ارائه مدل چندوجهی MAGNET برای تشخیص بدافزار اندروید با استفاده از شبکه‌های عصبی ترنسفورمر و گراف

علیرضا ایرانمنش[۱]، دکتر حمید میروزیری[۲]
[۱]دانشجوی کارشناسی ارشد، گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شهید باهنر کرمان
[۲]استاد، گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شهید باهنر کرمان
h.mirvaziri@gmail.com alirezairanmanesh78@gmail.com,

چکیده

با گسترش روزافزون استفاده از دستگاه‌های اندرویدی، تشخیص بدافزار به یکی از چالش‌های حیاتی امنیت سایبری تبدیل شده است. روش‌های سنتی که بر تحلیل تک‌وجهی متکی‌اند، در مواجهه با بدافزارهای پیچیده و تهدیدات روز صفر کارایی محدودی دارند. این پژوهش مدل نوین چندوجهی MAGNET را ارائه می‌دهد که با ترکیب هوشمندانه داده‌های جدولی، گرافی و ترتیبی و بهره‌گیری از معماری‌های پیشرفته شبکه‌های عصبی ترنسفورمر و گراف، دقت تشخیص بدافزار را بهبود می‌بخشد. مدل پیشنهادی شامل سه ماژول تخصصی En-hancedTabTransformer، GraphTransformer و SequenceTransformer به همراه مکانیزم توجه پویا و لایه ادغام چندوجهی است. ارزیابی بر روی مجموعه داده DREBIN شامل ۱۰,۵۶۰ نمونه نشان می‌دهد که مدل MAGNET با دقت ۹۷.۲۴٪، معیار F1 برابر ۹۸۲۳.۰ و AUC برابر ۹۸۱.۰، عملکرد برتری نسبت به روش‌های مرجع از جمله SVM (۹۰.۶٪)، CNN (۹۲.۸٪) و LSTM (۹۱.۵٪) ارائه می‌دهد. نتایج تأیید می‌کند که رویکرد چندوجهی و استفاده از معماری‌های نوین یادگیری عمیق، پتانسیل قابل‌توجهی در مقابله با تهدیدات پیچیده اندرویدی دارد.

کلیدواژه‌ها: تشخیص بدافزار اندروید، شبکه‌های عصبی ترنسفورمر، داده‌های چندوجهی، شبکه‌های عصبی گراف، یادگیری عمیق، امنیت سایبری

۱ مقدمه

استفاده گسترده از سیستم‌عامل اندروید در دستگاه‌های هوشمند، این پلتفرم را به هدف اصلی حملات سایبری و انتشار بدافزار تبدیل کرده است [۱]. آمارها نشان می‌دهند که بیش از ۸۵٪ از دستگاه‌های هوشمند جهان از سیستم‌عامل اندروید استفاده می‌کنند، که این امر اهمیت توسعه سیستم‌های کارآمد تشخیص بدافزار را دوچندان می‌کند.

روش‌های سنتی تشخیص بدافزار، شامل تحلیل ایستا و پویا، هرکدام با محدودیت‌های خاص خود مواجه‌اند. تحلیل ایستا که بر بررسی ویژگی‌های برنامه بدون اجرای آن متکی است، در شناسایی بدافزارهایی که از تکنیک‌های مبهم‌سازی یا رمزگذاری استفاده می‌کنند، ناتوان است [۲]. در مقابل، تحلیل پویا که رفتار برنامه را در حین اجرا بررسی می‌کند، علی‌رغم دقت بالاتر، دارای محاسبات سنگین و زمان‌بر است و پوشش کامل مسیرهای اجرایی را تضمین نمی‌کند.

چالش اصلی در تشخیص بدافزارهای اندرویدی، پیچیدگی و تنوع آن‌هاست. بدافزارهای مدرن از تکنیک‌های پیشرفته‌ای همچون چندشکلی، تغییر شکل و تکنیک‌های فرار برای گریز از سیستم‌های تشخیص استفاده می‌کنند.

علاوه بر این، ظهور تهدیدات روز صفر که الگوهای جدید و ناشناخته‌ای دارند، کارایی روش‌های مبتنی بر امضا و قوانین از پیش تعریف‌شده را محدود کرده است.

این پژوهش با هدف غلبه بر محدودیت‌های فوق، رویکردی نوین مبتنی بر ادغام داده‌های چندوجهی ارائه می‌دهد. فرضیه اصلی این است که ترکیب اطلاعات از منابع مختلف داده‌ای—جدولی، گرافی و ترتیبی—می‌تواند بازنمایی جامع‌تر و دقیق‌تری از رفتار برنامه‌های اندرویدی فراهم کند. مدل پیشنهادی Analysis (Multi-modal MAGNET NEtwork Graph-based for Threats) با بهره‌گیری از معماری‌های پیشرفته یادگیری عمیق نظیر ترنسفورمرها و شبکه‌های عصبی گراف، قابلیت استخراج الگوهای پیچیده و تعمیم‌پذیری به تهدیدات جدید را دارد.

۲ پیشینه تحقیق

۱.۲ تحلیل ایستا در تشخیص بدافزار

تحقیقات اولیه در زمینه تشخیص بدافزار اندروید بر تحلیل ویژگی‌های ایستا متمرکز بودند. مطالعه Arp و همکاران [۲] که منجر به توسعه سیستم Drebin شد، از ویژگی‌هایی نظیر مجوزها، فراخوانی‌های API و اجزای برنامه استفاده کرد. این رویکرد با استفاده از الگوریتم SVM، دقت ۹۴٪ در تشخیص بدافزار حاصل کرد. با این حال، این روش در مواجهه با بدافزارهای مبهم‌سازی‌شده محدودیت داشت.

Grace و همکاران بر تحلیل ساختاری کد متمرکز شدند و از گراف‌های جریان کنترل برای شناسایی الگوهای مخرب استفاده کردند [۵]. نتایج نشان داد که ترکیب ویژگی‌های ساختاری با ویژگی‌های سطح بالا می‌تواند دقت تشخیص را بهبود بخشد.

۲.۲ رویکردهای مبتنی بر یادگیری عمیق

با پیشرفت یادگیری عمیق، محققان شروع به استفاده از شبکه‌های عصبی برای تشخیص بدافزار کردند. Zhang و همکاران از شبکه‌های عصبی کانولوشنی برای تحلیل توالی‌های فراخوانی API استفاده کردند و دقت ۹۶.۷٪ حاصل کردند [۳]. Vinayakumar و همکاران نشان دادند که شبکه‌های عمیق می‌توانند الگوهای پیچیده‌تری را نسبت به روش‌های سنتی شناسایی کنند [۴].

Hou و همکاران از شبکه‌های بازگشتی LSTM برای مدل‌سازی رفتار پویای برنامه‌ها استفاده کردند و نشان دادند که این روش در تشخیص بدافزارهای پیچیده مؤثرتر است [۶]. با این حال، این مطالعات عمدتاً بر یک نوع داده (تک‌وجهی) متمرکز بودند.

۳.۲ شبکه‌های عصبی گراف در امنیت

استفاده از شبکه‌های عصبی گراف (GNN) در حوزه امنیت سایبری نسبتاً جدید است. Li و همکاران از GNN برای تحلیل گراف‌های فراخوانی توابع استفاده کردند و نشان دادند که این روش می‌تواند روابط پیچیده بین اجزای برنامه را مدل‌سازی کند [۷]. Kipf و Welling معماری Network Convolutional Graph را معرفی کردند که بنیان بسیاری از کاربردهای GNN شد [۸].

۴.۲ رویکردهای چندوجهی

تحقیقات محدودی در زمینه استفاده از داده‌های چندوجهی برای تشخیص بدافزار انجام شده است. Kim و همکاران از ترکیب ویژگی‌های ایستا و پویا استفاده کردند اما معماری آن‌ها ساده بود و از پیشرفت‌های اخیر یادگیری عمیق

بهره نمی‌برد [۹]. Xu و همکاران روشی برای ترکیب ویژگی‌های مختلف ارائه دادند اما مکانیزم ادغام آن‌ها کارآمد نبود [۱۰].

۵.۲ شکاف تحقیقاتی

علی‌رغم پیشرفت‌های حاصله، شکاف‌های مهمی در تحقیقات موجود وجود دارد:

۱. عدم استفاده مؤثر از داده‌های چندوجهی: اکثر مطالعات بر یک نوع داده متمرکز بوده‌اند

۲. محدودیت در معماری‌های ادغام: روش‌های موجود برای ترکیب انواع مختلف داده ساده و غیرکارآمد هستند

۳. عدم استفاده از معماری‌های نوین: استفاده محدود از ترنسفورمرها و GNN های پیشرفته

۴. ضعف در تعمیم‌پذیری: اکثر مدل‌ها در مواجهه با تهدیدات جدید عملکرد ضعیفی دارند

این پژوهش با ارائه مدل MAGNET که از معماری چندوجهی پیشرفته و مکانیزم‌های نوین یادگیری عمیق استفاده می‌کند، به دنبال پر کردن این شکاف‌هاست.

۳ روش پیشنهادی

۱.۳ کلیات مدل MAGNET

مدل MAGNET یک معماری چندوجهی است که از سه نوع داده مختلف برای تشخیص بدافزار اندروید استفاده می‌کند:

- **داده‌های جدولی**: ویژگی‌های ایستا شامل مجوزها، تعداد فایل‌ها، اندازه برنامه
- **داده‌های گرافی**: گراف‌های فراخوانی توابع که روابط ساختاری را نمایش می‌دهند
- **داده‌های ترتیبی**: توالی‌های فراخوانی API که الگوهای رفتاری زمانی را منعکس می‌کنند

هسته اصلی مدل شامل سه ماژول تخصصی، یک مکانیزم توجه پویا و یک لایه ادغام چندوجهی است که در ادامه به تفصیل شرح داده می‌شوند.

۲.۳ ماژول‌های تخصصی

۱.۲.۳ EnhancedTabTransformer

این ماژول برای پردازش داده‌های جدولی طراحی شده است. هر ویژگی به عنوان یک توکن در نظر گرفته شده و روابط بین ویژگی‌ها از طریق معماری ترنسفورمر مدل‌سازی می‌شود:

$$\mathbf{e}_i = \text{ReLU}(\text{LayerNorm}(W_{\text{emb}} x_i + b_{\text{emb}})) \quad (۱)$$

که در آن $\mathbf{e}_i$ جاسازی ویژگی iام است و $W_{\text{emb}} \in \mathbb{R}^{1\times 64}$ ماتریس وزن جاسازی است.

GraphTransformer ۲.۲.۳

این ماژول گراف‌های فراخوانی توابع را پردازش می‌کند. با استفاده از TransformerConv، اطلاعات در سراسر گراف منتشر می‌شود:

$$\mathbf{h}_v = W_{\text{node}}\mathbf{x}_v + b_{\text{node}} \tag{۲}$$

جایی که $\mathbf{h}_v$ بازنمایی گره v و $W_{\text{node}} \in \mathbb{R}^{64\times64}$ ماتریس وزن گره است.

SequenceTransformer ۳.۲.۳

این ماژول توالی‌های فراخوانی API را با استفاده از کدگذاری موقعیت سینوسی پردازش می‌کند:

$$PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d}}\right), \quad PE(pos, 2i+1) = \cos\left(\frac{pos}{10000^{2i/d}}\right) \tag{۳}$$

۳.۳ مکانیزم توجه پویا

یکی از نوآوری‌های کلیدی این پژوهش، استفاده از مکانیزم توجه پویا است که با پارامتر قابل یادگیری γ وزن‌دهی بهینه ویژگی‌ها را انجام می‌دهد:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\gamma \cdot \frac{QK^T}{\sqrt{d_k}}\right)V \tag{۴}$$

۴.۳ لایه ادغام چندوجهی

خروجی‌های سه ماژول از طریق مکانیزم توجه متقاطل ادغام می‌شوند:

$$\mathbf{z}_{\text{fused}} = \text{DynamicAttention}([\mathbf{z}_{\text{tab}}, \mathbf{z}_{\text{ġraph}}, \mathbf{z}_{\text{seq}}]) \tag{۵}$$

۴ پیاده‌سازی و نتایج

۱.۴ تنظیمات آزمایش

پیاده‌سازی مدل با استفاده از PyTorch ۰.۹.۱ و Geometric PyTorch ۰.۷.۱ انجام شد. آزمایش‌ها بر روی سیستمی با مشخصات زیر اجرا شدند:

- GPU: NVIDIA RTX ۳۰۹۰ با ۲۴ گیگابایت VRAM
- CPU: Intel Xeon E۵-۲۶۹۰ v۴ با ۳۲ هسته
- RAM: ۱۲۸ گیگابایت DDR۴

۲.۴ مجموعه داده

از مجموعه داده استاندارد DREBIN شامل ۵،۵۶۰ نمونه بدافزار و ۵،۰۰۰ نمونه سالم استفاده شد. ویژگی‌های استخراجی شامل:

- ۱۲۸ ویژگی جدولی (مجوزها، اندازه فایل، تعداد اجزا)
- گراف‌های فراخوانی با میانگین ۱،۲۴۵ گره و ۳،۸۷۲ یال
- توالی‌های API با میانگین طول ۸۷ فراخوانی

۳.۴ تنظیم ابرپارامترها

بهینه‌سازی ابرپارامترها با استفاده از Optuna انجام شد. پارامترهای بهینه عبارتند از:

- تعداد سرهای توجه: ۸
- تعداد لایه‌ها: ۴
- نرخ dropout: ۰.۲
- اندازه دسته: ۳۲
- نرخ یادگیری: ۰.۰۰۱

۴.۴ نتایج ارزیابی

ارزیابی با استفاده از اعتبارسنجی متقاطع ۵ـتایی انجام شد. نتایج حاصله در جدول ۱ نشان داده شده است:

جدول ۱: مقایسه عملکرد مدل‌های مختلف

مدل	دقت (%)	F۱-Score	AUC
SVM	۶.۹۰	۸۹۵.۰	۹۱۲.۰
CNN	۸.۹۲	۹۲۱.۰	۹۳۴.۰
LSTM	۵.۹۱	۹۰۸.۰	۹۲۵.۰
MAGNET	**۲۴.۹۷**	**۹۸۲۳.۰**	**۹۸۱.۰**

نتایج نشان می‌دهد که مدل MAGNET با دقت ۵٪.۲۴±۰.۹۷، معیار F۱ برابر ۰۰۲.۹۸۲۳±۰.۰ و AUC برابر ۰۰۳.۹۸۱±۰.۰، عملکرد برتری نسبت به روش‌های مرجع دارد.

۵.۴ تحلیل اجزا

برای درک تأثیر هر جزء مدل، مطالعه ablation انجام شد. نتایج نشان می‌دهد:

- حذف مکانیزم توجه پویا: کاهش ۱٪.۲ دقت
- استفاده از داده تک‌وجهی: کاهش ۸٪.۳ دقت
- حذف لایه ادغام: کاهش ۷٪.۱ دقت

۵ بحث و تحلیل نتایج

۱.۵ برتری مدل چندوجهی

نتایج به‌دست‌آمده تأیید می‌کند که استفاده از داده‌های چندوجهی نسبت به رویکردهای تک‌وجهی برتری قابل‌توجهی دارد. ترکیب اطلاعات ساختاری، رفتاری و آماری برنامه‌ها، امکان شناسایی الگوهای پیچیده‌تر و مقاوم‌تر در برابر تکنیک‌های فرار را فراهم می‌کند.

۲.۵ تأثیر معماری ترنسفورمر

استفاده از معماری ترنسفورمر با مکانیزم توجه پویا، کارایی مدل را به‌طور چشمگیری بهبود بخشیده است. این معماری قابلیت مدل‌سازی وابستگی‌های بلندمدت و تمرکز بر ویژگی‌های مهم را دارد.

۳.۵ پایداری و تعمیم‌پذیری

انحراف معیار پایین نتایج ($\pm ۰.۵\%$ برای دقت) نشان‌دهنده پایداری بالای مدل است. علاوه بر این، ارزیابی بر روی زیرمجموعه‌های مختلف داده نشان می‌دهد که مدل قابلیت تعمیم مناسبی به نمونه‌های جدید دارد.

۴.۵ محدودیت‌ها

با وجود نتایج مثبت، مدل با محدودیت‌هایی مواجه است:

- پیچیدگی محاسباتی بالا که زمان آموزش را افزایش می‌دهد
- وابستگی به کیفیت داده‌های ورودی
- نیاز به منابع محاسباتی قابل‌توجه برای استقرار

۶ تشکر و قدردانی

نویسندگان از حمایت‌های علمی و تخصصی دانشگاه شهید باهنر کرمان و دسترسی به امکانات محاسباتی لازم برای انجام این پژوهش تشکر می‌کنند. همچنین از راهنمایی‌های ارزشمند استاد محترم دکتر حمید میروزیری در طول مراحل مختلف این تحقیق صمیمانه سپاسگزاری می‌نمایند.

۷ نتیجه‌گیری و پیشنهادها برای تحقیقات آینده

۱.۷ نتیجه‌گیری

این پژوهش مدل چندوجهی MAGNET را برای تشخیص بدافزار اندروید معرفی کرد که با ترکیب هوشمندانه داده‌های جدولی، گرافی و ترتیبی و استفاده از معماری‌های پیشرفته یادگیری عمیق، دقت تشخیص را به‌طور قابل‌توجهی بهبود بخشید. نتایج تجربی نشان می‌دهد که:

- رویکرد چندوجهی برتری قابل‌توجهی نسبت به روش‌های تک‌وجهی دارد

- استفاده از معماری ترنسفورمر و مکانیزم توجه پویا کارایی مدل را بهبود می‌بخشد
- مدل پیشنهادی پایداری و تعمیم‌پذیری مناسبی دارد

۲.۷ کاربردهای عملی

مدل MAGNET می‌تواند در سناریوهای زیر مورد استفاده قرار گیرد:

- سیستم‌های امنیتی فروشگاه‌های نرم‌افزار
- راه‌حل‌های امنیتی سازمانی برای دستگاه‌های موبایل
- ابزارهای تحلیل بدافزار برای محققان امنیت

۳.۷ پیشنهادها برای تحقیقات آینده

بر اساس یافته‌های این پژوهش، تحقیقات آتی می‌تواند در جهات زیر گسترش یابد:

۱. **بهینه‌سازی عملکرد**: توسعه نسخه‌های سبک‌تر مدل برای استقرار در دستگاه‌های موبایل

۲. **ارزیابی گسترده‌تر**: آزمون مدل بر روی مجموعه داده‌های بزرگ‌تر و متنوع‌تر

۳. **مقاومت تخاصمی**: بررسی و بهبود مقاومت مدل در برابر حملات تخاصمی

۴. **تفسیرپذیری**: توسعه مکانیزم‌هایی برای تفسیر تصمیمات مدل

۵. **تشخیص بلادرنگ**: تطبیق مدل برای کاربردهای بلادرنگ

مراجع

[۱] Ra- & M.، Conti، S.، M. Gaur، V.، Ganmoor، V.، Laxmi، A.، Bharmal، P.، Faruki، and penetration، malware issues، of survey a security: Android .(۲۰۱۵) M. jarajan، *tutorials* & *surveys communications IEEE* defenses، ۱۷ (۲)، ۹۹۸-۱۰۲۲.

[۲] E. C. Siemens، & K.، Rieck، H.، Gascon، M.، Hubner، M.، Spreitzenbarth، D.، Arp، in malware android of detection explainable and Effective DREBIN: .(۲۰۱۴) T. R. *Ndss* pocket. your، ۱۴، ۲۳-۲۶.

[۳] mal- android Semantics-aware .(۲۰۱۴) Z. Zhao، & H.، Yin، Y.، Duan، M.، Zhang، *Proceedings* graphs. dependency api contextual weighted using classification ware ،*Security Communications and Computer on Conference SIGSAC ACM* ۲۰۱۴ *the of* ۱۱۰۵-۱۱۱۶.

[۴] & A.، Al-Nemrat، P.، Poornachandran، P.، K. Soman، M.، Alazab، R.، Vinayakumar، detection intrusion intelligent for approach learning Deep .(۲۰۱۹) S. Venkatraman، *Access IEEE* system، ۷، ۴۱۵۲۵-۴۱۵۵۰.

[۵] scalable RiskRanker: .(۲۰۱۲) X. Jiang، & S.، Zou، Q.، Zhang، Y.، Zhou، M.، Grace، inter- ۱۰th the of Proceedings detection. malware android zero-day accurate and .۲۹۴-۲۸۱ ،services and applications، systems، Mobile on conference national

[۶] learn- deep A Deep۴MalDroid: .(۲۰۱۶) Y. Ye، & L.، Chen، A.، Saas، S.، Hou، call system kernel linux on based detection malware android for framework ing Intelligence Web on Conference International IEEE/WIC/ACM ۲۰۱۶ sequences. .۱۱۱-۱۰۴ ،(WIW) Workshops

[۷] & ... D.، Octeau، A.، Bartel، S.، Rasthofer، M.، Papadakis، F.، T. Bissyandé، L.، Li، review. literature systematic A apps: android of analysis Static .(۲۰۱۷) E. Bodden، .۹۵-۶۷ ،۸۸ ،technology software and Information

[۸] con- graph with classification Semi-supervised .(۲۰۱۶) M. Welling، & N.، T. Kipf، .arXiv:۱۶۰۹.۰۲۹۰۷ preprint arXiv networks. volutional

[۹] learn- deep multimodal A .(۲۰۱۹) G. E. Im، & S.، Sezer، M.، Rho، B.، Kang، T.، Kim، Transactions IEEE features. various using detection malware android for method ing .۷۸۸-۷۷۳ ،(۳)۱۴ ،Security and Forensics Information on

[۱۰] for analysis Hybrid HADM: .(۲۰۱۶) J. Cavazos، & N.، Jayasena، D.، Zhang، L.، Xu، .۷۲۴-۷۰۲ ،(IntelliSys) conference systems intelligent SAI malware. of detection