“AI/ML based monitoring of security logs/alerts for intelligent threat management in modern networks”

Καμπάνης Αλέξανδρος Ιάσονας
Κωνσταντίνος Τσαγκάρης

Background Μηχανική Μάθηση

Υποκλάδος της Τεχνητής Νοημοσύνης.

Σχετίζεται με πολλούς ερευνητικούς κλάδους.

Στόχος είναι να εκπαιδευτεί ένας αλγόριθμος από ιστορικά δεδομένα και να γενικεύσει την γνώση που απέκτησε.

Οι κύριες κατηγορίες είναι supervised και unsupervised αλλά δεν είναι απόλυτο.

Στόχος και των δύο η εξαγωγή γνώσης.

Βρίσκει εφαρμογή σε πολλές περιπτώσεις

Βαθιά μηχανική μάθηση. Υποκλάδος

Background Μηχανική Μάθηση

Αλγόριθμοι ML/DL SVC,Naive Bayes,Random Forest, Decision Tree,CNN,RNN,Autoencoders, GANs,GNNs

Βήματα που ακολουθούνται Get/Prepare, Exploration, Model, Communicate

Για την επιλογή αλγορίθμου υπάρχουν πολλά κριτήρια όπως ταχύτητα εκπαίδευσης, ταχύτητα inference, είδος δεδομένων και άλλα.

Μετρικές αξιολόγησης. Εδώ πρέπει να επιλέξουμε την κατάλληλη μετρική σε σχέση με το πρόβλημα που λύνουμε και τι μας ενδιαφέρει να πετύχουμε.

Συνηθισμένα προβλήματα είναι κάποια όπως under fitting, over fitting, class imbalance

Background Ασφάλεια

Ένας πολύ γενικός ορισμός για την ασφάλεια είναι ότι στόχο έχει να προστατέψει το Confidentiality, Availability, Integrity.

Υπάρχουν πολλές βασικές ένοιες που αφορούν την ασφάλεια και περιγράφουμε στην εργασία.

Είδη επιθέσεων και πως κατηγοριοποιούνται. Πολλά taxonomies με διαφορετικά κρίτήρια

Το πιο σύνηθες ποια αρχή παραβιάζουν.

Εδώ κρατάμε την κατηγοριοποίηση με βάση τις επιπτώσεις καθώς ταιριάζει και με το dataset που έχουμε στην διάθεση μας

Resource and event montitoring. Βασική προ απαιτούμενη διαδικασία για να πετύχουμε τους στόχους

Background Complex event processing

Πρόκειται για κλάδο που επίσης σχετίζεται με πολλούς άλλους κλάδους.

Σκοπός είναι παρακολουθώντας πολλά event streams να αναγνωρίζει πολύπλοκες καταστάσεις που παρουσιάζονται.

Αυτές οι καταστάσεις ονομάζονται complex events.

Διαφορές με απλά repositories είναι ότι εδώ έχουμε στατικά queries και δυναμικά δεδομένα αντί για δυναμικά queries και στατικά δεδομένα.

Πολλά τα requirements. Τα βασικότερα είναι high throughtput, low latency και complex computations.

Πολλοί τύποι γλωσσών. Πιο διαδεδομένες αυτές που μοιάζουν με SQL

Background Ανίχνευση Απειλών

Είναι ο στόχος αυτής της εργασίας

Πολλά τα εργαλεία που χρησιμοποιούνται για κάθε περίπτωση και συνδυαστικά όπως IDS/Firewall/antivirus

Σε μεγάλης κλίμακας πληροφοριακά συστήματα όμως υπάρχουν ιδιατερότητες.

Συνήθως τα συστήματα αυτά χρησιμοποιούν λογισμικό που προσφέρει περισσότερη λειτουργικότητα

Βασικές τεχνολογίες που χρησιμοποιούνται είναι τα SIEM, EDR, NDR, XDR

Βασικές διαφορές στο από που αντλούν πληροφορία και επιπλέον λειτουργίες από ανίχνευση απειλών.

Background IDS

Πρόκειται για διαδεδομένη τεχνολογία που αποτελεί υποσύνολο των προηγουμένων.

Υπάρχουν πολλοί τρόποι να κατηγοριοποιηθούν αλλά οι πιο κοινοί είναι Network IDS, Host IDS και Hybrid IDS. Αυτό το taxonomy βασίζεται στο που μπορούν να βρεθούν τα δεδομένα.

Με βάση τις κατηγορίες προκύπτουν υποκατηγορίες όπως packet, flow, session analysis log based detection ML Hybrid rules και άλλα.

Επίσης μπορούν να κατηγοριοποιηθούν με βάση τον τρόπο που ανιχνέυουν την επίθεση. Misuse Detection(signatures) και anomaly detection(baseline συμπεριφοράς).

Απαιτήσεις για Destruction Resistance, Incident Recording, Adaptability, Scalability,

Σχετικές εργασίες

Πολλές σχετικές εργασίες.

Συνήθως κατηγοριοποιούνται με βάση την τεχνολογία που χρησιμοποιούν για το μοντέλο.

Όλες μένουν στην δημιουργία ενός μοντέλου και δεν ασχολούνται με όλη την διαδικασία και τα προβλήματα που προκύπτουν

Ξεχωρίζουμε μία εργασία η οποία πραγματοποιεί feature selection και πετυχαίνει state of the art επίδοση με μόνο δέκα features.

Μεθοδολογία

Όλες οι εργασίες απλά δημιουργούν ένα μοντέλο ML/DL

Η δική μας προσέγγιση είναι να αναδείξουμε και να προσπαθήσουμε να αντιμετωπίσουμε και όλα τα προβλήματα πέρα από το μοντέλο.

Χρήση κατάλληλων τεχνολογιών για να επιτευχθεί αυτό

Αλγόριθμοι ML, μηχανισμός Logging, Complex Event Proccessing

Στόχος να ικανοποιηθούν τα requirements για το IDS

Για το είδος του IDS μας οδηγούν τα δεδομένα

Μεθοδολογία, Βήματα

Το πρώτο βήμα είναι η δημιουργία μηχανισμού καταγραφής

Δεύτερο βήμα δημιουργία μηχανισμού συσχέτισης

Τρίτο βήμα δημιουργία μηχανισμού προ επεξεργασίας

Τέταρτο βήμα δημιουργία μηχανισμού inference

Πέμπτο βήμα Deployment

Μεθοδολογία Επιλογή Λύσεων

Για τον μηχανισμό καταγραφής επιλέγουμε την χρήση wrapper functions λόγω ετερογένειας στα logs και άλλων δυσκολιών. Επίσης θα μπορούσε να αξιοποιηθεί ο μηχανισμός syslog.

Για τον μηχανισμό συσχέτισης επιλέγουμε Complex Event Processing λόγω λειτουργικότητας και optimizations.

Ο μηχανισμός αυτός θα πραγματοποιεί λειτουργίες που πρέπει να γίνουν πριν το inference

Ο μηχανισμός inference θα βασίζεται σε shallow ML μοντέλο

Υλοποίηση Dataset

NSK-KDD dataset 20 ετών αποτελείται από 100000 εγγραφές.

Τα features κατατάσονται στις κατηγορίες Intrinsic, Content, Time based, Host based

Intrinsic Αυτά τα features αφορούν πληροοφορία από τις κεφαλίδες του πακέτου. Duration, bytes sent etc

Content. Αυτά αφορούν πληροφορία γενικά από την σύνδεση. Hot login su attempted etc

Time Based. Αφορά πληροφορία συσχετισμένη με τον χρόνο(2sec). Αριθμός συνδέσεων προς το ίδιο host etc

Host Based. Αφορά πληροφορία που αφορά ποσοστά

Υλοποίηση Dataset

Τα features χωρίζονται με βάση τον τύπο τους σε categorical, numeric και binary

Τα categorical είναι protocol type, Flag και service. Το protocol type παίρνει τρεις τιμές το Flag παίρνει δέκα και το service σαράντα.

Binary. Εδώ έχουμε έξι features που έχουν τιμή 0 ή 1. Αυτά είναι τα Land, logged_in, root_shell, su_attempted, is_host_login, is_guest_login.

Numeric. Όλα τα υπόλοιπα από τα 41 είναι αριθμητικά features

Υλοποίηση Dataset

Τα είδη επιθέσεων είναι πολλά αλλά κατατάσσονται σε 4 κατηγορίες. Αυτές είναι DoS, Probe, U2R, R2L

DoS. Αυτές οι επιθέσεις στόχο έχουν το availability.

Probe. Αυτές οι επιθέσεις στόχο έχουν την συλλογή πληροφοριών.

U2R. Αυτές οι επιθέσεις στόχο έχουν να αποκτήσουν δυνατότητες υπερ χρήστη από απλό χρήστη.

R2L. Αυτές οι επιθέσεις στόχο έχουν να αποκτήσει ο επιτιθέμενος πρόσβαση από απομακρυσμένη τοποθεσία.

Υλοποίηση Δημιουργία μοντέλου

Απαίτηση για απλό μοντέλο και χαμηλό latency. Λειτουργία πραγματικού χρόνου.

Απαίτηση για explainability. Πρέπει να μπορούμε να δικαιολογήσουμε τις αποφάσεις. Με βάση αυτά επιλέγουμε Decision Tree. Επιπλέον λόγος επιλέγουμε shallow μοντέλο διότι έχουμε καλά δομημένα δεδομένα.

Class imbalance. Επιλέγουμε και δοκιμάζουμε τρόπους για να αντιμετωπιστεί. SMOTE, Resampling, class weights.

Δημιουργία dataset. Επιλογή 1000 samples ανά κλάση όπου είναι εφικτό

Υλοποίηση Δημιουργία μοντέλου

Το πρώτο στάδιο data preprocessing

Normalization όπου χρειάζεται.

Αντιμετώπιση categorical features

Για το feature protocol one hot

Για το feature flag buisness logic

Για το feature service hash trick

Στην συνέχεια δοκιμάζουμε τρία σενάρια με βάρη,χωρίς βάρη, με smote. Τα αποτελέσματα συγκρίσιμα

Υλοποίηση Δημιουργία μοντέλου

Δεύτερο στάδιο είναι το feature selection.

Με βάση τις απαιτήσεις πρέπει να μειώσουμε την διάσταση των δεδομένων.

Δύο τρόποι Decission Tree και correlation Matrix

Διαφορετικά αποτελέσματα.

Κρατάμε τα 10 καλύτερα με βάση το correlation matrix

Εδώ θα μπορούσαν να γίνουν παραπάνω πειράματα.

Μας ενδιαφέρει ιδιαίτερα το recall της minority class U2R εκτός από την μέση απόδοση

Υλοποίηση Δημιουργία μοντέλου

Στην συνέχεια έχουμε hp tunning. Το κάνουμε με χρήση του GridSearch και δοκιμάζουμε βασικές υπερπαραμέτρους.

Επίσης δοκιμάζουμε και άλλους αλγορίθμους SVC Naive Bayes, MLP, Radom Forest, KNN

Η απόδοση όλων εκτός του Random Forest είναι πολύ χειρότερη από τον Decission Tree.

Τέλος πραγματοποιήσαμε ένα πείραμα για να βρούμε τον χρόνο του inference. Όλο το dataset σε πολύ λιγότερο από το 1/10 sec

Υλοποίηση Features Retrieval

Τα features που προέκυψαν από την προηγούμενη φάση είναι τα dest_bytes, src_bytes, land, is hot login, duration, num failed logins, su attempted, wrong fragment ,srv_Diff_host_rate, srv_count.

Για τα περισσότερα μπορούμε να τα βρούμε παρακολουθώντας logs.

Τα δύο τελευταία έχουν σχέση με πληροφορία που αφορά όλο το δίκτυο.

Με βάση αυτά το IDS που θα δημιουργήσουμε κατηγοριοποιείται ώς hybrid.

Τα features αυτά θεωρούμε ότι τα δίνει κάθε proccess που παρακολουθείται.

Υλοποίηση Event Corelation

Εδώ στόχος και χρησιμοποιώντας τα streams δεδομένων να ανιχνεύσουμε τα complex Events.

Χρήση Esper JAVA/.NET Framework

Πολλοί τρόποι που μπορεί να γίνει η υλοποίηση καθώς παρέχεται πλούσια λειτουργικότητα

Παρουσιάζουμε πιθανές επιλογές και δικαιολογούμε αυτήν που εφαρμόσαμε.

Στόχος να εκμεταλλευτούμε στο έπακρο την λειτουργικότητα.

Υλοποίηση Event Corelation

- Τα βήματα για την τελική λύση είναι
- Ορισμός των simple events. Πολλοί τρόποι POJOS, XML, AVRO etc
- Για την δημιουργία του query διαλέγουμε την λειτουργικότητα context και output clause.
- Το context παρέχει λειτουργικότητα παρόμοια με sessions. Για την δική μας περίπτωση κάθε session αναφέρεται σε κάθε process που παρακολουθούμε. Είδος parition είναι key segmented.
- Το output clause χρησιμοποιεί joins τα οποία καταγράφουν πληροφορία αλλά ενεργοποιούνται όταν συμβεί ένα γεγονός
- Strongly typed για subscribers αποτελεί περιορισμό

Συμπεράσματα

- Βλέπουμε ότι το use case που αναλύουμε είναι πολυδιάστατο και δεν τελειώνει με την εκπαίδευση ενός μοντέλου ML.
- Υπάρχουν αυστηρά requirements που πρέπει να ικανοποιηθούν σε όλο το pipeline της διαδικασίας αλλά και για κάθε component.
- Η χρήση της Μηχανικής μάθησης βελτιώνει πολύ την διαδικασία αλλά έχει και αδυναμίες.
- Το σύστημα που αναπτύσσεται εξαρτάται από τα δεδομένα που υπάρχουν.
- Τα δεδομένα αυτά δεν είναι πάντα διαθέσιμα

Συμπεράσματα

- Κάποια από τα ζητήματα που πρέπει να αντιμετωπιστούν είναι
- Ανάγκη για ασφάλεια του IDS. Εδώ πρέπει το σύστημα να τρέχει σε ασφαλές και απομονωμένο περιβάλλον και τα logs να αποθηκεύονται σε ασφαλές μέρος.
- Ασφάλεια μοντέλου. Το μοντέλο μπορεί να γίνει στόχος επίθεσης.
- Ανάγκη για authentication. Τα processes που παράγουν δεδομένα θα πρέπει να αυθεντικοποιούνται. Ένα κακόβουλο process θα μπορούσε να αχρηστεύσει το σύστημα.
- Ανάγκη για πλήρη και επικαιροποιημένα datasets

Ερωτήσεις