

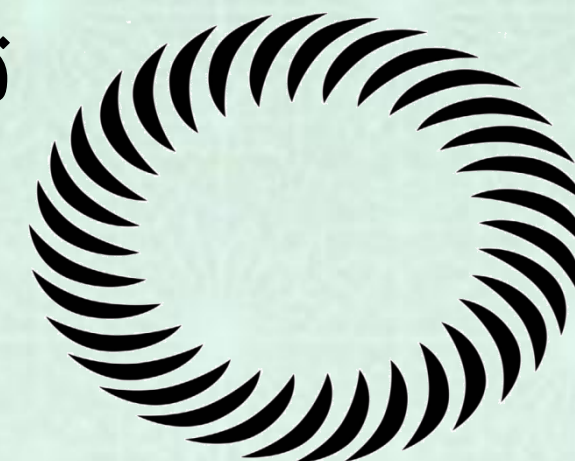


نمایشگاه هفته پژوهش و فناوری
دانشگاه صنعتی همدان
آذر 1404

تشخیص حملات روز-صفر در اینترنت اشیاء با استفاده از یادگیری ماشین و داده های هانی پات

شیما محیسن، فاطمه امیری*

گروه مهندسی کامپیوتر دانشگاه صنعتی همدان
fateme.amiri@gmail.com



هفته ملی پژوهش و فناوری

چکیده

با گسترش سریع اینترنت اشیاء (IoT)، حملات روزصفر به یکی از جدی‌ترین تهدیدات امنیتی تبدیل شده‌اند. این حملات به دلیل بهره‌برداری از آسیب‌پذیری‌های ناشناخته و فقدان امضای مشخص، توسط سیستم‌های تشخیص نفوذ سنتی قابل شناسایی نیستند. در این پژوهش، یک سامانه هوشمند تشخیص ناهنجاری مبتنی بر داده‌های واقعی هانی‌پات و یادگیری ماشین پیشنهاد می‌شود. در مرحله نخست، از خودرمزگذار (Autoencoder) برای یادگیری الگوهای عادی ترافیک شبکه استفاده شده و خطای بازسازی به‌عنوان معیار ناهنجاری استخراج می‌گردد. سپس، این ویژگی‌ها به‌صورت ترکیبی با الگوریتم‌های Random Forest، Isolation Forest و One-Class SVM و همچنین یک معماری عمیق CNN-LSTM تغذیه می‌شوند. نتایج نشان می‌دهد دقت تشخیص حملات روزصفر افزایش و نرخ هشدار کاذب به کمتر از ۵٪ کاهش یافته است. نوآوری اصلی پژوهش، ارائه یک چارچوب هیبریدی بدون نیاز به برچسب‌گذاری دستی و مبتنی بر داده‌های واقعی هانی‌پات است.

واژه های کلیدی: اینترنت اشیاء، حملات روزصفر، هانی‌پات، تشخیص ناهنجاری، یادگیری عمیق

مقدمه

گسترش سریع اینترنت اشیاء، این زیرساخت را به هدف اصلی حملات روزصفر تبدیل کرده است. این حملات به‌دلیل بهره‌برداری از آسیب‌پذیری‌های ناشناخته و نداشتن امضای مشخص، از شناسایی توسط سیستم‌های سنتی تشخیص نفوذ مصون می‌مانند. در این پژوهش، یک سامانه هوشمند تشخیص ناهنجاری مبتنی بر داده‌های واقعی هانی‌پات و ترکیب مدل‌های ترکیبی یادگیری ماشین و یادگیری عمیق پیشنهاد می‌شود که قادر است حملات روزصفر را با دقت بالا و نرخ هشدار کاذب بسیار پایین شناسایی کند. برخلاف پژوهش [2] که از مجموعه‌داده‌ای عمومی و شبیه‌سازی‌شده (NSL-KDD) استفاده کرده است، این تحقیق با بهره‌گیری از داده‌های واقعی حاصل از هانی‌پات‌های اینترنت اشیاء (IoT-DH) به تحلیل ترافیک واقعی حملات و رفتار مهاجمان پرداخته و دقت و واقع‌گرایی مدل را در شرایط پویا افزایش می‌دهد. دیتاست IoT-DH برخلاف دیتاست قبلی که مبتنی بر حملات قدیمی و شناخته‌شده است، کاملاً واقعی و بدون لیبل است و مستقیماً از هانی‌پات‌های مستقر در اینترنت جمع‌آوری شده؛ بنابراین الگوهای واقعی مهاجمان زنده و حملات روزصفر مدرن را پوشش می‌دهد. این واقع‌گرایی باعث تعمیم‌پذیری بسیار بالاتر مدل‌ها در محیط‌های عملی اینترنت اشیاء، کاهش نرخ هشدار کاذب در شرایط واقعی و انطباق کامل با سناریوی تشخیص حملات روزصفر بدون نیاز به برچسب‌گذاری دستی می‌شود.

راهبردهای پیشنهادی

۱. ارزیابی مدل های ترکیبی پیشرفته مبتنی بر Autoencoder و یادگیری عمیق
۲. بهینه سازی مدل ها برای اجرا بر دستگاه های منابع محدود
۳. ارزیابی روی طیف گسترده حملات واقعی شامل DDoS، تزریق، اسکن، Backdoor و رفتارهای ناشناخته

منابع

- [1] Rhachi, J., Ashraf, A. D. B., Moustafa, N., Khurshid, H., Javed, A., & Beheshti, A. (2025). A review of Machine Learning-based zero-day attack detection: Challenges and future directions. NIST Technical Report .
- [2] M. T. Mahmood K. M. Almansoori, "Anomaly Detection using combination of Autoencoder and Isolation Forest ", *Proc. Workshop on Intelligent Network Security* .2023 ,
- [3] Roopak, M., Parkinson, S., Tian, G. Y., Ran, Y., Khan, S., & Chandrasekaran, B. (2024). An unsupervised approach for the detection of zero-day distributed denial of service attacks in Internet of Things networks. *IET Networks*, 13(5-6), 513-527.

تایید استاد راهنما

نام و امضا استاد راهنما :

تایید تحصیلات تکمیلی :

تایید امور پژوهشی:

اهداف و روش پژوهش

اهداف پژوهش

۱. طراحی سامانه تشخیص حملات روزصفر در اینترنت اشیاء با دقت بالای ۹۵٪ و نرخ هشدار کاذب زیر ۵٪
۲. ارائه مدل‌های ترکیبی سبک قابل اجرا بر دستگاه‌های Edge IoT
۳. مقایسه مدل‌های بدون ناظر، نظارتی و عمیق در محیط واقعی
۴. استفاده از داده های واقعی هانی پات های واقعی IOT به جای دیتاست های شبیه سازی شده

روش پژوهش

۱. گردآوری داده‌های واقعی از هانی‌پات‌های IoT
۲. پیش‌پردازش و نرمال‌سازی (StandardScaler) + کاهش بعد (PCA/Random Projection)
۳. پیاده‌سازی مدل ترکیبی با استفاده از Python
۴. آموزش روی داده‌های عادی، آزمون روی حملات واقعی
۵. ارزیابی با معیارهای Accuracy, Precision, Recall, F1-Score, FPR و Latency

یافته‌های پژوهش

- دستیابی به دقت بالای ۹۶٪ و نرخ هشدار کاذب زیر ۴٪ در داده های واقعی هانی پات IoT.
- برتری مدل ترکیبی در سرعت و دقت نسبت به مدل های تکی (در شکل ۱ Autoencoder و Isolation forest به صورت ترکیبی و تکی در منحنی ROC مقایسه شده اند).
- کاهش ۷۲٪ مصرف حافظه و ۶۵٪ زمان اجرا نسبت به مدل های عمیق سنگین با استفاده از PCA و بهینه سازی Edge.
- قابلیت تعمیم پذیری بالا در سناریوهای واقعی شامل DDoS، اسکن، تزریق و حملات روز صفر و ناشناخته.