BAB I
PENDAHULUAN

1.1 Latar Belakang

Seiring dengan kemajuan teknologi ancaman serangan siber semakin canggih, hal ini menjadi perhatian bukan hanya tentang memastikan kerahasiaan data tetapi juga menghadapi dari serangan robot otomasi *spam* dan *crawlers.* Peningkatan kebutuhan informasi dan kemajuan teknologi mendorong beberapa perusahaan untuk mengembangkan teknologi baru agar *website* pengolahan data dan informasi dapat terhindar dari adanya serangan robot. *web scraping* dan *Robotic Process Automation* adalah beberapa metode yang digunakan untuk pengambilan data secara cepat dan terstruktur yang berguna dalam melakukan pengujian secara efisien dan efektif terhadap serangan robot [1].

Dikutip dari CNBC Indonesia dan AWS Amazon, pengunjung internet dikuasai oleh robot otomatis yang dapat menjalankan berbagai tugas secara berulang seperti *web crawlers bots*, *spam bots*, dan *transaction bots*. Dampak buruknya adalah ancaman keamanan data yang semakin meningkat dan membebani penyedia layanan. *web crawlers bot* dapat mengekstrak data sensitif pada halaman *website*, melanggar ketentuan layanan dan menyebabkan beban berlebih pada *website* tujuan dan *spam bots* dapat melakukan otomasi untuk berinteraksi dengan *website*, melakukan pendaftaran akun secara berulang, memindai konten *website* yang dapat menyebabkan gangguan bahkan kerugian karena menciptakan kesan popularitas palsu agar menarik pelanggan manusia nyata [2].

Web Scraping adalah proses ekstraksi data dan menganalisis dokumen dalam sebuah *website* untuk mengambil informasi dari elemen HTML atau JSON yang diinginkan, *web scraping* banyak digunakan untuk melakukan pencarian produk di *e-commerce* bahkan mengambil informasi di sebuah *website* tertentu [3]. Beberapa kasus yang melibatkan *web scraping* dalam *e-commerce* adalah mengambil data produk *flash sale* dan melakukan pembelian produk secara sepihak yang menyebabkan produk yang seharusnya dijual ke banyak pengguna menjadi

tidak tersebar secara merata [1]. Proses kerja *Web Scraping* ini akan mengakses *Website* dengan protokol *Hypertext Transfer Protocol* (HTTP) atau *Application Programming Interface* (API) yang kemudian mengekstrak data yang diperlukan dan menyimpannya di *database* atau *file* .CSV [2].

Robotic Process Automation atau RPA adalah sebuah teknologi otomatisasi sebuah pekerjaan dengan melakukan interaksi ke *website* dan *mobile application* yang digunakan untuk melakukan *task* terstruktur dan berulang sehingga tidak memerlukan manusia dalam melakukan pekerjaan yang sama secara bersamaan dengan akurat. RPA memungkinkan otomatisasi dalam sebuah *Website* dengan waktu yang lebih cepat dan tingkat akurasi lebih tinggi dan lebih menghemat *cost* dibanding dengan manusia [4]. RPA sangat cepat diaplikasikan dibandingkan dengan otomatisasi tradisional, RPA menggunakan *Graphical User Interface* (GUI) sebagai acuan untuk melakukan proses otomatisasi dengan mengambil *xpath* atau *locator* pada sebuah *element* di *Website* tersebut [4].

Berdasarkan penelitian terdahulu yang berjudul Implementasi Metode Proteksi Situs Web Dari Web Scraping menyatakan bahwa metode yang efektif untuk mencegah *Web* scraping dan botnet melibatkan peningkatan keamanan menggunakan captcha dan pembatasan laju (rate limit) pada situs *Web* demo [1]. Penelitian yang berjudul Analisis Keamanan Website Repository Institut Teknologi Telkom Purwokerto Menggunakan Metode Vulnerability Assessment menyatakan terdapat *port* yang terbuka dan kerentanan dengan status medium pada *Website* tersebut, yang memiliki potensi merusak sistem. [5]. Penelitian yang ketiga berjudul Rancangan Sistem Keamanan Jaringan dari serangan DDoS Menggunakan Metode Pengujian Penetrasi menunjukkan bahwa WAF berhasil mencegah serangan *HTTP Request* [6]. Penelitian berjudul Implementasi Web Scraping untuk Pengambilan Data Pada Website E-Commerce menunjukkan bahwa implementasi *web scraping* dengan HTML Parsing dan CSS Selector dapat berhasil mengambil data dari beberapa *website e-commerce* yang diteliti [7]. Penelitian yang berjudul Perbandingan Tingkat Keamanan Website Menggunakan Nmap Dan Nikto Dengan Metode Ethical Hacking berhasil mengidentifikasi

informasi penting terkait dengan *hostname*, *port* yang terbuka, dan jenis server yang digunakan oleh *Website* yang diuji. Penelitian ini juga membandingkan efektivitas Nmap dan Nikto dalam mengidentifikasi kerentanan keamanan pada *Website* [8]. Penelitian yang berjudul Analisis Keamanan Layanan E-Learning Terhadap Serangan Dos Dan Implementasi Mitigasi Pada Universitas Budi Luhur menunjukkan perlunya meningkatkan kapabilitas sistem untuk menghadapi serangan kategori *Application Layer Attack*. [9].

Berdasarkan latar belakang yang telah diuraikan, maka diperlukan sistem analisis untuk mengukur tingkat keamanan sebuah *website* dari serangan robot *web crawlers bot* dan *spam bots* dengan melakukan beberapa proses menggunakan metode *web scraping* dan *Robotic Process Automation.* Penelitian ini bertujuan untuk membuat sistem yang mampu melakukan otomatisasi proses untuk meningkatkan efisiensi dan meminimalisir permasalahan dalam proses pengujian keamanan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah pada penelitian ini adalah *website* bybit (bybit.spservices.my.id) dan pemesanan hotel (hotel.spservices.my.id) rentan terhadap ancaman robot seperti *web crawler bots* dan *spam bots* yang dapat mencuri data sensitif, membebani *server*, serta mengganggu operasional *website* oleh karena itu dibangun sistem untuk mengidentifikasi dan mengukur tingkat keamanan *website* dengan metode *web scraping* dan *robotic process automation*

1.3 Pertanyaan Penelitian

Berdasarkan uraian rumusan masalah diatas, penulis merumuskan beberapa pertanyaan terkait yang akan diteliti, antara lain:

1. Bagaimana sistem analisis otomatis dapat dibangun untuk mengukur tingkat keamanan sebuah *website* terhadap serangan robot dengan metode *equivalence partitioning*?
2. Bagaimana melakukan pengujian *website* dengan menggunakan sistem analisis

keamanan *website* menggunakan metode *web scraping* dan *Robotic Process Automation*?

3. Bagaimana hasil pengujian *website* dengan menggunakan sistem analisis keamanan situs web menggunakan metode *web scraping* dan *Robotic Process Automation* untuk mendeteksi serangan robot *Web Crawlers Bot, SQL Injection, XSS Injection* dan *Spam Bots*?

1.4 Batasan Masalah

Berdasarkan uraian diatas untuk memenuhi penelitian yang selaras dengan rumusan masalah yang diuraikan terdapat batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Pengembangan sistem berfokus pada metode deteksi keamanan *XSS Injection, Brute Force, File Injection* dan *SQL Injection*
2. Pengembangan sistem berfokus pada *platform windows*
3. Pengembangan sistem menggunakan bahasa pemrograman Python3

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan diatas berikut adalah tujuan dari penelitian ini:

1. Mengembangkan sistem analisis keamanan *website* menggunakan metode *Web Scraping* dan *Robotic Process Automation*
2. Mengetahui hasil pengujian sistem dengan menggunakan *Black-box Testing*

1.6 Manfaat Penelitian

Adapun dengan adanya penelitian ini, Penulis berharap terdapat manfaat yang dapat diambil sebagai berikut:

1. Penelitian ini dapat memberikan informasi dan gambaran mengenai hasil pengujian pada *website* tujuan, sehingga pemilik *website* dapat lebih meningkatkan keamanan berdasarkan metode *test* yang digunakan.
2. Penelitian ini dapat membantu Pembaca mendapatkan informasi mengenai *Web Scraping* dan *Robotic Process Automation* menggunakan bahasa pemrograman Python3.