Χαροκοπειο Πανεπιστήμιο
Τμημα Πληροφορικής και τηλεματικής

Μαυροπουλος Ανδρεας (217129)
11ο Εξάμηνο 6ο ετος

16/09/2021

Εργασία Διαχείρησης Δικτύων

Machine Learning and Software Defined Networks

Περιχόμενα

Τιτλος Εργασίας ……1
Περιεχόμενα……...2
Software Defined Networks……...3
Port based IP traffic classification..……..…..4
Payload based IP traffic classification..……...…5
Machine learning classification.……6
Using SDN Data for ML Classification.……7
Αρχιτεκτονική SDN Εφαρμογής.……....8
Deployment……9
Collected Data……...10
Data proccesing……11
Αποτελέσματα……...12
Machine Learning Approach for Predicting DdoS……....14
Machine Learning Algorithms used for Ddos Detection…...15
Αποτελέσματα Μοντέλων σε επιθέσεις……19
Machine Learning-Based Multipath Routing……....21
Machine Learning-based Multipath Routing Framework….23
Data Analytics and Training Management Module……24
Heuristic-Based Routing Module……....26
ML-Based Routing Module……...…26β
Future Aspects of SDN……....27
Συμπεράσματα……..……28
Αναφορές……...…30

Software Defined Networks (SDNs)

Τα Software Defined Networks (SDNs) παρέχουν έναν διαχωρισμό μεταξύ του control plane και του forwarding plane των δικτύων. Η εφαρμογή λογισμικού του επιπέδου ελέγχου και οι ενσωματωμένοι μηχανισμοί συλλογής δεδομένων του πρωτοκόλλου OpenFlow υπόσχονται να είναι εξαιρετικά εργαλεία για την εφαρμογή ελέγχου δικτύων μέσω Machine Learning.Το πρώτο βήμα προς αυτή την κατεύθυνση είναι να κατανοήσουμε τον τύπο των δεδομένων που μπορούν να συλλεχθούν σε SDNs και πώς μπορούν να αντληθούν πληροφορίες από αυτά τα δεδομένα.

Η ποσότητα των δεδομένων που ρέει μέσω τηλεπικοινωνιακών δικτύων και η πολυπλοκότητα των εφαρμογών που τα παράγουν συνεχίζουν να αυξάνονται.Η εξαγωγή γνώσεων από αυτά τα δεδομένα για την κατανόηση και την πρόβλεψη του αντικτύπου τους γίνεται όλο και πιο σημαντική για την αποτελεσματική διαχείριση του δικτύου.Το νέο πρότυπο των SDNs, ξεκλειδώνει νέες διαδρομές σε αυτόν τον τομέα με τους ενσωματωμένους μηχανισμούς συλλογής πληροφοριών και την ευελιξία και τηνπρογραμματικότητα που έχει φέρει στα δίκτυα. Το OpenFlow είναι το τυπικό API για τα SDNs. Καθορίζει τον τρόπο επικοινωνίας μεταξύ controller plane το οποίο εκτελείται σε λογισμικό και σε data plane υλικό.
Θα διευρευνήσουμε λοιπόν τι είδους γνώσεις είναι δυνατόν να πάρουμε από τις πληροφορίες που διατίθενται από το OpenFlow σχετικά με τα δεδομένα που ρέουν σε ένα δίκτυο SDN και με ποιους τρόπους μπορούν να χρησιμοποιηθούν αυτές οι πληροφορίες για τη διαχείριση / έλεγχο του δικτύου.

Η μελέτη αυτή αποτελεί το πρώτο βήμα προς μια εφαρμογή SDN για τον έλεγχο των δικτύων εταιρικών υπολογιστών με τη χρήση

μηχανικής μάθησης και επικεντρώνεται στην παρακολούθηση και την ταξινόμηση της κυκλοφορίας χρησιμοποιώντας δεδομένα που λαμβάνονται με το πρωτόκολλο OpenFlow.Υπήρξαν αρκετές προσεγγίσεις για να εντοπίσουν την κυκλοφορία που ρέει σε ένα δίκτυο.Παραδοσιακά τα transport layer port information και οι άμεση επιθεώρηση του packet payload έχουν χρησιμοποιηθεί για τον σκοπό αυτό.Ωστόσο, και οι δύο προσεγγίσεις έχουν αρκετές ελλείψεις και αλγόριθμοι τεχνητής νοημοσύνης , πιο συγκεκριμένα, μηχανικής μάθησης , έχουν εξεταστεί πιο πρόσφατα.Οι δυνατότητες που προκύπτουν από τη χρήση μηχανικής μάθησης είναι διάφορες οπως πχ: πρόβλεψη του όγκου της κυκλοφορίας,παρακολούθηση ασφαλείας και εντοπισμός επιθέσεων DoS και DdoS. Τα Software defined networks εχουν τη δυνατότητα να διευκολύνουν την εφαρμογή τέτοιων εφαρμογών σε δίκτυα μέσω των παρακάτω τεχνικών.

A) Port based IP traffic classification

Η ταξινόμηση βάσει θυρών ήταν μία από τις πιο χρησιμοποιούμενες τεχνικές στο παρελθόν και ήταν επιτυχής, σε κάποιο βαθμό, επειδή πολλές εφαρμογές χρησιμοποιούσαν σταθερούς αριθμούς θυρών που είχαν εκχωρηθεί από την Internet Assigned Numbers Authority (IANA). Παρόλλα αυτα με την πάροδο του χρόνου ορισμένοι περιορισμοί σε αυτήν την προσέγγιση έγιναν προφανείς. Εμφανίστηκε ένας μεγάλος αριθμός εφαρμογών που δεν έχουν καταχωρημένους αριθμούς θύρας και πολλές από αυτές χρησιμοποιούν δυναμικούς μηχανισμούς διαπραγμάτευσης θύρας για να κρυφτούν από τείχη προστασίας.Επίσης, η χρήση κρυπτογράφησης επιπέδου IP, συσκότισης και πληρεξουσίων μπορεί επίσης να αποκρύψει την κεφαλίδα TCP ή UDP, καθιστώντας αδύνατη την ανακάλυψη των αρχικών αριθμών θύρας.

B) Payload based IP traffic classification

Η προσέγγιση payload που είναι επίσης γνωστή ως Deep Packet Inspection (DPI), είναι σήμερα μία από τις πιο χρησιμοποιούμενες τεχνικές.Τα συστήματα DPI επιθεωρούν το payload των φορτίων και αναζητούν μοτίβα που προσδιορίζουν των τύπο της εφαρμογής.Αυτά τα μοτίβα, που ονομάζονται επίσης υπογραφές, ορίζονται από κανονικές εκφράσεις που αξιολογούνται από αυτοματοποιημένα συστήματα που λειτουργούν διαδοχικά και απαιτούν μεγάλη ποσότητα μνήμης.Αυτό μπορεί να προκαλέσει ένα πρόβλημα επεκτασιμότητας που είναι ακόμη πιο σοβαρό, καθώς το DPI εκτελείται στη διαδρομή επικοινωνίας.Για να μειωθεί το πρόβλημα της επεκτασιμότητας, έχουν αναπτυχθεί διάφορες τεχνικές κατά μήκος δύο διαφορετικών κατευθύνσεων. Το ένα είναι η τροποποίηση του αυτοματοποιημένου συστήματος όπως το Fast Finite Automata (FFA) ,το Deterministic finite Automata (DFA),Delayed input Automata και Differential encoding of Automata (DFAs). Το άλλο χρησιμοποιεί συγκεκριμένο υλικό, όπως Field Programmed Gate Array (FPGA) ή πιο πρόσφατα Graphics Processing Units (GPUs).Η χρήση των GPU για την επεξεργασία δεν είναι μια ασήμαντη εργασία, επειδή αυτές οι μονάδες έγιναν για Single Instruction Multiple Threads (SIMT) και η λειτουργία του αυτοματισμού είναι διαδοχική(sequential).Εκτός από αυτά τα προβλήματα, οι συσκευές που πρέπει να εισαχθούν στο δίκτυο έχουν υψηλό κόστος και αυξάνουν την πολυπλοκότητα του δικτύου.Το DPI μπορεί επίσης να είναι πολύ δύσκολο ή αδύνατο όταν ασχολείται με κρυπτογραφημένη κυκλοφορία που αυξάνεται.Τέλος, υπάρχουν ανησυχίες σχετικά με την προστασία της ιδιωτικής ζωής με την εξέταση των δεδομένων των χρηστών.

C) Machine learning classification

Τεχνικές μηχανικής μάθησης (ML) μπορούν επίσης να χρησιμοποιηθούν για την ταξινόμηση της κυκλοφορίας και έχουν προταθεί διάφορα έργα .Η ταξινόμηση της κυκλοφορίας μπορεί να πραγματοποιηθεί με τη χρήση supervised learning με αλγορίθμους όπως οι Support Vector Machines (SVM), neural networks και decision trees.Στο supervised learning υπάρχει ανάγκη απόκτησης χαρακτηρισμένων συνόλων δεδομένων εκπαίδευσης, κάτι που μπορεί να είναι δύσκολο στα δίκτυα υπολογιστών λόγω της δυσκολίας απόκτησης δειγμάτων ροής δικτύου με ακριβή σχολιασμό σε ένα ευρύ φάσμα εφαρμογών .Αυτό έχει ως ως αποτελέσματα πολλές από αυτές τις προσεγγίσεις να περιορίζονται σε χονδροειδείς ταξινομήσεις όπως web, P2p και VoiP εφαρμογές.
Μια εναλλακτική λύση είναι η χρήση unsupervised machine learning που τα δεδομένα που δίνονται στο εκπαιδευόμενο μοντέλο δεν έχουν σχεδιαστεί.Τα unsupervised machine learning μοντέλα χρησημοποιηούνται συνήθως για εργασίες ομαδοποίησης, όπου οι αλγόριθμοι ομαδοποιούν τα δεδομένα σε διαφορετικά συμπλέγματα σύμφωνα με ομοιότητες στις τιμές .Υπάρχουν διάφοροι αλγόριθμοι που μπορούν να χρησιμοποιηθούν σε μη επιτηρημένο ML και υπάρχει άφθονη βιβλιογραφία σχετικά με την εφαρμογή τους στην ταξινόμηση κυκλοφορίας.Προσεγγίσεις που βασίζονται σε K-Means , Self-Organizing Maps(SOM), και Density Based Spatial Clustering έχουν προταθεί για διάφορα σενάρια ταξινόμησης της κυκλοφορίας.Ένας συνδυασμός unsupervised και supervised μεθόδων μπορούν επίσης να χρησιμοποιηθούν τόσο με δεδομένα με ετικέτα όσο και χωρίς. Αυτή η προσέγγιση προσπαθεί να ξεπεράσει τις δυσκολίες στην απόκτηση δεδομένων με ετικέτα.Μπορεί να λειτουργήσει με ένα σύνολο δεδομένων όπου η πλειοψηφία των δεδομένων δεν έχει ετικέτα.

D) Using SDN Data for ML Classification

Οι προσεγγίσεις μέσω μηχανικής μάθησης στην ταξινόμηση της κυκλοφορίας εφαρμόζονται συνήθως σε δεδομένα που λαμβάνονται από ειδικά πρωτόκολλα συλλογής δεδομένων που βασίζονται στο πρότυπο Του Πρωτοκόλλου Διαδικτύου (IPFIX).Αυτές οι προσεγγίσεις χρησιμοποιούνται συνήθως σε ISP όπου οι δρομολογητές λειτουργούν ως σημεία μέτρησης για τη συλλογή δεδομένων για αποθήκευση και επεξεργασία.Το IPFIX συγκεντρώνει πακέτα σε ροές μέσω κανόνων συνάθροισης και παρέχει πληροφορίες ροής όπως η διάρκεια ροής, ο ρυθμός μετάδοσης bit και οι χρόνοι άφιξης μεταξύ πακέτων ροής.Στη συνέχεια, οι πληροφορίες αυτές υποβάλλονται σε επεξεργασία σε χωριστή ενότητα για την ταξινόμηση κυκλοφορίας βάσει ML.Το SDN παρέχει νέους τρόπους για την εκτέλεση αυτών των λειτουργιών. Η ανοικτή ροή έχει ενσωματωμένα στατιστικά στοιχεία ροής και οι κανόνες προώθησης ου είναι εγκατεστημένοι στα switches ταιριάζουν πακέτα σε ροές.πάρχουν επίσης πολλαπλά επίπεδα στα οποία μπορούν να συλλεχθούν αυτές οι πληροφορίες, από access switches όπου ο αριθμός των ροών είναι μικρότερος έως gateway switches όπου μπορεί να υπάρχει μεγάλος αριθμός ροών κυκλοφορίας.Μια άλλη διαφορά είναι ότι τα στατιστικά στοιχεία λαμβάνονται απευθείας στον ελεγκτή, όπου βρίσκεται το επίπεδο ελέγχου, ανοίγοντας τη δυνατότητα άμεσης χρήσης των πληροφοριών που έχουν ληφθεί.Υπάρχει επίσης μεγάλη ευελιξία στον τρόπο συλλογής των δεδομένων.Μπορούν να προσαρμοστούν να καθορίσουν διαφορετικούς τύπους ροής για διαφορετικούς τύπους κυκλοφορίας. Ενας προσαρμόσιμος αριθμός πακέτων μπορεί να αναλυθεί στον ελεγκτή πριν από την εγκατάσταση ενός κανόνα ροής στο διακόπτη για τη συλλογή των στατιστικών του.

Αρχιτεκτονική SDN Εφαρμογής

Αναπτύχθηκε μια εφαρμογή SDN που συλλέγει στατιστικά στοιχεία OpenFlow από τους ελεγχόμενους διακόπτες.Η εφαρμογή χρησιμοποιεί ενεργά το OpenFlow για να εγκαταστήσει ενα Flow Entry που δίνει εντολή στους ελεγχόμενους διακόπτες να προωθήσουν όλη την κυκλοφορία στον ελεγκτή.Όταν ένα πακέτο φτάσει στον ελεγκτή ο ελεγκτής το αναλύει για να προσδιορίσει το πρωτόκολλο μεταφοράς του.Για πακέτα TCP, αναλύονται οι σημαίες TCP για τον εντοπισμό των πακέτων που εμπλέκονται στην αρχική χειραψία TCP.Μετά τη χειραψία, τα πρώτα πέντε πακέτα που ανταλλάσσονται μεταξύ του υπολογιστή-πελάτη και του διακομιστή εξακολουθούν να λαμβάνονται στον ελεγκτή.Για το καθένα απο τα πακέτα αποθηκεύεται η σφραγίδα μεγέθους και η ώρας άφιξης.Στο πέμπτο πακέτο, ο ελεγκτής εγκαθιστά μια καταχώρηση ροής στο διακόπτη με ένα idle timeout και τα υπόλοιπα πακέτα που ανήκουν σε αυτήν τη σύνδεση υποβάλλονται σε επεξεργασία τοπικά στο διακόπτη μέχρι τη λήξη του idle timeout υποδεικνύοντας ετσι ότι τα πακέτα που ταιριάζουν με αυτήν τη ροή δεν φτάνουν πλέον στο διακόπτη.Εάν η εφαρμογή προορίζεται μόνο για παρακολούθηση κυκλοφορίας, η οδηγία επεξεργασίας πακέτων στις καταχωρήσεις ροής μπορεί να είναι για το διακόπτη να προωθήση τα πακέτα ως regular switch.Όταν μια καταχώρηση ροής καταργείται λόγω timeout τα στατιστικά στοιχεία της, δηλαδή ο αριθμός πακέτων και ο αριθμός byte, συλλέγονται στον ελεγκτή. Για εμάς η εφαρμογή χρησιμεύει μόνο ως εφαρμογή συλλογής δεδομένων και μπορεί να αναπτυχθεί με δύο διαφορετικούς τρόπους..Στην πρώτη περίπτωση, μόνο σε δίκτυα SDN, και στην δεύτερη περίπτωση, που μπορεί επίσης να εφαρμοστεί σε παλαιού τύπου δίκτυα εκτός SDN.έχουμε έναν μόνο

διακόπτη με δυνατότητα OpenFlow και χρησιμοποιούμε κατοπτρισμό θύρας για να λάβουμε ένα αντίγραφο της κυκλοφορίας του σημείου μέτρησης που παραδίδεται σε μια θύρα αυτού του διακόπτη.

Deployment

Για να αποκτήσουμε το σύνολο δεδομένων που μελετήθηκε προηγουμένως θα αναπτύξουμε εναν ενιαίο διακόπτη OpenFlow σε ένα δίκτυο παραγωγής επιχειρήσεων εκτός SDN.Ο διακόπτης OpenFlow λαμβάνει ένα αντίγραφο της κίνησης του συνδέσμου που συνδέει τον εποπτευόμενο χώρο εργασίας με τον κεντρικό δρομολογητή, αυτό το αντίγραφο παρέχεται μέσω κατοπτρισμού θύρας στο διακόπτη .Χρησιμοποιούμε έναν διακόπτη hp E3800 Openflow ενεργοποιημένο και τον ελεγκτή HP VAN SDN για την εκτέλεση της εφαρμογής παρακολούθησης κυκλοφορίας που χρησιμοποιεί το OpenFlow έκδοση 1.3 για τον έλεγχο του διακόπτη.Αυτή η ρύθμιση είναι πολύ ελαφριά και δεν παρεμβαίνει στην κανονική ροή της κυκλοφορίας στο ελεγχόμενο δίκτυο.Αποκτήσαμε δύο διαφορετικά σύνολα δεδομένων, ένα σύνολο δεδομένων χωρίς ετικέτα που περιέχει πληροφορίες από όλη την κυκλοφορία από το ελεγχόμενο δωμάτιο και ένα μικρότερο σύνολο δεδομένων που παράγεται υπό ελεγχόμενες συνθήκες και επισημαίνεται με τις αντίστοιχες εφαρμογές που δημιούργησαν τα δεδομένα.Τα δεδομένα με ετικέτα λαμβάνονται με απομόνωση της κυκλοφορίας ενός μόνο κεντρικού υπολογιστή στο διακόπτη. Ελέγχοντας τον κεντρικό υπολογιστή που παράγει την κυκλοφορία μπορούμε να αντιστοιχείσουμε τα δεδομένα με τις εφαρμογές.

Collected Data

Για όλες τις ροές TCP που εντοπίστηκαν στον ελεγκτή, η εφαρμογή SDN αποθηκεύει, σε ένα αρχείο CSV:το μέγεθος, τις χρονικές σημάνσεις και την ώρα άφιξης των πρώτων πέντε πακέτων,τις διευθύνσεις προέλευσης και προορισμού MAC και IP,της θύρες μεταφοράς προέλευσης και προορισμού,την διάρκεια, των αριθμό byte και των αριθμό πακέτων της σύνδεσης.

Οι ροές TCP εντοπίζονται ελέγχοντας τις σημαίες στο payload των Πακέτων σε OpenFlow που λαμβάνονται απο τον ελεγκτή.Το μέγεθος και οι χρονικές σημύσεις των πρώτων πέντε πακέτων λαμβάνονται επίσης απευθείας με την αποθήκευση του μεγέθους του payload Packet In οπως και επίσης η ώρας άφιξης των μηνυμάτων. Οι χρόνοι άφιξης των πρώτων πέντε πακέτων υπολογίζονται στη συνέχεια από αυτές τις πληροφορίες.Οι διευθύνσεις και οι θύρες λαμβάνονται επίσης μέσω αυτών των αρχικών μηνυμάτων Packet In και χρησιμοποιούνται για τη δημιουργία του match clause για το OpenFlow Flow Entry που είναι εγκατεστημένο στο διακόπτη.Ο αριθμός byte και ο αριθμός πακέτων λαμβάνονται μέσω των στατιστικών στοιχείων της καταχώρησης ροής.Τέλος, η διάρκεια λαμβάνεται αφαιρώντας τη χρονική σήμανση των αρχικών πακέτων από τη χρονική σήμανση του μηνύματος που αφαιρέθηκε από τη ροή που έλαβε ο ελεγκτής όταν λήξει το Flow Entry.

Παράδειγμα Πίνακα Αποθηκευμένων δυνατοτήτων για ροη TCP

Feature	Description
Packet Size (1 to 5)	Size of the first five packets payload in bytes
Packet Time stamp (1 to 5)	Arrival time of the first five packets in milliseconds (controller system time)

Data Processing

Τα δεδομένα που συλλέγονται έχουν πολύ διαφορετικά χαρακτηριστικά με τιμές σε διαφορετικές κλίμακες.Για να αποφευχθούν χαρακτηριστικά με τιμές μεγάλης κλίμακας που επικαλύπτουν τιμές μικρότερης κλίμακας που προκαλούν κακή απόδοση στους αλγόριθμους ML, τα δεδομένα που συλλέγονται αντιμετωπίζονται για να κλιμακωθούν και να κεντραρευτούν τις τιμές δυνατοτήτων.Αυτό γίνεται υπολογίζοντας τις τυπικές βαθμολογίες για κάθε δυνατότητα δεδομένων. Τα τυποποιημένα features έχουν περίπου zero mean και unit standard deviationε ξαλείφοντας έτσι την υψηλή μεταβλητότητα και τα την κλιμάκωση δεδομένων.

Ένα άλλο σημαντικό ζήτημα είναι η συσχέτιση μεταξύ των δυνατοτήτων. Τα χαρακτηριστικά που συσχετίζονται ιδιαίτερα είναι κάπως περιττά και προσθέτουν λίγα στην απόδοση των αλγορίθμων ταξινομητή.Η εξάλειψη τέτοιων περιττών χαρακτηριστικών είναι σημαντική, καθώς μειώνει το κόστος υπολογισμού διατηρώντας παράλληλα την ακρίβεια ταξινόμησης.Ο αλγόριθμος Principal Component Analysis (PCA) χρησιμοποιείται στο σύνολο δεδομένων για την εύρεση των κύριων στοιχείων.Αυτό θα μειώσει τον αριθμό των features σε περιπτώσεις όπου υπάρχουν συσχετιζόμενα δεδομένα.

Αποτελέσματα

Ο στόχος της εργασίας ήταν να ελεγχθεί η απόδοση των supervised classification αλγορίθμων σε τέτοια σενάρια.Για να γίνει αυτό, αποκτήσαμε ένα σύνολο δεδομένων με ετικέτα χρησιμοποιώντας έναν μόνο κεντρικό υπολογιστή για τη δημιουργία συνδέσεων TCP. Η κυκλοφορία που δημιουργείται από τον κεντρικό υπολογιστή αποστέλλεται στη συνέχεια μέσω κατοπτρισμού θύρας στο διακόπτη OpenFlow όπου η εφαρμογή SDN στον ελεγκτή εκτελεί τη συλλογή δεδομένων.Αυτό έχει ως αποτέλεσμα ένα σύνολο δεδομένων όπου είναι γνωστή η εφαρμογή που είναι υπεύθυνη για τις δυνατότητες που μετρώνται σε κάθε σύνδεση TCP.Σύλλεξαν πάνω απο 500 feature set samples για καθε μια απο της ακόλουθες εφαρμογές: Youtube, Vimeo, Facebook, Linkedin, Skype, Bittorrent,Web Browsing(HTTP) και Dropbox.

Για να ελεγχθεί η ακρίβεια ταξινόμησης, το σύνολο δεδομένων με ετικέτα πρέπει να διαιρείται σε σύνολα εκπαίδευσης και δοκιμών.Οι επαναλαμβανόμενες περιπτώσεις αγνοούνται αφήνοντας μόνο τις μοναδικές που αποτελούν το σύνολο εκπαίδευσης.Οι υπόλοιπες τιμές χρησιμοποιούνται ως σύνολο δοκιμών.Δοκιμάστηκαν ensemble learning classifiers,οι ταξινομητές χρησιμοποιούν ένα σύνολο απο regular classifiers και ταξινομούν δεδομένα λαμβάνοντας μια ζυγισμένη ψήφο από κάθε μεμονωμένη πρόβλεψη.Αποδίδουν καλύτερα από κάθε έναν από τους μεμονωμένους αλγόριθμους ταξινόμησης που

χρησιμοποιουνται εως τώρα. Πιο συγκεκριμένα δοκιμάστηκαν οι εξής αλγόριθμοι:Random Forests (RF), gradient boosting classifiers Stochastic Gradient Boosting καιxtreme Gradient Boosting (EGB).Οι ταξινομητές εκπαιδεύονται χρησιμοποιώντας το σύνολο εκπαίδευσης και αξιολογούνται με βάση το σύνολο δοκιμών. Επαναλαμβάνουμε τη διαδικασία εκπαίδευσης και δοκιμών 30 φορές για κάθε αλγόριθμο και υπολογίζουμε τη συνολική ακρίβεια του ταξινομητή ως τον μέσο όρο των τιμών ακρίβειας σε κάθε δοκιμή.Η ακρίβεια μετράται συγκρίνοντας τις ετικέτες που αποδίδονται από τον ταξινομητή με τις πραγματικές ετικέτες που συλλέχθηκαν στην πειραματική εγκατάσταση που εξηγήθηκε προηγουμένως. Τα αποτελέσματα ήταν παρόμοια σε όλους τους δοκιμασμένους αλγορίθμους που δείχνουν ότι τα δεδομένα που λαμβάνονται μέσω μηχανισμών SDN είναι κατάλληλα για ML supervised traffic classification.

A Machine Learning Approach for Predicting DdoS

Το παράδειγμα των Software Defined Networks (SDN) εισήχθη για να ξεπεραστούν οι περιορισμοί του παραδοσιακού δικτύου, όπως εξαρτήσεις προμηθευτών, πολιτικές ασυνέπειας κ.λπ.Γίνεται μια πολλά υποσχόμενη αρχιτεκτονική δικτύου που παρέχει στους φορείς εκμετάλλευσης μεγαλύτερο έλεγχο της υποδομής του δικτύου. Ο ελεγκτής που καλείτε επίσης το λειτουργικό σύστημα του SDN έχει τον κεντρικό έλεγχο του δικτύου. Παρά τις δυνατότητές της, η εισαγωγή διαφόρων αρχιτεκτονικών οντοτήτων αποτελεί πολλές απειλές για την ασφάλεια των επιπέδων SDN.Μεταξύ πολλών τέτοιων ζητημάτων ασφάλειας, η Distributed Denial of Services (DDoS) επίθεση είναι μια ταχέως αναπτυσσόμενη επίθεση που αποτελεί τεράστια απειλή για το SDN.Στοχεύει στη διαθεσιμότητα του δικτύου, πλημμυρίζοντας τον ελεγκτή με πλαστά πακέτα.Προκαλεί την παράλυση του ελεγκτή, και έτσι ολόκληρο το δίκτυο αποσταθεροποιείται.Ως εκ τούτου, είναι σημαντικό να σχεδιαστεί έναν ισχυρό μηχανισμό ανίχνευσης DDoS για να αποτραπεί η επίθεση. Θα δούμε πέντε τεχνικές μηχανικής μάθησης για να ταξινομήσουμε και να προβλέψουμε με ακρίβεια διαφορετικές επιθέσεις DDoS όπως: Smurf, UDP flood και HTTP flood.
Ο στόχος των Software Defined Networks (SDN)είναι να αποσυνδέσουν το επίπεδο ελέγχου από το επίπεδο προώθησης επιπλέον, αυτή η αρχιτεκτονική επιτρέπει πιο

ευέλικτη διαχείριση δικτύου στον φορέα εκμετάλλευσης δικτύου. Όλη η απόφαση δρομολόγησης και ο μηχανισμός ελέγχου ελέγχονται από μια κεντρική συσκευή που ονομάζεται ελεγκτής.Ο ελεγκτής στέλνει εντολή στις συσκευές προώθησης, όπως ο δρομολογητής και ο διακόπτης για τη διαχείριση των πακέτων δεδομένων.Στο SDN, το επίπεδο ελέγχου μπορεί να αποτελείται από έναν ή περισσότερους ελεγκτές ανάλογα με το μέγεθος και τη χρήση του δικτύου.Στο επίπεδο ελέγχου, ο ελεγκτής παρέχει κατανεμημένες πληροφορίες πολιτικής

Machine Learning Algorithms used for Ddos Detection

Υπάρχουν δύο βασικού τύποι τεχνικών μηχανικής μάθησης που χρησιμοποιούνται οι εποπτευόμενη μάθηση (supervised learning) και οι αλγόριθμοι μάθησης χωρίς επίβλεψη (unsupervised learning).Στους supervised learning αλγορίθμους κάθε δεδομένα εισόδου σχετίζονται με μια τάξη που ονομάζεται ετικέτα.Κατά τη διάρκεια της δοκιμής το μηχάνημα προβλέπει την κατηγορία των δεδομένων εισόδου με βάση το δείγμα εκπαίδευσης.Αυτό ονομάζεται εποπτευόμενο επειδή γνωρίζουμε την τάξη του δείγματος κατάρτισης κατά τη διάρκεια της φάσης εκμάθησης του μηχανήματος και η έξοδος του αλγορίθμου είναι οι εκπαιδευμένες τάξεις Οι εξης ML αλγόριθμοι χρησιμοποιήθηκαν:

• k-Nearest Neighbor (kNN):

Είναι ένας αλγόριθμος ταξινόμησης μη παραμετρικών και lazy learning classification αλγόριθμος. Ο όρος lazy learning υποδεικνύει ότι δεν κάνει γενίκευση χρησιμοποιώντας τα δεδομένα εκπαίδευσης. ιδέα πίσω από το kNN είναι να εντοπιστούν δείγματα k στο εκπαιδευτικό σύνολο των οποίων οι ανεξάρτητες μεταβλητές (x) σχετίζονται με νέα δείγματα (u).Στη συνέχεια, χρησιμοποιήστε αυτά τα δείγματα k για να ταξινομήσετε το νέο δείγμα σε μια κατηγορία (v).Όταν συζητάμε για τους γείτονες σημαίνει ότι υπάρχει ένα μέτρο απόστασης (ανομοιογένειας) που μπορεί να υπολογιστεί μεταξύ ανεξάρτητων μεταβλητών.

• Naïve Bayes (NB):

Αυτός ο αλγόριθμος βασίζεται στο θεώρημα του Bayes.Υποθέτει ότι η παρουσία ενός χαρακτηριστικού σε μια κλάση δεν έχει καμία σχέση με τα άλλα χαρακτηριστικά που υπάρχουν στην κατηγορία.Είναι εύκολο να κατασκευαστεί το μοντέλο Naïve Bayes και είναι χρήσιμο σε μεγάλα σύνολα δεδομένων.Αποδίδει καλά από τις εξαιρετικά εξελιγμένες τεχνικές ταξινόμησης. Μπορεί να εφαρμοστεί σε συμπεράνοντα στατιστικά στοιχεία και προβλήματα λήψης αποφάσεων που ασχολούνται με το συμπέρασμα των πιθανοτήτων

• Support Vector Machine (SVM):
Το SVM θεωρείται ως ο ταξινομητής με την υψηλότερη ακρίβεια στον τομέα του ML.Είναι ένα σύνολο σχετικών εποπτευόμενων μεθόδων μάθησης που χρησιμοποιούνται για την ταξινόμηση δεδομένων.Δεδομένου ενός συνόλου δειγμάτων κατάρτισης, κάθε δείγμα επισημαίνεται ως διαφορετικές κατηγορίες.Ένας αλγόριθμος SVM αναπτύσσει ένα μοντέλο που προβλέπει εάν ένα νέο δείγμα εμπίπτει σε μία από τις κατηγορίες. Ο αλγόριθμος βασίζεται στην εύρεση του υπερπλάνου που δίνει τη μέγιστη απόσταση διαχωρισμού μεταξύ των δειγμάτων εκπαίδευσης χρησιμοποιώντας την ακόλουθη λειτουργία.

• Random Forest (RF):
Οι εν λόγω ταξινομητές ενώνουν διαφορετικά δέντρα αποφάσεων για να προβλέψουν νέες πληροφορίες χωρίς ετικέτα, κάθε δέντρο αποφάσεων είναι διαθέσιμο στο δάσος και η ποιότητά του υπόκειται στην ποσότητα των δέντρων στο δάσος.Για κάθε δέντρο επιλέγονται τυχαία χαρακτηριστικά, κάθε αριθμός δέντρων μιλάει σε ένα μόνο δάσος και κάθε δάσος είναι μια κατηγορία πρόβλεψης για νέες πληροφορίες χωρίς ετικέτα.Σε αυτόν τον αλγόριθμο, γίνεται τυχαία επιλογή δυνατοτήτων για κάθε μεμονωμένο δέντρο. Στη συνέχεια, ένας αλγόριθμος μάθησης και συλλογής χρησιμοποιείται για την ταξινόμηση και την πρόβλεψη των εκροών υπό το φως ενός μεμονωμένου αριθμού δέντρων.Χρησιμοποιώντας αυτή τη στρατηγική,

δημιουργούνται πολυάριθμα δέντρα ταξινόμησης και κάθε ελεύθερο δέντρο χτίζεται από ένα εναλλακτικό κομμάτι του γενικού συνόλου δεδομένων.Αφού κάθε δέντρο ταξινομηθεί σε μια τάξη χωρίς ετικέτα μια άλλη ερώτηση θα πραγματοποιείται κάτω από κάθε δέντρο το οποίο στην συνέχεια θα "ψηφίζει" υπέρ της επιλογής του.Το δάσος που επιλέγεται ως νικητής εξαρτάται από τον πιο αξιοσημείωτο αριθμό ψήφων που έχουν καταγραφεί.

• Linear Regression (LR):

Το Linear Regression ρησιμοποιείται συνήθως στην προγνωστική ανάλυση. Είναι ένα μοντέλο που προϋποθέτει μια γραμμική σχέση μεταξύ της εισόδου και μιας μεταβλητής εξόδου.Συγκεκριμένα, η έξοδος μπορεί να υπολογιστεί από έναν γραμμικό συνδυασμό των μεταβλητών εισόδου.Για μία μόνο μεταβλητή εισόδου, το LR αναφέρεται ως απλή γραμμική ανάκρουση, ενώ σε περίπτωση πολλαπλών μεταβλητών εισόδου, αναφέρεται ως πολλαπλή γραμμική παλινδρόμηση.Στο LR παίρνουμε την έξοδο της γραμμικής λειτουργίας και στερεώνουμε την τιμή εντός του εύρους [0, 1] χρησιμοποιώντας τη σιγμοειδή λειτουργία.

Αποτελέσματα Μοντέλων σε επιθέσεις

Η απόδοση του ταξινομητή έχει αξιολογηθεί με βάση τους κύριους δείκτες απόδοσης με βάση τα confusion matrix, Accuracy,Precision και Recall.

- Confusion Matrix:Ενας Confusion matrix είναι ένας πίνακας N x N που χρησιμοποιείται για την αξιολόγηση των επιδόσεων ενός μοντέλου ταξινόμησης, όπου N είναι ο αριθμός των κλάσεων-στόχων.Ο πίνακας συγκρίνει τις πραγματικές τιμές-στόχους με αυτές που προβλέπονται από το μοντέλο μηχανικής μάθησης.
- Accuracy: Η ακρίβεια είναι μια μέτρηση για την αξιολόγηση μοντέλων ταξινόμησης.Ανεπίσημα, η ακρίβεια είναι το κλάσμα των προβλέψεων που έκανε σωστά το μοντέλο μας.

- Precision:Ειναι παρόμοιο με την ακρίβεια μονο που αλλαζει ο τύπος της σε: True Positives / (True Positives + False Positives)

- Recall: Η ανάκληση προσπαθεί να απαντήση στην ακόλουθη ερώτηση:Ποιο ποσοστό των πραγματικών θετικών στοιχείων εντοπίστηκε σωστά;

Μεταξύ των πέντε αλγορίθμων ταξινόμησης ο LR πέτυχε υψηλή ακρίβεια,υψηλό precision και υψηλό recall ενώ ο NB παρουσιάζει τα χειρότερα αποτέλεσματα. Τα NB

παρουσιάζουν χαμηλό αποτέλεσμα για την επίθεση Στρουμφ. Απο την αλλη πλευρά ομως το LR και το SVM πέτυχαν υψηλό ρυθμό ακρίβειας σε ολλα τα άλλα.Επιπλέον οι RF παίρνουν λιγότερο χρόνο σε σύγκριση με τα LR.Σε γενικές γραμμές, μπορεί να αποδοθεί ότι υπάρχει συμβιβασμός μεταξύ της ακρίβειας πρόβλεψης και του συνολικού χρόνου εκτέλεσης, ο οποίος θα πρέπει να λαμβάνεται υπόψη κατά την επιλογή μιας τεχνικής ML για τον εντοπισμό κυκλοφορίας DDoS στο SDN.Αξιοποιώντας το αποτέλεσμα του ταξινομητή, μπορούμε να ορίσουμε διαφορετικούς κανόνες ασφαλείας στον ελεγκτή για να ελέγξουμε τους πιθανούς εισβολείς αποκλείοντας ένα υποδίκτυο του δικτύου. Εχουμε χρησιμοποιήσει προσεγγίσεις ML για να προβλέψουμε την επίθεση DdoS σε ενα δίκτυο SDN. Αξιοποιώντας τη χρήση διαφορετικών αλγορίθμων ML, οι κανόνες ασφαλείας που ορίζονται από τον ελεγκτή μπορούν να ελέγξουν την κακόβουλη επίθεση.Τα πειραματικά αποτελέσματα έδειξαν ότι η κατάλληλη επιλογή αλγορίθμων ML θα μπορούσε να βοηθήσει στην ακριβή πρόβλεψη της επίθεσης και στον καθορισμό των κανόνων ασφαλείας.Η μέση ακρίβεια πρόβλεψης που επιτυγχάνεται από το LR είναι 98.652%, πράγμα που σημαίνει ότι αυτός ο ταξινομητής μπορεί να προβλέψει με ακρίβεια την κακόβουλη κυκλοφορία.Από την άλλη, το RF πέτυχε 98.409% με λιγότερο χρόνο εκτέλεσης από το LR. Θα πρέπει δηλαδή να συνεχιστούν οι δοκιμασίες περεταίρω μέχρι να βγεί καταλητικό αποτέλεσμα.

Machine Learning-Based Multipath Routing

Η κίνηση του δικτύου έχει αυξηθεί εκθετικά λόγω της ταχείας ανάπτυξης εφαρμογών τεχνολογίας πληροφοριών και επικοινωνιών, όπως οι υπηρεσίες cloud services, τα μέσα κοινωνικής δικτύωσης, τα έξυπνα τηλέφωνα, το Διαδίκτυο των Πραγμάτων (IoT) και οι εφαρμογές διαδικτύου.Η Cisco προβλέπει τριπλάσια αύξηση της κυκλοφορίας IP από το 2017 έως το 2022.Οι παραδοσιακές λύσεις δρομολόγησης και κυκλοφοριακής μηχανικής θα πρέπει να επανεξεταστούν για να ανταποκριθούν, σε πραγματικό χρόνο, σε αυτήν την εκρηκτική αύξηση της κυκλοφορίας.Ωστόσο, η άκαμπτη αρχιτεκτονική των παλαιών δικτύων δεν επιτρέπει την καινοτόμο βελτιστοποίηση του δικτύου λόγω της στενής σύζευξης μεταξύ του επιπέδου δεδομένων και του επιπέδου ελέγχου.Εναλλακτικά, η δικτύωση που ορίζεται από το λογισμικό (SDN) έχει πρόσφατα αναδειχθεί σε αρχιτεκτονική δικτύου που επιτρέπει τη δυναμική διαχείριση δικτύου και τη λειτουργικότητα του προγράμματος, και ως εκ τούτου τη βελτιστοποίηση της απόδοσης. Το SDN διευκολύνει τη δημιουργία πιο καινοτόμων προγραμματιστικων λύσεων ελέγχου δικτύου και δρομολόγησης με βάση μια παγκόσμια προβολή του δικτύου sta-tus και τον λεπτό λεπτομερή έλεγχο της κυκλοφορίας δικτύου και των πόρων δικτύου.Οι πρόσφατες εξελίξεις στους αλγόριθμους βαθιάς μάθησης και η πρόσφατη εμφάνιση του SDN, έχουν γεφυρώσει το χάσμα μεταξύ μηχανικής μάθησης (ML) και δικτύωσης.Επιπλέον, ο κεντρικός λογικός έλεγχος μετατοπίζει τη νοημοσύνη από το

επίπεδο υποδομής στο επίπεδο ελέγχου όπου μπορούν να χρησιμοποιηθούν αποτελεσματικές τεχνολογίες πληροφορικής όπως η μονάδα επεξεργασίας γραφικών (GPU) και, ως εκ τούτου, παρέχουν την απαραίτητη υπολογιστική ισχύ για τη νοημοσύνη της μηχανής.Επιπλέον, οι δυνατότητες επιθεώρησης πακέτων των στοιχείων προώθησης και η σφαιρική προβολή της κατάστασης δικτύου που είναι διαθέσιμη στον κεντρικό ελεγκτή (CC) διευκολύνουν την εκπαίδευση σε πραγματικό χρόνο των αλγόλιθων ML που καθοδηγούνται από δεδομένα.Επιπλέον, η προγραμματισιμότητα και η ευελιξία του δικτύου στην εγκατάσταση νέων κανόνων προώθησης πακέτων επιτρέπει την εκτέλεση αποτελεσμάτων δρομολόγησης σε πραγματικό χρόνο που υπολογίζονται από αλγορίθμους ML.Αρκετές πρόσφατες προσπάθειες επικεντρώθηκαν στην εφαρμογή τεχνικών ML για τη δρομολόγηση της κυκλοφορίας στο δίκτυο. Κάθε ενότητα ML εφαρμόζει έναν εποπτευόμενο αλγόριθμο εκμάθησης για την επίλυση του προβλήματος δρομολόγησης μεταξύ ενός ζεύγους κόμβων στο δίκτυο.Οι ενότητες εκπαιδεύονται με βάση ένα καθορισμένο σύνολο δεδομένων που είναι καθιερωμένα απο :set of the network topology, QoS requirements, και optimal integral paths που δημιουργούνται απο εναν ευριστικό αλγόριθμο.
Σε αντίθεση με το multipath routing, το integral routing δεν επιτρέπει τη διαίρεση της κυκλοφορίας σε πολλές διαδρομές. Οι εκπαιδευμένες ενότητες ML μπορούν να δώσουν ευρετικές λύσεις δρομολόγησης σε πραγματικό χρόνο για

κάθε ζεύγος κόμβων

Machine Learning-based Multipath Routing Framework

θα δώσουμε μια επτομερή περιγραφή του προτεινόμενου πλαισίου MLMR.Ο κύριος στόχος του πλαισίου MLMR είναι η πραγματοποίηση ευρετικής δρομολόγησης πολλαπλών διαδρομών και πολλαπλών διαδρομών σε πολλά χρονικά πλαίσια σε δίκτυα που ορίζονται από λογισμικό με σαγηνευμένους συνδέσμους και πλευρικούς περιορισμούς απόδοσης.Παρόλο που το πλαίσιο MLMR αντικαθιστά τους παραδοσιακούς αλγορίθμους που επιλύουν το πρόβλημα MCNF με πλευρικούς περιορισμούς, εξαρτούνται πό τέτοιους αλγόριθμους για τη δημιουργία ενός συνόλου δεδομένων με ετικέτα για την εκπαίδευση του DNN(Deep Neural Network).Ως εκ τούτου, η βέλτιστη λύση που προβλέπεται από το πλαίσιο MLMR είναι συγκρίσιμη με τη λύση που υπολογίζεται με την ευρετική μέθοδο.Ωστόσο, το πλαίσιο MLMR είναι σημαντικά πιο υπολογιστικά αποδοτικό από τους παραδοσιακούς αλγορίθμους.
Το πλαίσιο MLMR μπορεί να εφαρμοστεί στο επίπεδο εφαρμογής της αρχιτεκτονικής SDN. Τα northbound και southbound APIs θεωρείται ότι εφαρμόζονται για να διευκολύνουν την ανταλλαγή πληροφοριών μεταξύ του προτεινόμενου πλαισίου, cc και των στοιχείων που προορίζονται για την παρακολούθηση.Τα στοιχεία προώθησης στο επίπεδο υποδομής αναφέρουν μήτρες

κυκλοφορίας στον ελεγκτή, το οποίο εξαρτάται από το προτεινόμενο πλαίσιο MLMR για τον υπολογισμό της λύσης δρομολόγησης πολλαπλών διαδρομών.Στη συνέχεια, ο ελεγκτής υλοποιεί αυτήν τη λύση δρομολόγησης στους πίνακες προώθησης των στοιχείων προώθησης.Αυτή η διαδικασία επαναλαμβάνεται είτε όταν ένα πακέτο μη διαδρομολογημένης ροής φτάσει σε οποιοδήποτε από τα στοιχεία προώθησης είτε εντοπιστεί μια σημαντική αλλαγή στην κατάσταση δικτύου.

Η αρχιτεκτονική MLMR αποτελείται από τις ακόλουθες πέντε ενότητες:

• data acquisition and results reporting module;
• data analytics and training management module;
• data warehouse module;
• heuristic-based routing module;
• ML-based routing module.
Θα αναπτύξουμε περεταίρω 3 σημεία.

1) **Data Analytics and Training Management Module**
Διαφορετικές τεχνικές ανάλυσης δεδομένων μπορούν να χρησιμοποιηθούν σε αυτή την ενότητα για την ταξινόμηση της κυκλοφορίας και τη ομαδοποίηση .Δύο εργασίες εκτελούνται σε αυτήν την ενότητα, προκειμένου να βελτιωθεί η ποιότητα και η ακρίβεια των λύσεων υπολογιστικής δρομολόγησης.Πρώτον, ο προσδιορισμός του κατάλληλου

συνόλου δεδομένων εκπαίδευσης από την αποθήκη δεδομένων για την κατάρτιση του DNN της μονάδας δρομολόγησης.Επειδή το προτεινόμενο πλαίσιο αποθηκεύει μετρήσεις κυκλοφορίας στη βάση δεδομένων NoSQL, διευκολύνει την εφαρμογή της ανάλυσης Big Data για τον προσδιορισμό του συνόλου δεδομένων εκπαίδευσης για την εκπαίδευση του DNN.Ο σχεδιασμός τέτοιων αναλυτικών στοιχείων Big Data είναι πέρα από το πεδίο εφαρμογής αυτού του έργου. Δεύτερον, ενεργοποίηση της εκπαιδευτικής διαδικασίας του DNN όταν εντοπίζεται σημαντική αλλαγή στην κατάσταση του δικτύου ή στις απαιτήσεις κυκλοφορίας.Η ενότητα αναλύει, σε απευθείας σύνδεση, τις πληροφορίες που προστίθενται στην αποθήκη δεδομένων για να ανιχνεύσει την ανομοιογένεια στις μήτρες κυκλοφορίας χρησιμοποιώντας τις στατιστικές τεχνικές όπως η principle component analysis (PCA), και τα Hurst exponent estimation methods. Η σημασία της αλλαγής της κατάστασης του δικτύου καθορίζεται από τη μείωση του κινούμενου μέσου όρου των εκθεμάτων Hurst για ένα παράθυρο μετρήσεων κυκλοφορίας, κάτω από ένα προκαθορισμένο όριο. Η σημασία της αλλαγής της κατάστασης του δικτύου καθορίζεται από τη μείωση του κινούμενου μέσου όρου των εκθεμάτων Hurst για ένα παράθυρο μετρήσεων κυκλοφορίας, κάτω από ένα προκαθορισμένο όριο.Επιπλέον, η διαδικασία κατάρτισης του DNN μπορεί να ενεργοποιηθεί εάν αναφερθεί αποτυχία σύνδεσης.

2) Heuristic-Based Routing Module

Δεδομένης της παγκόσμιας άποψης του δικτύου και των traffic matrices ενας ευριστικός αλγόριθμος μπορεί να χρησιμοποιηθεί για την επίλυση του προβλήματος MCNF με πλευρικούς περιορισμούς.Ο αλγόριθμος υπολογίζει ένα σύνολο διαδρομών απο traffic demands δρομολογημένων μέσω των paths για να βγάλει αποτέλεσματα. Επιπλέον, τα αποτελέσματα αυτά επιστρέφονται στην ενότητα απόκτησης δεδομένων και αναφοράς αποτελεσμάτων, η οποία τα διαβιβάζει στην CC.Μέχρι τη δημιουργία αυτών των αποτελεσμάτων, η κατάσταση του δικτύου ενδέχεται να έχει αλλάξει, γεγονός που επιδεινώνει την ποιότητα τέτοιων λύσεων.Ωστόσο, τα αποτελέσματα αυτά εφαρμόζονται προσωρινά μόνο κατά την προετοιμασία και την κατάρτιση του DNN.

3) ML-Based Routing Module

Το σύνολο δεδομένων που είναι αποθηκευμένο στη μονάδα αποθήκης δεδομένων περιέχει μια ακολουθία απο time stamped traffic matrices Εισάγουμε έναν εκθέτη για να υποδηλώσει το χρονικό ευρετήριο στο οποίο λαμβάνεται αυτός ο πίνακας κυκλοφορίας.Το μέγεθος του πίνακα κυκλοφορίας είναι $N \times N$. Μπορεί να αναδιαμορφωθεί σε διάνυσμα του μεγέθους $1 \times N^2$ στοιβάζοντας τις μεταθέσεις των γραμμών.Το σύνολο δεδομένων περιέχει επίσης τη λύση

δρομολόγησης που υπολογίζεται από τη μονάδα δρομολόγησης που βασίζεται σε ευριστικό αλγόριθμο και αντιστοιχεί σε έναν συγκεκριμένο πίνακα κυκλοφορίας.

Το ML-based routing module δημιουργεί μια ακολουθία tuples των διανυσμάτων που αντιπροσωπεύουν τον πίνακα κυκλοφορίας.Στην εποπτευόμενη μάθηση, το DNN μαθαίνει μια λειτουργία M που χαρτογραφεί καλύτερα την είσοδο στην έξοδο με βάση ένα σύνολο δεδομένων εκπαίδευσης με ετικέτα.Μόλις ο ML αλγόριθμος μάθει την λειτουργία αντιστοίχισης M, η λειτουργική θα μπορεί να βγάζει εκτιμήσης σε real-time.

Συμπέρασμα

Τα Software Defined Networks έιναι μια αναδυόμενη τεχνολογία που έχει φέρει την καινοτομία στη δικτύωση, με την αποσύνδεση του control plane και του data plane , αφαιρώντας το ιδιόκτητο κομμάτι της αρχιτεκτονικής δικτυων κάνωντας το ετσι ενα ανοιχτό και προγραμματιζόμενο δίκτυο.Λόγω του πολυάριθμου πλεονεκτήματος αυτής της αρχιτεκτονικής, πολλές εταιρείες μετατοπίζονται από την παραδοσιακή αρχιτεκτονική δικτύου στη νέα αρχιτεκτονική SDN.Ωστόσο, το SDN ως νέα τεχνολογία έχει προκύψει ζητήματα που αποτελούν πρόκληση για το μέλλον της τεχνολογίας.Η ασφάλεια είναι ένα από τα κύρια θέματα που απειλεί το μέλλον της

τεχνολογίας SDN. Για αυτόν τον λόγο λοιπόν παρουσιάσαμε πως η χρήση Machine Learning μπορεί ακομά και σε μεγάλο βαθμό να αντιμετωπίσει αυτα τα προβλήματα ασφάλειας με γρήγορο και ακριβή τρόπο. Το πρόβλημα που παρουσιάζεται ομως με τους Machine Learning αλγορίθμους , είναι ο μεγάλος όγκος δεδομένων που απαιτούνται οπως και η υπολογιστική δύναμη. Η υπολογιστική δύναμη που απαιτήται για την εκμάθηση τέτοιων μοντέλων είναι επεκτατική και ακριβή, και για αυτον τον λόγο δεν μπορεί ο καθένας να τα "αντιγράψει" στο σπίτι του. Ίσως με την πάροδο του χρόνου η υπολογιστική δύναμη και τα δεδομένα που απαιτούνται για να φτιάξει οποιοσδήποτε ενα Machine Learning αλγορίθμο σε ενα scale δικτύου, γίνουν πιο προσίτα.

Future Aspects of SDN

Εδώ και πολλά χρόνια, βλέπουμε τον κλάδο των τηλεπικοινωνιών να αλλάζει ραγδαία.Με την εξέλιξη της τεχνολογίας 4G, εξακολουθούσε να υπάρχει ανάγκη για ένα πιο ευέλικτο, πιο γρήγορο και ευέλικτο δίκτυο, το οποίο θα αποτελεί τη βάση της τεχνολογίας 5G.Η τεχνολογία 5G δεν έχει ακόμη αξιοποιηθεί πλήρως και θα διαδραματίσει ζωτικό ρόλο στο μέλλον.Στο μέλλον, το 5G είναι βέβαιο ότι θα βασιστεί στο SDN.Αυτή η κίνηση θα καταστήσει τα δίκτυα

5G σε ανοιχτού κώδικα κλάδο των τηλεπικοινωνιών.Η παροχή lower latency και ανταγωνιστικής τιμής τεχνολογίας 5G με συνδυασμό SDN θα βοηθήσει στη βελτίωση της εμπειρίας των πελατών σε πολλές πλατφόρμες.Το SDN βρίσκεται ακόμα σε πρώιμο στάδιο, καθώς πολύ λίγοι άνθρωποι είναι εκπαιδευμένοι γι 'αυτό.Θα κερδίσει ακόμη μεγαλύτερη δημοτικότητα καθώς όλο και περισσότερο εργατικό δυναμικό εκπαιδεύεται γι 'αυτό.Όντας μία από τις τεχνολογίες που θα σχεδιαστούν για τον έλεγχο της δικτύωσης, το μέλλον του SDN έγκειται στην τρέχουσα κατάστασή του.Το SDN εξακολουθεί να είναι μια αναδυόμενη τεχνολογία και πρέπει να ξεπεράσει τις προκλήσεις, ώστε να μην παραβιαστεί από κάποιον. Με την ανάπτυξη εφαρμογών Machine Learning και Neural Network μοντέλων μπορεί να καταφέρουμε να δούμε ενα αδιαπέραστο δίκτυο SDN

Αναφορές:

1.Machine Learning in Software Defined Networks: Data Collection and Traffic Classification (Pedro Amaral,João Dinis,Paulo Pinto,Luis Bernardo,João Tavares,Henrique S. Mamede)

2.A Machine Learning Approach for Predicting DDoS Traffic in Software Defined Networks (Kshira Sagar Sahoo,Prasenjit Maiti,Amaan Iqbal,Bibhudatta Sahoo)

3. Machine Learning Based Intrusion Detection System for Software Defined Networks (Atiku Abubakar , Bernardi Pranggono)

4.Machine Learning-Based Multipath Routing for Software Defined Networks (Mohamad Khattar Awad Imtiaz, Ahmad Marwa Hassan Hafez Ahmed, Ali F. Almutairi)