BAB I
PENDAHULUAN

1.1. Latar Belakang

Pertumbuhan Data digital global diprediksi mencapai 175 zettabytes pada 2025, dengan 80% diantaranya berupa data pribadi [1]. Sektor pendidikan menjadi salah satu sektor yang terdampak oleh pertumbuhan data ini [2] dan menghadapi tantangan baru dalam melindungi data pribadi [3]. Badan Siber dan Sandi Negara BSSN pada tahun 2023 melaporkan 34.9 juta data pribadi termasuk milik mahasiswa bocor [4]. Hal ini sejalan dengan temuan Setiawan dan Najicha (202) yang menemukan 279 juta data pribadi bocor dan diperdagangkan di *darkweb* [5]. Sementara itu menurut SOC Radar pada 2023 hingga pertengahan 2024 terdapat 89 aktor ancaman yang menargetkan Indonesia dengan 234 aktivitas perdagangan data di *darkweb* [6].

Kebocoran data di sektor pendidikan tidak hanya terjadi di Indonesia, tetapi juga menjadi masalah serius secara global [7]. Hylender et al. (2024) dalam laporan Verizon (DBIR) menyebut sektor pendidikan sebagai target terbesar kedua serangan siber, dengan 1.537 insiden kebocoran data yang dikonfirmasi, 98% di antaranya bermotif finansial [8]. Di Amerika Serikat, tercatat sebanyak 3.713 insiden kebocoran data di institusi pendidikan, dengan 60% di antaranya berasal dari universitas [9]. Pada 2021, BSSN mencatat 5.574 kasus peretasan di Indonesia, 36,49% di antaranya menargetkan situs pendidikan tinggi, menjadikannya salah satu sektor rentan [10].

Kebocoran data sering terjadi akibat kurangnya kontrol keamanan yang tepat dan data yang tidak terenkripsi atau dienkripsi dengan buruk [11], [12]. Salah satu contohnya adalah kebocoran data Universitas Indonesia, di mana informasi pribadi seperti nomor registrasi, alamat, dan pengalaman kerja diduga dijual di forum gelap *BreachForums* tanpa enkripsi [13]. Universitas Tanjungpura Pontianak juga mengalami hal yang sama dimana 52 ribu data berupa ID, email, nama pengguna, kata sandi dan nomor telepon bocor tanpa enkripsi [14]. Universitas

Diponegoro juga mengalami kebocoran data, di mana *password* dienkripsi dengan buruk menggunakan MD5 yang mudah didekripsi [15].

Situasi tersebut menegaskan pentingnya pendekatan keamanan yang lebih efektif untuk melindungi informasi pribadi, terutama pada sistem informasi akademik, seperti Sistem Informasi Fakultas Teknik (SimFT) di sebuah universitas swasta di Keresidenan Madiun, sering menyimpan data akademik penting, termasuk informasi pribadi mahasiswa [16]. SimFT menggunakan kombinasi algoritma *Caesar Cipher* dan *Base64* sebagai solusi keamanan. *Caesar Cipher* adalah metode kriptografi klasik yang mengganti karakter *plaintext* dengan menggesernya sejumlah posisi tertentu [17], [18], sedangkan *Base64* adalah skema *encoding* yang mengubah data biner menjadi teks berbentuk karakter yang dapat dicetak, sehingga memudahkan penyimpanan dan transmisi data [19]. Namun, kedua metode ini memiliki keterbatasan dalam menghadapi serangan modern, menekankan perlunya algoritma enkripsi yang lebih kuat untuk melindungi data akademik [20], [21].

Namun, kombinasi *Caesar Cipher* dan *Base64* memiliki kelemahan signifikan. *Base64* meningkatkan ukuran data hingga 33% karena mengubah tiga *byte* biner menjadi empat karakter ASCII [19]. Selain itu, *Caesar Cipher* lemah secara keamanan karena rentan terhadap serangan analisis frekuensi, yang memungkinkan pola kemunculan karakter dalam *ciphertext* dianalisis untuk memecahkan *cipher* dengan mudah [18]. Kombinasi kedua metode ini, meskipun memberikan lapisan perlindungan dasar, tidak cukup kuat untuk menghadapi ancaman siber yang lebih maju, terutama dalam konteks data sensitif seperti data akademik yang ada pada SimFT. Kelemahan-kelemahan ini menunjukkan bahwa diperlukan pendekatan yang lebih kuat dan lebih efisien dalam melindungi data pribadi dan akademik.

Salah satu pendekatan yang menjanjikan adalah mengintegrasikan teknik kompresi dan enkripsi secara bersamaan, metode efisiensi penyimpanan sekaligus keamanan data secara signifikan [22], [23]. Salah satu solusi yang potensial adalah algoritma kompresi *Lempel-Ziv* 4 (LZ4), dikembangkan Yann Collet pada 2011 sebagai varian LZ77 dengan pendekatan berbasis kamus (*dictionary-based*

compression) [24]. LZ4 dirancang untuk kompresi dan dekompresi cepat dengan latensi rendah, menjadikannya ideal bagi aplikasi yang membutuhkan efisiensi waktu dan sumber daya [25]. LZ4 menawarkan keseimbangan optimal antara kecepatan dan efisiensi sumber daya [24], menjadikannya ideal untuk sistem seperti SimFT yang memerlukan performa tinggi dalam menangani data akademik secara *real-time*.

Dalam konteks keamanan data yang lebih komprehensif, *Advanced Encryption Standard* (AES) hadir sebagai solusi enkripsi yang sangat superior dibandingkan dengan *Caesar Cipher* [26]. AES, distandarisasi NIST pada 2001, adalah algoritma enkripsi simetrik 256 bit yang menawarkan tingkat keamanan tinggi [27]. Dibandingkan dengan *Caesar Cipher* yang hanya memiliki 25 kemungkinan pergeseran [17], AES memiliki kompleksitas kriptografis yang praktis tidak mungkin dipecahkan dengan serangan *brute force* [27]. Algoritma ini telah digunakan secara luas dalam sistem keamanan kritikal, termasuk perlindungan data pemerintah, perbankan, dan infrastruktur teknologi informasi [28]. Pada implementasi SimFT, penggunaan AES-256 dapat memberikan perlindungan yang lebih baik terhadap potensi serangan siber, menjamin kerahasiaan dan integritas data akademik dengan tingkat keamanan yang tinggi.

Berdasarkan permasalahan dan solusi diatas penelitian ini mengusulkan implementasi algoritma kompresi LZ4 pada SIMFT sebagai tahap pra-enkripsi, diikuti dengan penggunaan AES untuk menggantikan *Caesar Cipher*. LZ4 akan mengurangi ukuran data dan meningkatkan efisiensi enkripsi, sementara AES menyediakan tingkat keamanan yang lebih baik dibandingkan *Caesar Cipher*. Dengan kombinasi ini, penyimpanan data menjadi lebih hemat tanpa mengorbankan kecepatan atau keamanan.

1.2. Perumusan Masalah

Berdasarkan permasalahan dari latar belakang yang telah diuraikan di atas, maka rumusan masalah dari penelitian ini adalah ‘bagaimana penerapan LZ4 untuk mengurangi ukuran data akademik sebelum proses enkripsi dengan AES 256 pada aplikasi SIMFT?’

1.3. Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan algoritma kompresi LZ4 pada file unggahan di aplikasi SimFT sebelum proses enkripsi yang telah tersedia. Pendekatan ini diharapkan dapat meminimalkan ukuran data akademik yang tersimpan, sehingga meningkatkan efisiensi penggunaan ruang penyimpanan.

1.4. Batasan Masalah

Agar penelitian ini lebih terfokus dan mencapai hasil yang sesuai dengan tujuan yang telah ditetapkan, diperlukan beberapa batasan masalah sebagai berikut:

1. Aplikasi ini akan menggunakan bahasa pemrograman TypeScript.
2. Penelitian ini dibatasi pada fitur praktikum pada aplikasi SimFT yang berjalan dalam lingkup kampus Universitas X.
3. Penelitian ini berfokus pada penerapan algoritma kompresi LZ4 pada file unggahan di aplikasi SIMFT. Aspek di luar penggunaan LZ4 dan proses unggah file tidak dibahas dalam penelitian ini.
4. Proses enkripsi menggunakan AES-256 merupakan bagian dari sistem yang sudah tersedia dan tidak menjadi objek kajian dalam penelitian ini.

1.5. Manfaat Penelitian

Adapun manfaat dan kontribusi dari penelitian ini adalah sebagai berikut:

1. Kontribusi Metodologi
   a. Menerapkan algoritma kompresi LZ4 sebagai tahap pra-enkripsi dalam sistem informasi yang sudah menggunakan AES-256.
   b. Merancang dan mengimplementasikan tahapan pra-enkripsi yang efektif dengan menggunakan algoritma LZ4 untuk mendukung efisiensi penyimpanan pada sistem informasi.
2. Kontribusi Praktis
   a. Memberikan solusi praktis untuk mengurangi ukuran data unggahan, sehingga membantu fakultas dalam mengelola dan menyimpan data secara lebih efisien.
   b. Mengoptimalkan pemanfaatan kapasitas penyimpanan tanpa mengubah mekanisme keamanan yang sudah ada pada sistem.

c. Memberikan informasi empiris mengenai pengaruh penerapan LZ4 terhadap ukuran data sebelum proses enkripsi.