είναι συγκλίσιμη. Ακόμη και όταν συγκλίνουν, αυτή η σύγκλιση είναι αργή, υπολογιστικά ακριβή και μπορεί να οδηγήσει σε κακά τοπικά ελάχιστα.

3.4.2 ESN (ECHO STATE NETWORK)

Μία διαφορετική μέθοδος εκπαίδευσης και χρήσης των ΑΝΔ είναι τα δίκτυα Echo State (ESNs) στην Μηχανική Μάθηση, και τα Liquid State Machines (LSMs) στην υπολογιστική νευροεπιστήμη. [4] Έχει δειχθεί ότι τα ΑΝΔ συχνά λειτουργούν αρκετά καλά ακόμη και χωρίς πλήρη προσαρμογή όλων των βαρών του δικτύου. Στην κλασική προσέγγιση ESN, το ΑΝΔ ονομάζεται και δεξαμενή (reservoir) δημιουργείται τυχαία, και εκπαιδεύεται μόνο η έξοδος από το ESN.

Η τάση που ξεκίνησε από τα ESNs και τα LSMs έγινε γνωστή ως reservoir computing (RC). Το RC είναι σήμερα μια παραγωγική περιοχή έρευνας, παρέχοντας σημαντικές πληροφορίες για τα ΑΝΔ, προμηθεύοντας πρακτικά εργαλεία μηχανικής μάθησης, καθώς και δυνατότητες υπολογισμού με μη συμβατικό hardware . Το RC συμπεριλαμβάνει αρκετές σχετικές μεθόδους και επεκτάσεις της αρχικής ιδέας, αλλά η κλασική προσέγγιση ESN εξακολουθεί να κρατάει τη θέση της για την απλότητα και τη δύναμή της. [14]

Ωστόσο, η εμφανής απλότητα των ESNs μπορεί μερικές φορές να είναι παραπλανητική. Η επιτυχής εφαρμογή των ESNs απαιτεί κάποια εμπειρία. Υπάρχει ένας αριθμός πραγμάτων που μπορούν να γίνουν λάθος. Ειδικότερα, η αρχική δημιουργία του ESN δικτύου επηρεάζεται από μερικές γενικές παραμέτρους, και αυτές πρέπει να οριστούν με φρόνηση. Ωστόσο, το ίδιο μπορεί να ειπωθεί για σχεδόν κάθε τεχνική Μηχανικής Μάθησης.

3.4.2.1 Βασικό Μοντέλο

Τα ESNs εφαρμόζονται σε επιβλεπόμενες χρονικές εργασίες Μηχανικής Μάθησης, όπου για ένα συγκεκριμένο σήμα εισόδου εκπαίδευσης $u(n) \in R^{N_u}$, γνωρίζουμε ένα επιθυμητό σήμα εξόδου $y^{target}(n) \in R^{N_y}$. Εδώ $n = 1, \dots, T$ είναι ο διακριτός χρόνος και T είναι ο αριθμός των σημείων δεδομένων στο σύνολο εκπαίδευσης. Στην πράξη, το σύνολο δεδομένων μπορεί να αποτελείται από πολλές ακολουθίες διαφορετικών μηκών, αλλά αυτό δεν αλλάζει τις αρχές. Η εργασία είναι να εκπαιδευτεί ένα μοντέλο με έξοδο $y(n) \in R^{N_y}$, όπου το $y(n)$ ταιριάζει όσο το δυνατόν καλύτερα με το $y^{target}(n)$, ελαχιστοποιώντας το σφάλμα $y^{target}(n)$, και, ακόμη πιο σημαντικό, την γενίκευση σε μη ορατά δεδομένα. Η μέτρηση σφάλματος E είναι συνήθως ένα Μέσο Τετραγωνικό Σφάλμα (MSE), για παράδειγμα το Σφάλμα Ρίζας-Μέσης-Τετραγωνικής (RMSE).

$$E(y, y^{target}) = \frac{1}{N_y} \sum_{i=1}^{N_y} \sqrt{\frac{1}{T} \sum_{n=1}^T (y_i(n) - y_i^{target}(n))^2}$$

Το RMSE μπορεί επίσης να κανονικοποιηθεί (διαιρεθεί) ανά διάσταση από τη διακύμανση του $y^{target}(n)$, παράγοντας ένα Κανονικοποιημένο Ρίζας μέσης Τετραγωνικής τιμής σφάλματος (NRMSE). Το NRMSE έχει απόλυτη ερμηνεία: δεν εξαρτάται από την αυθαίρετη κλιμάκωση του στόχου $y^{target}(n)$ και η τιμή του 1 μπορεί να επιτευχθεί με μία απλή σταθερή έξοδο $y(n)$ που

ορίζεται στη μέση τιμή του $y^{target}(n)$. Αυτό υποδηλώνει ότι ένα λογικό μοντέλο μίας σταθεράς διεργασίας θα έπρεπε να επιτυγχάνει ακρίβεια NRMSE μεταξύ μηδέν και ένα.

ESNs χρησιμοποιούν έναν τύπο ΑΝΔ με μονάδες διακριτού χρόνου που έχουν διαρροή και είναι συνεχής τιμές. Οι τυπικές εξισώσεις ενημέρωσης είναι:

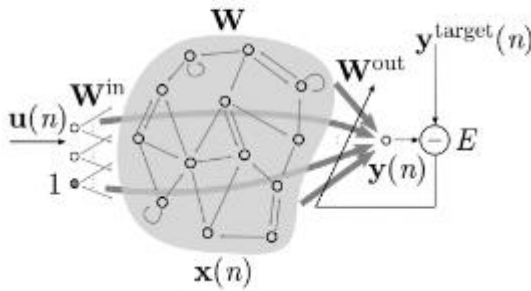
$$\begin{aligned}\tilde{x}(n) &= \tanh(W^{in}[1; u(n)] + Wx(n-1)), \\ x(n) &= (1-a)x(n-1) + a\tilde{x}(n),\end{aligned}$$

Όπου $x(n)$ ανήκει R^{N_x} είναι ένα διάνυσμα ενεργοποίησης νευρώνων και η $\tilde{x}(n)$ ανήκει R^{N_x} είναι η ενημέρωσή της, το οποίο συμβαίνει στο χρονικό βήμα n , όπου η συνάρτηση $\tanh(\cdot)$ εφαρμόζεται στοιχείο-στοιχειακά, Το $[\cdot; \cdot]$ αναπαριστά την κάθετη συνένωση διανυσμάτων (ή πινάκων). Το $W^{in} \in R^{N_x \times (1+N_u)}$ και $W \in R^{N_x \times N_x}$ αντιπροσωπεύουν τους πίνακες βαρών εισόδου και αναδρομικής ανάδρασης αντίστοιχα, ενώ το $a \in (0, 1]$ αντιπροσωπεύει τον συντελεστή διαρροής. Πέρα από την $\tanh$, μπορούν να χρησιμοποιηθούν και άλλες συναρτήσεις εκτός από αυτήν, αλλά η $\tanh$ είναι η πιο συνηθισμένη επιλογή. Κάποιες φορές το μοντέλο χρησιμοποιείται χωρίς τον συντελεστή διαρροής όπου είναι μία ειδική περίπτωση $a=1$ και επομένως $\tilde{x}(n) \equiv x(n)$

Η γραμμική έξοδος του επιπέδου εξόδου ορίζεται ως

$$y(n) = W^{out} [1; u(n); x(n)],$$

όπου $y(n) \in R^{N_y}$ η έξοδος του δικτιού, $W^{out} \in R^{N_y \times (1+N_u+N_x)}$ η έξοδος του πίνακα βαρών και $[\cdot; \cdot; \cdot]$ αναπαριστά την κάθετη συνένωση διανυσμάτων (ή πινάκων).



Εικόνα 12 Εικόνα ενός ESN

Μπορεί να εφαρμοστεί μια επιπλέον μη γραμμικότητα στο $y(n)$ στην εξίσωση, καθώς και συνδέσεις ανάδρασης W^{fb} από $y(n-1)$ έως $\tilde{x}(n)$.

Η αρχική μέθοδος των RC που εισήχθη με τα ESNs [16] ήταν:

1. Δημιουργία ενός μεγάλου τυχαίου δικτύου ESN ΑΝΔ (W^{in} , W , a).
2. Εκτέλεσή του, χρησιμοποιώντας τα δεδομένα εκπαίδευσης $u(n)$ και συλλογή των αντίστοιχων καταστάσεων ενεργοποίησης των ESN $x(n)$.
3. Υπολογισμός των γραμμικών βαρών εξόδου W^{out} από το ESN χρησιμοποιώντας γραμμική παλινδρόμηση, ελαχιστοποιώντας το MSE μεταξύ $y(n)$ και $y_{target}(n)$.

Χρήση του εκπαιδευμένου δικτύου σε νέα δεδομένα εισόδου $u(n)$ υπολογίζοντας το $y(n)$ χρησιμοποιώντας τα εκπαιδευμένα βάρη εξόδου W_{out} .

3.4.2.2 Φτιάχνοντας ένα ESN

3.4.2.2.1 Λειτουργία

Στην πράξη είναι σημαντικό να έχουμε υπόψη ότι το ESN δρα (i) ως μη γραμμική επέκταση και (ii) ως μνήμη της εισόδου $u(n)$ ταυτόχρονα. Υπάρχει μια παραλληλία μεταξύ RC και μεθόδων πυρήνα στην Μηχανική Μάθηση. Το ESN μπορεί να θεωρηθεί ως (i) μη γραμμική υψηλό διαστατική επέκταση $x(n)$ του σήματος εισόδου $u(n)$.

Για εργασίες ταξινόμησης, τα δεδομένα εισόδου $u(n)$ που δεν είναι γραμμικά διαχωρίσιμα στον αρχικό χώρο R^{N_u} , συχνά γίνονται γραμμικά διαχωρίσιμα στον επεκταμένο χώρο R^{N_x} του $x(n)$, όπου διαχωρίζονται από το W_{out} .

Ταυτόχρονα, (ii) το ESN λειτουργεί ως μνήμη, παρέχοντας ένα προσωρινό πλαίσιο. Αυτός είναι ένα κρίσιμος λόγος για τη χρήση των ΑΝΔ. Στις εργασίες όπου η μνήμη δεν είναι απαραίτητη, οι μη-προσωρινές τεχνικές μηχανικής μάθησης που υλοποιούν λειτουργικές αντιστοιχίες από την τρέχουσα είσοδο στην τρέχουσα έξοδο πρέπει να χρησιμοποιούνται.

3.4.2.2.2 Γενικές Παράμετροι από το ESN

Δεδομένου του μοντέλου ΑΝΔ το ESN καθορίζεται από το σύνολο (W_{in} , W , α). Οι πίνακες σύνδεσης εισόδου και ανατροφοδότησης W_{in} και W παράγονται τυχαία σύμφωνα με κάποιες παραμέτρους που συζητούνται αργότερα και ο ρυθμός διαρροής α (α) επιλέγεται ως ελεύθερη παράμετρος.

Αναλογικά με άλλες προσεγγίσεις Μηχανικής Μάθησης, και ιδιαίτερα Νευρωνικών Δικτύων, αυτά που εδώ ονομάζουμε "παράμετροι" θα μπορούσαν εξίσου καλά να ονομάζονται "μετα-παράμετροι" ή "υπερ-παράμετροι", καθώς δεν είναι συγκεκριμένα βάρη σύνδεσης αλλά παράμετροι που διέπουν τις κατανομές τους. Θα τις αποκαλούμε "γενικές παραμέτρους" για να αντικατοπτρίσουμε καλύτερα τη φύση τους, ή απλά "παραμέτρους" για συντομία. Οι καθοριστικές γενικές παράμετροι του ESN είναι: το μέγεθος N_x , η αραιότητα, η κατανομή των μη μηδενικών στοιχείων, και η φασματική ακτίνα του W ; η κλιμάκωση του W_{in} ; και ο ρυθμός διαρροής α .

3.4.2.2.3 Μέγεθος του ESN

Ένα εμφανώς κρίσιμο χαρακτηριστικό του μοντέλου είναι το N_x , ο αριθμός των μονάδων στο ESN. Η γενική ιδέα είναι ότι όσο μεγαλύτερο είναι το ESN, τόσο καλύτερη είναι η επίδοση που μπορεί να επιτευχθεί στο ΑΝΔ. Δεδομένου ότι η εκπαίδευση και η εκτέλεση ενός ESN είναι υπολογιστικά φθηνή σε σχέση με άλλες προσεγγίσεις εκπαίδευσης στο ΑΝΔ, τα μεγέθη των ESN της τάξης των 10^4 δεν είναι σπάνια. Όσο μεγαλύτερος είναι ο χώρος των σημάτων του ESN $x(n)$, τόσο ευκολότερο είναι να βρεθεί μια γραμμική συνδυαστική αναπαράσταση των σημάτων για να προσεγγιστούν τα $y^{target}(n)$.

Εμπειρικά το ESN μπορεί να είναι υπερβολικά μεγάλο μόνο όταν η εργασία είναι ασήμαντη και δεν υπάρχουν αρκετά δεδομένα διαθέσιμα $T < 1 + N_u + N_x$.

Οπότε μπορεί να βγει το συμπέρασμα ότι για προκλητικές εργασίες, χρησιμοποιήστε το μεγαλύτερο δυνατό ESN που μπορείτε να υποστηρίξετε υπολογιστικά.

Σε περιβάλλοντα ακαδημαϊκής έρευνας, όταν συγκρίνουμε διαφορετικές προσεγγίσεις αντί να επιδιώκουμε την απόλυτη βέλτιστη απόδοση, συχνά περιορίζουμε το μέγεθος του ESN για λόγους ευκολίας και συμβατότητας των αποτελεσμάτων. Ακόμη και όταν επιδιώκουμε την απόλυτη απόδοση, είναι δύσκολο να ξεκινήσουμε με το μεγαλύτερο δυνατό ESN από την αρχή.

Οπότε προκύπτει ότι καλό είναι να επιλέγονται γενικές παράμετροι με μικρότερα ESN, και στη συνέχεια το μέγεθος να κλιμακώνεται σε μεγαλύτερα.

Ένα κάτω του ορίου για το μέγεθος του ESN N_x μπορεί να εκτιμηθεί περίπου λαμβάνοντας υπόψη τον αριθμό των ανεξάρτητων πραγματικών τιμών που το ESN πρέπει να θυμάται από την είσοδο για να εκτελέσει με επιτυχία την εργασία του. Το μέγιστο πλήθος αποθηκευμένων τιμών, που ονομάζεται χωρητικότητα μνήμης, στα ESN δεν μπορεί να υπερβεί το N_x .

Το N_x θα πρέπει να είναι τουλάχιστον ίσο με την εκτίμηση των ανεξάρτητων πραγματικών τιμών που το ESN πρέπει να θυμάται από την είσοδο για να λύσει την εργασία του.

Για τις εισόδους $u(n)$, αυτή η εκτίμηση είναι N_u φορές μια προσεγγιστική εκτίμηση του αριθμού των χρονικών βημάτων που οι εισοδοί πρέπει να θυμούνται για να επιλύσουν την εργασία τους. Στην πράξη συχνά υπάρχουν χρονικές και διακριτικές συσχετίσεις στο $u(n)$, που το καθιστούν κάπως "συμπιεσμένο". Επίσης, οι μορφές των "καμπυλών λήθης" των ESN συνήθως δεν είναι ορθογώνιες (εξαρτώνται από άλλες παραμέτρους), δηλαδή η λήθη δεν είναι αυτοστιγμαία αλλά σταδιακή.

3.4.2.2.4 Αραιότητα του ESN

Στις αρχικές δημοσιεύσεις που αφορούσαν τα ESN, συνιστάται να γίνουν οι συνδέσεις του ESN αραιές, δηλαδή να καθίσταται η πλειονότητα των στοιχείων στο W_{in} ίσα με το μηδέν. Επίσης, οι αραιές συνδέσεις τείνουν να παρέχουν ελαφρώς καλύτερη απόδοση. Γενικά, η αραιότητα του ESN δεν επηρεάζει σημαντικά την απόδοση και αυτή η παράμετρος έχει χαμηλή προτεραιότητα για

βελτιστοποίηση. Ωστόσο, η αραιότητα επιτρέπει γρήγορες ενημερώσεις του ESN άρα και πιο γρήγορους ρυθμούς εάν χρησιμοποιούνται αραιές αναπαραστάσεις πινάκων.

Εάν, ανεξάρτητα από το μέγεθος του ESN, επιλεγεί ένας σταθερός αριθμός εξόδων, τότε το υπολογιστικό κόστος των ενημερώσεων κατάστασης του δικτύου αυξάνεται μόνο γραμμικά με το μέγεθος του δικτύου αντί για τετραγωνικά. Αυτό μειώνει σημαντικά το κόστος της λειτουργίας μεγάλων ESN. Οι υπολογιστικές εξοικονομήσεις δεν απαιτούν σχεδόν καθόλου επιπλέον προσπάθεια όταν το περιβάλλον προγραμματισμού υποστηρίζει αποτελεσματικές αναπαραστάσεις και λειτουργίες με αραιούς πίνακες, όπως κάνουν πολλοί.

3.4.2.3 Διανομή των Μη-Μηδενικών Στοιχείων

Ο πίνακας W συνήθως παράγεται αραιός, με τα μη μηδενικά στοιχεία να έχουν είτε μια συμμετρική ομοιόμορφη, διακριτή ή κανονική κατανομή με κέντρο το μηδέν. Συνήθως προτιμάτε μια ομοιόμορφη κατανομή λόγω της συνεχούς. Οι κανονικές κατανομές είναι επίσης δημοφιλείς. Και οι δύο κατανομές δίνουν σχεδόν την ίδια απόδοση, η οποία εξαρτάται από άλλες παραμέτρους. Η διακριτή κατανομή τείνει να δίνει ένα ελαφρώς λιγότερο πλούσιο χώρο σημάτων (υπάρχει μη μηδενική πιθανότητα ίδιων νευρώνων), αλλά ενδέχεται να διευκολύνει την ανάλυση του τι συμβαίνει στο ESN. Το πλάτος των κατανομών δεν έχει σημασία, καθώς επαναφέρεται με έναν τρόπο που εξηγείται στο επόμενο τμήμα. Ο πίνακας εισόδου W_{in} συνήθως παράγεται σύμφωνα με τον ίδιο τύπο κατανομής όπως και το W , αλλά συνήθως είναι πυκνός.

3.4.2.4 Φασματική ακτίνα

Μία από τις βασικές παραμέτρους ενός ESN είναι η φασματική ακτίνα του πίνακα σύνδεσης W του ESN. Αυτό κάνει κλιμάκωση τον πίνακα W στο εύρος της κατανομής των μη-μηδενικών στοιχείων του.

Συνήθως, δημιουργείται ένας τυχαίος και αραιός πίνακας W . Υπολογίζεται η φασματική ακτίνα $\rho(W)$ του W , και έπειτα διαιρείται ο W με το $\rho(W)$ για να παραχθεί ένας πίνακας με μοναδιαίο φασματικό ακτίνομα. Στη συνέχεια, αυτός κλιμακώνεται εύκολα με το τελικό επιθυμητό φασματικό ακτίνομα, το οποίο καθορίζεται σε μια διαδικασία ρύθμισης. Για να λειτουργήσει σωστά ένα ESN, πρέπει να πληροί την ιδιότητα της "επαναλαμβανόμενης κατάστασης": η κατάσταση του ESN $x(n)$ πρέπει να καθορίζεται μοναδικά από τη φθίνουσα ιστορία της εισόδου $u(n)$. Αυτό σημαίνει ότι, για αρκετά μεγάλη είσοδο $u(n)$, η κατάσταση του ESN $x(n)$ δεν εξαρτάται από τις αρχικές συνθήκες πριν την είσοδο.

Οι μεγάλες τιμές του $\rho(W)$ μπορούν να οδηγήσουν σε ESN σε πολλαπλά σταθερά σημεία, περιοδικού χαρακτήρα, παραβιάζοντας την ιδιότητα του echo state.

Η $\rho(W) < 1$ εξασφαλίζει την ιδιότητα του echo state σε σχεδόν όλες τις καταστάσεις. Αν και είναι δυνατό να παραβιαστεί η ιδιότητα του echo state, ακόμα και με $\rho(W) < 1$, αυτό είναι απίθανο να συμβεί στην πράξη. Το πιο σημαντικό, η ιδιότητα του echo state συχνά ισχύει για $1/\rho(W) \geq 1$ για μη μηδενικές εισόδους $u(n)$. Αυτό μπορεί να εξηγηθεί από το γεγονός ότι οι ισχυρές εισόδους $u(n)$ παρέχουν ενεργοποιήσεις των νευρώνων μακριά από το 0, όπου η μη γραμμικότητά τους $\tanh()$ έχει μοναδιαία κλίση, προς περιοχές όπου αυτή η κλίση είναι μικρότερη, μειώνοντας έτσι τα κέρδη των νευρώνων και την αποτελεσματική δύναμη των συνδέσεων ανάδρασης. Εν γένει, για μη μηδενικές $u(n)$ το $\rho(W) < 1$ δεν είναι απαραίτητη συνθήκη για την ιδιότητα του echo state και οι βέλτιστες τιμές του $\rho(W)$ μπορεί να είναι μερικές φορές σημαντικά μεγαλύτερες από το 1.

Στην πράξη, η τιμή του $\rho(W)$ πρέπει να επιλεγεί ώστε να μεγιστοποιηθεί η απόδοση, με την τιμή 1 να λειτουργεί ως αρχικό σημείο αναφοράς. Το φασματικό ακτίνωμα θα πρέπει να είναι μεγαλύτερο σε εργασίες που απαιτούν μεγαλύτερη μνήμη της εισόδου.

Η κλιμάκωση του πίνακα βαρών εισόδου W_{in} είναι μια ακόμα κρίσιμη παράμετρος προς βελτιστοποίηση σε ένα ESN. Για ομοιόμορφα κατανεμημένα W_{in} συνήθως ορίζουμε την κλίμακα εισόδου α ως το εύρος του διαστήματος $[-\alpha; \alpha]$ από το οποίο επιλέγονται οι τιμές του W_{in} . Για κανονικά κατανεμημένα βάρη εισόδου μπορεί να ληφθεί η τυπική απόκλιση ως μέτρο κλιμάκωσης. Για να μειώσουμε τον αριθμό των ελεύθερα αριθμήσιμων παραμέτρων, συχνά κλιμακώνουμε όλες τις στήλες του W_{in} μαζί χρησιμοποιώντας ένα μόνο κλιμακωτικό παράγοντα. Εάν τα υπόλοιπα "ενεργά" κανάλια εισόδου συμβάλλουν στην εργασία με πολύ διαφορετικούς τρόπους, τότε συνιστάται επίσης η βελτιστοποίηση των κλιμακώσεων τους ξεχωριστά.

Κλιμάκωση ομοιόμορφα ολόκληρου του W_{in} για να έχει λίγες γενικές παραμέτρους στο ESN. Ωστόσο, για αύξηση της απόδοσης:

- κλιμακώστε ξεχωριστά την πρώτη στήλη του W_{in} (δηλαδή, τις βαθμονομημένες εισόδους).
- κλιμακώστε ξεχωριστά τις υπόλοιπες στήλες του W_{in} αν τα κανάλια του $u(n)$ συμβάλλουν διαφορετικά στην εργασία.

Αυτό διαφοροποιεί τον αριθμό των ελεύθερων γενικών παραμέτρων που πρέπει να οριστούν για το W_{in} από 1 έως $N_u + 1$.

Η κανονικοποίηση των δεδομένων εισόδου είναι συμφέρουσα για τα ESN όπως και για οποιαδήποτε άλλη μέθοδο μηχανικής μάθησης. Αυτό τοποθετεί κάθε εργασία μάθησης σε ένα πιο τυποποιημένο πλαίσιο. Είναι χρήσιμο να περιοριστεί το εύρος των τιμών των δεδομένων εισόδου. Για παράδειγμα, εφαρμόστηκε το squashing $\tanh(\cdot)$ στο $u(n)$ αν η κατανομή του είναι απεριόριστη. Διαφορετικά, τα ακραία σημεία μπορούν να οδηγήσουν την κατάσταση του δεξαμενοποιημένου $x(n)$ σε "άγνωστες" περιοχές που δεν καλύπτονται καλά από τις συνήθεις εργασιακές τροχιές του $x(n)$ για τις οποίες έχουν βελτιστοποιηθεί οι γενικές παράμετροι ή έχουν μάθει οι έξοδοι. Αυτό μπορεί να οδηγήσει σε εικονική απώλεια χρήσιμης μνήμης (λόγω κορεσμών στις μη γραμμικότητες ενεργοποίησης) ή απρόβλεπτες εξόδους σε αυτά τα σημεία, αντίστοιχα.

Η κλιμάκωση της εισόδου καθορίζει πόσο μη γραμμικές είναι οι αντιδράσεις του reservoir. Ενώ το $\rho(W)$ επίσης επηρεάζει τη μη γραμμικότητα, οι ενεργοποιήσεις του reservoir γίνονται ασταθείς κατά την αύξηση του $\rho(W)$, είναι προφανές ότι η κλιμάκωση του W_{in} , μαζί με την κλιμάκωση του W (δηλαδή, το $\rho(W)$) καθορίζει το ποσοστό του πόσο η τρέχουσα κατάσταση $x(n)$ εξαρτάται από την τρέχουσα είσοδο $u(n)$. Η κλιμάκωση της εισόδου ρυθμίζει:

- το ποσό της μη γραμμικότητας της αναπαράστασης reservoir $x(n)$ (αυξάνεται επίσης με το $\rho(W)$);
- τον σχετικό αντίκτυπο της τρέχουσας εισόδου στο $x(n)$ σε σύγκριση με την ιστορία (ανάλογα με το $\rho(W)$).

Έχει παρατηρηθεί εμπειρικά ότι η αναπαράσταση των διαφορετικών κύριων συνιστωσών του $u(n)$ στο $x(n)$ είναι περίπου ανάλογη της τετραγωνικής ρίζας των μεγεθών τους στο $u(n)$. Με άλλα λόγια, το reservoir τείνει να εξομαλύνει το φάσμα των κύριων συνιστωσών του $u(n)$ στο $x(n)$ - κάτι που πρέπει να ληφθεί υπόψιν κατά την επιλογή της κατάλληλης αναπαράστασης ή προ επεξεργασίας των δεδομένων. Για παράδειγμα, αν οι μικρότερες κύριες συνιστώσες δεν μεταφέρουν καμία χρήσιμη πληροφορία, μπορεί να είναι χρήσιμο να τις αφαιρέσετε από τα δεδομένα με την μέθοδο Ανάλυση Κύριων Συνιστωσών (PCA) πριν τις τροφοδοτήσετε σε ένα reservoir.

3.4.2.5 Ρυθμός διαρροής

Ο ρυθμός διαρροής συμβολίζεται με το α μπορεί να θεωρηθεί ως η ταχύτητα της δυναμικής ενημέρωσης του reservoir που διακριτοποιείται στο χρόνο.

$$\dot{x} = -x + \tanh(W^{in}[1; u] + Wx)$$

Αν φτιάξουμε την διακριτή παραγωγό του Euler

$$\frac{\Delta x}{\Delta t} = \frac{x(n+1) - x(n)}{\Delta t}$$

Έτσι, το α μπορεί να θεωρηθεί ως το χρονικό διάστημα στο συνεχές χώρο μεταξύ δύο διαδοχικών χρονικών βημάτων στη διακριτή υλοποίηση. Επίσης, εμπειρικά, η επίδραση της ρύθμισης του α είναι συγκρίσιμη με αυτήν του επαναδειγματοληπτικού $u(n)$ και $y_target(n)$ όταν τα σήματα είναι αργά. Η ρύθμιση του leaking rate α μπορεί ακόμη να προσαρμοστεί δυναμικά για να αντιμετωπίσει το χρονικό wrapping των σημάτων.

Έτσι, το α μπορεί να θεωρηθεί ως το χρονικό διάστημα στο συνεχές χώρο μεταξύ δύο διαδοχικών χρονικών βημάτων στη διακριτή υλοποίηση. Επίσης, εμπειρικά, η επίδραση της ρύθμισης του α είναι συγκρίσιμη με αυτήν του επαναδειγματοληπτικού $u(n)$ και $y_target(n)$ όταν τα σήματα είναι αργά. Η ρύθμιση του leaking rate α μπορεί ακόμη να προσαρμοστεί δυναμικά για να αντιμετωπίσει το χρονικό wrapping των σημάτων.

Σε ορισμένες περιπτώσεις, η ρύθμιση ενός μικρού α , και συνεπώς η εισαγωγή αργών δυναμικών για το $x(n)$, μπορεί δραματικά να αυξήσει τη διάρκεια βραχυπρόθεσμης μνήμης στο ESN.

3.5 Πρακτική Προσέγγιση για την Παραγωγή *reservoir*

3.5.1 Προτεραιότητα Παραμέτρων

Ενώ όλες οι παράμετροι του ESN που συζητήθηκαν έχουν καθοδήγηση στην ρύθμισή τους, η επίλυση ορισμένων από αυτές είναι πιο απλή από άλλες.

Οι τρεις βασικές παράμετροι που πρέπει να βελτιστοποιηθούν σε μια δεξαμενή ESN είναι:

- κλιμάκωση(-εις) της εισόδου.
- φασματική ακτίνα.
- ρυθμός διαρροής(-ων).

Όπως εξηγήθηκε προηγουμένως, η απόδοση μπορεί επιπλέον να βελτιωθεί σε πολλές περιπτώσεις με το "διαχωρισμό" μιας μεμονωμένου παραμέτρου σε πολλές. Η ρύθμιση διαφορετικών κλιμακώσεων στις στήλες του Win (που αντιστοιχούν στη βαθμίδα εισόδου και ενδεχομένως σε διαφορετικές διαστάσεις εισόδου αν είναι διαφορετικής φύσης) μπορεί να προχωρήσει μακριά. Επίσης, η ρύθμιση των ρυθμών διαρροής α , διαφορετικά για διαφορετικές μονάδες (για παράδειγμα, διαχωρίζοντάς τους σε διάφορες υποπληθυσμιακές ομάδες με σταθερή τιμή) μπορεί να βοηθήσει πολύ σε πολυχρονοσειριακές εργασίες.

3.5.2 Η διάταξη για την επιλογή παραμέτρων

Ένα από τα κύρια πλεονεκτήματα των ESNs είναι ότι η εκμάθηση των εξόδων είναι γρήγορη. Αυτό μπορεί να μας δείξει κατά την αξιολόγηση πόσο καλό είναι ένα *reservoir* που δημιουργήθηκε από ένα συγκεκριμένο σύνολο παραμέτρων.

Ο πιο πρακτικός τρόπος αξιολόγησης είναι να εκπαιδευτεί η έξοδος και να μετρηθεί το σφάλμα της.

Μπορεί να χρησιμοποιηθεί είτε ο έλεγχος ή το σφάλμα εκπαίδευσης. Ο έλεγχος είναι, φυσικά, προτιμώμενος αν υπάρχει κίνδυνος υπερπροσαρμογής. Το σφάλμα εκπαίδευσης έχει το πλεονέκτημα της χρήσης λιγότερων δεδομένων και σε μερικές περιπτώσεις δεν υπάρχει ανάγκη να επανεκκινηθεί ένα εκπαιδευμένο δίκτυο.

Αν ένα σύνολο δεδομένων επικύρωσης είναι απαραίτητο για την εκπαίδευση της εξόδου το σφάλμα σε αυτό μπορεί να χρησιμοποιηθεί χωρίς επιπλέον κόστος, ως μια συμβιβαστική λύση μεταξύ της εκπαίδευσης και ενός ακόμα ξεχωριστού συνόλου επικύρωσης. Αν η εκπαίδευση της εξόδου και η επικύρωση δεν είναι αρκετά γρήγορες, μπορούν να χρησιμοποιηθούν μικρότερα αρχικά *reservoir*, ή ένα μειωμένο αντιπροσωπευτικό σύνολο δεδομένων. Για τον ίδιο λόγο, συχνά είναι υπερβολικό να χρησιμοποιείται ένα *k-fold cross-validation* στην παραμετροποίηση των γενικών παραμέτρων, τουλάχιστον στα αρχικά στάδια, εκτός αν τα δεδομένα είναι πραγματικά περιορισμένα.

Είναι ουσιώδες να ληφθεί υπόψη ότι οι δεξαμενές που παράγονται τυχαία, ακόμα και με τις ίδιες παραμέτρους, εμφανίζουν ελαφρά διαφορετικές επιδόσεις. Αυτή η διακύμανση είναι διαρκώς παρούσα, αλλά συνήθως είναι πιο εμφανής σε μικρότερες δεξαμενές σε σύγκριση με μεγαλύτερες. Οι τυχαίες διακυμάνσεις μέσα σε μια μεγάλη δεξαμενή τείνουν να "μεσοστροωθούν". Παρ' όλα αυτά, είναι σημαντικό να διατηρείται αυτή η τυχειότητα της επίδοσης ξεχωριστά από εκείνη που οφείλεται σε διαφορετικές τιμές παραμέτρων.

Για να εξαλειφθεί η τυχειότητα της διακύμανσης της απόδοσης, διατηρήστε τον τυχαίο σπόρο σταθερό και/ή κάντε μέσο όρο πάνω από αρκετά δείγματα ESN.

Η διατήρηση ενός τυχαίου seed στο περιβάλλον προγραμματισμού πριν από την παραγωγή των reservoir καθιστά το τυχαίο χαρακτηριστικό των δεξαμενών πανομοιότυπο σε δοκιμές και επομένως τα επαναληπτικά πειράματα προβλέψιμα. Η χρήση ενός μοναδικού ESN είναι πιο γρήγορη, αλλά υπάρχει ο φόβος της κάτω του μέσου όρου απόδοσης και/ή της υπερεκπαίδευσης των παραμέτρων σε ένα συγκεκριμένο παράδειγμα ενός τυχαίου παραγόμενου ESN.

Οι καλές παράμετροι που έχουν βρεθεί ενδέχεται να μη λειτουργούν το ίδιο καλά σε μια διαφορετική υλοποίηση λογισμικού ή, για παράδειγμα, διαφορετικό μέγεθος μιας δεξαμενής.

3.5.3 Επιλογή Παραμέτρων με Χειροκίνητο Τρόπο

Όταν ρυθμίζονται χειροκίνητα οι παράμετροι της δεξαμενής, αλλάζετε μία παράμετρο κάθε φορά. Οι αλλαγές σε πολλές παραμέτρους ταυτόχρονα συχνά έχουν αντίθετες επιδράσεις στην απόδοση, μιας και είναι αδύνατον να γνωρίζουμε ποια συνέβαλε τι. Μια λογική προσέγγιση είναι να οριστεί μια μόνο παράμετρος σε μια επαρκώς καλή τιμή πριν αρχίσει να αλλάζει μια άλλη, και να επαναλαμβάνετε αυτή τη διαδικασία μέχρι η απόδοση να είναι ικανοποιητική.

Είναι επίσης συνιστώμενο να κρατάτε σημειώσεις ή να καταγράφετε αυτόματα την απόδοση κατά τη διάρκεια επεκτεταμένων βελτιστοποιήσεων. Ωστόσο, οι τύποι σφαλμάτων είναι συχνά μη-κυρτοί και η δοκιμή απομακρυσμένων τιμών των παραμέτρων μπορεί μερικές φορές να οδηγήσει σε σοβαρές βελτιώσεις.

Πάντα να σχεδιάζετε δείγματα των σημάτων ενεργοποίησης του reservoir $x(n)$ για να αποκτήσετε μια αίσθηση του τι συμβαίνει μέσα στο reservoir.

Αυτό μπορεί να αποκαλύψει εάν τα $x(n)$ είναι υπερκορεσμένα, υποενεργοποιημένα, εμφανίζουν αυτόνομη κυκλική ή χαοτική συμπεριφορά κ.λπ. Συνολικά, η παρουσίαση πληροφοριών πέρα από το ρυθμό σφάλματος βοηθά πολύ στην κατανόηση πώς οι παράμετροι θα πρέπει να προσαρμοστούν.

Συνήθως, η καλή μέση απόδοση δεν βρίσκεται σε ένα πολύ στενό εύρος παραμέτρων, επομένως η λεπτομερής ρύθμιση των παραμέτρων δεν προσφέρει σημαντική βελτίωση και δεν είναι απαραίτητη.

3.5.4 Αυτοματοποιημένη Επιλογή Παραμέτρων

Δεδομένου ότι η χειροκίνητη βελτιστοποίηση παραμέτρων μπορεί να γίνει γρήγορα κουραστική, συχνά προτιμώνται αυτοματοποιημένες προσεγγίσεις. Δεδομένου ότι τα ESN έχουν μόνο λίγες παραμέτρους που απαιτούν πιο προσεκτική ρύθμιση, η μέθοδος grid search είναι πιθανότατα η πιο απλή επιλογή. Είναι αρκετά εύκολο να υλοποιηθεί με μερικές εμφωλευμένες επαναλήψεις, και υψηλού επιπέδου βιβλιοθήκες προγραμματισμού ML.

Μια λογική προσέγγιση είναι να γίνει μια πιο γενική αναζήτηση σε ένα πιο ευρύ φάσμα παραμέτρων για να εντοπιστούν περιοχές υποσχόμενες και στη συνέχεια να γίνει μια λεπτομερέστερη αναζήτηση (μικρότερα βήματα) σε αυτές τις περιοχές. Όπως αναφέρθηκε, συνήθως το πλέγμα δεν πρέπει να είναι πολύ πυκνό για να επιτευχθεί καλή απόδοση.

Η καλύτερη απόδοση σε ένα όριο ενός καλυμμένου grid είναι ένα καλό σημάδι ότι η βέλτιστη απόδοση ενδέχεται να βρίσκεται έξω από το πλέγμα. [14]

3.5.4.1 NSGA ALGORITHM

Στην ενότητα αυτή περιγράφονται μετρικές που χρησιμοποιούνται για την αξιολόγηση της απόδοσης των δικτύων echo state (ESN) και παρουσιάζεται μια μεθοδολογία που χρησιμοποιεί το NSGA (Non-dominated Sorting Genetic Algorithm) για την καθιέρωση ενός καθολικού χώρου συμπεριφορών για την προσομοίωση της συμπεριφοράς των ESN. Εδώ είναι ένα περιληπτικό παράδειγμα:

1. **Κατάταξη Kernel (KR):** Μετρά την ικανότητα του δικτύου να διακρίνει μεταξύ διαφορετικών εισόδων.
2. **Γενική Κατάταξη (GR):** Αξιολογεί τη γενίκευση του δικτύου, με χαμηλότερο GR να υποδεικνύει μεγαλύτερη ανθεκτικότητα του δικτύου.
3. **Χωρητικότητα Μνήμης (MC):** Ποσοτικοποιεί τη γραμμική μακροπρόθεσμη αποθήκευση του δικτύου, αντικατοπτρίζοντας τα χαρακτηριστικά της κατάστασης echo.

Το NSGA χρησιμοποιείται για τη δημιουργία ενός καθολικού χώρου συμπεριφορών για την προσομοίωση διαφόρων συμπεριφορών των ESN. Η νομοτέλεια, που μετρά τη συστηματοποίηση των εξελικτικών κόμβων, χρησιμοποιείται για να διασφαλίσει ότι ο καθολικός χώρος συμπεριφοράς αντιπροσωπεύει κατάλληλα τη συνολική συμπεριφορά των ESN.

Η μεθοδολογία περιλαμβάνει την τυχαία αρχικοποίηση ενός πληθυσμού ατόμων (Ps). Μέσω μιας διαδικασίας που περιλαμβάνει "διασταύρωση" και "μετάλλαξη", τα άτομα που νικούν παραμένουν αναλλοίωτα, ενώ τα άτομα που χάνουν ανταλλάσσουν γονίδια με τα νικητήρια άτομα, παράγοντας νέα άτομα. [17]

1. Αρχικοποίηση:

- Δημιουργία ενός αρχικού πληθυσμιακού συνόλου των συνολικών παραμέτρων του ESN.
- Αναθέστε σε κάθε σύνολο παραμέτρων μια αρχική τιμή 0.
- Αναθέστε σε κάθε σύνολο παραμέτρων μια αρχική απόσταση πληθυσμού 0.

2. Αξιολόγηση:

- Για κάθε σύνολο παραμέτρων στην πληθυσμιακή, εκπαίδευση ενός ESN και αξιολόγηση την απόδοσή του χρησιμοποιώντας κατάλληλες μετρικές (π.χ., RMSE, NRMSE).
- Υπολογισμός του αριθμού των κυριαρχούμενων συνολικών παραμέτρων και με βάση της απόδοσής τους.

3. Ταξινόμηση μη-κυριαρχούμενων:

- Πραγματοποιήστε ταξινόμηση μη-κυριαρχούμενων για να χωριστεί η πληθυσμός σε διάφορα μέτωπα βάσει των σχέσεων κυριαρχίας.
- Αναθέστε σε κάθε σύνολο παραμέτρων έναν βαθμό ανάλογα με το μέτωπό του.

4. Ανάθεση Απόστασης Πληθυσμού:

- Υπολογίστε την απόσταση πληθυσμού για κάθε σύνολο παραμέτρων σε κάθε μέτωπο.
- Ταξινομήστε τα σύνολα παραμέτρων σε κάθε μέτωπο βάσει των μετρικών τους απόδοσης.
- Αναθέστε αποστάσεις πληθυσμού σε σύνολα παραμέτρων βάσει της θέσης τους στα ταξινομημένα μέτωπα.

5. Επιλογή Γονέων:

- Επιλέξτε σύνολα παραμέτρων για αναπαραγωγή με βάση τον βαθμό και τις αποστάσεις πληθυσμού.
- Χρήση της μεθόδου επιλογής γονέων για αναπαραγωγή.

6. Συνδυασμός και Μετάλλαξη:

- Πραγματοποίηση τις λειτουργίες συνδυασμού και μετάλλαξης για τη δημιουργία απογόνων από τους επιλεγμένους γονείς.
- Εφαρμογή στους τελεστές συνδυασμού και μετάλλαξης με καθορισμένες πιθανότητες.

7. Αξιολόγηση Απογόνων:

- Για κάθε απόγονο σύνολο παραμέτρων που προέκυψε από τη διαδικασία αναπαραγωγής, εκπαιδεύεται ένα ESN και αξιολογείτε η απόδοσή του.

8. Επιλογή Επιζώντων:

- Συνδυασμός του πληθυσμού και των απογόνων.
- Εφαρμογή περιβαλλοντικής επιλογής για να επιλεγθεί η επόμενη γενιά συνόλων παραμέτρων βάσει της μη-κυριαρχίας και της απόστασης πληθυσμού.

9. Κριτήριο Τερματισμού:

- Έλεγχος αν έχει επιτευχθεί ένα κριτήριο τερματισμού (π.χ., συνθήκη σύγκλισης, αριθμός επαναλήψεων).

10. Επανάληψη:

- Επανάληψη τα βήματα 3 έως 9 μέχρι να επιτευχθεί το κριτήριο τερματισμού. [6]

4

ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

4.1 Εισαγωγή

Το πειραματικό μέρος ασχολείται με την προσομοίωση ενός ΦΜΚΣ με την χρήση ενός echo state network ESN σε python.

4.1.1 ESN ως ΦΜΚΣ

Όπως αναφέρθηκε και στο κεφάλαιο 2.7.5 μία νευρομορφική ΦΜΚΣ μπορεί να πραγματοποιηθεί χρησιμοποιώντας ένα οποιοδήποτε τύπο νευρωνικού δικτύου, το ESN θα μπορούσε να λειτουργήσει ως ΦΜΚΣ μίας και είναι ένα νευρωνικό δίκτυο και σαν απόκριση παίρνουμε τον πίνακα των βαρών του.

4.2 Υλοποίηση

Όλοι οι κώδικες της διπλωματικής είναι ανεβασμένοι στο gitlab στον σύνδεσμο https://gitlab.com/GBones/esn_as_puf με περιγραφή του τι κάνουν.

Για αρχή το esn που έχει γίνει στην συγκεκριμένη εργασία βασίζεται στον κώδικα [18] (ο οποίος βρίσκεται και στο παράρτημα II) έχει δύο βασικές λειτουργίες generative και predictive.

```
# generative mode:
u = y
## this would be a predictive mode:
u = data[trainLen+t+1]
```

Εικόνα 13 Generative and Predictive Mode

Στην generative λειτουργία αφού έχει αρχικοποιηθεί με κάποια αρχικά δεδομένα στην είσοδο παράγει έξοδο αυτόνομα και αναδρομικά με βάση τις εσωτερικές δυναμικές του δικτύου.

Στην predictive λειτουργία που είναι και αυτή που μας ενδιαφέρει το δίκτυο δέχεται εξωτερική είσοδο μέχρι ένα σημείο και καλείται να προβλέψει τις επόμενες καταστάσεις με βάση το σήμα στην είσοδο.

Αφού το δίκτυο εκπαιδευτεί στα δεδομένα, παίρνει νέα δεδομένα στην είσοδο και η έξοδος του συγκρίνεται με τις πραγματικές μελλοντικές καταστάσεις του σήματος για να αξιολογηθεί η απόδοση των προβλέψεων του δικτύου.

Στο gitlab στον φάκελο RC το αρχείο main_narma.py έχει το esn με είσοδο μία NARMA 10 (περισσότερα για το τι είναι η NARMA στο παράρτημα 1) οι βασικές μεταβλητές που αν αλλάξουν επηρεάζουν το esn είναι για τα δεδομένα εκπαίδευσης.

Το μήκος εκπαίδευσης (trainLen) δηλαδή πόσα σημεία του σήματος θα χρησιμοποιηθούν για την εκπαίδευση.

Το μήκος τεσταρίσματος (testLen) δηλαδή πόσα σημεία θα χρησιμοποιηθούν για τον έλεγχο

Το initLen, δηλαδή πόσα σημεία θα χρησιμοποιηθούν για την αρχικοποίηση του δικτύου

```
# load the data
trainLen = 3000 #4000 #20000
testlen = 1000
initlen = 100 #the first 100 steps would be used for solely for initializing the internal state
data = np.loadtxt('NARMA_10')[3050:]
```

Εικόνα 14 Κομμάτι του κώδικα για την ρύθμιση των δεδομένων εκπαίδευσης

Για το ESN

```
resSize = 1000 #700 #200
a = 0.8131811018689734 #0.9 # leaking rate
```

Εικόνα 15 Μέγεθος Reservoir και ρυθμού Διαρροής

Το μέγεθος του reservoir resize, δηλαδή, από πόσα reservoir θα αποτελείται.
α. ο ρυθμός διαρροής

```
np.random.seed(49)
```

Εικόνα 16 Random Seed

Το random seed χρησιμοποιείται όταν θέλουμε να διασφαλίσουμε ότι τα αποτελέσματα που βασίζονται σε τυχαία δεδομένα μπορούν να επαναληφθούν ακριβώς με την ίδια ακολουθία τυχαίων αριθμών σε διαφορετικές εκτελέσεις του προγράμματος.

```
W *= 0.6550174608604511 / rhoW #1.25
```

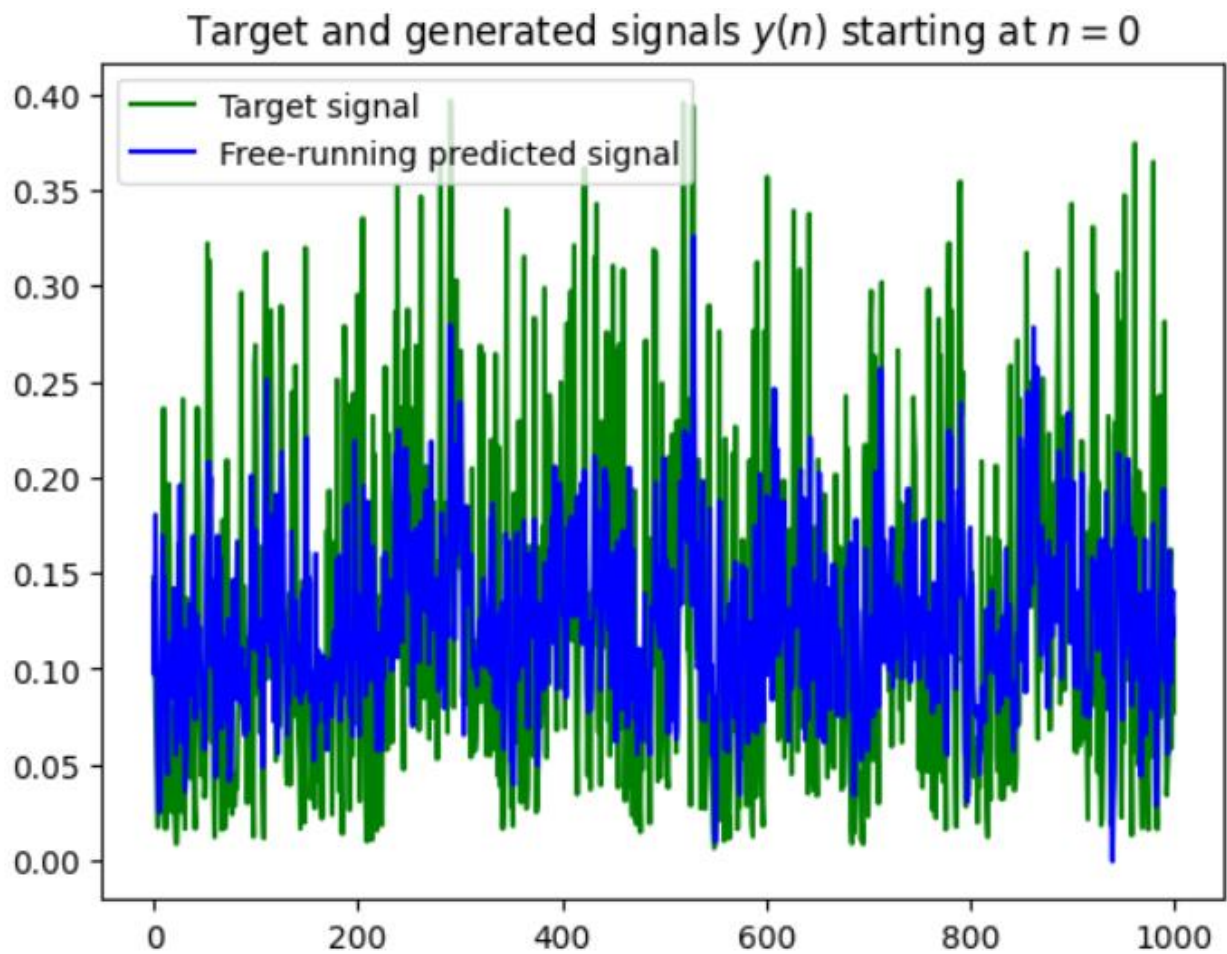
Εικόνα 17 Φασματική ακτίνα

Η φασματική ακτίνα ρυθμίζεται από εδώ

Σαν έξοδο ο κώδικας εξάγει το MSE (mean square error) το οποίο προκύπτει από την απόκλιση που έχει το πραγματικό σήμα εξόδου με το επιθυμητό

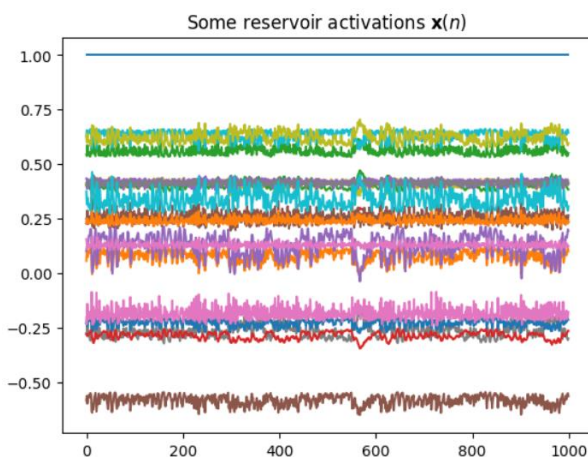
MSE = 0.00456758284816694

Εικόνα 18 MSE Mean Square Error



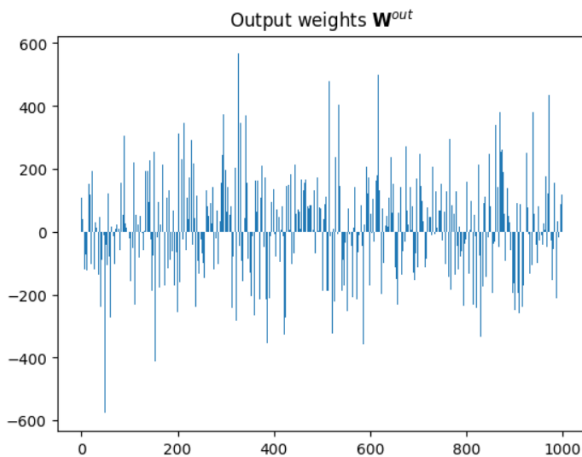
Εικόνα 19 Μπλε χρώμα σήμα στην έξοδο , πράσινο χρώμα επιθυμητή έξοδος.

Ακόμη δείχνει την ενεργοποίηση των reservoir η οποία δεν πρέπει να ξεπερνάει το 1 για να διατηρείται το echo state.



Εικόνα 20 Ενεργοποίηση βαρών εξόδου

Καθώς επίσης δείχνει και τις τιμές των βαρών



Εικόνα 21 Ιστόγραμμα Τιμών Βαρών

Αφού τρέξουμε τον NSGAII αλγόριθμο

```
Reservoir Size: 176.60952529561342, Leaking Rate: 0.827244240977505, Spectral Radius: 0.665298875598733  
MSE (ESN): 0.004352703166083727, MSE (Ridge): 0.016164261119303833
```

Εικόνα 22 Έξοδος NSGAII

Με βάση το MSE, λαμβάνουμε τις καλύτερες παραμέτρους για το ESN μας

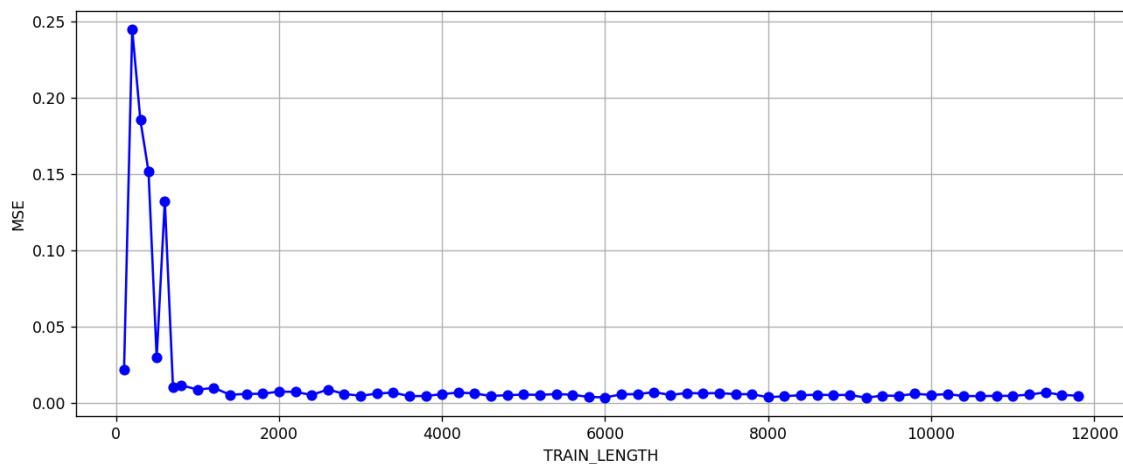
Το reservoir size πρέπει να είναι πάντα ακέραιος αριθμός μιας και δεν γίνεται σε πραγματικές συνθήκες να υπάρχουν υποδιαίρέσεις του, οπότε επειδή το 176 είναι μικρός αριθμός για reservoir θα προτιμηθεί ένας πολύ μεγαλύτερος αριθμός όπως 1000 προκειμένου να μην είναι τόσο εύκολα αναστρέψιμη η διαδικασία για το πως προκύπτουν τα βάρη.

Πρακτικά όσο περισσότερα βάρη υπάρχουν, τόσο περισσότεροι αριθμοί προκύπτουν που μετά θα γίνουν δυαδικοί. Η διαδικασία από τα βάρη ως το σημείο που θα γίνουν δυαδικοί είναι αναστρέψιμη.

Για να βρεθεί το βέλτιστο μήκος των σημείων εκπαίδευσης θα πρέπει με τις βέλτιστες παραμέτρους που έχουμε βρεί να τρέξουμε το ίδιο ESN για πολλές διαφορετικές επαναλήψεις προκειμένου να το δούμε αυτό το κάνει ο κώδικας main_rand_seed.py κάνοντας το για train length μέχρι 11800 με βήμα 200 βλέπου ότι με βάση το MSE η απόδοσή του σταθεροποιείται μετά το 1000 σημείο. Εδώ να σημειωθεί ότι όσο περισσότερα δείγματα έχουμε τόσο περισσότερος ο χρόνος εκπαίδευσης.

Ο άξονας Y δείχνει το MSE

Και ο άξονας X το train_length



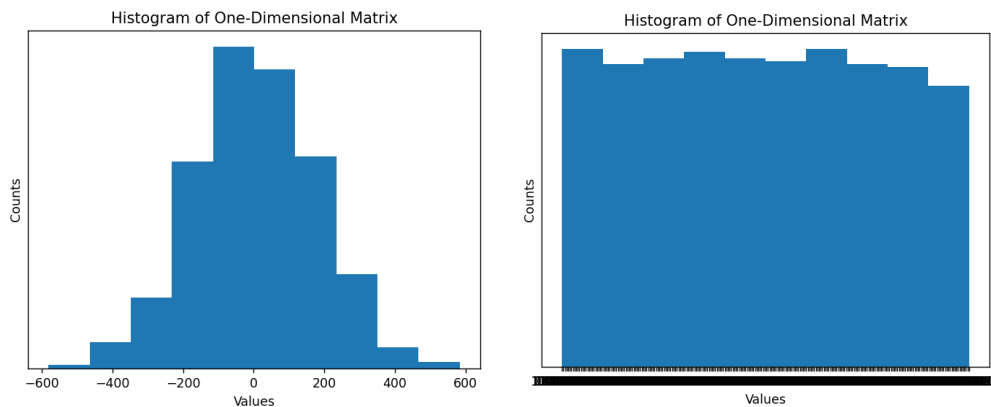
Εικόνα 23 MSE ανά μήκος εκπαίδευσης

4.2.1 Inter distance

Όπως αναφέρθηκε και στο κεφάλαιο 3.3.2 το "inter-distance" στις ΦΜΚΣ είναι η απόσταση μεταξύ των αποκρίσεων που παράγονται από διαφορετικές ΦΜΚΣ στιγμιότυπα υπό τις ίδιες προκλήσεις.

Με τις ρυθμίσεις που προέκυψαν από το NSGAIΙ στον φάκελο OneRC_for_allNarma εκτελούμε τον κώδικα στον φάκελο OneNarma_for_allRC με το όνομα main.py ο οποίος τρέχει για μία πρόκληση (NARMA10) το ίδιο ESN αλλά για πολλά διαφορετικά seed value με αυτό τον τρόπο πετυχαίνουμε να προσομοιάσουμε την απόκριση για πολλά διαφορετικά αντίγραφα της ίδια ΦΜΚΣ.

Έπειτα πρέπει να γίνει η ομογενοποίηση των βαρών προκειμένου όλα τα βάρη να έχουν την ίδια πιθανότητα. Στον φάκελο CodeForWeightPlot άμα τρέξουμε το αρχείο weight_normal_to_uniform.py πραγματοποιεί την ομογενοποίηση σε ένα βάρος μία μικρή τροποποίηση πραγματοποιεί ομογενοποίηση σε όλα τα βάρη. Μετά με την weight Istogram βλέπουμε το ιστόγραμμα των βαρών για να δούμε αν έγινε σωστά η μετατροπή.



Εικόνα 24 Ραβδόγραμμα Βαρών πριν και μετά την ομογενοποίηση

Προκειμένου να εξετάσουμε κατά πόσο τυχαία είναι η απόκριση των ΦΜΚΣ παίρνουμε τον ομογενοποιημένο πίνακα βαρών του κάθε challenge και θα υπολογίσουμε το mean hamming distance για όλες τις προκλήσεις το οποίο θα πρέπει να είναι όσο πιο κοντά γίνεται στο 0.5 . Ο κώδικας που θα κάνει τον υπολογισμό του hamming distance βρίσκεται στον φάκελο Hamming_Distance εκτελώντας το αρχείο main.py. Πριν γίνει αυτό θα πρέπει να εκτελεστεί ο κώδικας scaling_binary_weights.py ο οποίος κάνει κβάντιση τριών bit στα βάρη

```
Mean Hamming distance between all pairs of matrices: 0.4995000410432072
Standard deviation of Hamming distances: 0.005592625365331501
Minimum Hamming distance: 0.47355289421157687
Maximum Hamming distance: 0.5245758483033932
```

Εικόνα 25 Hamming distance

4.2.2 Inter challenge

Όπως αναφέρθηκε και στη θεωρία το inter chalange αναφέρεται στην hamming distance μεταξύ των διαφορετικών προκλήσεων στην ίδια ΦΜΚΣ.

Οι προκλήσεις στην συγκεκριμένο ΦΜΚΣ είναι η NARMA 10 για διαφορετικά seed οπότε κάθε narma10 με διαφορετικό seed αποτελεί και διαφορετικό challenge στο συγκεκριμένο πειραματικό έχουμε 1000 narma10 με seed τιμές από 0 έως 999 (δυστυχώς λόγω όγκου δεν μπόρεσαν να ανέβουν στο gitlab οι χίλιες Narma10)

Με τις ρυθμίσεις που προέκυψαν από το NSGAI στον φάκελο OneRC_for_allNarma εκτελούμε το main.py αρχείο το οποίο εκτελεί ένα ESN (με τις ρυθμίσεις που αναφέραμε πιο πριν τις οποίες τις έχουμε περάσει χειροκίνητα μέσα στον κώδικα) με όλα τα Narma που υπάρχουν στο narma_10.npz και τα αποθηκεύει σε διαφορετικούς φακέλους ανάλογα με την τιμή του seed κάθε Narma10, και μέσα στον κάθε φάκελο αποθηκεύει τον πίνακα βαρών και το MSE.

Αυτό θα μας δώσει 1000 φακέλους

Προκειμένου να εξετάσουμε κατά πόσο τυχαία είναι η απόκριση της ΦΜΚΣ παίρνουμε τον πίνακα βαρών του κάθε challenge και θα υπολογίσουμε το mean hamming distance για όλες τις προκλήσεις το οποίο θα πρέπει να είναι όσο πιο κοντά γίνεται στο 0.5 .

Ο κώδικας που θα κάνει των υπολογισμό του hamming distance βρίσκεται στον φάκελο Hamming_Distance εκτελώντας το αρχείο main.py

```
Mean Hamming distance between all pairs of matrices: 0.4996458073020573
Standard deviation of Hamming distances: 0.002173042597500594
Minimum Hamming distance: 0.48727834665334663
Maximum Hamming distance: 0.5104895104895105
```

Εικόνα 26 Mean Hamming

Οπότε με αυτά τα αποτελέσματα σημαίνει ότι δεν μπορούμε να αναπαράγουμε τις αποκρίσεις της συγκεκριμένης ΦΜΚΣ χωρίς να ξέρουμε την ακριβή εσωτερική δομή του (δηλ. των πίνακα βαρών για την εσωτερική διασύνδεση των κόμβων)

4.2.3 Intra reservoir

Το Intra reservoir αναφέρεται στο hamming distance μεταξύ της ίδιας πρόκλησης στην ίδια ΦΜΚΣ αλλά με την εισαγωγή θορύβου στο σύστημα.

Στο κώδικα Win_With_Rand_distr.py περιέχεται η αλλαγή που έγινε στο ESN προκειμένου να εισαχθεί ο θόρυβος. Αυτό επιτευχθεί αποθηκεύοντας την αρχική τιμή των βαρών εισόδου και έπειτα προσθέτοντας μια τυχαία κατανομή της τάξης του 0.001

```
# Add normal distribution noise to Win
noise = np.random.normal(0, 0.001, Wstart.shape)
```

Εικόνα 27 Τυχαία κατανομή

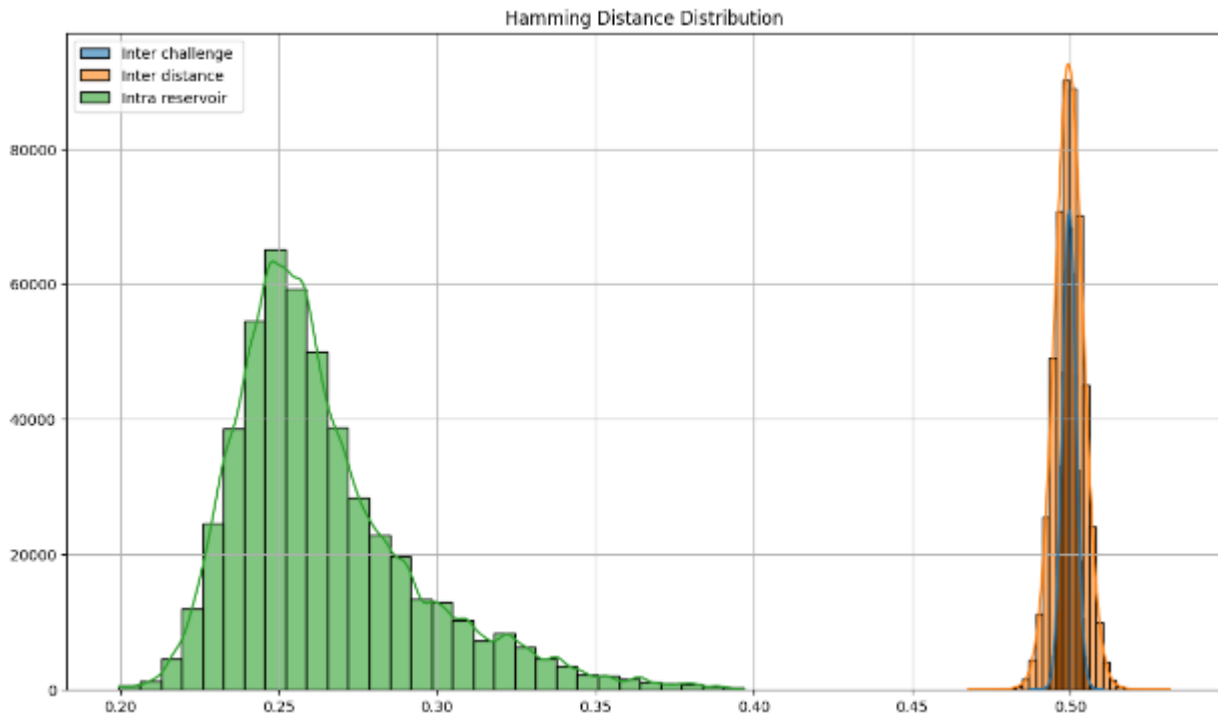
Αλλάζοντας το seed της τυχαίας κατανομής και προσθέτοντάς το στον αρχικό πίνακα βαρών το κάνουμε για 1000 φορές και αποθηκεύονται σε φακέλους με βάση το seed της τυχαίας κατανομής και μέσα στον κάθε φάκελο αποθηκεύει τον πίνακα βαρών και το MSE.

Προκειμένου να εξετάσουμε κατά πόσο τυχαία είναι η απόκριση της ΦΜΚΣ παίρνουμε τον πίνακα βαρών του κάθε σεναρίου και θα υπολογίσουμε το mean hamming distance για όλα τα σενάρια τα οποία θα πρέπει να έχουν ένα χαμηλό hamming distance. Ώστε έχοντας την συγκεκριμένη ΦΜΚΣ και την ίδια πρόκληση να μπορείς να αναπαράγεις το αποτέλεσμα.

```
Mean Hamming distance between all pairs of matrices: 0.2638288406352278
Standard deviation of Hamming distances: 0.029041822477493155
Minimum Hamming distance: 0.1996007984031936
Maximum Hamming distance: 0.39687292082501663
```

Εικόνα 28 Mean Hamming

Τέλος εκτελώντας το plot_hamming_distance.py μας δίνει την γραφική όλων των hamming distance



Εικόνα 29 Γράφημα Hamming Distance κατανομών

4.3 Συμπεράσματα

Οι τιμές του Hamming distance (0.4995) για το inter distance και (0.4996) το inter challenge δείχνουν ότι το ESN παράγει διαφορετικές αποκρίσεις για διαφορετικά challenges και για διαφορετικά ESN. Αυτό είναι επιθυμητό, καθώς δείχνει ότι το ESN μπορεί να λειτουργήσει ως μία αξιόπιστη ΦΜΚΣ.

Η χαμηλότερη τιμή του Hamming distance (0.2638) για το intra reservoir δείχνει ότι το ESN έχει κάποια σταθερότητα στις αποκρίσεις του για το ίδιο challenge, αλλά υπάρχει και κάποιος θόρυβος. Μια χαμηλότερη τιμή είναι επιθυμητή για να εξασφαλιστεί η αξιοπιστία του PUF.

Παράρτημα I [NARMA]

Τα μοντέλα ARX, ARMAX χρησιμοποιούνται ευρέως για την πρόβλεψη χρονοσειρών. Τα Αυτοπαλινδρομικά μοντέλα με Κινούμενο Μέσο και Εξωγενή Είσοδο (ARMAX) και οι παραλλαγές τους AR, ARX, ARMA, είναι μια από τις πιο χρησιμοποιούμενες μαθηματικές αναπαραστάσεις για την αναγνώριση γραμμικών συστημάτων. Η οικογένεια μοντέλων ARX χρησιμοποιείται συνήθως για την μοντελοποίηση γραμμικών συστημάτων. Για μη γραμμικά σενάρια έχουμε την κλάση NARMAX. Το NARMAX ξεκίνησε ως όνομα μοντέλου, αλλά σύντομα έγινε φιλοσοφία όσον αφορά τον εντοπισμό μη γραμμικών συστημάτων. [19]

Το NARMAX έχει μελετηθεί εκτενώς στη θεωρία και σταδιακά υιοθετήθηκε σε εφαρμογές [5] [6].

Η απόκτηση μοντέλων NARMAX συνίσταται στην εκτέλεση των παρακάτω βημάτων:

- ✓ Δυναμική δοκιμή και συλλογή δεδομένων
- ✓ Επιλογή μαθηματικής αναπαράστασης
- ✓ Ανίχνευση της δομής του μοντέλου
- ✓ Εκτίμηση παραμέτρων
- ✓ Επικύρωση
- ✓ Ανάλυση του μοντέλου.

ΑΝΑΛΥΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

Όπως ήδη αναφέρθηκε τα μοντέλα NARMAX δεν είναι μια απλή επέκταση των μοντέλων ARMAX, είναι σε θέση να αντιπροσωπεύουν τα πιο διαφορετικά και πολύπλοκα μη γραμμικά συστήματα και περιγράφεται ως:

$$y_k = F^l \left[y_{k-1}, \dots, y_{k-n_y}, x_{k-d}, x_{k-d-1}, \dots, x_{k-d-n_x}, e_{k-1}, \dots, e_{k-n_e} \right] + e_k$$

όπου $n_x \in N$ $n_y \in N$ $n_e \in N$ είναι οι μέγιστες καθυστερήσεις για την έξοδο και την είσοδο του συστήματος αντίστοιχα, $x_{kx} \in R^{n_x}$ είναι η είσοδος του συστήματος και $y_{ky} \in R^{n_y}$ είναι η έξοδος στον διακριτό χρόνο $k \in N^n$, $e_k \in R^{n_e}$ δηλώνει τις αβεβαιότητες και πιθανό θόρυβο στον διακριτό χρόνο k . F^l είναι κάποια μη γραμμική συνάρτηση των συντελεστών παλινδρόμησης εισόδου και εξόδου με βαθμό μη γραμμικότητας $l \in N$ και d είναι μια χρονική καθυστέρηση που συνήθως ορίζεται σε $d=1$.

Εάν δεν συμπεριλάβουμε όρους θορύβου, e_{k-n_e} , έχουμε μοντέλα NARX. Αν ρυθμίσουμε $l=1$ τότε ασχολούμαστε με τα μοντέλα ARMAX. Αν $l=1$ και δεν συμπεριλαμβάνουμε όρους εισόδου και θορύβου, έχουμε μοντέλο AR (ARX αν συμπεριλάβουμε εισόδους, ARMA αν συμπεριλάβουμε όρους θορύβου), αν $l>1$ και δεν υπάρχουν όροι εισόδου, έχουμε το NARMA μοντέλο. Εάν δεν υπάρχουν όροι εισόδου ή θορύβου, έχουμε το NAR μοντέλο, γενικά υπάρχουν πολλές παραλλαγές.

ΠΑΡΟΥΣΙΑΣΗ NARMAX

Υπάρχουν πολλές μη γραμμικές αναπαραστάσεις συναρτήσεων για την προσέγγιση της άγνωστης αντιστοίχισης $f(\cdot)$ με μεθόδους NARMAX που βασίζονται σε:

- Νευρωνικά δίκτυα
- Μοντέλα βασισμένα σε ασαφή λογική
- Ακτινικές συναρτήσεις βάσης.
- Wavelet βάση
- Πολυωνυμική βάση
- Γενικευμένα προσθετικά μοντέλα

Τα δομημένα μοντέλα μπλοκ και πολλές αρχιτεκτονικές νευρωνικών δικτύων μπορούν όλα να θεωρηθούν ως υποσύνολα του μοντέλου NARMAX. Από τότε που εισήχθη το NARMAX, αποδεικνύοντας ποια κατηγορία μη γραμμικών συστημάτων μπορεί να αναπαρασταθεί από αυτό το μοντέλο, πολλά αποτελέσματα και αλγόριθμοι έχουν προκύψει με βάση αυτήν την περιγραφή. Το μεγαλύτερο μέρος της πρώιμης εργασίας βασίστηκε σε πολυωνυμικές επεκτάσεις του μοντέλου NARMAX. Αυτές εξακολουθούν να είναι οι πιο δημοφιλείς μέθοδοι σήμερα, αλλά άλλες πιο πολύπλοκες μορφές που βασίζονται σε wallets και άλλες επεκτάσεις (ορθολογικό μοντέλο - rational model) έχουν εισαχθεί για να αντιπροσωπεύουν σοβαρά μη γραμμικά και εξαιρετικά πολύπλοκα μη γραμμικά συστήματα.

ΠΟΛΥΩΝΥΜΑ NARMAX

Το πολυώνυμο NARMAX είναι ένα μαθηματικό μοντέλο που βασίζεται σε εξισώσεις διαφορών και συσχετίζει την τρέχουσα έξοδο ως συνάρτηση των προηγούμενων εισόδων και εξόδων.

Το πολυωνυμικό μοντέλο NARMAX με ασυμπτωτικά σταθερά σημεία ισορροπίας μπορεί να περιγραφεί ως:

$$\begin{aligned}
 y_k = & \sum_0^p + \sum_{i=1}^p \Theta_y^i y_{k-i} + \sum_{j=1}^q \Theta_e^j e_{k-j} + \sum_{m=1}^r \Theta_x^m x_{k-m} \\
 & + \sum_{i=1}^p \sum_{j=1}^q \Theta_{ye}^{ij} y_{k-i} e_{k-j} + \sum_{i=1}^p \sum_{m=1}^r \Theta_{yx}^{im} y_{k-i} x_{k-m} + \sum_{j=1}^q \sum_{m=1}^r \Theta_{ex}^{jm} e_{k-j} x_{k-m} \\
 & + \sum_{i=1}^p \sum_{j=1}^q \sum_{m=1}^r \Theta_{yex}^{ijm} y_{k-i} e_{k-j} x_{k-m} + \sum_{m_1=1}^r \sum_{m_2=m_1}^r \Theta_{x^2}^{m_1 m_2} x_{k-m_1} x_{k-m_2} \dots \\
 & + \sum_{m_1=1}^r \dots \sum_{m_l=m_{l-1}}^r \Theta_{x^l}^{m_1, \dots, m_l} x_{k-m_1} x_{k-m_l}
 \end{aligned}$$

όπου

$$\sum_0^p, c_y^i, c_e^j, c_x^m, c_{ye}^{ij}, c_{yx}^{im}, c_{ex}^{jm}, c_{yex}^{ijm}, c_{x^2}^{m_1 m_2} \dots c_{x^l}^{m_1, \dots, m_l}, \text{ είναι σταθεροί παράμετροι.}$$

Πιο κάτω παρατίθεται το μοντέλο NARMAX ως παράδειγμα μοντέλου για κατανόηση. Το παρακάτω είναι ένα μοντέλο NARMAX τάξης 2 που προσδιορίστηκε από πειραματικά δεδομένα ενός κινητήρα/γεννήτριας συνεχούς ρεύματος χωρίς προηγούμενη γνώση του μοντέλου [20]

$$y_k = 1.7813y_{k-1} - 0.7962y_{k-2} + 0.0339x_{k-1} - 0.1597x_{k-1}y_{k-1} + 0.0338x_{k-2} \\ + 0.1297x_{k-1}y_{k-2} - 0.1396x_{k-2}y_{k-1} + 0.1086x_{k-2}y_{k-2} \\ + 0.0085y_{k-2}^2 + 0.1938e_{k-1}e_{k-2}$$

Αρχικά η "δομή" ενός μοντέλου είναι το σύνολο των όρων (συντελεστές παλινδρόμησης) που περιλαμβάνονται στο τελικό μοντέλο. Οι παράμετροι είναι οι τιμές που πολλαπλασιάζουν κάθε έναν από αυτούς τους όρους. Στο παραπάνω παράδειγμα μπορούμε να παρατηρήσουμε ότι τα πολυωνμικά μοντέλα NARMAX δηλαδή έχουν μια μη γραμμική δομή, αλλά είναι γραμμικά στις παραμέτρους. Αυτή η παρατήρηση είναι πολύ σημαντική για την εκτίμηση των παραμέτρων.

Ας εξετάσουμε την περίπτωση όπου έχουμε τα δεδομένα εισόδου και εξόδου κάποιου συστήματος. Για λόγους απλότητας θεωρούμε μία είσοδο και μία έξοδο. Έχουμε τα δεδομένα χωρίς όμως να γνωρίζουμε ποιες καθυστερήσεις να επιλέξουμε για την είσοδο ή την έξοδο. Επιπλέον δεν γνωρίζουμε τίποτα για τη μη γραμμικότητα του συστήματος. Για αυτό τον λόγο απαιτείται να ορίσουμε κάποιες τιμές για τις μέγιστες καθυστερήσεις των όρων εισόδου, εξόδου και θορύβου. Αν επιλέξουμε $n_y = n_x = l = 2$ έχουμε τις εξής δυνατότητες για τη σύνθεση του τελικού μοντέλου:

$$\text{constant}, y_{k-1}, y_{k-2}, y_{k-1}^2, y_{k-2}^2, x_{k-1}, x_{k-2}, x_{k-1}^2, x_{k-2}^2, y_{k-1}y_{k-2},$$

$$y_{k-1}x_{k-1}, y_{k-1}x_{k-2}, y_{k-2}x_{k-1}, y_{k-2}x_{k-2}, x_{k-1}x_{k-2}.$$

Από τα παραπάνω φαίνεται ότι υπάρχουν 15 υποψήφιοι όροι για να συνθέσουμε το τελικό μοντέλο. Υπάρχει η επιλογή να χρησιμοποιηθούν όλοι οι όροι επειδή υπάρχουν μόνο 15. Αυτό μπορεί να οδηγήσει σε μια πολύ λανθασμένη αναπαράσταση του προς μοντελοποίηση συστήματος. Έχουμε την δυνατότητα να τρέξουμε έναν αλγόριθμο με σκοπό για να ελέγξουμε τους υποψήφιους συντελεστές παλινδρόμησης, ώστε να επιλέξουμε μόνο τους σημαντικούς. Σε αυτή την περίπτωση, έχουμε 32768 πιθανές δομές μοντέλων προς δοκιμή. [21]

NARMA 10 Time Series:

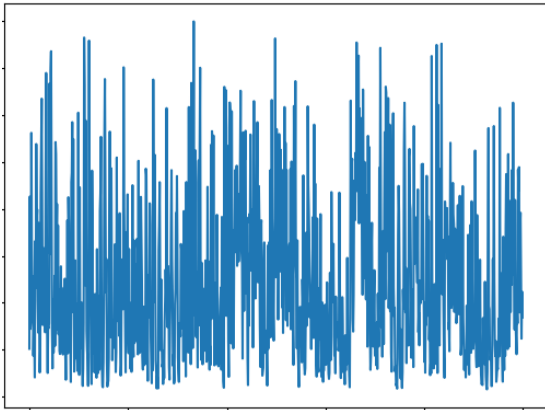
Ο μη γραμμικός αυτοπαλινδρομικός κινητός μέσος όρος (NARMA) είναι μια χρονική διεργασία διακριτού χρόνου με χρονική υστέρηση 10ης τάξης. Για να απλοποιηθεί ο συμβολισμός χρησιμοποιείται το y_t για να δηλώσουμε το $y(t)$. Η χρονοσειρά NARMA 10 δίνεται από:

$$y_t = \alpha y_{t-1} + \beta y_{t-1} \sum_{i=1}^n y_{t-i} + \gamma u_{t-n} u_{t-1} + \delta,$$

Όπου $n=10$, $\alpha=0.3$, $\beta=0.05$, $\gamma=1.5$, $\delta=0.1$. Η είσοδος u_t θεωρείται ομοιόμορφη κατανομή στο διάστημα $[0,0.5]$. Ο NARMA 10 χρησιμοποιείται σε μη γραμμικά προβλήματα που παρουσιάζουν εξάρτησή από μεγάλες χρονικές καθυστερήσεις. Χρησιμοποιείται συχνά για την αξιολόγηση της

χωρητικότητας της μνήμης και της υπολογιστικής ισχύος ενός Recurrent Neural Network – RNN και ειδικότερα του Echo State Network -ESN .

Το μοντέλο NARMA10 περιέχει μια ασταθή περιοχή στο χώρο των φάσεων που κάνει το σύστημα να αποκλίνει ανάλογα με την επιλογή του εύρους εισόδου και τις αρχικές συνθήκες. Επιπλέον, η ικανότητα επεξεργασίας πληροφοριών του μοντέλου ποικίλλει ανάλογα με το εύρος εισόδου. Αυτές οι ιδιότητες εμποδίζουν την ασφαλή εφαρμογή αυτού του μοντέλου και τις δίκαιες συγκρίσεις μεταξύ των πειραμάτων. [22]



Εικόνα 30 Κυματομορφή NARMA10

Παράρτημα II [ΒΑΣΙΚΟΣ ΚΩΔΙΚΑΣ ESN]

```
# -*- coding: utf-8 -*-
```

```
''''
```

```
Run the Esn For different train lengths  
output MSE as a text as Title  
and the weight matrix for each train length 800 - 12000 step 200  
''''
```

```
''''
```

```
A minimalistic Echo State Networks demo with Mackey-Glass (delay 17) data  
in "plain" scientific Python.  
from https://mantas.info/code/simple\_esn/  
(c) 2012-2020 Mantas Lukoševičius  
Distributed under MIT license https://opensource.org/licenses/MIT  
''''
```

```
import matplotlib  
import numpy as np  
import os  
import matplotlib.pyplot as plt  
from sklearn.metrics import mean_squared_error  
#matplotlib.use('Qt5Agg')  
from scipy import linalg
```

```
# numpy.linalg is also an option for even fewer dependencies
```

```
# load the data  
#trainLen = 3000 #20000  
testLen = 1000  
initLen = 100 #the first 100 steps would be used for solely for initializing the internal state  
data = np.loadtxt('NARMA_10')[3050:] #here you can load your data and next you can declare from  
which point to start  
##### Normalize data #####
```

```
# Define the desired range for normalization (0 to 1 in this case)
#min_value = 0.0
#max_value = 1.0

# Min-Max scaling
#normalized_data = (data - data.min()) / (data.max() - data.min()) * (max_value - min_value) +
min_value #na kano standarization adi gia scaling afero to meso oro kai diero me to standard
deviation
#to standard

#data = normalized_data
#####

# Standardize the input data
#mean = np.mean(data)
#std = np.std(data)
#data = (data - mean) / std
#####

# plot some of it
plt.figure(10).clear()
plt.plot(data[:1000])
plt.title('A sample of data')

# generate the ESN reservoir
inSize = outSize = 1
resSize = 322
a = 0.9991752770086721 # leaking rate
#In the context of this code, a is the leaking rate, and it determines the extent to which the previous
state of each
# neuron in the reservoir influences its current state. The leaking rate is a hyperparameter of the
ESN that controls the
# memory and dynamics of the reservoir. A smaller a results in longer memory, while a larger a
results in faster decay of
# past information in the reservoir states.
```

```
#seed_matrix = [3 , 14 ,18 ,52 ,42, 76, 23, 8, 92, 10]

np.random.seed( 3 )#42
Win = (np.random.rand(resSize, 1 + inSize) - 0.5) * 1
#Win = np.zeros((resSize, 1 + inSize))
W = np.random.rand(resSize, resSize) - 0.5
# normalizing and setting spectral radius (correct, slow):
print('Computing spectral radius...')
rhoW = max(abs(linalg.eig(W)[0]))
print('done.')
W *= 0.9234212477909265 / rhoW #1.25

for trainLen in range(800,12000, 200):
# allocated memory for the design (collected states) matrix
    X = np.zeros((1 + resSize, trainLen - initLen))
    #X = np.zeros((1 + inSize + resSize, trainLen - initLen))
    # set the corresponding target matrix directly
    Yt = data[None, initLen + 1:trainLen + 1]

    # run the reservoir with the data and collect X
    x = np.zeros((resSize, 1))
    for t in range(trainLen):
        u = data[t]
        x = (1 - a) * x + a * np.tanh(np.dot(Win, np.vstack((1, u))) + np.dot(W, x))
        #x = (1 - a) * x + a * np.tanh(np.dot(W, x))
        if t >= initLen:
            X[:, t - initLen] = np.vstack((1, x))[:, 0] ###1 u

    # train the output by ridge regression
    reg = 1e-8 # regularization coefficient
    # direct equations from texts:
    # X_T = X.T
    # Wout = np.dot( np.dot(Yt,X_T), linalg.inv( np.dot(X,X_T) + \
    #   reg*np.eye(1+inSize+resSize) ) )
    # using scipy.linalg.solve:
    Wout = linalg.solve(np.dot(X, X.T) + reg * np.eye(1 + resSize),
```

np.dot(X, Yt.T)).T

```
# run the trained ESN in a generative mode. no need to initialize here,
# because x is initialized with training data and we continue from there.
Y = np.zeros((outSize, testLen))
u = data[trainLen]
for t in range(testLen):
    x = (1 - a) * x + a * np.tanh(np.dot(Win, np.vstack((1, u))) + np.dot(W, x))
    y = np.dot(Wout, np.vstack((1, x)))
    Y[:, t] = y
    # generative mode:
    #u = y
    ## this would be a predictive mode:
    u = data[trainLen+t+1]

# compute MSE for the first errorLen time steps
errorLen = 100

mse = sum(np.square(data[trainLen + 1:trainLen + errorLen + 1] -
                    Y[0, 0:errorLen])) / errorLen
#mse = np.square
#y_true = data[trainLen + 1:trainLen + errorLen + 1]
#mse = mean_squared_error(data[trainLen + 1:trainLen + errorLen + 1] , Y[0, 0:errorLen])
#mean MSE
#mse = np.sqrt(mse)
#mse = mse / (np.max(y_true) - np.min(y_true))

print('MSE = ' + str(mse))

# plot some signals
# plt.figure(1).clear()
# plt.plot(data[trainLen + 1:trainLen + testLen + 1], 'g')
# plt.plot(Y.T, 'b')
# plt.title('Target and generated signals $y(n)$ starting at $n=0$')
# plt.legend(['Target signal', 'Free-running predicted signal'])
```

```
# plt.figure(2).clear()
# plt.plot(X[0:20, 0:1000].T)
# plt.title(r'Some reservoir activations $\mathbf{x}(n)$')

# plt.figure(3).clear()
# plt.bar(np.arange(1 + resSize), Wout[0].T)
# plt.title(r'Output weights $\mathbf{W}^{\{out\}}$')

# plt.show()
# Dynamically generate file names
# Create a folder to save files
output_folder = f'results_trainLen_{trainLen}'
os.makedirs(output_folder, exist_ok=True)

# Dynamically generate file names within the folder
weights_filename = os.path.join(output_folder, f'Wout_trainLen_{trainLen}.npy')
mse_filename = os.path.join(output_folder, f'MSE_trainLen_{trainLen}.txt')
mse_filename1 = os.path.join(output_folder, f'{mse}.txt')

# Save weights and MSE
np.save(weights_filename, Wout) # Save the Wout matrix in binary format
np.savetxt(mse_filename, np.array([mse]), delimiter=',') # Save the MSE in text format
np.savetxt(mse_filename1, np.array([mse]), delimiter=',') # Save the MSE in text format
```

5

Bibliography

- [1] D. E. H. Ulrich Ruhrmair, «PUFs at a Glance,» *This work was supported in part by C-FAR,*
- [2] A. B. a. R. S. C. Pranesh Santikellur, «Deep Learning based Model Building Attacks on Arbiter PUF Compositions,» 2019.
- [3] N. Karimi, J.-L. Danger και S. Guilley, «Impact of Aging on the Reliability of Delay PUFs.,» *Journal of Electronic Testing*, 2018.
- [4] Y. M. J. L. S.-k. W. J. Y. Ke-li Li, «Research and design of a high-security configurable RO-PUF based,» σε *10th International Conference of Information and Communication Technology (ICICT-2020)*, China, 2021.
- [5] K. S. D. & K. T. Shimizu, «Glitch PUF: extracting information from,» *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, αρ. E95.A(1), p. 223–233, 2012.
- [6] [Ηλεκτρονικό]. Available: <https://www.embedded.com/a-comparison-of-sram-vs-quantum-derived-semiconductor-pufs/>.
- [7] A. B. ,. P. R. a. C. M. Dimitris Dermanis, Photonic Physical Unclonable Function Based on Integrated Neuromorphic Devices, *JOURNAL OF LIGHTWAVE TECHNOLOGY*, VOL. 40, NOVEMBER 15 2022.
- [8] R. Maes, Physically Unclonable Function, 2013.
- [9] R. V. Maes, « Intra-and inter-distance in fuzzy commitment using physically unclonable functions,» 2007.
- [10] P. & K. T. A. Tuyls, «A new security primitive based on physical unclonable functions,» σε *International Workshop on Cryptographic Hardware and Embedded Systems (CHES)*, 2006.
- [11] J. Andrew Rukhin, A Statistical Test Suite for Random Number Generators for, *C O M P U T E R S E C U R I T Y*, 2010.
- [12] H. Jaeger, A tutorial on training recurrent neural networks, covering BPTT, RTRL, EKF and the "echo state network" approach, Fraunhofer Institute for Autonomous Intelligent Systems (AIS), 2003.
- [13] «simplilearn,» [Ηλεκτρονικό]. Available: <https://www.simplilearn.com/recursive-neural-network-in-deep-learning-article>.
- [14] M. Lukoševičius, A Practical Guide to Applying Echo State Networks, Bremen, 2012.
- [15] T. N. . a. H. M. Wolfgang Maass, «Real-time computing without stable states:a new framework for neural computation based on perturbations.,» *Neural Computation*, p. 14(11):2531–2560, 2002.

- [16] Y. C. . Z. B. . D. Devegowda, «Multi-objective optimization for rapid and robust optimal,» *Comput Geosci*, 2014.
- [17] ZhaoZhao Zhang, YingQin Zhu, XiaoHui Wang, Wen Yu, «Optimal echo state network parameters based on behavioural spaces,» *Neurocomputing*, 2022.
- [18] M. Lukoševičius. [Ηλεκτρονικό]. Available: https://mantas.info/code/simple_esn/.
- [19] «Nonlinear System Identification: NARMAX Methods in the Time,» *Frequency, and Spatio-Temporal Domains*, 2013.
- [20] [Ηλεκτρονικό]. Available: [7] <https://sysidentpy.org/landing-page/ch0-narmax-intro/#fnref:1>.
- [21] S. a. B. Chen, «Representations of nonlinear system: the NARMAX,» 1989.
- [22] S. Ma, «Short-wavelength reverberant wave systems for physical realization of reservoir computing,» 2022.
- [23] G. E. & D. S. Suh, « Physical unclonable functions for device authentication and secret key generation.,» σε *In Proceedings of the 44th annual Design Automation Conference*, 2007.
- [24] A. F. Y. L. a. J. D. R. Farhan Bin Tarik, «Robust optical physical unclonable function using,» 2020.
- [25] B. C. D. V. D. M. & D. S. Gassend, «Silicon physical random functions. In Cryptographic hardware and embedded systems,» Springer, Berlin, Heidelberg, 2002.
- [26] J. K. S. S. S. G. J. & T. Guajardo, «. FPGA intrinsic PUFs and their use for IP protection. In Cryptographic hardware and embedded systems,» Springer, Berlin, Heidelberg, 2007.
- [27] S. S. & D. Sakhare, «A Review—Hardware Security Using PUF (Physical Unclonable Function),» 2019.
- [28] I. P. a. N. Sklavos, «Physical Unclonable Functions (PUFs) Design Technologies: Advantages and Trade Offs,» pp. 427-442.
- [29] P. W. Y. Z. G. L. L. D. J. Z. Li Ni, «PI PUF: A Processor-Intrinsic PUF for IoT,,» *Computers and Electrical Engineering*, τόμ. 105, 2023.
- [30] [Ηλεκτρονικό]. Available: <https://community.fs.com/article/single-mode-cabling-cost-vs-multimode-cabling-cost.html>.
- [31] N. O., «Nonlinear System Identification: From Classical Approaches to Neural Networks,» 2001.

