مقاله مروری
الگوریتم‌های رمزگذاری چندلایه بر روی داده‌های مراقبت‌های بهداشتی

سمیه کرباسی[1]

[1]دانشگاه پیام‌نور تهران پست الکترونیکی: Somayeh.Karbasy@gmail.com

چکیده

در مورد محرمانگی داده‌های مراقبت‌های بهداشتی (PHI) هنگامی که در محیط ابر ذخیره می‌شوند، سیاست‌های درستی باید اعمال شود. این اطلاعات می‌تواند به دلیل ذخیره‌سازی در قالب ساده یا با استفاده از الگوریتم‌های رمزگذاری ضعیف، به خطر بیفتد. بعضی از اطلاعات مربوط به بیمار ضرورت دارد که محافظت شده باشد و توسط افراد غیرمجاز دیده نشده یا تغییر داده نشوند. هدف اصلی این گزارش ارائه یک الگو و روش مطمئن برای حفظ محرمانگی داده‌های بیماران است. چون داده‌ها، به صورت همیشگی در محیط ابری در دسترس هستند. این امر با رمزگذاری و رمزگشایی داده‌ها به صورت چند لایه به دست می‌آید. الگوریتم‌هایی که برای رمزگذاری استفاده خواهیم کرد؛ الگوریتم‌های استاندارد هستند که توسط NIST توصیه می‌شوند. در این‌جا هدف استفاده ار الگوریتم‌های چندگانه برای حفظ محرمانگی داده‌ها است. در پایان‌نامه مورد بررسی، برای اطمینان از محرمانه بودن اطلاعات ذخیره شده در محیط ابر، یک روش رمزگذاری چند لایه را پیشنهاد شده است. این تکنیک پیشنهادی در صورت استفاده در قالب چند لایه، امنیت تکنیک های رمزنگاری را بهبود می‌بخشد. یک سیستم محلی برای پژوهش تنظیم کرده است. از پایگاه داده رابطه‌ای و فریم‌ورک ۴,۵ استفاده کرده است. مجموعه‌ای از ۵۰۰ پرونده ساختگی بیمار برای استفاده از روش‌های پیشنهادی استفاده می‌شود. این پژوهش برای بررسی محرمانگی روش‌های پیشنهادی انجام شده است. این پژوهش به ما نشان می‌دهد که وقتی داده‌ها در محیط ابری هستند، تکنیک‌های رمزگذاری چند لایه برای بخش‌های بهداشت عمومی مناسب‌ترند.[۴]

کلمات کلیدی: رایانش ابری ، داده های مراقبت‌های بهداشتی ، رمزگذاری، کلید

۱. مقدمه

رایانش ابری، سرویس محاسباتی مورد تقاضا است.(منابع محاسباتی در صورت تقاضا و در حد نیاز در دسترس هستند.) که بیشترین امکانات را در اختیار بخش‌های مراقبت‌های بهداشتی مورد نیاز بیماران قرار می‌دهد. داده‌ها به ساده‌ترین صورت و حتی از راه دور ذخیره و بازیابی می‌شوند و امکان تغییر دارند. در واقع بیمارستان‌ها نیازی به ذخیره‌سازی محلی داده‌ها ندارند و فقط کافی است سرور مورد نیاز جهت دسترسی به اطلاعات را بخرند. هدف اصلی رایانش ابری، به اشتراک گذاری منابع و دسترسی بهینه به آن‌ها است.

۲. پیشینه تحقیق

در اوایل دهه ۱۹۶۰ معماری سرور مشتری فقط برای رایانه‌های اصلی و کلاینت مورد استفاده قرار گرفت. در آن زمان ذخیره اطلاعات بسیار گران بود. هزینه CPU نیز بسیار زیاد بود. به همین دلیل از Mainframe برای ذخیره سازی و پردازش استفاده می‌شد. برای دسترسی به داده‌ها و پردازش، از ترمینال‌های تخلیه استفاده می‌شد.

در سال ۲۰۰۶ آمازون شروع به فعالیت خود در زیر شاخه‌ای به نام خدمات وب آمازون کرد.

گوگل نسخه آزمایشی Google App Engine را در آوریل ۲۰۰۸ منتشر کرد. در همان سال ناسا OpenNebula را نیز معرفی کرد. این اولین پروژه منبع آزاد بود که برای خصوصیات ابرهای ترکیبی به کار گرفته شد.

در سال ۲۰۱۰ مایکروسافت Azure توسط مایکروسافت منتشر شد.

در سال ۲۰۱۲، موتور محاسبه Google قبل از اینکه در دسامبر ۲۰۱۳ در دسترس عمومی قرار بگیرد، در حالت پیش‌نمایش منتشر شد

در سال ۲۰۱۱ دکتر سلیمان و همکاران در مقاله‌ای، روش چندلایه‌ای را برای خدمات سلامت الکترونیکی مطابق با سند ISO۱۷۷۹۹ تعریف کرده و اطلاعات را به سه دسته اطلاعات سری، بسیار محرمانه و خصوصی تقسیم کرده؛ الگوریتم‌های رمزگذاری متقارن، DES۳ و تابع مقدار هش را معرفی کرده اند. نویسندگان از اندازه کلید ۱۹۳ بیت برای لایه ۱ و ۱۲۹ بیت تا ۱۹۲ بیت برای لایه ۲ و ۱۱۲ تا ۱۲۸ برای لایه ۳ و ۸۰ تا ۱۱۱ بیت برای لایه ۴ استفاده کرده‌اند. کار اصلی نویسندگان روی الگوریتم DES۳ است. و از یک الگوریتم برای رمزگذاری و رمزگشایی استفاده می‌شود.[۶]

در سال ۲۰۱۳ کیا و همکاران در مقاله‌ای با استفاده از SOAP / XML داده‌ها را با AES رمزگذاری کرده اند.[۲]

در سال ۲۰۱۶ ژو و همکاران نویسندگان به خوبی مدل مراقبت‌های بهداشتی جدیدی را برای ذخیره داده‌های ابری در نظر گرفته‌اند. آن‌ها RBE (رمزنگاری مبتنی بر نقش) را اعمال کرده‌اند. ابتدا، آن‌ها مدل PCEHR (سوابق الکترونیکی کنترل الکترونیکی شخصی) را که توسط دولت استرالیا معرفی شده شرح داده‌اند. سپس PCEHR در RBE برای امنیت داده استفاده می‌شود. آن‌ها ساختار آرم داده‌ها و ویژگی‌هایش را بر اساس رمزگذاری طراحی می‌کنند و ادعا کردند که رویکرد آنها کنترل انعطاف‌پذیری در ذخیره‌سازی داده‌ها را فراهم می‌کند.[۸]

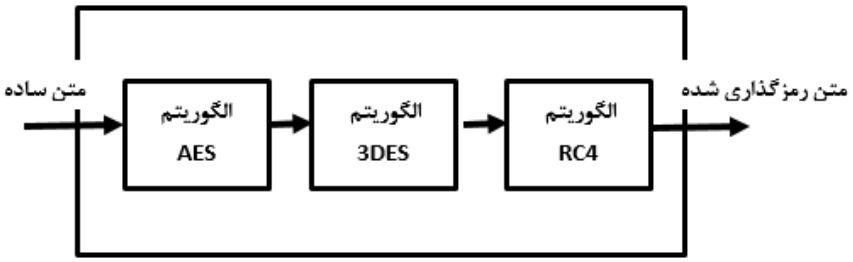
در سال ۲۰۱۹ سودهیپ و همکار در مقاله‌ای رمزگذاری مبتنی بر ویژگی سیاست رمزگذاری (CP-ABE) را معرفی کرده‌اند. کلید رمزگذاری شامل خط‌مشی‌هایی است و آن‌ها می‌گویند اگر کلید هک شده‌باشد، آندسته از سوابق رمزگشایی می‌شوند که کلید آن‌ها هک می‌شود اما بقیه موارد همچنان محافظت می‌شوند.[۵]

در سال ۲۰۱۹ هما و همکار ، درباره روش رمزنگاری منحنی بیضوی بحث کردند و روش‌های تولید کلید اصلی شخص ثالث را معرفی کردند. مالک داده را برای درخواست کلید و رمزگذاری سند به صورت آنلاین به بخش دیگر ارسال می‌کند. شخص ثالث رمز را رمزگذاری و به صاحب داده ارسال و مالک تاریخ را در سرور ابری بارگذاری می‌کند و کلید را برای استفاده در آینده نگه می‌دارد.[۱]

در سال ۲۰۱۹ پارا و همکاران از تکنیک‌هایی استفاده کردند که در آن، آرم داده‌ها با استفاده از درون‌یابی خطی ایجاد شده و سپس مستطیل جادویی با استفاده از الگوریتم LSB ایجاد و با استگانوگرافی، داده‌ها را رمزگذاری کردند.[۳]
در سال ۲۰۱۹ وزید و همکاران نویسندگان در مورد مدل تهدید و احراز هویت برای دستگاه های مبتنی بر Iot در محیط ابر بحث کرده‌اند و سعی کرده‌اند چالش‌های فعلی امنیت و داده‌های مبتنی بر اینترنت اشیا در ابر را بررسی کنند. تمرکز اصلی آنها در تحقیق، سازوکار احراز هویت است و مفهوم مجازی تکنیک جدید را ارائه داده‌اند. در مقاله خود یک مطالعه تطبیقی در مورد هزینه‌های ارتباطی و فنی و حرفه ای انجام داده‌اند. محاسن و معایب تکنیک‌های احراز هویت موجود نیز در دست بررسی است اما راه‌حل مشخصی پیشنهاد نمی‌شود.[۷]

۳. راه‌اندازی آزمایشی طرح پیشنهادی

ابتدا به تکنیک چندلایه رمزگذاری (شکل ۱) توجه کنید؛ در این حالت متنی که باید محافظت شود؛ وارد سیستم الگوریتم چندگانه می‌شود. داخل سیستم الگوریتم چندگانه در این مورد خاص، ۳ الگوریتم AES و DES۳ و RC٤ وجود دارد. خروجی هر الگوریتم، به صورت رمزگذاری شده و دوباره توسط الگوریتم بعدی طبق شکل، رمزگذاری می‌شود و خروجی نهایی با رمزگذاری چندلایه و ایمنی بالا ایجاد می‌شود.



شکل ۱ : تکنیک محافظت از چند لایه

برای توسعه طرح یک مجموعه داده ساختگی(برای ایمنی بیمار) از حدود ۵۰۰ بیمار انتخاب کرد؛ روی داده‌ها، الگوریتم‌های رمزگذاری موجود در RDBMS اعمال می‌شوند.و داده‌ها در محیط ابری ذخیره می‌شوند.

در آغاز ثبت نام بیماران، شماره پرونده پزشکی با با رمز عبور پیچیده‌ای که به طور تصادفی ایجاد شده است، برای اطلاعات بیمار اختصاص می‌یابد. برای دسترسی در وب سایت، بیمار شماره پرونده پزشکی(MR No) را به عنوان نام کاربری و رمز ورود وارد می‌کند و روی Login کلیک می‌کند. اگر نام کاربری معتبر باشد و رمز عبور آن درست باشد، پس از رمزگشایی نسخه پزشک برای وی نمایش داده می‌شود. الگوریتم رمزگذاری متقارن AES را با ترکیب کلیدهای مختلف و ۳DES بر روی داده‌ها اعمال می‌کنیم زیرا کلید در RDBMS ذخیره می‌شود و توسط سرور Microsoft SQL محافظت می‌شود و از رمز عبور محافظت می‌کند.

۴. فرآیند رمزگذاری و رمزگشایی در RDBMS

فرایند رمزگذاری کلی با استفاده از SQL Server بر روی یک ستون مطابق مراحل زیر است. کلید اصلی در پایگاه‌داده، براساس روش متقارن محافظت می‌شود.

مرحله ۱ : ایجاد کلید اصلی

ابتدا باید کلید اصلی با رمز عبور مناسب ایجاد شود و سپس گواهینامه بر اساس کلید اصلی تولید می‌شود.

مرحله ۲ : ایجاد گواهی

گواهینامه دیجیتالی جهت محافظت از کلید اصلی پایگاه‌داده ایجاد می‌شود.

مرحله ۳ : کلید متقارن
یک کلید متقارن برای رمزگذاری و رمزگشایی بر اساس الگوریتم‌های رمزگذاری در سرور sql ، ساخته می‌شود؛ به عنوان مثال AES۱۲۸ ، AES۱۹۲ و AES۲۵۶ .
مرحله ۴ : رمزگذاری ستون‌ها
شمای جدول و فیلدهای را رمزگذاری می‌کنیم.

۵. تجزیه و تحلیل طرح پیشنهادی

مجموعه داده ساختگی از ۵۰۰ بیمار را برای هدف آزمایش آماده شده است. برخی از ویژگی‌ها با در نظر گرفتن *GDPR* و *HIPAA* (قانون حفاظت اطلاعات آمریکا و اروپا)گرفته شده است. به عنوان مثال *MR No* (شماره پرونده پزشکی)، نام، نام نسبی، جنسیت، آدرس، تاریخ تولد، تاریخ ثبت، *NIC*، شماره تلفن همراه و شماره حساب/اطلاعات کارت اعتباری. این ویژگی‌ها به ویژه هنگامی که داده ها در فضای ابری قرار دارند، نیاز به مراقبت بیشتری دارند. برای افزایش سطح محرمانگی، خصوصیات ویژه *PHI* را برای رمزگذاری و رمزگشایی در نظر گرفته شده است.

۶. نصب پیکربندی سخت‌افزار و نرم‌افزار

سخت‌افزار مورد نیاز

- پردازنده *Intel Core i7-6500U Processor*
- *Ram* ۸ گیگابایتی
- هارد دیسک ۵۰۰ گیگابایتی

سیستم عامل و نرم‌افزار مورد نیاز

- ویندوز ۱۰ یا بالاتر
- *Visual Studio* ۱۲ یا ۱۵
- *SQL Server 2014* یا بالاتر
- *Framework* ۴,۵

۷. ایجاد کلیدها و گواهینامه‌ها و رمزگذاری داده‌ها

ایجاد کلید و گواهینامه‌ها با دستورات SQL انجام می‌شود. داده‌های مربوط به بیمار که باید روی ابر بارگذاری شوند؛ تهیه شده و فرآیند رمزگذاری روی آن اعمال می‌شود. داده‌ها در جدول پایگاه‌داده با فرمت رمزگذاری‌شده، ذخیره می‌شوند.

۸. رمزگشایی داده‌ها

بیمار وارد سایت مورد نظر می‌شود که در برگه ثبت نام چاپ شده است و از شماره پرونده پزشکی به عنوان نام کاربری و رمز عبور استفاده می‌کند و بر روی ورود کلیک می‌کند. جزئیات پرونده پزشکی بیمار، نشان داده شده است.

۹. تحلیل نتایج

- اگر الگوریتم رمزگذاری منفرد AES با اندازه کلید ۱۲۸ بیتی اعمال شود، کل زمان سپری شده ۲۵۰ میلی‌ثانیه خواهدبود.
- اگر AES با اندازه کلید ۱۹۲ بیتی اعمال شود، کل زمان سپری شده ۲۱ میلی‌ثانیه خواهد بود.
- اگر AES با اندازه کلید ۲۱۸ بیتی اعمال شود، کل زمان سپری شده ۳۲۱ میلی‌ثانیه خواهد بود.
- اگر ۳ DESبه تنهایی اعمال شود، کل زمان سپری شده ۲۶۰ میلی‌ثانیه خواهد بود.

نتایج حاصل از ترکیب چند الگوریتم با مجموعه داده‌های مشابه نیز در این شکل نشان داده شده است.

- اگر ترکیبی ازAES ۱۲۸ و ۱۹۲ استفاده شود؛ زمان سپری شده ۵۲۱ میلی‌ثانیه خواهد بود.
- اگر ترکیبی ازAES ۱۲۸ و ۲۵۶ استفاده شود؛ زمان سپری شده ۴۷۴ میلی‌ثانیه خواهد بود.
- اگر ترکیبی ازAES ۱۲۸ و ۳ DESاستفاده شود، زمان سپری شده ۲۹ میلی‌ثانیه خواهد بود.
- اگر ترکیبی ازAES ۱۹۲ وAES ۲۵۶ استفاده شود، مدت زمان سپری شده ۳۵ میلی ثانیه خواهد بود.
- اگر ترکیبی ازAES ۱۹۲ و ۳ DESاستفاده شود، مدت زمان سپری شده ۸۶ میلی‌ثانیه خواهد بود.
- اگر ترکیبی ازAES ۲۵۶ و ۳ DESاستفاده شود، مدت زمان سپری شده ۱۶۳ میلی ثانیه خواهدبود.

زمان پردازنده را در میلی‌ثانیه برای ۵۰۰ رکورد با الگوریتم‌های رمزگذاری منفرد مقایسه می‌کنیم.

- ۱۲۸AES زمان پردازنده ۲۰۳ میلی ثانیه را می‌گیرد.
- ۱۹۲AES زمان پردازنده ۷۸ میلی‌ثانیه است.
- ۲۵۶AES ، ۷۲میلی‌ثانیه طول می‌کشد.
- ۳DES، ۷۸ میلی‌ثانیه طول می‌کشد.

نتایج زمان پردازنده برای چندین ترکیب الگوریتم با مجموعه داده‌های مشابه نیز نشان داده شده است.

- ۱۲۸AES و ۱۹۲، ۱۴۱ میلی‌ثانیه طول می‌کشد.
- ۱۲۸AES وAES ۲۵۶ ، ۹۴ میلی‌ثانیه طول می‌کشد.
- ۱۹۲AES وAES ۲۵۶ ۳۱ میلی‌ثانیه طول می‌کشد.
- ۱۹۲AES و ۳ DES، ۱۵ میلی‌ثانیه طول می‌کشد.
- ۲۵۶AES و ۳ DES، ۱۱۰ میلی‌ثانیه طول می‌کشد.

اگرچه زمان پردازنده با ۲۵۶AES و ۳DES به زمان CPU بیشتری نیاز دارد اما زمان سپری شده ۲۵۶AES و ۳DES طرح بهتری را برای رویکردهای چندلایه ارائه می‌دهد زیرا رمزگذاری فقط بارگذاری اطلاعات را انجام می‌دهد. برای بهترین سطح محرمانگی ۲۵۶AES با ۳DES مناسب است.

۱۰. دستورالعمل‌های موجود

قانون HIPAA و GDPR

HIPAA یک کلمه اختصاری به مفهوم " قابلیت حمل و پاسخگویی بیمه درمانی است". این مصوبه قوانین مختلفی را درباره حفاظت از داده‌های بیمار ارائه می‌دهد.

۱۸ ویژگی زیر اعمالی که باید محافظت شوند را مشخص می‌کند:

- نام و نام خانوادگی بیمار
- آدرس شامل کد پستی، شهر، کشور
- همه تاریخ‌ها
- شماره تلفن
- نمابر

- شناسه ایمیل
- شماره بیمه
- سوابق پزشکی شماره
- اطلاعات کارت سلامت
- حساب بانکی بدون/ اطلاعات کارت اعتباری
- گواهینامه یا گواهینامه رانندگی
- شماره خودرو
- شناسه دستگاه و شماره سریال
- آدرس وب
- آدرس پروتکل اینترنت
- بیومتریک
- هر نوع تصویری
- هر مشخصه دیگری که بتواند منحصراً فرد را شناسایی کند.

GDPR (مقررات عمومی حفاظت از داده ها) مقررات اتحادیه اروپا است که در سال ۲۰۱۶ پذیرفته شده است. پس از سال ۲۰۱۸ این قانون برای کلیه سازمان های کشورهای اتحادیه اروپا اجباری شده است که ذخیره اطلاعات شخصی فرد را باید مطابق با GDPR باشد.

۱۱. پیشنهادها

زمانی که به عنوان سمینار دانشجویی، کار بر روی این پایان‌نامه را شروع کردم؛ جذب موضوع آن شدم. امنیت داده‌های مربوط به بیمار و حفظ و نگه‌داری اطلاعات برای استفاده مجدد بیمار و پزشک. دانشجویان ارشد الگوریتم‌های رمزگذاری را به خوبی می‌دانند اما ترکیب آن‌ها و رسیدن به حالت ایده‌آل کاری است که راستای فکری این پایان‌نامه بوده است. طراحی نرم‌افزار به کاررفته در این سیستم، تقریبا راحت بوده و نوشتن برنامه رمزنگاری به توجه به ترکیب الگوریتم‌های راحت است. موردی که نیاز به دقت دارد و در واقع نیرو محرکه کار و عامل برتری طرح می‌باشد؛ ژیدا کردن بهترین ترکیب‌های الگوریتم‌های رمزگذاری است که بتوان به صورت چند لایه استفاده کرد.

در مورد مشخصات سخت‌افزاری مورد نیاز طرح، اکثر سیستم‌های موجود شرایط لازم را دارند و نیاز به توسعه خاصی نیست ولی مشخصات نرم‌افزاری ممکن است مشکلاتی از لحاظ هزینه برای توسعه وجود آورد؛ به طور مثال ویندوز ۱۰ ممکن است روی بعضی سیستم‌ها قابل نصب نباشد. لذا یکی از پیشنهادهای من تبدیل دستورات به صورت است که روی مشخصات پایین‌تر نرم‌افزاری قابل نصب و اجرا باشد. یک حالت موثر می‌تواند وجود چند نسخه با تاکید بر روی مشخصات بالاتر باشد که احیانا در صورت وجود مشکلات زیر ساختی قابل اجرا باشد. موردی که جای کار بیشتر دارد و مورد نیاز است توجه به داده‌های غیرمتنی بیمار است که با تکنیک‌های بروز رمزگذاری، ایمن‌تر باشند.

پیشنهاد دیگری که مربوط به نگارش پایان‌نامه است؛ استفاده از شیوه ارجاع به منابع می‌باشد. به نظر بنده، شیوه ارجاع بهتر است مطابق راهنماری نگارش پایان‌نامه معاونت آموزشی و تحصیلات تکمیلی دانشگاه پیام‌نور باشد. شیوه ارجاع در این پایان‌نامه به صورت لینک به منابع بود و شخصا کار مشکلی در درک مفاهیم آن داشتم.

۱۲. ارائه ایده برای پایان‌نامه‌های جدید تکمیلی

- بررسی و تست بهترین ترکیب الگوریتم‌های رمزگذاری از لحاظ سرعت،هزینه،کارایی و ... روی داده‌های مراقبت‌های بهداشتی برای رسیدن به مطلوب‌ترین نتیجه
- پیاده‌سازی الگوریتم‌های رمزگذاری بومی ایران روی داده‌های مراقبت‌های بهداشتی
- رمزگذاری چندلایه کارا روی داده‌های تصویری مراقبت‌های بهداشتی
- ساخت برنامه امنیتی Open Source رمزگذاری چند لایه روی انواع سیستم ها

۱۳. نتیجه‌گیری

در این مقاله به موضوع اجرای رمزگذاری چند لایه داده‌های مراقبت‌های بهداشتی پرداخته شد؛ انواع ترکیب‌ها تست شد و با ذکر سخت‌افزار و نرم‌افزار مورد نیاز، طرح پیاده‌سازی شده؛ بررسی شد. در ادامه پیشنهادات جهت ادامه طرح داشتم و موضوعات پایان‌نامه‌ها و سمینارهای بعدی در راستای این مقاله مطرح شد.

مراجع

١. V. S. V. Hema and R. Kesavan, "Ecc based secure sharing of healthcare data in the health cloud environment," Wireless Personal Communications, vol. ١٠٨, no. ٢, pp. ١٠٢١–١٠٣٥, ٢٠١٩.

٢. M. M. Kiah, M. S. Nabi, B. Zaidan, and A. Zaidan, "An enhanced security solution for electronic medical records based on aes hybrid technique with soap/xml and sha-١," Journal of medical systems, vol. ٣٧, no. ٥, p. ٩٩٧١, ٢٠١٣.

٣. S. A. Parah, A. Bashir, M. Manzoor, A. Gulzar, M. Firdous, N. A. Loan, and J. A. Sheikh, "Secure and reversible data hiding scheme for healthcare system using magic rectangle and a new interpolation technique," in Healthcare Data Analytics and Management. Elsevier, ٢٠١٩, pp. ٢٦٧–٣٠٩

٤. H.A.Shah. (٢٠٢٠). "A Multilayer Encryption Model To Protect Healthcare Data in Cloud Environment". (Unpublished master's thesis). University of Islamabad

٥. K. Sudheep and S. Joseph, "Review on securing medical big data in healthcare cloud," in *٢٠١٩ ٥th International Conference on Advanced Computing & Communication Systems (ICACCS)*. IEEE, ٢٠١٩, pp. ٢١٢–٢١٥

٦. R. Sulaiman, D. Sharma, W. Ma, and D. Tran, "A new security model using multilayer approach for e-health services," Journal of Computer Science, vol. ٧, no. ١١, pp. ١٦٩١–١٧٠٣, ٢٠١١.

٧. M. Wazid, A. K. Das, R. Hussain, G. Succi, and J. J. Rodrigues, "Authentication in cloud-driven iot-based big data environment: Survey and outlook," Journal of Systems Architecture, vol. ٩٧, pp. ١٨٥-١٩٦, ٢٠١٩.

٨. L. Zhou, V. Varadharajan, and K. Gopinath, "A secure role-based cloud storage system for encrypted patient-centric health records," *The Computer Journal*, vol. ٥٩, no. ١١, pp. ١٥٩٣–١٦١١, ٢٠١٦.