Quantitative Analysis of SHA-256 Security Margin Based on Historical Cryptanalysis Trends for Long-Term Resilience Evaluation of the Bitcoin Ecosystem

Ikhwan Al Hakim
Department of Informatics
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung, Indonesia
13522147@std.stei.itb.ac.id

Dr. Ir. Rinaldi Munir, M.T.
Department of Informatics
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung, Indonesia
rinaldi@informatika.org

***Abstract*—The security of the Bitcoin ecosystem heavily relies on the integrity of the SHA-256 hash function. Although the full version (64 rounds) is currently considered secure, advancements in cryptanalysis on reduced-round versions are gradually thinning the algorithm's security margin. This study aims to evaluate the long-term resilience of SHA-256 through a quantitative analysis of historical cryptanalysis trends and a simulation of attack impacts via a Proof-of-Concept (PoC). The methodology involves compiling data from over 15 scientific publications to model the decay rate of the security margin using linear, logarithmic, and exponential regression functions under a Weighted Least Squares approach. The results show that the logarithmic regression model provides the highest accuracy (MAPE of 3.34% - 6.97%) in describing the deceleration of the cryptanalysis progress due to the algorithm's internal resistance. This model projects that the security margin against standard collision attacks remains secure beyond the year 2100, while non-standard attacks are predicted to reach the 64-round threshold by the year 2073. PoC simulations demonstrate that the Merkle Tree structure and the double-hashing mechanism in the Bitcoin protocol provide a significant additional layer of defense, where logic-solver-based preimage attacks remain impractical to match the current computational power of the network. This study concludes that SHA-256 still possesses an adequate security buffer, but monitoring degradation at the compression function level remains crucial as an early indicator for future algorithm transitions.**

***Index Terms*—SHA-256, Bitcoin, Security Margin, Cryptanalysis, Trend Analysis Model, Proof-of-Concept**

I. Introduction

Advancements in the field of cryptanalysis continuously test the security boundaries of existing cryptographic algorithms [1]. The SHA-256 hash function, despite being designed with a more complex structure than its predecessors, is not immune to this testing. Since its publication, researchers have consistently demonstrated attacks on reduced-round versions (i.e., versions with a fewer number of rounds). This progress has evolved from initial attacks on 21 rounds [2], increasing to 28 rounds, and reaching a record of 38 rounds by Mendel et al. [6].

Following these discoveries, cryptanalysis progress against SHA-256 rounds entered a prolonged period of stagnation lasting eleven years (2013–2023). However, this deadlock was broken through a breakthrough attack on 39 rounds using SAT/SMT-based sparsity optimization by Y. Li et al. [3]. This breakthrough in 2024 provided a new data point, thereby enabling a quantitative projection of the security margin decay rate of SHA-256.

Nevertheless, it must be emphasized that all collected historical attack data, including the latest record in 2024, only succeeded in attacking reduced-round versions rather than the full 64-round version of SHA-256 [4]. Bitcoin, as the pioneer of crypto assets, relies on the integrity of the SHA-256 hash function for mining, address generation, and ensuring the integrity of transaction data through the Merkle Tree structure [5].

To measure the resilience of an algorithm quantitatively, the cryptographic community uses a concept called the *security margin*. This metric is defined as the difference between the total number of rounds in the full algorithm and the maximum number of rounds successfully breached by the best-known cryptanalytic attack [7]. For SHA-256, a successful attack on 39 rounds leaves a security margin of 25 rounds. However, merely measuring the security margin numerically is insufficient to illustrate the reality of the threat faced. An experimental approach based on a Proof-of-Concept (PoC) is required to verify the feasibility of implementing such attacks on Bitcoin's operational infrastructure [8].

This study aims to model the decay rate of the SHA-256 security margin quantitatively based on historical trends of reduced-round attacks, project its long-term resilience within the context of the Bitcoin ecosystem, and demonstrate the theoretical impacts of these attacks through PoC simulations.

II. Literature Review

A. Cryptographic Hash Functions and SHA-256

A hash function H is a mathematical function that takes an input of variable size and produces a fixed-size output [1]. The operational structure of SHA-256 is built using the Merkle-Damgård construction [9], where the security of the hash function depends on the resilience of its compression function against preimage, second preimage, and collision attacks [4].

B. Internal Mechanism of SHA-256

SHA-256 produces a 256-bit hash value by processing a 512-bit message block and a 256-bit chaining variable through a message schedule and a state update [4].

The 512-bit message block is divided into 16 32-bit words (W_0 to W_{15}) and expanded into 64 words (W_0 to W_{63}) recursively for $16 \leq t < 64$:

$$W_t = \sigma_1(W_{t-2}) + W_{t-7} + \sigma_0(W_{t-15}) + W_{t-16} \quad (1)$$

where:

$$\sigma_0(X) = (X \ggg 7) \oplus (X \ggg 18) \oplus (X \gg 3) \quad (2)$$

$$\sigma_1(X) = (X \ggg 17) \oplus (X \ggg 19) \oplus (X \gg 10) \quad (3)$$

State updates are performed in 64 rounds using eight 32-bit working registers (A, B, C, D, E, F, G, H). In each round t, the registers are updated through:

$$T_1 = H_t + \Sigma_1(E_t) + \text{Ch}(E_t, F_t, G_t) + K_t + W_t \quad (4)$$

$$T_2 = \Sigma_0(A_t) + \text{Maj}(A_t, B_t, C_t) \quad (5)$$

$$\begin{aligned} A_{t+1} &= T_1 + T_2 & E_{t+1} &= D_t + T_1 \\ B_{t+1} &= A_t & F_{t+1} &= E_t \\ C_{t+1} &= B_t & G_{t+1} &= F_t \\ D_{t+1} &= C_t & H_{t+1} &= G_t \end{aligned}$$

The SHA-256 compression function adopts the Davies-Meyer scheme with a feed-forward step at the end of the process to disrupt the mathematical reversal that could compromise the preimage resistance property [10]:

$$f(H_{i-1}, M_i) = E_{M_i}(H_{i-1}) \oplus H_{i-1} \quad (6)$$

C. Historical Cryptanalysis of SHA-256

Based on degrees of freedom and the attacker's control over input parameters [11], attacks are classified into several types, including *Full Collision*, *Full Preimage*, *Semi-Free-Start (SFS) Collision*, and *Pseudo-Collision* [**?**].

The development of collision attacks on SHA-256 began with 21 rounds by Nikolić et al. [2], increased to 24 rounds [11], and reached a record of 38 rounds by Mendel et al. [6] in 2013. This record stood for a decade before being validated using a hybrid SAT+CAS approach by Alamgir et al. [12] and subsequently broken to 39 rounds by Y. Li et al. [3] in 2024 using SAT/SMT-based sparsity control. Meanwhile, progress in preimage attacks has been slower, with the best record at 41 rounds by Sasaki et al. [13].

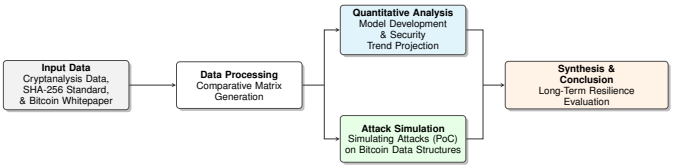

Fig. 1. General Research Architecture Diagram

III. Problem Analysis and Solution Design

A. Quantitative Security Trend Modeling

To compare attack capabilities objectively, the exponential attack complexity of 2^k is normalized into a metric of efficiency per round (*Security Bits per Round*, B) based on differential cryptanalysis theory [14]:

$$B = -\log_2 p = \frac{k}{R_{broken}} \quad (7)$$

The projected theoretical complexity on the full 64 rounds is calculated by:

$$k_{64} = B \times 64 \quad (8)$$

Not all literature data are included in the regression calculations. A data point is chosen as part of the *frontier* if it successfully breaches a higher number of rounds than previous years or possesses a lower complexity per round (B) than other attacks on the same number of rounds in that year. Additionally, a 2001 Anchor Point is introduced under the assumption that $R_{broken} = 0$ [4].

The Weighted Least Squares (WLS) method is applied by assigning weights of $w = 1.0$ for standard conditions, $w = 0.5$ for SFS, and $w = 0.25$ for other non-standard categories. Three hypothesis models are tested:

1) *Linear*: $f(x) = a(x - 2001) + b$
2) *Logarithmic*: $f(x) = a \ln(x - 2000) + b$
3) *Exponential*: $f(x) = b \cdot e^{a(x-2001)}$

B. Proof-of-Concept Simulation Design

To validate the practical impact on the Bitcoin ecosystem, a decentralized protocol simulator was designed with several operational simplifications (Table I) to accommodate local personal computing limits [15].

TABLE I
Simplification of the Proof-of-Concept Simulation Code

Aspect	Original Bitcoin Protocol	PoC Simulation
Transaction Size	Variable (> 100 bytes)	*Fixed* (64 bytes)
I/O Structure	Usually 1-in, 2-out	1-in, 1-out
Fee Mechanism	Block reward + Tx Fees	Only block reward
Balance Calculation	UTXO system	Ledger accumulation
Difficulty Target	Adaptive and extremely high	Static and low
Attribute Rules	Strict verification	Flexible (collision vector)

Four testing scenarios are implemented:

1) *Scenario 1*: Manipulation of the transaction *nValue* attribute to alter the transfer nominal without compromising the *Merkle Root* integrity using a 39-round collision vector [3].

2) *Scenario 2*: Redirection of the destination of funds by modifying the recipient address in the *scriptPubKey* attribute [17].
3) *Scenario 3*: Evaluation of attack failure due to semantic rule violations of the ECDSA signature in the *scriptSig* [11].
4) *Scenario 4*: Exploitation of the PoW consensus mechanism through preimage search using the Z3 *Logic Solver*, followed by simulations of *Shadow Mining* and *Chain Reorganization*.

IV. Implementation and Testing

A. Security Trend Projection Analysis

The accuracy of the three tested regression models is evaluated using R^2, MAPE, and RMSE metrics (Table II).

The logarithmic regression model consistently provides the best fit for most of the standard data groups. For the academic *Standard Collision* group, this model records an R^2 of 0.9742, MAPE of 6.97%, and RMSE of 1.89 *rounds*. The asymptotic nature of the logarithmic function successfully captures the deceleration of the cryptanalysis progress rate, which is driven by the strengthening of the internal confusion and diffusion barriers of SHA-256 as it approaches full rounds [16].

Based on the long-term extrapolation of the standard logarithmic model, the theoretical security margin of SHA-256 against standard collision attacks remains secure beyond the year 2100.

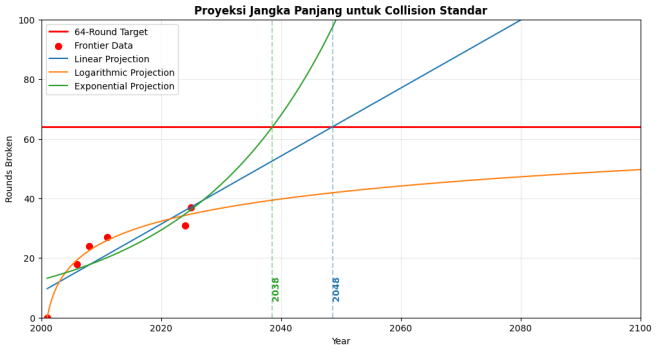

Fig. 2. Long-Term Projection of the Security Margin on the Standard Collision Group up to the Year 2100.

However, for the *Non-Standard Collision* group which allows initial value modifications (SFS), the linear model yields the best accuracy ($R^2 \geq 0.93$, MAPE $\leq 14.07\%$). This reflects that the advancement of modern solver automation tools is expanding constantly to bypass the internal limits of the compression function. Based on the non-standard logarithmic model, the 64-round threshold is predicted to be reached by the year 2073.

For standard *preimage* attacks, although the logarithmic model projects the depletion of the theoretical security margin by 2030, the practical security status of the Bitcoin ecosystem remains secure because the complexity of academic attacks is still at an extreme level ($> 2^{200}$), which physically exceeds the global computational capacity.

B. Proof-of-Concept Simulation Results

The PoC simulation successfully demonstrates the impact of each attack scenario:

1) *Scenario 1*: The attacker successfully inserts a manipulated transaction (M2) with a smaller nominal value into the blockchain. Because the *TxID* of M1 (legitimate) and M2 are identical, the *Merkle Root* integrity of the block is still deemed valid by the network, resulting in a direct financial loss for the recipient.

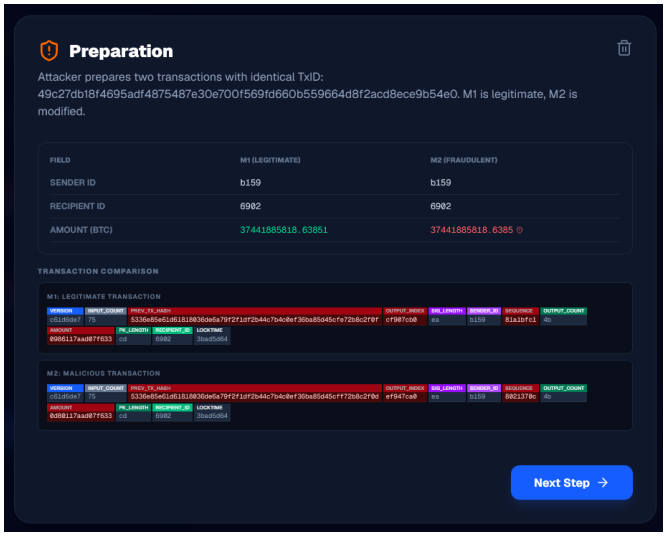

Fig. 3. Visualization of binary differences in the nValue attribute between the legitimate (M1) and manipulated (M2) transactions.

2) *Scenario 2*: Bit modification in the *scriptPubKey* field successfully redirects the destination of funds to the attacker's wallet address without disrupting the matching of the *Merkle Root* value in the block header.
3) *Scenario 3*: A collision attack using the vector by Indesteege et al. [11] fails to enter the simulator network's *mempool*. The bit corruption in the *scriptSig* causes the ECDSA digital signature verification to fail. This proves the efficacy of the multi-layered defense of the Bitcoin protocol, where the validity of the hash function alone is insufficient to manipulate consensus.
4) *Scenario 4*: The *Z3 Logic Solver* successfully finds a fraudulent *nonce* on the reduced-round version. The attacker successfully launches a *Shadow Mining* attack and forces a *Chain Reorganization* using the *longest chain rule*, which relegates the honest block to an *orphaned* status. However, on the double-hashing version (SHA-256d), the explosion of logical clauses in the solver makes the search remain impractical.

V. Conclusion and Future Work

A. Conclusion

This study concludes that the SHA-256 algorithm in the Bitcoin ecosystem is currently in an exceptionally secure state. The security margin decay rate can be modeled most accurately using a logarithmic regression function, reflecting the

TABLE II
RECAPITULATION OF ACCURACY (R^2), ERROR (MAPE %), AND RMSE OF THE TREND MODELS

Data Group	Model	Academic			Practical			Quantum		
		R^2	MAPE	RMSE	R^2	MAPE	RMSE	R^2	MAPE	RMSE
Collision Std.	Linear	0.76	15.70	5.77	0.66	17.36	6.01	0.13	37.83	16.28
	Logarithmic	**0.97**	**6.97**	**1.89**	**0.98**	**6.23**	**1.56**	**0.64**	**18.63**	**10.53**
	Exponential	0.66	14.97	6.82	0.52	15.81	7.19	-0.21	43.59	19.22
Collision Non-Std.	**Linear**	**0.94**	**12.87**	**4.47**	**0.93**	**14.07**	**4.48**	-	-	-
	Logarithmic	0.76	25.57	9.22	0.78	18.89	7.74	-	-	-
	Exponential	0.91	18.81	5.63	0.88	21.71	5.75	-	-	-
Preimage Std.	Linear	0.95	12.63	4.68	-	-	-	0.99	6.02	2.36
	Logarithmic	**0.99**	**3.84**	**1.38**	-	-	-	**0.99**	**3.34**	**1.28**
	Exponential	0.81	18.35	8,96	-	-	-	0.88	14.15	7.70
Preimage Non-Std.	Linear	0.99	6.02	2.36	-	-	-	-	-	-
	Logarithmic	**0.99**	**3.34**	**1.28**	-	-	-	-	-	-
	Exponential	0.88	14.15	7.70	-	-	-	-	-	-

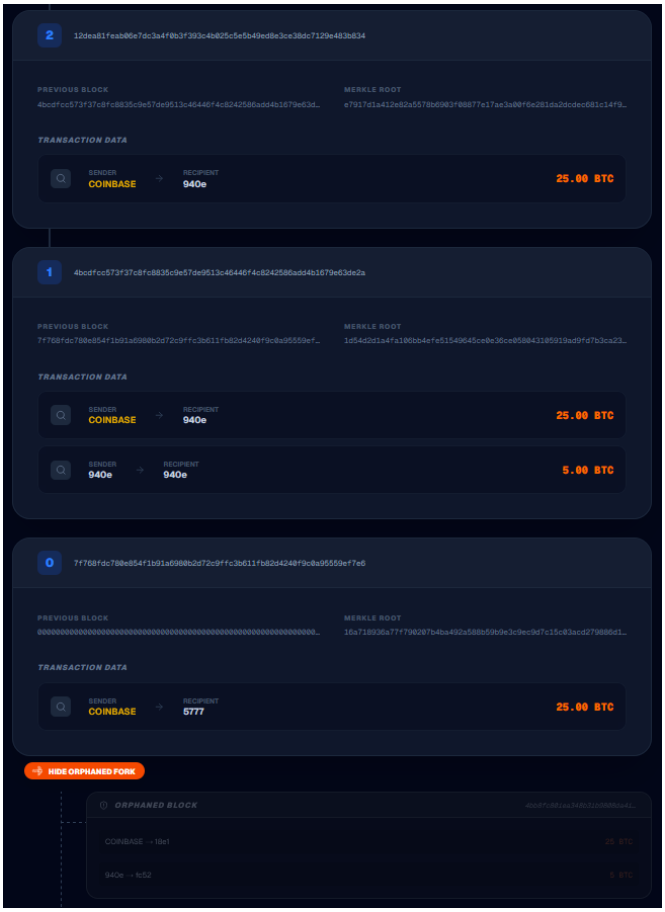

Fig. 4. Interface view of the simulator during a Chain Reorganization caused by the attacker's shadow chain being longer than the honest chain.

algorithm's internal resistance properties. Projections indicate that the security margin against standard collision attacks will persist beyond the year 2100.

Although the theoretical impact of the attacks is successfully demonstrated through PoC simulations, the Bitcoin protocol is equipped with a multi-layered defense system. The ECDSA signature mechanism mitigates the execution of semantically manipulated transactions, and the double-hashing structure increases the complexity of algebraic formulations on the logic solver during preimage searches.

B. Future Work

Future research could enhance the representation level of the PoC simulation protocol by integrating more complex features, such as Segregated Witness (SegWit) transactions, or modeling the dynamic growth trends of the global hash rate in the future to acquire more precise estimates of the algorithm transition timeframe.

REFERENCES

[1] B. Preneel, "The First 30 Years of Cryptographic Hash Functions and the NIST SHA-3 Competition," in *International Conference on Cryptology in Africa (AFRICACRYPT 2009)*, pp. 1–14, 2009.

[2] I. Nikolić and A. Biryukov, "Collisions for Step-Reduced SHA-256," in *International Workshop on Fast Software Encryption (FSE 2008)*, vol. 5086, pp. 1–15, 2008.

[3] Y. Li, F. Liu, and G. Wang, "New Records in Collision Attacks on SHA-2," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT 2024)*, pp. 158–186, 2024.

[4] *National Institute of Standards and Technology*, "Secure Hash Standard (SHS)," FIPS PUB 180-4, August 2015.

[5] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," Online, 2008.

[6] F. Mendel, T. Nad, and M. Schläffer, "Improving Local Collisions: New Attacks on Reduced SHA-256," in *EUROCRYPT 2013*, pp. 262–278, 2013.

[7] E. Homsirikamol, P. Morawiecki, M. Rogawski, and M. Srebrny, "Security Margin Evaluation of SHA-3 Contest Finalists Through SAT-Based Attacks," in *CISIM 2012*, pp. 56–67, 2012.

[8] C. E. Kendig, "What is Proof of Concept Research and How Does It Generate Epistemic and Ethical Categories for Future Scientific Practice?," *Science and Engineering Ethics*, vol. 22, no. 3, pp. 735–753, 2015.

[9] H. Tiwari, "Merkle-Damgård Construction Method and Alternatives: A Review," *Journal of Information and Organizational Sciences*, vol. 41, no. 2, pp. 283–304, 2017.

[10] R. C. Merkle, "One Way Hash Functions and DES," in *CRYPTO '89*, vol. 435, pp. 428–446, 1989.

[11] S. Indesteege, F. Mendel, B. Preneel, and C. Rechberger, "Collisions and Other Non-random Properties for Step-Reduced SHA-256," in *SAC 2008*, vol. 5381, pp. 276–293, 2008.

[12] N. Alamgir, S. Nejati, and C. Bright, "SHA-256 Collision Attack with Programmatic SAT," *arXiv preprint arXiv:2406.20072*, 2024.

[13] Y. Sasaki, L. Wang, and K. Aoki, "Preimage Attacks on 41-Step SHA-256 and 46-Step SHA-512," in *FSE 2009*, pp. 352–371, 2009.

[14] E. Biham and A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems," *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

[15] K. Okupski, "Bitcoin Developer Reference," *Technische Universiteit Eindhoven*, Working Paper, July 2016.

[16] J. S. Armstrong, *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Boston, MA: Springer US, 2001.
[17] Z. Zhang, M. Li, L. Gao, and M. Wang, "Collision Attacks on SHA-256 up to 37 Steps with Improved Trail Search," in *EUROCRYPT 2025*, 2025.
[18] Blockchain.com, "Total Hash Rate (TH/s)," 2026. [Online]. Available: https://www.blockchain.com/charts/hash-rate
[19] Chick3nman, "Hashcat v6.2.6 Benchmark on the NVIDIA RTX 4090," 2022. [Online]. Available: https://gist.github.com/Chick3nman/32e662a5bb63bc4f51b847bb422222fd