



ارائه یک رویکرد مبتنی بر بلاک چین برای خودکارسازی تضمین صحت و محرمانگی داده‌های ثبت رخداد

فاطمه آملی^۱، مصطفی بستام^{۲*}، احسان عطائی^۳

دانش‌آموخته کارشناسی ارشد، گروه مهندسی کامپیوتر، دانشکده فناوری و مهندسی، دانشگاه مازندران، بابلسر، ایران^۱

استادیار، گروه مهندسی کامپیوتر، دانشکده فناوری و مهندسی، دانشگاه مازندران، بابلسر، ایران^{۲*}

دانشیار، گروه مهندسی کامپیوتر، دانشکده فناوری و مهندسی، دانشگاه مازندران، بابلسر، ایران^۳

چکیده

با افزایش تهدیدات سایبری، داده‌های ثبت رخداد به عنوان منبعی کلیدی برای شناسایی و تحلیل حوادث امنیتی شناخته می‌شوند و بینش‌های ارزشمندی از فعالیت‌های سامانه ارائه می‌دهند؛ با این حال، هرگونه دست‌کاری در این داده‌ها می‌تواند دقت تحلیل‌های امنیتی را کاهش داده و تصمیم‌گیری‌های مرتبط با امنیت را تحت تأثیر قرار دهد. فناوری بلاک چین با ویژگی‌هایی نظیر غیرمتمرکز بودن، شفافیت و تغییرناپذیری بستری قابل اعتماد برای تضمین صحت داده‌ها فراهم می‌کند. در این پژوهش، به جای ذخیره‌سازی مستقیم داده‌های خام که پرهزینه و محدودکننده است، چارچوبی خودکار معرفی شده است که از بلاک چین عمومی اتریوم و قراردادهای هوشمند برای ذخیره‌سازی رمزنگاری شده داده‌های ثبت رخداد استفاده می‌کند. این روش با کاهش هزینه‌های ذخیره‌سازی، در عین حفظ محرمانگی و امکان راستی‌آزمایی داده‌ها، کارایی بالایی ارائه می‌دهد. فرایند تضمین صحت داده‌ها در دو مرحله انجام می‌شود: ثبت و مقایسه دوره‌های هش‌ها و اعتبارسنجی دسته‌ای در بازه‌های زمانی بلندتر برای کشف هرگونه دست‌کاری احتمالی. ارزیابی این مدل در شبکه آزمایشی سیپولیا نشان داده است که هزینه‌های عملیاتی و سربار پردازشی بهینه شده و امکان استفاده از این روش در مقیاس وسیع فراهم است. این پژوهش، روشی نوآورانه و عملی برای خودکارسازی تضمین صحت داده‌های ثبت رخداد ارائه می‌دهد و راه‌کاری قابل اتکا برای ارتقای امنیت اطلاعات در کاربردهای واقعی پیشنهاد می‌کند.

واژگان کلیدی: مدیریت داده‌های ثبت رخداد، صحت داده‌ها، بلاک چین، اتریوم، قرارداد هوشمند.

A Blockchain-Driven Approach to Automating Event Log Data Integrity and Confidentiality

Fatemeh Amoli¹, Mostafa Bastam^{2*}, Ehsan Ataie³

M.Sc. Graduate, Department of Computer Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran¹

Assistant Professor, Department of Computer Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran^{2*}

Associate Professor, Department of Computer Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran³

Abstract

With the rapid rise of cybersecurity threats and the increasing complexity of digital security, event log data serves as a critical source for identifying and analyzing cyberattacks and threats. This data provide key insights into system activities, essential for detecting unauthorized intrusions, analyzing suspicious behaviors, and conducting security investigations. However, any alteration or tampering with the data can disrupt the analysis and detection processes, leading to incorrect security decisions.

* Corresponding author

* نویسنده عهده‌دار مکاتبات

Blockchain technology, with its unique features such as decentralization, immutability, and transparency, has been recognized as a reliable and secure platform for storing and protecting data. This technology enables the storage of data hashes in a way that any changes can be easily detected. However, directly storing the vast volume of event log data on the blockchain faces challenges such as high costs and storage space limitations.

In this research, an innovative model has been presented to automate the assurance of event log data integrity and confidentiality using the public Ethereum blockchain and smart contracts. Instead of storing event log data directly, only their hashes have been saved on the blockchain. This approach not only reduces storage costs but also ensures data confidentiality.

The automated data integrity assurance process in this model occurs in two stages:

1. Stage One: Event log data hashes have been periodically stored on the blockchain and compared with previous hashes.
2. Stage Two: Over longer intervals, all stored hashes have been reviewed and validated to prevent any potential tampering.

In this study, the costs associated with implementing this model on the Ethereum Sepolia test network had been precisely calculated. The analysis indicates that operational costs and computational overhead have been optimized across different time intervals, demonstrating the model's feasibility for large-scale deployment.

Ultimately, this research tries to introduce a novel and practical model, taking a significant step toward automating the assurance of event log data integrity and confidentiality, providing a reliable solution for real-world applications.

Keywords: Log management, Data integrity, Blockchain, Ethereum, Smart Contract.

مشکلاتی را در استفاده گسترده از این فناوری ایجاد کرده است [۴، ۵]. در بسیاری از پژوهش‌ها، راه‌حلهایی برای کاهش هزینه‌ها و بهینه‌سازی فرایند ذخیره‌سازی بلاک‌چین ارائه شده است؛ اما بیشتر این مطالعات به جنبه‌های خودکارسازی و خودکفاکردن فرایندهای ذخیره‌سازی و تضمین صحت داده‌ها توجه نکرده‌اند.

این پژوهش با هدف رفع این کمبود، مدلی نوآورانه برای تضمین صحت و یکپارچگی داده‌های ثبت رخداد ارائه می‌دهد که از ویژگی‌های بلاک‌چین عمومی اتریوم^۳ و قراردادهای هوشمند^۴ بهره می‌برد. این مدل فرایند ثبت هش^۵ داده‌های رخداد را به‌صورت خودکار و غیرمتمرکز انجام داده و از این طریق، نه تنها صحت داده‌ها را به‌طور مؤثر تضمین می‌کند، بلکه هزینه‌های ذخیره‌سازی و زمان پردازش را نیز به کمینه می‌رساند. در این پژوهش، علاوه بر ارائه مدل پیشنهادی، پیاده‌سازی این مدل و تحلیل هزینه‌های آن نیز مورد بررسی قرار می‌گیرد. مقاله حاضر به شرح زیر ساختار یافته است: در بخش دوم، پژوهش‌های مرتبط بررسی می‌شود. بخش سوم به ارائه مدل پیشنهادی و بخش چهارم به بررسی چگونگی پیاده‌سازی روش پیشنهادی و ارزیابی نتایج حاصله معطوف شده است؛ در نهایت، در بخش پنجم، نتیجه‌گیری و پیشنهادات برای پژوهش‌های آینده آورده شده است.

۲- کارهای مرتبط

مطالعات پیشین به‌طور گسترده‌ای به ارائه راه‌کارهای مختلف برای استفاده از بلاک‌چین به‌منظور تضمین صحت داده‌های

۱- مقدمه

با گسترش روزافزون تهدیدات سایبری و پیچیدگی‌های امنیتی در دنیای دیجیتال امروزی، سازمان‌ها بیش‌ازپیش نیازمند ابزارهای مؤثر و قابل اعتماد برای حفاظت از داده‌های حساس خود هستند؛ در این راستا، داده‌های ثبت رخداد^۱ به‌عنوان یکی از منابع کلیدی برای شناسایی و مقابله با تهدیدات نقش حیاتی دارند. این داده‌ها شامل جزئیاتی از فعالیت‌های سامانه‌ها هستند که می‌توانند اطلاعات مهمی در خصوص حملات سایبری، نقض‌های امنیتی و رفتارهای مشکوک ارائه دهند؛ با این حال، دست‌کاری یا حذف این داده‌ها می‌تواند به راحتی منجر به از بین رفتن شواهد ضروری و ایجاد چالش‌های جدی در فرایندهای پژوهش‌های امنیتی و قضایی شود [۱، ۲]؛ در این میان، فناوری بلاک‌چین^۲ به دلیل ویژگی‌های منحصر به فرد خود مانند غیرمتمرکز بودن، شفافیت و تغییرناپذیری توجه بسیاری از پژوهش‌گران را به خود جلب کرده است. بلاک‌چین به‌عنوان یک بستر ذخیره‌سازی امن و قابل اعتماد می‌تواند فرایند ثبت و حفظ داده‌ها را به شکلی که امکان تغییر یا دست‌کاری آن‌ها وجود نداشته باشد، تضمین کند؛ به این ترتیب، این فناوری قادر است، نقش مهمی در محافظت از یکپارچگی داده‌های ثبت رخداد ایفا و از اعتبار این داده‌ها در فرایندهای تحلیل تهدیدات و حملات سایبری اطمینان حاصل کند [۳]؛ با این حال، استفاده از بلاک‌چین در ذخیره‌سازی داده‌های ثبت رخداد با چالش‌های خاصی روبه‌رو است؛ حجم بالای داده‌های تولیدی و سرعت بالای ایجاد آن‌ها، به‌علاوه هزینه‌ها و محدودیت‌های ناشی از ذخیره‌سازی بر روی بلاک‌چین

³ ETHEREUM

⁴ Smart Contract

⁵ Hashing

¹ Log Files

² Blockchain

سلامت، جاوید و همکارانش [۱۶] یک سامانه امنیتی مبتنی بر بلاک چین برای جلوگیری از دست کاری داده‌های پزشکی پیشنهاد داده‌اند و همچنین خور و همکارانش [۱۷] روشی برای تأیید امنیت داده‌های دستگاه‌های اینترنت اشیا کم‌مصرف ارائه کرده‌اند.

به‌طور خاص، پژوهش‌های [۱۸، ۱۹] رویکردهای مختلفی را برای ذخیره‌سازی ایمن و کارآمد داده‌های ثبت رخداد بر روی بلاک چین پیشنهاد داده‌اند. همه این مقالات بر افزایش امنیت داده‌ها و جلوگیری از دست کاری آن‌ها تمرکز دارند و از روش‌های مختلفی مانند قراردادهای هوشمند، درخت مرکب^۷، IPFS و یادگیری ماشین برای بهبود عملکرد سامانه استفاده می‌کنند. به‌طور خاص، پوتز و همکارانش [۱۸] بر عملکرد بالا و مقاومت در برابر حملات تأکید دارند. علی و همکارانش [۲۰] از Elasticsearch برای ذخیره اسناد JSON استفاده کرده‌اند و از یادگیری ماشین برای شناسایی تهدیدات بهره برده‌اند. لی و همکارانش [۲۱] از قراردادهای هوشمند^۸ ABAC برای کنترل دسترسی به داده‌ها استفاده کرده‌اند و بر جست‌وجوی سریع داده‌ها متمرکز شده‌اند. جین و همکاران [۷] از IPFS برای ذخیره داده‌های خام و بلاک چین برای ذخیره هش‌ها استفاده کرده‌اند و بر کاهش هزینه‌های تراکنش متمرکز شده‌اند؛ همچنین، جیانگ و همکاران [۲۲] از درخت مرکب چهارگانه برای تأیید یک‌پارچگی داده‌ها استفاده کرده‌اند و به کاهش هزینه‌های ذخیره‌سازی در بلاک چین پرداخته‌اند. همه این مطالعات نشان می‌دهند که بلاک چین می‌تواند به‌عنوان یک فناوری کلیدی برای ایجاد سامانه‌های ذخیره‌سازی داده‌های امن و قابل اعتماد عمل کند.

در این پژوهش، روشی نوآورانه برای تضمین امنیت داده‌های ثبت رخداد با استفاده از بلاک چین اتریوم و قراردادهای هوشمند ارائه شده‌است. این روش با خودکارسازی فرایند ثبت هش داده‌های ثبت رخداد از هرگونه دست کاری در داده‌ها جلوگیری می‌کند و به‌طور قابل اعتماد، صحت و اصالت آن‌ها را تأیید می‌کند؛ علاوه بر این، با تحلیل دقیق هزینه‌ها، کارایی و مقرون‌به‌صرفه بودن این روش بررسی شده‌است. در مقایسه با مطالعات پیشین روش پیشنهادی با ارائه یک راه‌کار عملی و قابل اجرا برای محیط‌های واقعی، گامی مهم در جهت کاربردی‌سازی فناوری بلاک چین در حوزه امنیت اطلاعات برداشته‌است. جدول (۱) مقایسه‌ای بین کارهای انجام‌شده و روش پیشنهادی را نشان می‌دهد.

از دیگر امتیازات برجسته این پژوهش می‌توان به پیاده‌سازی عملی مدل پیشنهادی و ارزیابی دقیق نتایج حاصل از آن اشاره کرد که اعتبار و قابلیت اطمینان نتایج پژوهش را به‌طور قابل توجهی افزایش داده‌است.

مختلف پرداخته‌اند و هر یک سعی داشته‌اند چالش‌های مربوط به حجم بالای داده‌ها و هزینه‌های ذخیره‌سازی در بلاک چین را برطرف کنند؛ با این حال، پژوهش‌های موجود بیشتر بر جنبه‌های نظری و فنی تمرکز داشته‌اند و به‌طور محدود به موضوع خودکارسازی فرایند تضمین امنیت داده‌ها و تحلیل هزینه‌های آن پرداخته‌اند. بسیاری از پژوهش‌گران به ترکیب بلاک چین با فناوری‌های مکمل مانند^۱ IPFS پرداخته‌اند تا چالش‌های مقیاس‌پذیری و هزینه‌های ناشی از ذخیره‌سازی داده‌ها در بلاک چین را کاهش دهند. IPFS یک پروتکل ذخیره‌سازی و اشتراک‌گذاری فایل غیرمتمرکز است که از آدرس‌دهی محتوا^۲ برای شناسایی داده‌ها استفاده می‌کند. این سامانه به‌صورت هم‌تابه‌متا عمل می‌کند و داده‌ها را به‌صورت توزیع‌شده بین گره‌ها ذخیره می‌کند که باعث افزایش کارایی، امنیت می‌شود^۳ [۴]. مطالعات [۶-۸] با ذخیره داده‌های حجیم در IPFS و ثبت هش آن‌ها در بلاک چین، به کاهش قابل توجه هزینه‌های ذخیره‌سازی دست یافته‌اند.

در حوزه محاسبات ابری، بلاک چین به‌عنوان یک لایه امنیتی برای حفاظت از داده‌ها مطرح شده‌است. مطالعات [۹، ۱۰] راه‌کارهایی برای استفاده از بلاک چین در تأمین امنیت و یک‌پارچگی داده‌های ثبت رخداد در محیط‌های ابری^۴ ارائه داده‌اند؛ همچنین تاگوچی و همکاران [۱۱] طراحی مبتنی بر بلاک چین برای ذخیره‌سازی داده‌های ثبت رخداد در ابرهای عمومی^۵ را پیشنهاد کرده‌اند. آوسون دیوید و همکاران [۱۲]، چهارچوبی برای کاهش چالش‌های پیش‌روی بازرسان در به‌دست‌آوردن شواهد قابل قبول از اکوسامانه‌های ابری ارائه کرده‌اند که از بلاک چین برای ذخیره‌سازی این شواهد استفاده می‌کند؛ علاوه بر این، تیان و همکاران [۱۳] یک طرح حسابرسی عمومی برای گزارش رفتار کاربران در محیط‌های ابری مشترک پیشنهاد داده‌اند که بر پایه بلاک چین بنا شده‌است.

ترکیب بلاک چین با سایر فناوری‌ها مانند یادگیری ماشین^۵ و اینترنت اشیا^۶ نیز مورد توجه پژوهش‌گران قرار گرفته است. آلتولیان و همکاران [۱۴] با استفاده از قراردادهای هوشمند، امنیت و یک‌پارچگی داده‌های تولیدشده به‌وسیله اینترنت اشیا را بهبود بخشیده‌اند؛ همچنین لی و همکارانش [۱۵] یک سامانه ذخیره‌سازی توزیع‌شده برای داده‌های آتش‌نشانی اینترنت اشیا ارائه داده‌اند که بر پایه بلاک چین و IPFS ساخته شده‌است. در حوزه

^۱ InterPlanetary File System

^۲ Content Addressing

^۳ Cloud

^۴ Public Cloud

^۵ Machine Learning

^۶ Internet of Things

^۷ Merkle tree

^۸ Attribute-Based Access Control contract

(Table-1): Work comparison table

پشتیبانی خودکار	پلتفرم بلاک چین	قرارداد هوشمند	الگوریتم هش	نوع بلاک چین	محرمانگی لاگ	صحت لاگ	مرجع
خیر	اتریوم	بله	SHA256-Keccak256	عمومی	بله	بله	[۱۶]
خیر	Fabric	بله	CID of IPFS	خصوصی	بله	بله	[۱۷]
خیر	EXONUM	بله	SHA256	مجاز	بله	بله	[۲۰]
خیر	Multichain	بله	SHA256	خصوصی	بله	بله	[۲۱]
خیر	اتریوم	بله	SHA256	عمومی	بله	بله	[۷]
خیر	اتریوم	بله	CID of IPFS	عمومی	خیر	بله	[۲۲]
بله	سپولیا اتریوم	بله	SHA256	عمومی	بله	بله	روش پیشنهادی

۳- روش پیشنهادی

در شکل (۱) نمای کلی از روش پیشنهادی آورده شده است. مدل پیشنهادی این مقاله برای تضمین خودکار صحت داده‌های ثبت رخداد از فناوری بلاک چین و قراردادهای هوشمند استفاده می‌کند. این مدل با ذخیره‌سازی هش داده‌ها در بلاک چین از یک پارچگی داده‌ها محافظت می‌کند و در عین حال، محرمانگی آن‌ها را حفظ کرده و امکان تأیید صحت داده‌ها را فراهم می‌آورد؛ علاوه بر این، در این مدل تلاش شده است تا هزینه‌های عملیاتی مرتبط با ذخیره‌سازی داده‌ها و پردازش محاسباتی به‌دقت محاسبه شود. در ادامه، جزئیات این روش شرح داده خواهد شد؛ با این حال، پیش از پرداختن به این جزئیات ضروری است که برخی از مفاهیم و ابزارهای مورد استفاده در این پژوهش تعریف و توضیح داده شوند.

۳-۱- تعاریف

در این بخش برخی از مفاهیم کلیدی و اساسی که در مدل پیشنهادی استفاده شده‌اند، توضیح داده شده است.

بلاک چین: یک فناوری غیرمتمرکز برای ذخیره‌سازی داده‌ها است که به‌طور عمده در ارزهای دیجیتال استفاده می‌شود. این فناوری از زنجیره‌ای از بلوک‌ها تشکیل شده که هر بلوک شامل مجموعه‌ای از تراکنش‌ها یا داده‌ها است. بلاک چین به دلیل ویژگی‌های امنیتی خود مانند تغییرناپذیری، شفافیت و توزیع داده‌ها در میان گره‌های مختلف شبکه برای بسیاری از کاربردهای امنیتی و ذخیره‌سازی داده‌ها مناسب است.

قرارداد هوشمند: یک برنامه رایانه‌ای است که بر روی بلاک چین اجرا می‌شود [۲۲] و به‌طور خودکار شرایط

قرارداد را بر اساس ورودی‌های داده‌شده اجرا می‌کند. این قراردادهای به‌طور معمول به زبان‌های برنامه‌نویسی خاص مانند سالیدیتی نوشته می‌شوند و برای انجام تراکنش‌ها یا پردازش اطلاعات در بلاک چین بدون نیاز به واسطه یا شخص ثالث طراحی شده‌اند [۲۳].

هش: فرایندی است که در آن داده‌های ورودی از هر اندازه به یک مقدار ثابت و کوتاه به نام «هش» تبدیل می‌شوند. هش‌ها برای تضمین یک پارچگی داده‌ها استفاده می‌شوند؛ به این معنی که اگر داده‌های ورودی تغییر کنند، هش خروجی به‌طور قابل توجهی تغییر خواهد کرد. الگوریتم‌هایی مانند SHA-256 برای تولید هش‌ها استفاده می‌شوند. در این مقاله برای تضمین صحت داده‌های ثبت رخداد از هش داده‌ها استفاده شده است.

ذخیره‌سازی داده‌ها در بلاک چین: بر حسب معمول به دلیل هزینه‌های بالا و محدودیت‌های فضایی صورت نمی‌گیرد. در این مقاله تنها هش داده‌های ثبت رخداد در بلاک چین ذخیره می‌شود تا ضمن حفظ محرمانگی داده‌ها از صحت آن‌ها نیز اطمینان حاصل شود. این روش از ذخیره‌سازی داده‌های خام جلوگیری می‌کند و در عین حال امکان تأیید صحت داده‌ها را فراهم می‌آورد.

سپولیا:^۱ یکی از شبکه‌های آزمایشی^۲ اتریوم است که برای آزمایش و ارزیابی قراردادهای هوشمند استفاده می‌شود. این شبکه مشابه شبکه اصلی اتریوم عمل می‌کند، اما تراکنش‌ها و هزینه‌ها در آن به‌صورت آزمایشی انجام می‌شود. سپولیا به توسعه‌دهندگان این امکان را می‌دهد که پیش از استقرار قراردادهای هوشمند در شبکه اصلی اتریوم آن‌ها را آزمایش کنند.

¹ Sepolia

² Testnet

متامسک:^۱ یک کیف پول دیجیتال برای تعامل با بلاک چین‌های اتریوم و سایر بلاک چین‌هاست. این کیف پول به کاربران این امکان را می‌دهد که به راحتی از طریق مرورگر یا اپلیکیشن موبایل خود، به بلاک چین متصل شوند و تراکنش‌ها را انجام دهند؛ همچنین متامسک ابزار مناسبی برای ارتباط با شبکه‌های آزمایشی مانند سیولیا فراهم می‌آورد و برای توسعه‌دهندگان به عنوان یک ابزار مفید برای استقرار و مدیریت قراردادهای هوشمند عمل می‌کند.

۳-۲- شرح روش پیشنهادی

روش پیشنهادی به گونه‌ای طراحی شده است که صحت داده‌های ثبت رخداد را به صورت خودکار و در دو مرحله به کاربر اطلاع می‌دهد. در این روش از بلاک چین به عنوان یک پایگاه داده امن برای ذخیره داده‌های ثبت رخداد استفاده شده است؛ باین حال، باید توجه داشت که ذخیره سازی داده‌ها روی بلاک چین هزینه بر است؛ همچنین به دلیل ویژگی‌های مشخص بلاک چین، این روش با هدف تضمین صحت داده‌های ثبت رخداد از طریق ذخیره سازی آن‌ها بر روی بلاک چین طراحی شده است. داده‌های ثبت رخداد که به عنوان اطلاعات اصلی فعالیت‌ها در سطح شبکه محسوب می‌شوند، نیازمند حفظ محرمانگی‌اند، اما از آنجا که داده‌های ذخیره شده در بلاک چین برای تمامی اعضای شبکه قابل مشاهده‌اند، این مسئله با لزوم حفظ محرمانگی این داده‌ها تناقض دارد؛ بنابراین با در نظر گرفتن دو عامل مهم، یعنی هزینه بالای ذخیره سازی داده در بلاک چین و ضرورت حفظ محرمانگی داده‌های ثبت رخداد پیشنهاد می‌شود به جای ذخیره مستقیم این داده‌ها تنها هش آن‌ها روی بلاک چین ذخیره شود.

در این روش از یک تابع که بر روی یک سرور در حال اجرا است استفاده می‌شود. این تابع از داده‌های ثبت رخداد سامانه در بازه‌های زمانی مشخص هش می‌گیرد و هش‌ها با استفاده از توابع قرارداد هوشمند در یک آرایه روی بلاک چین ذخیره می‌شوند.

برای تضمین خودکار صحت داده‌ها، تابع مورد نظر در بازه‌های زمانی مشخص اجرا شده و با فراخوانی توابع قرارداد هوشمند، هش جدید محاسبه شده از داده‌های ثبت رخداد بر روی شبکه بلاک چین ذخیره می‌شود. در زمان ذخیره هش جدید آخرین هش پیشین با هش مجددی که از داده‌های ثبت رخداد در بازه پیشین گرفته شده است، مقایسه می‌شود. در صورت مغایرت، به کاربر

اطلاع داده می‌شود و به این ترتیب، به صورت خودکار از تغییرات صورت گرفته در داده‌های ثبت رخداد آگاه می‌شود. از طرفی، آرایه مورد نظر در بازه طولانی‌تری (که در ارزیابی‌ها یک روز در نظر گرفته شده است) بازنشانی می‌شود. در زمان بازنشانی، دوباره از داده‌های ثبت رخداد سامانه در همان بازه‌ها هش گرفته شده و با هش‌های ذخیره شده در بلاک چین مقایسه می‌شود و در صورت مغایرت به کاربر اطلاع داده می‌شود با استفاده از این روش دو مرحله‌ای، اطمینان بیشتری از صحت داده‌ها به دست می‌آید. این رویکرد به ویژه در برابر نفوذگرانی مؤثر است که امکان دارد، پس از انجام اقدامات مخرب، تنها داده‌های ثبت رخداد مربوط به بازه زمانی ورود خود را حذف کنند؛ داده‌هایی که ممکن است متعلق به بازه‌های زمانی متعددی پیش از بازه زمانی فعلی باشند. در صورتی که صحت سنجی تنها به صورت یک مرحله‌ای و محدود به داده‌های ثبت رخداد بازه زمانی قبلی انجام شود، روش پیشنهادی در چنین مواردی عملکرد مطلوبی نخواهد داشت؛ به همین دلیل، صحت سنجی باید طی دو مرحله انجام شود. در مرحله نخست، داده‌های ثبت رخداد مربوط به بازه زمانی پیشین مورد صحت سنجی قرار می‌گیرند؛ سپس در انتهای بازه زمانی دوم، صحت داده‌های ثبت رخداد متعلق به چندین بازه زمانی به صورت هم زمان بررسی می‌شود. این رویکرد باعث افزایش دقت و اطمینان در تضمین صحت داده‌ها می‌شود. انتخاب این بازه‌های زمانی، شامل: (۱) بازه زمانی کوتاه تر و مشخص برای ذخیره هش داده‌های ثبت رخداد روی بلاک چین و (۲) زمان بازنشانی آرایه حاوی هش‌ها بسیار حائز اهمیت است. این انتخاب باید به گونه‌ای باشد که هم هزینه مالی کمتری به همراه داشته باشد و هم از هدف خود، یعنی تضمین صحت داده‌های ثبت رخداد دور نشود.

۳-۳- پارامترهای مهم در روش پیشنهادی

در روش پیشنهادی برای تضمین صحت داده‌های ثبت رخداد به صورت خودکار در دو مرحله چندین پارامتر وجود دارد که تأثیر مستقیمی بر دستیابی به اهداف این روش دارند. این پارامترها در ادامه توضیح داده می‌شوند:

نوع الگوریتم هش: یکی از پارامترهای کلیدی است. این پارامتر هم بر مدت زمان مورد نیاز برای هش گرفتن از داده‌ها تأثیر دارد و هم بر طول رشته هش. از آنجا که طول رشته هش بر میزان هزینه ذخیره سازی آن در بلاک چین تأثیر می‌گذارد، نوع الگوریتم تأثیر زیادی بر دو جنبه هزینه، یعنی سربار زمانی و هزینه مالی خواهد داشت.

¹ MetaMask

حجم داده‌های ثبت رخداد: در مدت زمان مشخص شده برای هش گرفتن از این داده‌ها به وسیله سرور تأثیر مستقیم دارد؛ به عبارت دیگر حجم بیشتر داده‌ها باعث افزایش سربار زمانی برای هش گرفتن از آن‌ها خواهد شد؛ بنابراین هرچه حجم داده‌ها بزرگ‌تر باشد، زمان بیشتری برای پردازش و هش گرفتن از آن‌ها نیاز است.

طول رشته‌های هش: که از هش گرفتن از داده‌های ثبت رخداد در بازه‌های زمانی مشخص به دست می‌آید، بر هزینه ذخیره‌سازی داده‌ها در بلاک‌چین تأثیر زیادی دارد. هر الگوریتم هش خروجی با طول متفاوتی تولید می‌کند؛ برای مثال خروجی الگوریتم SHA-256 یک رشته ۳۲ بیتی است. این تفاوت طول رشته‌های هش می‌تواند تأثیر زیادی بر هزینه ذخیره‌سازی در بلاک‌چین داشته باشد.

بازه زمانی برای هش گرفتن از داده‌ها: یکی دیگر از پارامترهای تأثیرگذار است؛ هرچه این بازه زمانی کوتاه‌تر باشد، فرایند تضمین صحت داده‌ها مؤثرتر خواهد بود، اما هزینه ذخیره‌سازی هش‌ها در بلاک‌چین بیشتر می‌شود؛ بنابراین انتخاب بازه زمانی بر عملکرد و هزینه‌گذاری سامانه تأثیر مستقیم دارد. هزینه استقرار قراردادهای هوشمند: که تنها یک بار انجام می‌شود، نیز جزو پارامترهای مهم است. این هزینه بستگی به نحوه نوشتن قرارداد هوشمند دارد. در این مدل، تلاش شده‌است تا قرارداد هوشمند به بهینه‌ترین شکل ممکن نوشته شود تا هزینه‌ها به حداقل برسد.

هزینه خواندن داده از بلاک‌چین: که در پایان هر دوره زمانی تعیین شده باید محاسبه شود، پارامتر دیگری است که باید در نظر گرفته شود. این هزینه به‌ویژه زمانی که نیاز به دسترسی به داده‌ها برای بررسی وضعیت هش‌ها وجود دارد، مهم است. **طول آرایه هش‌ها:** هر چه بیشتر باشد، هزینه بازنشانی آن نیز افزایش خواهد یافت. این پارامتر باید در طراحی سامانه دقیق مورد توجه قرار گیرد تا از افزایش بی‌مورد هزینه‌ها جلوگیری شود.

بازه زمانی برای بازنشانی آرایه هش‌ها: نیز تأثیر زیادی بر هزینه حذف داده‌ها از بلاک‌چین دارد؛ هرچه این بازه زمانی طولانی‌تر باشد، تعداد هش‌هایی که باید ذخیره شوند، بیشتر خواهد بود و در نتیجه آرایه بزرگ‌تر می‌شود. این افزایش اندازه آرایه هزینه‌های اضافی برای بازنشانی و حذف داده‌ها از بلاک‌چین را به همراه خواهد داشت.

۳-۴- رویه اجرای مدل پیشنهادی

روش پیشنهادی با هدف بررسی خودکار تغییرات در داده‌های ثبت رخداد طراحی شده‌است. در این روش، هش این داده‌ها در فواصل زمانی معین محاسبه شده و در بلاک‌چین ذخیره می‌شوند تا امکان تشخیص هرگونه تغییر فراهم شود. این پروتکل شامل دو بازه زمانی مجزا است:

بازه زمانی کوتاه‌مدت: اجرا در فواصل پانزده دقیقه‌ای
بازه زمانی بلندمدت: اجرا در انتهای هر دوره ۲۴ ساعته در ابتدا، قرارداد هوشمند مستقرشده بر بستر بلاک‌چین، همچنین آرایه‌ای به نام logHashes به منظور ذخیره‌سازی هش‌ها مقداردهی اولیه می‌شود؛ سپس داده‌های ثبت رخداد به‌طور دوره‌ای هش شده و به‌صورت ایمن در بلاک‌چین ذخیره می‌شوند. در هر بازه زمانی کوتاه‌مدت، تابع (StoreHash) اجرا می‌شود. مراحل آن به‌صورت زیر است:

۱. دریافت داده‌های ثبت رخداد مربوط به پانزده دقیقه گذشته
۲. محاسبه هش جدید از داده‌ها
۳. بازیابی هش پیشین ذخیره‌شده در بلاک‌چین و مقایسه با هش جدید
۴. در صورت مغایرت هش‌ها، صدور هشدار تغییر غیرمجاز داده
۵. در صورت یکسان بودن هش‌ها بررسی تکراری بودن هش برای جلوگیری از حملات Replay Attack
۶. ذخیره هش جدید در آرایه logHashes در بلاک‌چین در انتهای هر بازه بلندمدت (۲۴ ساعته)، تابع (ResetAndVerifyHashes) اجرا می‌شود که مراحل آن به‌قرار زیر است:

۱. محاسبه مجدد هش‌ها از داده‌های ثبت رخداد ۲۴ ساعت گذشته
۲. مقایسه هش‌های جدید با هش‌های ذخیره‌شده
۳. هشدار در صورت تشخیص مغایرت (نشانه‌ای از دست‌کاری داده‌های تاریخی)
۴. در صورت تطابق کامل، پاک‌سازی آرایه logHashes برای دوره بعدی

این ساختار با تلفیق قابلیت‌های تغییرناپذیری بلاک‌چین و منطق بررسی چندمرحله‌ای تضمین می‌کند که هیچ‌گونه تغییر یا دست‌کاری در داده‌های ثبت رخداد بدون شناسایی باقی نماند؛ همچنین با طراحی دقیق توابع و زمان‌بندی اجرا حملاتی مانند Replay، Reentrancy و Front-running به‌طور مؤثری کنترل می‌شوند و امکان تأثیرگذاری آن‌ها بر عملکرد سامانه از بین می‌رود.

۳-۵- تضمین صحت و محرمانگی در مدل

هدف اصلی این مدل تضمین صحت و تضمین محرمانگی داده‌های ثبت رخداد است. به‌طور جزئی‌تر:

تضمین صحت: برای اطمینان از صحت داده‌های ثبت رخداد از ویژگی تغییرناپذیری بلاک‌چین استفاده می‌شود. با ذخیره هش داده‌ها بر روی بلاک‌چین هرگونه تغییر در داده‌های اصلی قابل شناسایی خواهد بود؛ همچنین بررسی صحت داده‌ها در دو بازه زمانی متفاوت (کوتاه‌تر و بلندتر) احتمال خطا را کاهش می‌دهد.

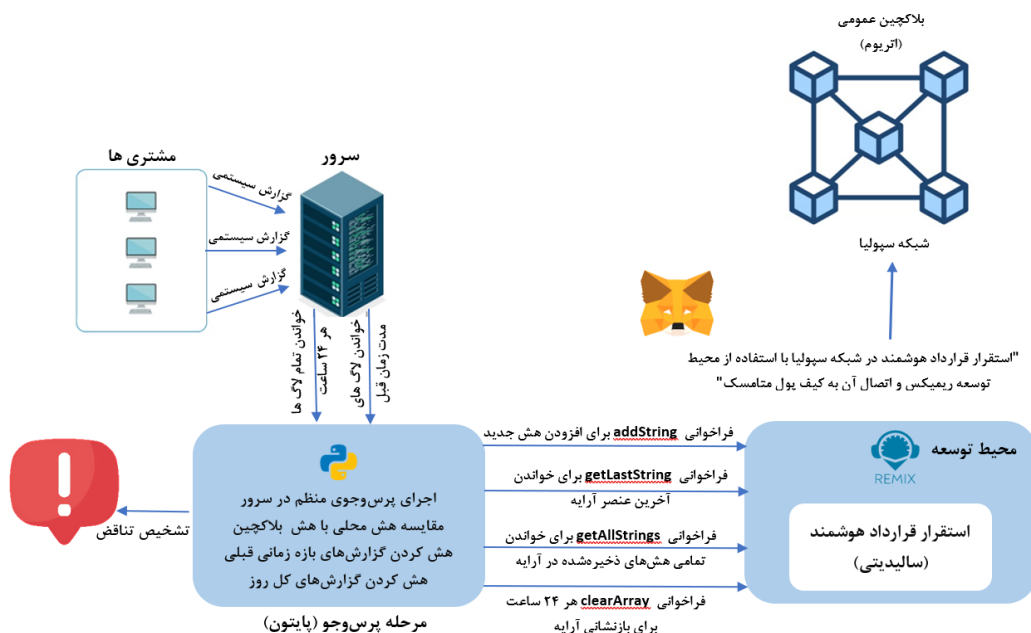
هش‌های جدید را خودکار شناسایی کند و به کاربر اطلاع دهد. استفاده از قرارداد هوشمند در این مدل به دلیل طراحی دقیق آن امنیت بالایی را فراهم می‌آورد. نخستین دلیل جلوگیری از Replay attack این است که هر هش به‌طور مستقل و برای هر بازه زمانی مشخص در بلاک‌چین ذخیره می‌شود و هیچ گونه تکرار هش‌های قدیمی وجود ندارد. برای Reentrancy attack، چون هیچ‌گونه فراخوانی به توابع خارجی یا تغییرات پیچیده وضعیت در قرارداد وجود ندارد، امکان سوءاستفاده از وضعیت قرارداد وجود ندارد؛ درنهایت، Front-running attack به دلیل زمان‌بندی خودکار و دقیق عملیات و ذخیره هش‌ها در بازه‌های زمانی مشخص، در عمل غیرممکن می‌شود؛ به این ترتیب، استفاده از قرارداد هوشمند در مدل باعث می‌شود که این حملات نتوانند به سامانه آسیب برسانند و عملیات به‌طور امن و دقیق انجام شود.

تضمین محرمانگی: به دلیل ماهیت عمومی بلاک‌چین، ذخیره مستقیم داده‌های خام (داده‌های ثبت رخداد) بر روی آن امن نیست؛ بنابراین برای ایجاد اثر انگشتی منحصر به فرد از داده‌ها، از الگوریتم هش استفاده می‌شود. این اثر انگشت (هش) به جای داده‌های اصلی بر روی بلاک‌چین ذخیره می‌شود؛ به این ترتیب محرمانگی داده‌ها حفظ می‌شود و تنها صحت آن‌ها قابل تأیید خواهد بود.

۳-۶- بررسی امنیت قراردادهای هوشمند

در مدل پیشنهادی استفاده از قراردادهای هوشمند نقش کلیدی در تأمین امنیت و صحت داده‌های ثبت رخداد داشت. قرارداد هوشمند به‌عنوان یک سازوکار مطمئن برای ذخیره و مقایسه هش‌های داده‌های ثبت رخداد در بلاک‌چین عمل و تضمین می‌کند که تغییرات غیرمجاز در داده‌ها شناسایی شوند؛ از آنجا که بلاک‌چین تغییرناپذیر است قرارداد هوشمند قادر است هر گونه مغایرت بین هش‌های ذخیره‌شده و

(شکل-۱): مدل پیشنهادی
(Figure-1): Proposed Model



اطلاق می‌شود که سرور برای هش گرفتن از داده‌های ثبت رخداد با حجم‌های مختلف نیاز دارد. این زمان تحت تأثیر مستقیم حجم داده‌ها و زمان پردازش آن‌ها قرار دارد. هزینه مالی به هزینه‌هایی اشاره دارد که برای ذخیره‌سازی رشته‌های هش در شبکه بلاک‌چین مورد نظر صرف می‌شود. این هزینه بسته به طول رشته هش و تعداد تراکنش‌ها متفاوت است.

۴-۱- محیط پیاده‌سازی

پیاده‌سازی‌های مورد نظر در این پژوهش بر روی یک سامانه با مشخصات پردازنده Intel Core i7 حافظه شازنده

۴- پیاده‌سازی و ارزیابی

در این بخش از مقاله پیاده‌سازی روش پیشنهادی که در بخش سوم شرح داده شد، مورد بررسی قرار می‌گیرد. در اینجا، جزئیات و اجزای مختلف روش مورد نظر به طور کامل بیان می‌شود و نتایج حاصل از اجرای آن ارائه می‌شود. هدف اصلی این روش تضمین صحت داده‌های ثبت رخداد به‌صورت خودکار و حفظ محرمانگی آن‌هاست. برای ارزیابی اثربخشی این روش، پیاده‌سازی آن انجام شده است. هدف از این پیاده‌سازی، محاسبه دقیق هزینه‌های مالی و سربار زمانی مرتبط با اجرای این روش است. سربار زمانی به مدت زمانی

نمایش داده شده است. این هزینه شامل هزینه های ذخیره سازی کد قرارداد و انجام نخستین تراکنش های آن در شبکه بلاک چین سپولیا می شود.

برای ارزیابی عملکرد سامانه زمان لازم برای هش گرفتن از داده های ثبت رخداد با حجم های مختلف محاسبه شد؛ به این منظور، تابع مورد نظر بر روی سرور اجرا شد.

این تابع داده های ثبت رخداد را با حجم های مختلف از مجموعه داده استخراج کرده و از آن ها هش تولید کرد. این هش ها با استفاده از الگوریتم SHA-256 محاسبه شدند؛ سپس با فراخوانی توابع قرارداد هوشمند مستقر شده در شبکه سپولیا، هش ها در آرایه ای ذخیره شدند. استفاده از قرارداد هوشمند در این فرایند به طور عمده برای تضمین یک پارچگی و صحت ذخیره سازی هش ها در بلاک چین بود. هر هش محاسبه شده از داده ها به طور ایمن در بلاک چین ذخیره شد که ویژگی تغییرناپذیری بلاک چین صحت داده ها را تضمین می کند.

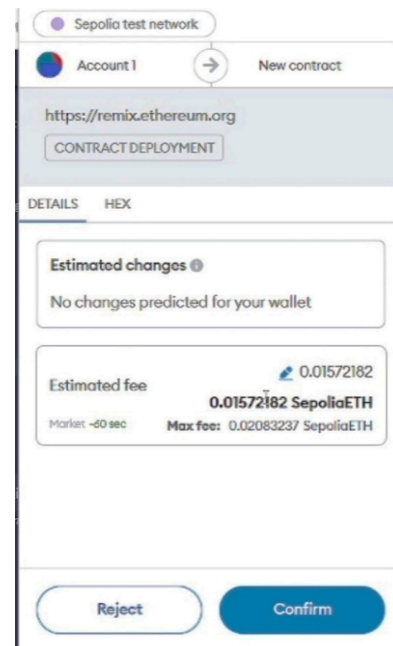
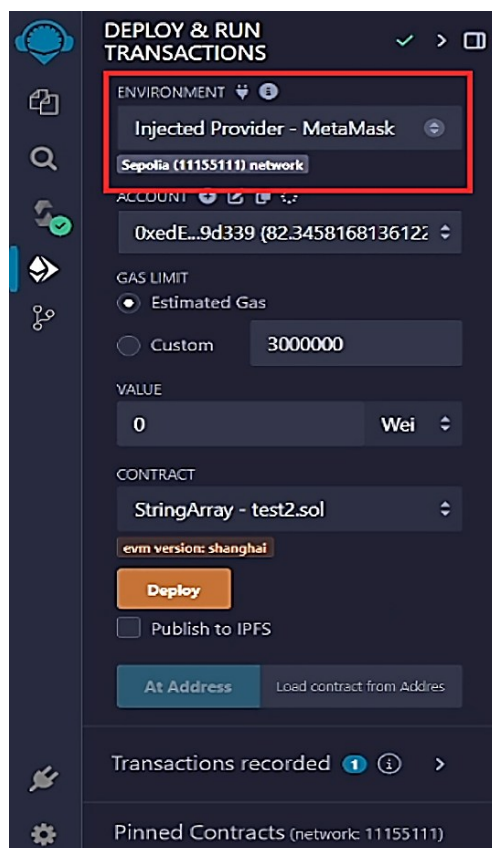
این فرایند شامل مراحل مختلفی از جمله اجرای تابع، محاسبه هش ها و ذخیره سازی آن ها در شبکه بلاک چین سپولیا بود که به طور مؤثری سربار زمانی و هزینه های مالی مرتبط با پیاده سازی را اندازه گیری کرد.

گیگابایت و سامانه عامل Windows 11 انجام شده است. تابع مورد استفاده برای هش گرفتن از داده های ثبت رخداد به زبان برنامه نویسی پایتون پیاده سازی شده است و از الگوریتم SHA-256 برای تولید هش از داده های ثبت رخداد با حجم های مختلف استفاده شده است. داده های ثبت رخداد از مجموعه داده [۲۴] استخراج شده اند. در این پژوهش، قرارداد هوشمند با استفاده از زبان برنامه نویسی Solidity، که زبان اصلی پشتیبانی شده به وسیله شبکه اتریوم است [۲۵]، طراحی و پیاده سازی شده است. برای ذخیره سازی داده ها، از تست نت Sepolia، جدیدترین شبکه آزمایشی اتریوم استفاده شده است که مبتنی بر Solidity عمل می کند.

۴-۲- نحوه پیاده سازی

جهت ارتباط با شبکه بلاک چین دو ابزار کلیدی مورد استفاده قرار گرفتند: محیط توسعه Remix IDE که یک ابزار قدرتمند و کاربر پسند برای توسعه قراردادهای هوشمند است، و کیف پول MetaMask که به دلیل امنیت و محبوبیت بالا برای مدیریت تراکنش ها و ارتباط با شبکه انتخاب شد. این ترکیب از ابزارها فرایند توسعه و آزمایش قرارداد هوشمند را ساده و ایمن کرده است.

در ابتدا، قرارداد هوشمند مورد نظر در شبکه تست نت سپولیا مستقر شد. شبکه سپولیا به عنوان یک شبکه آزمایشی بلاک چین مبتنی بر اتریوم برای آزمایش و ارزیابی قراردادهای هوشمند مورد استفاده قرار می گیرد و این امکان را فراهم می آورد که عملکرد قراردادها بدون نیاز به پرداخت هزینه های تراکنش در شبکه اصلی بررسی شود. هزینه استقرار این قرارداد هوشمند حدود ۰.۰۱۶ اتر بود که در شکل (۲)



(شکل ۲-): اتصال به شبکه سپولیا از طریق کیف پول متاماسک
(Figure-2): Connecting to the Sepolia network via Metamask wallet

برای ارزیابی روش پیشنهادی ابتدا لازم است، معیارهای مطرح شده در مدل تعیین شوند. در این پژوهش، از الگوریتم SHA-256 برای هش گرفتن از داده‌های ثبت رخداد استفاده شده است. خروجی این الگوریتم رشته‌های ۳۲ بیتی است. طول این رشته‌ها تأثیر مستقیم بر هزینه ذخیره‌سازی آن‌ها در شبکه سپولیا دارد؛ هر چه طول رشته‌های خروجی بیشتر باشد، هزینه ذخیره‌سازی آن‌ها در شبکه بلاک‌چین افزایش می‌یابد؛ همچنین بازه زمانی نخست برابر با پانزده دقیقه در نظر گرفته شده است که انتخاب این زمان بستگی کامل به کمترین زمان مورد نیاز برای حملات سایبری دارد؛ یعنی اگر این زمان برابر حداقل زمان لازم برای حملات سایبری در نظر گرفته شود این مدل به صورت صددرصد می‌تواند صحت داده‌ها را تضمین بدهد، اما با پژوهش‌های انجام شده و فاکتورهای پیچیده‌ای که در شبکه‌های مختلف متفاوت است نمی‌توان کمترین زمانی را برای حملات سایبری مختلف در نظر گرفت؛ از طرفی این بازه زمانی تأثیر مستقیمی بر هزینه ذخیره‌سازی رشته‌ها در شبکه سپولیا دارد؛ به‌طوری که هرچه بازه زمانی کوتاه‌تر باشد، هزینه ذخیره‌سازی بیشتر خواهد بود؛ در حالی که دقت مدل بالاتر می‌رود؛ بنابراین میزان زمان پانزده دقیقه برای بازه زمانی نخست در این مدل می‌تواند با توجه به اهمیت بین میزان دقت مدل و میزان هزینه مورد انتظار متفاوت باشد. باید توجه داشت که مدل پیشنهادی برای تضمین صحت داده‌های ثبت رخداد، زمانی که حملات در بازه زمانی کوتاه‌تر از پانزده دقیقه رخ دهند، نمی‌تواند به‌طور کامل پاسخگو باشد؛ زیرا این زمان برابر کمینه زمان مورد نیاز برای حملات مختلف سایبری نیست و اگر نفوذگری به‌طور دقیق در این بازه زمانی وارد سامانه شده و عملیات مخرب انجام دهد و در نهایت داده‌های ثبت رخداد حاصل از کار خود را حذف کرده و از سامانه خارج شود هشی که روی شبکه ذخیره می‌شود یک مقدار تغییر یافته است که صحت آن معنایی ندارد. قرارداد هوشمند مورد استفاده در این پژوهش به زبان سالی‌دیتی نوشته شده و در محیط Remix کامپایل شده است؛ سپس با استفاده از کیف پول MetaMask قرارداد در شبکه تست‌نت سپولیا مستقر شده است؛ همچنین، برای ارزیابی دقیق‌تر، بازه زمانی دوم برابر با یک روز در نظر گرفته شده است. داده‌های ثبت رخداد با حجم‌های مختلف از ۲۸۰ کیلوبایت تا یک گیگابایت به تابع هش داده شده‌اند تا مدت زمان لازم برای هش گرفتن از این داده‌ها ثبت و تحلیل شود.

۴-۳-۱- ارزیابی سربار زمانی

شکل (۳) زمان مورد نیاز برای تولید هش از داده‌های ثبت رخداد با حجم‌هایی بین ۲۸۰ کیلوبایت تا یک گیگابایت را

نمایش می‌دهد؛ همان‌طور که در شکل مشاهده می‌شود، در پیاده‌سازی انجام شده فایلهای ثبت رخداد با اندازه‌های مختلف به تابع مربوطه به‌عنوان ورودی داده شد و در نتیجه ۳۱۸ ثانیه زمان برد تا هش یک فایل با حجم یک گیگابایت محاسبه شود. این زمان قابل‌توجهی محسوب نمی‌شود؛ زیرا این تابع بر روی یک سامانه معمولی اجرا شده است؛ علاوه بر این حجم داده‌های ثبت رخداد تولید شده با سامانه‌های معمولی در شبکه به‌طور متوسط بسیار کمتر از یک گیگابایت است؛ همچنین در صورت استفاده از سرورهایی با سخت‌افزار قدرتمندتر این زمان به مراتب کاهش خواهد یافت. با استناد به پژوهش انجام شده توسط بارتولتی [۲۵] حجم داده‌های تولید شده در یک سرور معمولی شبکه متوسط طی یک ربع، بسیار کمتر از یک گیگابایت است. با در نظر گرفتن اینکه این سرورها به‌طور معمول از سخت‌افزار بسیار قوی‌تری نسبت به سامانه مورد استفاده در این آزمایش بهره می‌برند، می‌توان انتظار داشت که زمان لازم برای هش‌گیری کاهش چشم‌گیری یابد.

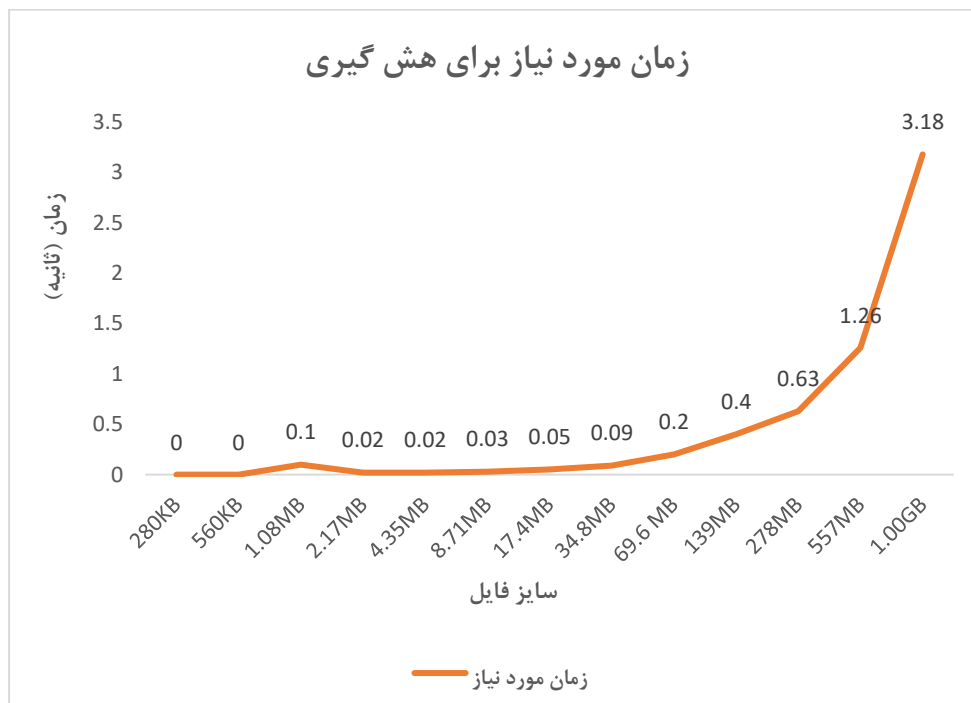
۴-۳-۲- ارزیابی هزینه مالی

هزینه مالی شامل ذخیره رشته‌های هش ۳۲ بیتی هر پانزده دقیقه بر روی شبکه سپولیا و همچنین هزینه بازنشانی آرایه حاوی هش‌ها در انتهای هر روز است. گفتنی است که خواندن داده‌ها از بلاک‌چین هزینه‌ای ندارد. شکل (۴) هزینه لازم برای ذخیره رشته‌های ۳۲ بیتی بر روی یک آرایه با طول‌های مختلف در شبکه سپولیا را نشان می‌دهد؛ این رشته‌ها همان داده‌هایی است که هر پانزده دقیقه بر روی شبکه بلاک‌چین ذخیره می‌شود و حاصل هش‌گیری داده‌های ثبت رخداد در همان بازه است. رفتار سینوسی نمودار به‌طور عمده به دلیل شرایط متغیر شبکه در زمان‌های مختلف است، نه به دلیل طول آرایه‌ای که هش‌ها در آن ذخیره می‌شوند. این هزینه بین ۰.۰۰۱ تا ۰.۰۰۳ اتر متغیر است، اما به‌طور متوسط حدود ۰.۰۰۲ اتر است. شکل (۵) هزینه ریست کردن آرایه با طول‌های مختلف را نشان می‌دهد که در انتهای بازه زمانی دوم انجام می‌شود و هرچه تعداد رشته‌هایی که در طول این بازه زمانی روی آرایه ذخیره شده‌اند، بیشتر باشد هزینه ریست کردن آن هم بیشتر می‌شود.

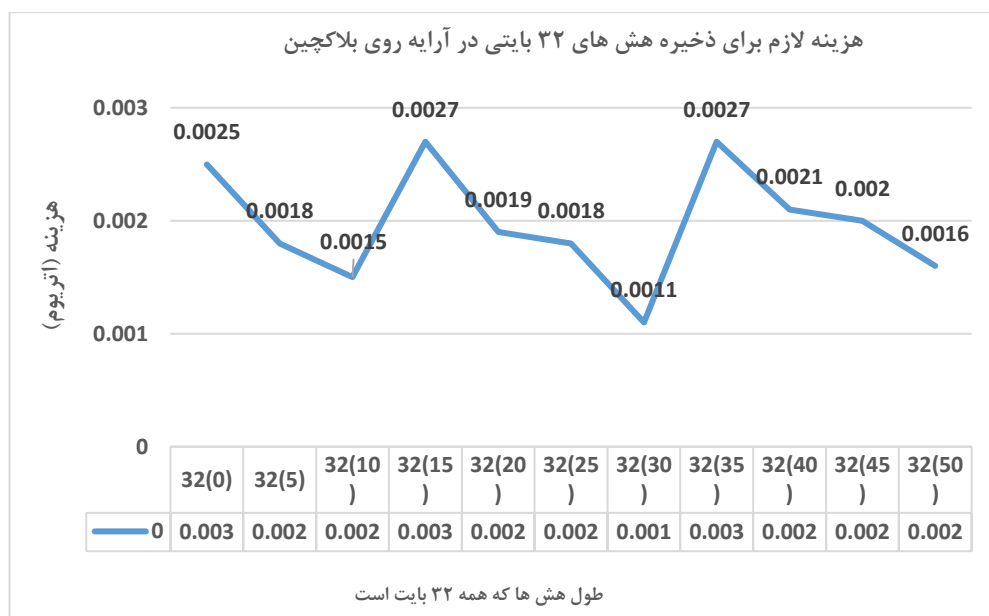
طبق روش پیشنهادی، برای به کمترین حد رساندن احتمال خطا، صحت داده‌ها باید طی دو مرحله تضمین شود: مرحله نخست: در انتهای بازه زمانی نخست انجام می‌شود، و در آن، تنها هش مربوط به بازه زمانی پیشین صحت‌سنجی می‌شود. مرحله دوم: در انتهای بازه زمانی دوم که بازه طولانی‌تر است، صحت هش داده‌های چندین بازه زمانی ذخیره شده در آرایه

توجه داشت که در زمانی که طول آرایه برابر با هشتاد بود، هزینه بازنشانی در بالاترین مقدار خود قرار داشت که این امر به دلیل شرایط خاص شبکه در آن لحظه است. در حالت کلی، این نمودار صعودی است، بدین معنا که با افزایش طول آرایه، هزینه بازنشانی آن نیز بیشتر می شود.

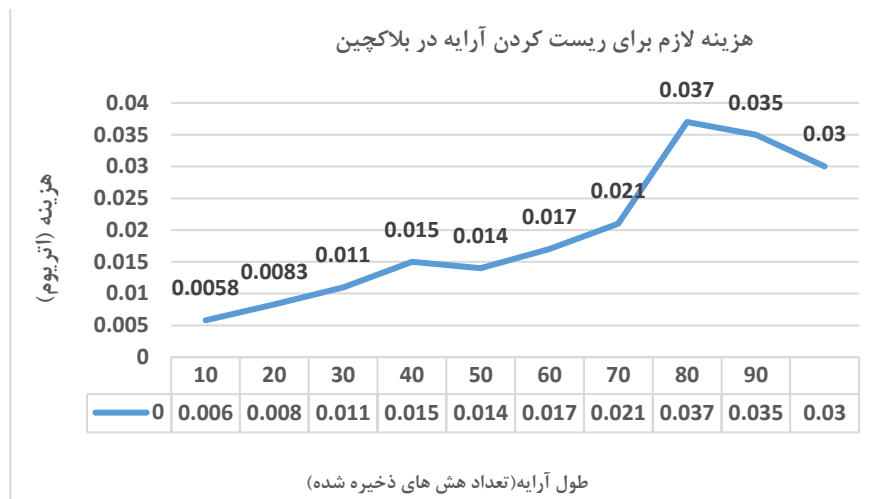
بررسی می شود. طول آرایه از تقسیم بازه زمانی دوم بر بازه زمانی نخست به دست می آید. در مرحله دوم، داده ها از آرایه ذخیره شده روی بلاک چین خوانده می شوند و پس از صحت سنجی آرایه ریست می شود. دقت شود که خواندن از آرایه هزینه ای به دنبال ندارد، اما نوشتن در آن هزینه بر است. همان طور که در شکل (۵) مشاهده می شود، هرچه طول آرایه بزرگ تر باشد، هزینه بازنویسی آن نیز بیشتر خواهد بود. باید



(شکل-۳): هزینه زمانی مورد نیاز برای هش گرفتن از داده های ثبت رخداد توسط کوئری با حجم های متفاوت
(Figure-3): The time cost required for hashing event log data using queries with varying sizes



(شکل-۴): هزینه لازم برای ذخیره هش ۳۲ بیتی بر روی آرایه با طول متفاوت روی بلاک چین سپولیا
(Figure-4): The cost required to store a 32-byte hash on an array with varying lengths on the Sepolia blockchain



(شکل-۵): هزینه لازم برای ریست کردن آرایه با طول‌های مختلف در شبکه سپولیا
(Figure-5): The cost required to reset an array with varying lengths on the Sepolia network

می‌شود، می‌توان مدت زمان مشخصی را برای هش‌گیری تعیین کرد.

با تمامی پارامترهای تعیین‌شده، این طرح قادر است صحت داده‌های ثبت رخداد را به صورت خودکار تضمین کند؛ همچنین با توجه به پیاده‌سازی انجام‌شده، هزینه‌های پیاده‌سازی این طرح، شامل هزینه مالی و سربرار زمانی، محاسبه شده‌است؛ بالاین‌حال، همان‌طور که در تمامی تمهیدات امنیتی پیش‌بینی شده‌است، هیچ تمهید امنیتی قادر به تضمین صددرصدی نسبت به هدف خود نیست. به‌طور خاص، طبق طرح پیشنهادی اگر حمله‌ای صورت گیرد که مدت زمان آن کمتر از مدت زمان هش‌گیری باشد و به‌طور دقیق در این بازه زمانی اتفاق بیفتد و نفوذگر اقدام به حذف داده‌های ثبت رخداد کند، طرح پیشنهادی قادر به تضمین صددرصدی صحت این نوع داده‌ها نخواهد بود.

۴-۳-۳- ارزیابی پیچیدگی زمانی مدل

با توجه به جدول (۲)، برای بررسی پیچیدگی زمانی روش‌های مختلف ذخیره‌سازی داده‌های ثبت رخداد در بلاکچین، نحوه پردازش، هش‌گذاری و ذخیره‌سازی اطلاعات در هر پژوهش مورد بررسی قرار گرفت. در روش پیشنهادی، تنها هش داده‌ها در بلاکچین ذخیره شده و داده‌های خام خارج از زنجیره باقی می‌مانند. این فرایند موجب کاهش چشم‌گیر هزینه و زمان پردازش می‌شود. از آنجا که طول هش همواره ثابت است، زمان لازم برای ذخیره‌سازی مستقل از حجم داده‌ها بوده و در نتیجه، پیچیدگی زمانی این روش $O(1)$ است. این ویژگی باعث می‌شود که روش پیشنهادی، یک راه‌کار بهینه و کم‌هزینه برای تضمین صحت داده‌ها باشد. در پژوهش [۱۶] نیز رویکردی مشابه اتخاذ شده و تنها هش داده‌ها در بلاکچین اتریوم ذخیره می‌شود؛ بنابراین، این روش نیز

با توجه به معیارهای مشخص‌شده در این پیاده‌سازی، برای محاسبه هزینه کل مالی که این مدل در طول یک ماه ممکن است داشته باشد، فرضیات زیر در نظر گرفته می‌شود:
فرض نخست: الگوریتم مورد استفاده SHA-256 است که طول رشته خروجی آن ۳۲ بایت است.
فرض دوم: در انتهای هر روز، آرایه حاوی هش‌ها بازنشانی می‌شود.

فرمول محاسبه هزینه کل در یک ماه به شرح زیر است:

$$T = A + 30 \times (Z + 24 \times (X \times Y)) \quad (۱)$$

که در آن:

هزینه کل در یک ماه = T

هزینه استقرار اولیه قرارداد هوشمند = A

هزینه مورد نیاز برای ذخیره هش ۳۲ بیتی در شبکه بلاکچین = X

تعداد دفعات هش‌گیری در ساعت = Y

هزینه بازنشانی کردن آرایه با طول مشخص = Z

گفتنی است که برای ماه‌های بعدی، هزینه استقرار اولیه قرارداد هوشمند تکرار نخواهد شد و این هزینه تنها یک‌بار در محاسبات اعمال می‌شود. بر اساس فرمول به‌دست‌آمده، هرچه تعداد دفعات هش‌گیری در طول روز کمتر باشد، هم هزینه ذخیره داده‌ها روی بلاکچین کاهش می‌یابد و هم هزینه بازنشانی آرایه. این موضوع با توجه به فرمول و شکل‌های (۴) و (۵) مشخص است. شکل (۴) بیان‌کننده این نکته است که در هر بار ذخیره‌سازی یک هش ۳۲ بیتی در هر یک از خانه‌های مختلف آرایه هزینه ثابتی وجود ندارد و به شرایط شبکه وابسته است. در این شکل برای مثال خانه‌هایی با اندیس مضرب پنج آرایه در نظر گرفته شده‌اند؛ بالاین‌حال، باید در نظر داشت که هش‌گیری از داده‌ها در بازه‌های زمانی کوتاه‌تر منجر به تضمین صحت مؤثرتر داده‌ها می‌شود؛ بنابراین، با توجه به بیشینه هزینه‌ای که در نظر گرفته

دارای پیچیدگی زمانی $O(1)$ است؛ با این حال، امکان وجود تأخیر جزئی در تأیید هش‌ها از طریق قراردادهای هوشمند وجود دارد که در برخی موارد، زمان پردازش را تحت تأثیر قرار می‌دهد.

برخی پژوهش‌ها مانند [۱۷] و [۲۲] برای کاهش هزینه‌های ذخیره‌سازی از فناوری IPFS استفاده کرده‌اند. در این روش‌ها، داده‌های اصلی در IPFS ذخیره شده و تنها هش آن‌ها در بلاک‌چین ثبت می‌شود. این روش باعث کاهش هزینه‌های ذخیره‌سازی در بلاک‌چین می‌شود، اما به دلیل ماهیت توزیع‌شده IPFS، بازیابی داده‌ها مستلزم جست‌وجو در یک ساختار درختی (Merkle DAG) است که پیچیدگی زمانی آن $O(\log N)$ خواهد بود. این موضوع می‌تواند منجر به افزایش زمان تأیید داده‌ها شود. در مقابل، روش‌هایی مانند [۲۰] و [۷] داده‌های ثبت رخداد را به‌طور کامل در بلاک‌چین ذخیره می‌کنند. در این روش‌ها، هرچه حجم داده‌های ثبت رخداد افزایش یابد، زمان پردازش و هزینه ذخیره‌سازی نیز افزایش می‌یابد؛ از این رو، پیچیدگی زمانی این روش‌ها $O(N)$ است؛ این بدان معناست که زمان ذخیره‌سازی به‌صورت خطی با افزایش حجم داده‌ها رشد می‌کند که موجب هزینه بالاتر و تأخیر بیشتر در پردازش داده‌ها

می‌شود؛ درنهایت، روش‌هایی مانند [۲۱] که از بلاک‌چین‌های خصوصی مانند Multichain استفاده می‌کنند، تنها هش داده‌ها را ذخیره کرده و نیازی به پردازش اضافی ندارند. این باعث می‌شود که پیچیدگی زمانی این روش نیز $O(1)$ باشد؛ با این حال، وابستگی به بلاک‌چین خصوصی موجب کاهش شفافیت و محدود شدن دسترسی عمومی به داده‌ها می‌شود.

در کل مقایسه روش‌های موجود نشان می‌دهد که روش پیشنهادی با پیچیدگی زمانی $O(1)$ ضمن کاهش هزینه ذخیره‌سازی از نظر زمان پردازش نیز بهینه‌ترین گزینه است؛ در مقابل، روش‌هایی که از IPFS استفاده می‌کنند، هزینه ذخیره‌سازی را کاهش می‌دهند، اما باعث افزایش زمان بازیابی داده‌ها می‌شوند ($O(\log N)$)؛ درنهایت روش‌هایی که داده‌های ثبت رخداد را مستقیم در بلاک‌چین ذخیره می‌کنند؛ اگرچه امنیت بالایی دارند، اما هزینه و زمان پردازش بالاتری را به همراه دارند ($O(N)$)؛ بر این اساس، روش پیشنهادی تعادلی بین امنیت، هزینه و سرعت پردازش ایجاد کرده و می‌تواند یک گزینه کاربردی و مقرون‌به‌صرفه برای تضمین صحت و محرمانگی داده‌های ثبت رخداد باشد.

(جدول ۲): مقایسه عملکرد با پژوهش‌های دیگر
(Table-2): Performance comparison with other studies

معایب	کارایی زمان	کارایی هزینه	پیچیدگی زمانی	محرمانگی	پلتفرم	پژوهش
در برابر حملات بسیار سریع محدودیت دارد	بهینه (ارزیابی شده برای سر بار محاسبات)	بهینه (ارزیابی شده برای کاهش هزینه ذخیره‌سازی)	$O(1)$	بله	اتریوم (سپولیا)	Proposed Method
تأخیر جزئی در اجرای قرارداد هوشمند	متوسط	بالا (به دلیل ذخیره‌سازی مستقیم بلاک‌چین)	$O(1)$	بله	اتریوم	Javed et al. [16]
کندی در بازیابی داده‌ها	سریع (بلاک‌چین خصوصی زمان پردازش را کاهش می‌دهد)	متوسط (IPFS هزینه را کاهش می‌دهد)	$O(\log N)$	بله	Hyperledger Fabric	Khor et al. [17]
هزینه و زمان ذخیره‌سازی بالا	متوسط	بالا (بلاک‌چین‌های مجاز هزینه‌های راه‌اندازی بالایی دارند)	$O(N)$	بله	EXONUM	Ali et al. [20]
نیاز به بلاک‌چین خصوصی	سریع	متوسط (چند زنجیره‌ای کارایی را بهبود می‌بخشد)	$O(1)$	بله	Multichain	Li et al. [21]
هزینه بالا، زمان ذخیره طولانی	متوسط	بالا (ذخیره‌سازی مستقیم در اتریوم)	$O(N)$	بله	اتریوم	Jain et al. [7]
کندی در بازیابی داده‌ها	آهسته (بازیابی IPFS تأخیر را اضافه می‌کند)	کم (IPFS هزینه را به میزان قابل توجهی کاهش می‌دهد)	$O(\log N)$	خیر	اتریوم	Jiang et al. [26]

داده‌ها و پردازش محاسباتی دقیق محاسبه شد؛ با این حال، این روش در برابر حملات بسیار کوتاه‌مدت (حملاتی که مدت زمان آن‌ها کمتر از بازه زمانی نخست است) قادر به ارائه تضمین صددرصدی نیست؛ لذا همواره نیاز است که بین هزینه مالی (زمان فراخوانی توابع رابطه مستقیم با هزینه دارد) و میزان ضمانت مورد انتظار موازنه‌ای تعیین شود.

در این پژوهش تنها هش داده‌ها در بلاک‌چین ذخیره شده‌است، نه خود داده‌های ثبت رخداد. این رویکرد برای حفظ

۵- نتیجه‌گیری و کارهای آینده

در این پژوهش، روشی نوین برای تأیید خودکار صحت داده‌های ثبت رخداد ارائه شده‌است. این روش با بهره‌گیری از فناوری بلاک‌چین و قراردادهای هوشمند، علاوه بر تضمین صحت داده‌ها، محرمانگی اطلاعات را نیز حفظ می‌کند. یکی از ویژگی‌های برجسته این کار، محاسبه دقیق هزینه‌های عملیاتی اجرای این روش است؛ در این راستا، با پیاده‌سازی عملی این روش بر روی شبکه سپولیا (یک شبکه آزمایشی اتریوم)، هزینه‌های ذخیره‌سازی

- [3] س. کدخدا ده خانی، ح. زنگی آبادی زاده، م. قاسمی، م. رحمانی، و. ف. وظیفه دوست، «فناوری بلاکچین: مروری بر مفاهیم، چالش‌های بلاکچین در خدمات عمومی و طبقه‌بندی توکن‌های بلاکچین»، دو فصلنامه محاسبات و سامانه‌های توزیع شده، شماره ۶، ۱۱۸-۱۳۶، ۱۴۰۲.
- [3] S.kadkhodadehkhani, H.Zangiabadi Zadeh, M.Ghasemi, F.Vazifehdoost, M.Rahmani. Blockchain Technology: A Review of Concepts, Blockchain Challenges in Public Services, and Blockchain Token Classification, *Journal of Distributed Computing and Systems (JDCS)*, Vol 6, Issue 1, Page 118-136, 2023.
- [4] ب. پورولیخان نوخندان، «یک طرح ذخیره‌سازی توزیع شده بر اساس بلاکچین و IPFS برای داده‌های IoT آتش‌نشانی»، ششمین همایش و نمایشگاه بین‌المللی آتش‌نشانی و ایمنی شهری، ۱۴۰۳.
<https://civilica.com/doc/2059774>
- [4] B. Pourvelikhhan Noukhandan, "A Distributed Storage Scheme Based on Blockchain and IPFS for Fire Department IoT Data," in *Proc. 6th International Conference and Exhibition on Fire Fighting and Urban Safety*, Iran, 2024. [in Persian] Available: <https://civilica.com/doc/2059774>
- [5] پورعسکری، حسن، خطیبی بردسیری، عمید، محمدی قنات قستانی، مختار «حفظ حریم خصوصی در اینترنت اشیا برای انتقال داده‌ها در حوزه سلامت با استفاده از زنجیره بلوکی»، پردازش علامت و داده‌ها، دوره ۲۱، شماره ۳، ص ۱۴۹-۱۷۸، ۱۴۰۳.
- [5] A. Hassan Pour Askari, A. Khatibi Bardsiri, and M. Mohammadi Ghanat Ghestani, "IoT privacy for the transmission of data in the field of health using blockchain," (in eng), *Signal and Data Processing*, Research vol. 21, no. 3, pp. 149-178, 2024.
doi: 10.61186/jsdp.21.3.149.
- [6] M. H. Rakib, S. Hossain, M. Jahan, and U. Kabir, "A blockchain-enabled scalable network log management system," *Journal of Computer Science*, vol. 18, no. 6, p. 496.508, 2022.
- [7] P. Jain, "Decentralize log file storage and integrity preservation using blockchain," *International Journal of Computer Science and Information Technologies*, vol. 11, no. 2, pp. 21-30, 2020.
- [8] N. Salunke, S. Sonawane, and D. Motwani, "Decentralized evidence storage system using blockchain and IPFS," in *International Conference on Information, Communication and Computing Technology*, 2023: Springer, pp. 259-280.
- [9] M. Kumar, A. K. Singh, and T. S. Kumar, "Secure log storage using blockchain and cloud infrastructure," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2018: IEEE, pp. 1-4.
- [10] W. Pourmajidi and A. Miranskyy, "Log chain: Blockchain-assisted log storage," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*, 2018: IEEE, pp. 978-982.
- [11] Y. Taguchi, A. Kanai, and S. Tanimoto, "A distributed log management method using a blockchain Scheme," in *2020 IEEE International Conference on Consumer Electronics (ICCE)*, 2020: IEEE, pp. 1-3.

محرمانگی داده‌ها بسیار مهم است؛ زیرا ذخیره‌سازی داده‌های خام بر روی بلاکچین به دلیل ماهیت عمومی آن امنیت را تهدید می‌کند؛ بنابراین، تنها هش‌های تولیدشده از داده‌های ثبت رخداد در بلاکچین ذخیره شدند. این هش‌ها به عنوان اثر انگشتی منحصربه‌فرد از داده‌ها عمل می‌کنند که امکان تأیید صحت داده‌ها را فراهم می‌آورد، بدون آنکه خود داده‌ها مستقیم در بلاکچین ثبت شوند. در کارهای آینده، به منظور کاهش بار مالی اجرای طرح، هدف ما بهره‌برداری از الگوریتم‌های یادگیری ماشین است. استفاده از این الگوریتم‌ها به ما این امکان را می‌دهد که داده‌های با اهمیت و اولویت‌دار را از مجموعه داده‌های ثبت رخداد شناسایی کنیم. در حال حاضر، تمام داده‌های ثبت رخداد مستقیم در بلاکچین ذخیره نمی‌شوند؛ بلکه هش‌های آن‌ها در بلاکچین ثبت می‌شود. با استفاده از الگوریتم‌های یادگیری ماشین می‌توان داده‌هایی را که از نظر امنیتی یا عملیاتی اهمیت بیشتری دارند، شناسایی و اولویت‌بندی کرد؛ سپس تنها این داده‌های اولویت‌دار از طریق هش در بلاکچین ذخیره خواهند شد. این رویکرد، به ویژه در مقیاس‌های بزرگ، منجر به کاهش چشم‌گیر هزینه‌های ذخیره‌سازی خواهد شد؛ چرا که به جای ذخیره‌سازی مستقیم تمام داده‌ها، تنها هش داده‌هایی که در فرایند تحلیل و تضمین صحت اهمیت دارند، در بلاکچین ذخیره خواهد شد؛ به این ترتیب، حجم داده‌های ذخیره‌شده در بلاکچین کاهش قابل ملاحظه‌ای می‌یابد و در نتیجه، هزینه‌های مربوط به ذخیره‌سازی داده‌ها نیز کاهش خواهد یافت.

علاوه بر این، با استفاده از الگوریتم‌های یادگیری ماشین امکان سفارشی‌سازی مدل برای انواع مختلف داده‌ها به وجود می‌آید. این امکان به مدل اجازه می‌دهد تا نحوه ذخیره‌سازی و تحلیل داده‌ها را هوشمند بهینه‌سازی کند. این بهبود در کارایی و دقت سامانه، به کاهش زمان پردازش و افزایش کارایی کلی سامانه خواهد انجامید.

6-References

۶-مراجع

- [1] پارسائیان، محمودرضا و صمیمی، حسین، «رویکردی نوین برای جلوگیری از آسیب‌پذیری در قراردادهای هوشمند و مقابله با حملات Reentrancy بر بستر بلاکچین»، یازدهمین کنگره ملی سراسری فناوریهای نوین در حوزه توسعه پایدار، ایران، تهران، ۱۴۰۰.
<https://civilica.com/doc/1469930>
- [1] M. R. Parsaeian and H. Samimi, "A Novel Approach to Prevent Vulnerabilities in Smart Contracts and Counter Reentrancy Attacks on Blockchain Platform," in *Proc. 11th National Congress of New Technologies in Sustainable Development*, Tehran, Iran, 2021. Available: <https://civilica.com/doc/1469930>
- [2] تیموری، احمد، دی پیر، محمود، «سامانه دو سطحی تشخیص نفوذ برای شبکه اینترنت اشیا مبتنی بر یادگیری عمیق»، پردازش علامت و داده‌ها، دوره ۲۱، شماره ۳، ص ۲۲-۱۴۰۳.
- [2] A. Teymouri and M. Deypir, "Two-level intrusion detection system for Internet of Things network based on deep learning", *Signal and Data Processing*, vol. 21, no. 3, pp. 3-22, 2024.

- [25] M. Bartoletti, F. Fioravanti, G. Matricardi, R. Pettinau, and F. Sainas, "Towards benchmarking of Solidity verification tools," *arXiv preprint arXiv:2402.10750*, 2024.
- [26] J. Jiang, X. Zhang, and Z. Yuan, "Feature selection for classification with Spearman's rank correlation coefficient-based self-information in divergence-based fuzzy rough sets," *Expert Systems with Applications*, vol. 249, p. 123633, 2024.



فاطمه آملی دانش‌آموخته کارشناسی‌ارشد مهندسی کامپیوتر گرایش شبکه‌های کامپیوتری از دانشگاه مازندران است. وی مدرک کارشناسی خود را در رشته مهندسی کامپیوتر گرایش سخت‌افزار از دانشگاه صنعتی نوشیروانی بابل دریافت کرده‌است. برخی زمینه‌های پژوهشی مورد علاقه ایشان شبکه و امنیت شبکه‌های کامپیوتری است.

نشانی رایانامه ایشان عبارت است از:

amoli.fateme.h.1996@gmail.com



مصطفی بستام استادیار دانشکده مهندسی کامپیوتر دانشگاه مازندران است. او دکترای خود را در رشته شبکه‌های کامپیوتری از دانشگاه صنعتی امیرکبیر (پلی‌تکنیک تهران)

در سال ۱۳۹۶ دریافت کرد. مدرک کارشناسی خود را در رشته مهندسی کامپیوتر از دانشگاه شهید باهنر کرمان در سال ۱۳۸۵ و مدرک کارشناسی‌ارشد را در رشته شبکه‌های کامپیوتری از دانشگاه صنعتی امیرکبیر در سال ۱۳۸۸ دریافت کرد. زمینه‌های پژوهشی مورد علاقه ایشان عبارت‌اند از: رایانش ابری و مه، شبکه‌های نرم‌افزار محور، اینترنت اشیا، یادگیری ماشین و بلاک‌چین.

نشانی رایانامه ایشان عبارت است از:

bastam@umz.ac.ir



احسان عطائی تحصیلات خود را در مقاطع کارشناسی، کارشناسی‌ارشد و دکترای رشته مهندسی کامپیوتر به‌ترتیب در سال‌های ۱۳۸۱، ۱۳۸۳ و ۱۳۹۶ از

دانشگاه صنعتی شریف تهران دریافت کرد و هم‌اکنون دانشیار گروه مهندسی کامپیوتر دانشکده مهندسی و فناوری دانشگاه مازندران است. زمینه‌های پژوهشی مورد علاقه ایشان عبارت‌اند از: رایانش ابری، سامانه‌های توزیع‌شده و مدل‌سازی کارایی و اتکاپذیری.

نشانی رایانامه ایشان عبارت است از:

ataie@umz.ac.ir

- [12] K. Awusun-David, T. Al-Hadhrami, M. Alazab, N. Shah, and A. Shalaginov, "BCFL logging: An approach to acquire and preserve admissible digital forensics evidence in cloud ecosystem," *Future Generation Computer Systems*, vol. 122, pp. 1-13, 2021.
- [13] H. Tian, J. Wang, C.-C. Chang, and H. Quan, "Public auditing of log integrity for shared cloud storage systems via blockchain," *Mobile Networks and Applications*, pp. 1-13, 2023.
- [14] M. Altulyan, L. Yao, S. Kanhere, and C. Huang, "A blockchain framework data integrity enhanced recommender system," *Computational Intelligence*, vol. 39, no. 1, pp. 104-120, 2023.
- [15] L. Li, D. Jin, T. Zhang, and N. Li, "A secure, reliable and low-cost distributed storage scheme based on blockchain and IPFS for firefighting IoT data," *IEEE Access*, vol. 11, pp. 97318-97330, 2023.
- [16] H. Javed et al., "Blockchain-based logging to defeat malicious insiders: The case of remote health monitoring systems," *IEEE Access*, vol. 12, pp. 12062-12079, 2023.
- [17] J. H. Khor, M. Sidorov, M. T. Ong, and S. Y. Chua, "Public blockchain-based data integrity verification for low-power IoT devices," *IEEE Internet of Things Journal*, vol. 10, no. 14, pp. 13056-13064, 2023.
- [18] B. Putz, F. Menges, and G. Pernul, "A secure and auditable logging infrastructure based on a permissioned blockchain," *Computers & Security*, vol. 87, p. 101602, 2019.
- [19] Z. Liu, L. Ren, Y. Feng, S. Wang, and J. Wei, "Data integrity audit scheme based on quad merkle tree and blockchain," *IEEE Access*, vol. 11, pp. 59263-59273, 2023.
- [20] A. Ali, A. Khan, M. Ahmed, and G. Jeon, "BCALS: Blockchain-based secure log management system for cloud computing," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 4, p. e4272, 2022.
- [21] W. Li, Y. Feng, N. Liu, Y. Li, X. Fu, and Y. Yu, "A secure and efficient log storage and query framework based on blockchain," *Computer Networks*, vol. 252, p. 110683, 2024.
- [22] P. Jiang, B. Qiu, and L. Zhu, "Toward reliable and confidential release for smart contract via ID-based TRE," *IEEE Internet of Things Journal*, vol. 9, no. 13, pp. 11422-11433, 2021.

[23] او شیلدز، رچی، ناصر، مهدی، صادقی، حسین.

«قراردادهای هوشمند: توافقات حقوقی در پرتو بلاکچین».

پژوهش‌های حقوقی، مجله ۱۸، شماره ۳۷، صص ۲۶۱-۱۳۹۸، ۲۸۸.

10.48300/jlr.2019.91607

- [23] O. Shields, R. Naser, and H. Sadeghi, "Smart Contracts: Legal Agreements for the Blockchain," *Journal of Legal Research*, vol. 18, no. 37, pp. 261-288, 2019. [in Persian] 10.48300/jlr.2019.91607
- [24] J. Zhu, S. He, P. He, J. Liu, and M. R. Lyu, "Loghub: A large collection of system log datasets for ai-driven log analytics," in *2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE)*, 2023: IEEE, pp. 355-366.