SECURING CLOUD STORAGE USING AES BASED ON GEO-KEY METHOD

NUR SYAFIQAH BINTI MOHD SHAMSUDDIN

UNIVERSITI SAINS ISLAM MALAYSIA

SECURING CLOUD STORAGE USING AES BASED ON GEO-KEY METHOD

Nur Syafiqah binti Mohd Shamsuddin

Thesis submitted in partial fulfilment for the degree of
MASTER OF SCIENCE

UNIVERSITI SAINS ISLAM MALAYSIA

October 2023

بسم الله الرحمن الرحيم

AUTHOR DECLARATION

I hereby declare that the work in this proposal is my own except for quotations and summaries which have been duly acknowledged.

Date: 10 October 2023

Signature:

Name: Nur Syafiqah Mohd Shamsuddin

Matric No: 3172548

Address: Selangor, Malaysia

ACKNOWLEDGEMENT

This study would not have been possible without the guidance and help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude to my supervisor Assoc. Prof. Dr. Sakinah Ali Pitchay whose sincerity and encouragement I will never forget. She has guided me with her patience and professionalism which greatly helps me and motivates me to keep moving forward and complete the thesis when sometimes the journey seems impossible to be continued. May Allah bless her. I also wish to extend my gratitude to Dr. Murtadha Arif Sahbudin for his technical guidance, valuable idea and suggestion throughout this project. My experiment could not been smoothly successful without his critical advice and skilful support.

Special thanks to my university, Universiti Sains Islam Malaysia, and the Malaysian government (Ministry of Higher Education) for the financial grant sponsorships. I also wish to extend my thankfulness to my parents and family members who have supported me throughout my Master's candidature with their love, patience, and understanding. Deepest thanks to my husband, who keep me grounded, remind me of what is important, lending his hand while I'm drowned in my experiment and are always supportive of my adventures.

I would also like to thank all my colleagues and my master-mate friends around me for always supporting me and encouraging me with their love and best wishes.

ABSTRAK

Apa-apa sahaja yang disambungkan ke internet terdedah kepada risiko siber termasuk risiko kebocoran data peribadi dalam storan awan. Penggunaan internet telah meningkat dengan ketara di seluruh negara semasa pandemik Covid-19 menyebabkan sejumlah besar insiden siber dilaporkan. Menurut Tinjauan Pengguna Internet (IUS), laporan 2020 oleh Suruhanjaya Komunikasi dan Multimedia Malaysia (SKMM), peratusan pengguna internet pada 2020 meningkat 1.3% daripada 87.4% pada 2018 kepada 88.7% pada 2020 manakala sejumlah 8,669 keselamatan siber. insiden telah dilaporkan kepada Cyber Security Malaysia pada tahun 2021. Hasil daripada peningkatan bilangan pengguna internet, terdapat tiga isu keselamatan data utama yang dikenal pasti perlu dipertingkatkan iaitu (i) penyimpanan fail selamat, (ii) kunci penyulitan mudah terjejas, dan (iii) akses fail jarak jauh tanpa kebenaran di storan awan. Salah satu cara terbaik untuk melindungi keselamatan fail dalam storan ialah dengan menggunakan kaedah kriptografi yang merangkumi proses penyulitan dan penyahsulitan. Walau bagaimanapun, kaedah penyulitan yang kukuh seperti kaedah AES masih mampu untuk diserang jika kunci penyulitan terdedah. Penyesuaian kepada persekitaran pasca-pandemik dengan bekerja luar dari pejabat fizikal juga telah mendedahkan akses jauh tanpa kebenaran terhadap fail dalam storan awan kerana kekurangan sekatan akses berasaskan lokasi. Oleh itu, untuk meningkatkan keselamatan kebolehcapaian fail dalam storan, penyelidikan ini bertujuan untuk membangunkan kaedah kriptografi yang dipertingkatkan dengan melaksanakan maklumat lokasi untuk menjana kunci penyulitan menggunakan kaedah AES. Kaedah AES dipertingkat dibangunkan dengan menggunakan gabungan koordinat longitud dan latitud, kata laluan pengguna dan alamat MAC peranti untuk menjana kunci penyulitan yang dikenali sebagai kunci geo. Kemudian, ia dinilai untuk mengesahkan prestasinya dengan menjalankan perbandingan prestasi masa antara kaedah sedia ada dan kaedah AES yang dipertingkatkan menggunakan pelbagai jenis format sebagai set data. Keputusan menunjukkan bahawa kaedah AES yang dipertingkatkan mengambil masa 2.73% lebih lama untuk dilaksanakan berbanding kaedah AES sedia ada kerana gabungan tambahan parameter utama untuk menjana kunci geo manakala kaedah AES sedia ada hanya memerlukan satu parameter untuk menjana penyulitan kunci iaitu kata laluan pengguna. Seterusnya, kaedah AES yang dipertingkatkan dinilai dengan mengesahkan kejayaan penyahsulitan fail di lokasi yang berbeza di mana ia telah menunjukkan bahawa hanya fail yang terletak di lokasi yang dimaksudkan boleh dinyahsulit tertakluk kepada julat toleransi jaraknya. Ujian lain telah dijalankan untuk menilai integriti fail yang dinyahsulit yang menunjukkan semua fail yang telah dinyahsulit mempunyai nilai cincang yang sama seperti cincang fail asal. Keputusan daripada penilaian ketiga ini telah membuktikan bahawa kaedah geo-kunci AES yang dipertingkatkan mengekalkan integriti data asal tanpa sebarang pengubahsuaian dan rasuah yang tidak diingini. Kaedah AES yang dipertingkatkan mempunyai sumbangan yang besar untuk melindungi fail dalam storan jika fail telah dicuri dalam situasi insiden kebocoran data keselamatan kerana fail yang disulitkan tidak boleh diakses oleh pemilik yang tidak dibenarkan dan hanya dinyahsulit dalam julat lokasi yang dimaksudkan. Penyelidikan ini juga sejajar dengan salah satu strategi tunggak dalam Strategi Keselamatan Siber Malaysia (MCSS) untuk mewujudkan mekanisme perlindungan kebocoran data dalam pengurusan organisasi dan operasi perniagaan.

Keywords: penjanaan kunci, geo-kunci, koordinat lokasi, alamat MAC, kaedah AES

ABSTRACT

Anything connected to the internet is exposed to cyber risks including the risk of a personal data breach in cloud storage. The use of the internet has risen significantly nationwide during the Covid-19 pandemic resulting in a huge number of cyber incidents being reported. According to the Internet Users Survey (IUS), 2020 report by the Malaysian Communications and Multimedia Commission (MCMC), the percentage of internet users in 2020 grew by 1.3% from 87.4% in 2018 to 88.7% in 2020 while a total of 8,669 cybersecurity incidents were reported to the Cyber Security Malaysia in 2021. As a result of the increasing number of internet users, there are three major data security issues identified that need to be enhanced which are (i) secure file storage, (ii) vulnerable encryption key, and (iii) lack of remote access restriction at cloud storage. One of the strongest options to provide security for file protection in the storage is by implementing a cryptographic method that includes encryption and decryption process. However, strong encryption method such as the AES method is still available to be attacked if the encryption key is vulnerable. Adaptation to a post-pandemic environment by working remotely from the physical office has also attracted unauthorized remote access to files in cloud storage due to a lack of location-based access restrictions. Therefore, to improve the security of accessibility files in storage, this research aims to develop an enhanced cryptographic method by implementing location information to generate the encryption key using the AES method. The enhanced AES method is developed by using the combination of longitude and latitude coordinates, user password, and device MAC address to generate the encryption key known as geo-key. Then, it is evaluated to verify its performance by conducting a time performance comparison between the existing method and the enhanced AES method using a variety of format types as the data set. The result shows that the enhanced AES method takes 2.73% longer time to execute than the existing AES method takes due to the additional combination of key parameters to generate the geo-key while the existing AES method only required a single parameter to generate the key encryption which is the user password. Next, the enhanced AES method is evaluated by validating file decryption successfulness at a different location where it has shown that only files located at an intended location can be decrypted subject to their toleration range of distance. The other testing has been conducted to evaluate the integrity of decrypted files which showing all files that have been decrypted are having the same hash value as the original file's hash. The results from this third evaluation have proved that the enhanced AES geo-key method keeps the integrity of the original data without any unintended modification and corruption. The enhanced AES method has a significant contribution to protect files in storage if the file has been stolen in the situation of security data breach incident as the encrypted file could never be accessed by an unauthorized owner and only be decrypted inside the intended range of location. This research also aligned with one of the pillar strategies in Malaysia Cyber Security Strategy (MCSS) to establish data leakage protection mechanisms in organization management and business operation.

Keywords: key generation, geo-key, location coordinate, MAC address, AES method.

الملخص

يتعرض أي شيء متصل بالإنترنت للمخاطر السيبرانية بما في ذلك مخاطر إختراق البيانات الشخصية في التخزين السحابي. ارتفع استخدام الإنترنت بشكل كبير على الصعيد الوطني خلال جائحة Covid-19 مما أدى إلى الإبلاغ عن عدد كبير من الحوادث السيبرانية. وفقًا لمسح مستخدمي الإنترنت (IUS)، تقرير 2020 الصادر عن لجنة الاتصالات والوسائط المتعددة الماليزية (MCMC)، نمت نسبة مستخدمي الإنترنت في عام 2020 بنسبة 1.3٪ من 87.4٪ في عام 2018 إلى 88.7٪ في عام 2020 بينما بلغ إجمالي عدد مستخدمي الإنترنت 8669. تم الإبلاغ عن الحوادث إلى Cyber Security Malaysia في عام 2021. نتيجة لتزايد عدد مستخدمي الإنترنت، تم تحديد ثلاث مشكلات رئيسية تتعلق بأمن البيانات والتي تحتاج إلى التعزيز وهي (1) تخزين الملفات الآمن، (2) مفتاح التشفير الضعيف و (3) الوصول غير المصرح به إلى الملفات عن بعد في التخزين السحابي. أحد أقوى الخيارات لتوفير الأمان لحماية الملفات في التخزين هو تنفيذ طريقة تشفير تتضمن عملية التشفير وفك التشفير. ومع ذلك، لا تزال طريقة التشفير القوية مثل طريقة AES متاحة للهجوم إذا كان مفتاح التشفير ضعيفًا. كما أدى التكيف مع بيئة ما بعد الجائحة من خلال العمل عن بُعد من المكتب الفعلي إلى جذب الوصول غير المصرح به عن بُعد إلى الملفات الموجودة في التخزين السحابي بسبب الافتقار إلى قيود الوصول القائمة على الموقع. لذلك، لتحسين أمان ملف الوصول في التخزين، يهدف هذا البحث إلى تطوير طريقة تشفير محسّنة من خلال تنفيذ معلومات الموقع لإنشاء مفتاح التشفير باستخدام طريقة AES. تم تطوير طريقة AES المحسّنة باستخدام مجموعة إحداثيات خطوط الطول والعرض وكلمة مرور المستخدم وعنوان MAC للجهاز لإنشاء مفتاح التشفير المعروف باسم المفتاح الجغرافي. بعد ذلك، يتم تقييمها للتحقق من أدائها من خلال إجراء مقارنة أداء الوقت بين الطريقة الحالية وطريقة AES المحسّنة باستخدام أنواع تنسيقات متنوعة كمجموعة بيانات. توضح النتيجة أن طريقة AES المحسّنة تستغرق وقتًا أطول في التنفيذ بنسبة 2.73٪ مما تستغرقه طريقة AES الحالية نظرًا للجمع الإضافي من المعلمات الرئيسية لإنشاء المفتاح الجغرافي بينما تتطلب طريقة AES الحالية معلمة واحدة فقط لإنشاء تشفير المفتاح وهي كلمة مرور المستخدم. بعد ذلك، يتم تقييم طريقة AES المحسّنة من خلال التحقق من نجاح فك تشفير الملف في موقع مختلف حيث أظهر أن الملفات الموجودة في الموقع المقصود فقط يمكن فك تشفيرها وفقًا لنطاق التسامح الخاص بها. تم إجراء الاختبارات الأخرى لتقييم سلامة الملفات التي تم فك تشفيرها والتي تبين أن جميع الملفات التي تم فك تشفيرها لها نفس قيمة التجزئة مثل تجزئة الملف الأصلي. أثبتت نتائج هذا التقييم الثالث أن طريقة المفتاح الجغرافي AES المحسّنة تحافظ على سلامة البيانات الأصلية دون أي تعديل أو تلف غير مقصود. تساهم طريقة AES المحسّنة بشكل كبير في حماية الملفات المخزنة في حالة سرقة الملف في حالة وقوع حادث خرق لبيانات الأمان حيث لا يمكن مطلقًا الوصول إلى الملف المشفر بواسطة مالك غير مصرح له وفك تشفيره فقط داخل النطاق المقصود من الموقع. يتماشى هذا البحث أيضًا مع إحدى الاستراتيجيات الأساسية في استراتيجية الأمن السيبراني في ماليزيا (MCSS) لإنشاء آليات حماية تسرب البيانات في إدارة المؤسسة وتشغيل الأعمال.

الكلمات المفتاحية: توليد المفتاح، المفتاح الجغرافي، إحداثيات الموقع، عنوان MAC، طريقة AES.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

n	Prime factorization
p,q	Prime numbers
$\emptyset\ (n)$	Totient function
a, b	Greatest common divisor
c	Cipher text
m	Original message
λ	Longitude in radians of the location to project
φ	Latitude in radians of the location to project
φ_1	Standard parallels
λ_0	Central meridian of the map
x	Horizontal coordinate on the map
y	Vertical coordinate on the map
R	Radius of the globe
D_T	Distance threshold limit
(x, y)	Coordinate location (lat, lon)
A	Original location
B	Requested location
$\Delta\ T$	Variant distance between 2 point of A and B
V	Distance validity

LIST OF EQUATIONS