BAB I
PENDAHULUAN

1.1 Latar Belakang Masalah

Sistem operasi Android telah mencapai dominasi yang tak tertandingi di pasar perangkat mobile global, sebuah pencapaian yang didorong oleh sifat sumber terbukanya (*open-source*) yang mendorong inovasi dan adopsi yang luas [3]. Namun, keterbukaan ini secara inheren juga melipatgandakan masalah keamanan, menjadikannya platform yang sangat menarik bagi para pengembang malware [12]. Akibatnya, lanskap ancaman siber untuk Android telah berkembang secara eksponensial, dengan ribuan sampel malware baru muncul setiap harinya, masing-masing dengan tingkat kecanggihan yang terus meningkat [3]. Malware modern telah berevolusi jauh dari sekadar gangguan menjadi ancaman kompleks yang mampu menyebabkan kerugian finansial yang signifikan, pencurian data pribadi berskala besar, dan pelanggaran privasi yang parah [5].

Menghadapi ancaman yang dinamis ini, industri keamanan siber secara historis bergantung pada solusi antivirus tradisional yang berbasis tanda tangan (*signature-based*). Pendekatan ini bekerja dengan mencocokkan hash atau pola unik dari sebuah file dengan database tanda tangan malware yang telah diketahui. Meskipun efektif untuk ancaman yang sudah dikenal, metode ini secara fundamental bersifat reaktif. Metode ini memiliki kelemahan signifikan saat dihadapkan pada varian malware baru (*zero-day*), malware *polimorfik* yang secara konstan mengubah kodenya untuk menghindari deteksi, atau ancaman yang menggunakan teknik penyamaran canggih. Pendekatan heuristik, yang menganalisis perilaku atau karakteristik mencurigakan, menawarkan perbaikan, namun model *machine learning* awal seringkali kesulitan menafsirkan kompleksitas semantik dan niat jahat yang tersirat dalam kode.

Perkembangan terkini dalam kecerdasan buatan, khususnya *Large Language Models (LLM)*, membuka peluang untuk sebuah paradigma baru. LLM, yang dilatih pada korpus data teks dan kode yang sangat besar, tidak hanya mampu mengenali sintaksis tetapi juga memahami semantik dan konteks.

Kemampuan ini memungkinkan pergeseran fundamental dalam analisis malware: dari pendekatan "seperti apa kode ini terlihat?" menjadi "apa yang coba dilakukan oleh kode ini?". Dengan menerapkan LLM sebagai mesin penalaran, analisis statis dapat ditingkatkan untuk mengidentifikasi niat jahat bahkan dari kode yang belum pernah terlihat sebelumnya, menawarkan solusi proaktif untuk lanskap ancaman yang terus berkembang.

Meskipun studi-studi awal telah mengeksplorasi LLM untuk tugas-tugas keamanan siber, banyak di antaranya masih menggunakan LLM sebagai alat bantu, generator data sintetis, atau sebagai bagian dari arsitektur *Retrieval-Augmented Generation (RAG)* yang kompleks. Terdapat celah penelitian yang jelas di mana penelitian sebelumnya belum menjadikan LLM sebagai mesin analitik utama dalam alur kerja analisis statis yang sederhana dan efisien.

Penelitian ini bertujuan untuk mengisi celah tersebut dengan memposisikan LLM secara langsung sebagai inti penalaran semantik, menguji hipotesis bahwa rekayasa prompt yang canggih dapat mengarahkan model untuk menganalisis niat jahat secara efektif hanya dari artefak statis.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, penelitian ini dirumuskan ke dalam beberapa pertanyaan kunci sebagai berikut:

1. Bagaimana kemampuan penalaran kontekstual dan pemahaman kode dari *Large Language Model (LLM)* yang sudah ada dapat dimanfaatkan untuk melakukan analisis statis yang efektif dalam mendeteksi malware pada aplikasi Android?
2. Apakah metodologi yang optimal untuk melakukan pra-pemrosesan pada paket aplikasi Android (.apk) agar dapat diubah menjadi format tekstual yang sesuai untuk dianalisis oleh LLM?
3. Bagaimana rekayasa prompt (prompt engineering) dapat diterapkan secara strategis untuk mengarahkan LLM agar berfungsi sebagai seorang ahli keamanan siber, sehingga mampu mengidentifikasi niat jahat secara akurat dari kombinasi cuplikan kode, daftar izin, dan fitur statis lainnya?

4. Bagaimana kinerja sistem deteksi berbasis LLM ini jika dibandingkan dengan metode deteksi tradisional, terutama dalam hal metrik akurasi, presisi, dan *recall*, saat diuji menggunakan dataset yang beragam yang terdiri dari aplikasi berbahaya dan aplikasi aman?

1.3 Tujuan Penelitian

Sejalan dengan rumusan masalah, Penelitian ini bertujuan untuk:

1. Merancang dan mengimplementasikan sebuah alur kerja (*fine-tuning*) yang lengkap untuk analisis statis otomatis pada file APK Android dengan menggunakan API dari *Large Language Model* (seperti Google Gemini atau OpenAI).
2. Mengembangkan dan menyempurnakan serangkaian prompt terstruktur yang dapat memaksimalkan akurasi analitis LLM dalam mengidentifikasi berbagai jenis ancaman keamanan.
3. Mengevaluasi kinerja sistem yang diusulkan secara sistematis dengan menggunakan metrik klasifikasi standar pasda dataset yang seimbang antara aplikasi berbahaya dan aplikasi aman.
4. Memberikan kontribusi berupa pendekatan heuristik baru dalam deteksi malware yang melampaui pencocokan pola sintaktis dan bergerak menuju analisis niat semantik.

1.4 Batasan Masalah

Untuk menjaga fokus dan kedalaman penelitian, ditetapkan batasan-batasan sebagai berikut:

1. Penelitian ini akan berfokus secara eksklusif pada analisis statis. Meskipun analisis dinamis memberikan data perilaku yang sangat berharga, integrasinya dianggap di luar cakupan skripsi ini dan akan dibahas sebagai area untuk penelitian di masa depan.
2. Sistem yang dikembangkan akan bergantung pada LLM pra-terlatih yang tersedia untuk umum melalui API mereka. Penelitian ini tidak akan mencakup proses pelatihan atau *fine-tuning* model LLM dari awal.

3. Studi ini mengakui adanya tantangan dari teknik *obfuscation* kode tingkat lanjut dan metode anti-analisis yang dapat menghambat proses dekompilasi. Meskipun demikian, penelitian ini tidak bertujuan untuk menyelesaikan sepenuhnya masalah tersebut, melainkan untuk menguji efektivitas LLM pada kode yang berhasil didekompilasi.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat yang signifikan, baik secara teoritis bagi komunitas akademis maupun secara praktis bagi industri keamanan siber.

1. Manfaat Teoritis
   - Memberikan kontribusi pada bidang keilmuan keamanan siber dengan memvalidasi sebuah pergeseran paradigma dari deteksi berbasis pencocokan pola sintaktis menuju analisis niat semantik.
   - Menunjukkan bagaimana kemampuan penalaran kontekstual dan pemahaman kode dari *Large Language Models (LLM)* dapat diterapkan secara efektif untuk mengidentifikasi niat jahat, yang melampaui keterbatasan metode heuristik tradisional.
   - Menyediakan landasan dan studi evaluatif bagi penelitian di masa depan yang ingin mengembangkan teknik analisis hibrida atau model LLM yang di-*fine-tuning* khusus untuk domain analisis malware.

2. Manfaat Praktis
   - Menghasilkan sebuah purwarupa alur kerja (*pipeline*) otomatis yang dapat diimplementasikan oleh analis keamanan siber untuk mempercepat proses triase dan analisis awal file APK yang mencurigakan.
   - Menawarkan pendekatan deteksi yang lebih tangguh dan proaktif terhadap ancaman baru (*zero-day*) dan malware *polimorfik*, di mana metode berbasis tanda tangan (*signature-based*) seringkali gagal.

- Memberikan wawasan tentang desain *prompt engineering* yang efektif untuk tugas-tugas keamanan siber, termasuk pertimbangan untuk mitigasi serangan *prompt injection* yang disematkan dalam kode malware itu sendiri.

1.6 Sistematika Penulisan

1. BAB I: Pendahuluan — berisi latar belakang, rumusan masalah, tujuan penelitian, Batasan, manfaat, dan juga sistematika penulisan.
2. BAB II: Tinjauan Pustaka — menyajikan tinjauan pustaka dan landasan teori yang relevan, termasuk taksonomi malware Android, teknik analisis malware, dan peran LLM dalam keamanan siber.
3. BAB III: Metodologi Penelitian — menjelaskan metodologi penelitian secara rinci, mulai dari pengumpulan dataset, perancangan sistem pra-pemrosesan, hingga implementasi analisis menggunakan LLM.
4. BAB IV: Hasil dan Pembahasan — menyajikan hasil pengujian dan evaluasi kinerja system dan juga hasil analisis.
5. BAB V: Kesimpulan dan Saran — menyajikan hasil pengujian dan evaluasi kinerja system.