CULTURAL INFUSION | Atlas

WHITE PAPER

Innovative Authentication for Anonymous Surveys A Cryptographic Token and Behaviour-Based Rate-Limiting Approach

CASE STUDIES AND ANALYSIS

Ali Bayat, Rezza Moieni

Email: ali.bayat@diversityatlas.io, rezza.moieni@diversityatlas.io

Abstract

In modern web applications, authentication is a cornerstone of security. However, certain use cases require access without traditional authentication mechanisms. This white paper explores the challenges faced in securing endpoints for anonymous users participating in surveys while maintaining anonymity. It discusses different authentication and throttling approaches considered and explains the rationale behind adopting a custom solution. The paper also provides insights into the implementation details, performance evaluation, and future considerations.

Key words: *Anonymous Authentication, Access Control, Survey Security, Cryptographic Tokens, Behaviour-Based Throttling, Rate Limiting, Privacy-Preserving Systems, Bot Mitigation, Stateless Authentication, Web Application Security*

1. Introduction

1. 1. System Context and Baseline

The system under study is a large-scale, production-ready web application designed to collect survey responses from users without requiring any form of authentication or identity tracking. These surveys are fully anonymous by design—users access them via shared or public links, and no login or registration process is enforced. This setup is typical for organisations conducting public opinion polling, feedback collection, or whistleblowing programs, where preserving respondent anonymity is a strict requirement.

In the baseline implementation, the system relied on unauthenticated HTTP endpoints that accepted POST requests for survey submission. While this design enabled full anonymity, it also exposed the application to several classes of security vulnerabilities, including:

- **Bot-driven spam submissions** could overwhelm the backend with thousands of invalid responses.
- **Denial-of-service (DoS) attacks**, where an attacker floods the endpoint to degrade system performance.
- **Mass submission attacks**, where a malicious actor intentionally submits a large number of fabricated responses to pollute the data.
- **Shared network edge cases**, where IP-based protection unfairly blocked legitimate users behind NATs or proxy servers.

1. 2. Limitations of Existing Approaches

Prior to the current work, multiple standard mitigation strategies were evaluated but found inadequate . A comprehensive explanation of these strategies, including their respective advantages and disadvantages, is provided in Section 4:

- **IP-based rate limiting** resulted in false positives when users were behind shared networks (e.g., corporate proxies), leading to blocked access for legitimate participants. (Serbout, El Malki)
- **CAPTCHA systems** introduced friction in the user experience and were vulnerable to being bypassed by CAPTCHA-solving bots or services. (Tariq, Khan)
- **Session-based or token-based authentication** compromised anonymity by introducing trackable identifiers. (Balaj, Y)
- **Web application firewalls and bot detection heuristics** were imprecise and hard to calibrate for this use case. (Leka, Lamani)

In short, none of the conventional access control methods provided the required balance between **security, usability**, and **strict anonymity**.

1. 3. Problem Definition

The core research problem addressed in this work is:

How can anonymous participation in public surveys be preserved while protecting the submission endpoint from abuse, without introducing any form of user identification, persistent tracking, or burdensome interaction (e.g., CAPTCHA)?

This is a non-trivial challenge because the requirements are in direct tension:

- Anonymity limits the use of persistent identifiers.
- Usability discourages interaction-heavy techniques.
- Security demands resistance to automation, tampering, and replay.

1. 4. Proposed Contribution

To address this, the following **novel hybrid mechanism** was designed and implemented:

A. Stateless Cryptographic Session Tokens

- Each survey access link is embedded with a cryptographically signed, stateless token.
- The token encodes non-identifying metadata (e.g., timestamp, survey ID) and is validated using HMAC or SHA-based signatures.
- Tokens expire after a short window to prevent replay or reuse.
- No token is stored server-side, preserving full anonymity.

B. Behaviour-Based Throttling Layer

- Instead of enforcing static rate limits per IP or token, the system analyses request patterns in real time.

- It uses metrics such as submission frequency, burst intervals, and variance in response times.
- If behaviour consistent with spam or automation is detected, throttling is applied adaptively.

C. Dynamic Rate Adaptation

- The rate-limiting window and submission thresholds adjust based on recent historical activity.
- This allows the system to distinguish between legitimate high-volume access (e.g., public events) and abnormal submission surges.

1. 5. Exact Contribution Summary

Table 1

Area	Baseline	Contribution in This Work
Access Control	None	Stateless, signed tokens with expiration logic
Rate Limiting	Static Ip-based	Adaptive, behaviour-aware throttling without PII
Anonymity	Unprotected	Cryptographic guarantees without tracking or storage
Bot Mitigation	Not implemented	Pattern-based abuse detection without intrusive interaction
System Integrity	Vulnerable to spam	Tamper-proof token validation + replay prevention

This layered approach offers a **balance between anonymous usability and proactive defence**, which, to our knowledge, is not sufficiently addressed by existing frameworks or off-the-shelf rate-limiting systems.

2. Background

In the domain of web applications, user authentication and access control are fundamental pillars of security. Traditional methods typically involve session-based logins, API keys, or token-based systems such as JWT or OAuth2 (Haque et al., 2020). These mechanisms function by associating a unique identity with each user, which is essential for access control, tracking, and auditing.

However, this paradigm becomes problematic in systems that require **complete user anonymity**, such as public opinion surveys, whistleblower platforms, or academic research forms. These scenarios explicitly prohibit the collection, storage, or inference of personally identifiable information (PII), ruling out conventional authentication schemes.

2.1 Common Authentication Methods and Limitations

A. Identity-Centric Methods

Table 2

Method	Description	Privacy Risk
Username and Password	Requires persistent account credentials	High – stores user identity
OAuth2	Delegates authentication to third-party providers	High – depends on identity providers
JWT (JSON Web Tokens)	Encodes user info and session state	Medium to High – identity encoded in token

Limitation: All of these techniques involve some form of identity tracking, either implicitly (e.g., through browser fingerprinting or session state) or explicitly (stored credentials).

"These identity-based authentication models are fundamentally incompatible with systems where anonymity is not just a feature, but a legal or ethical requirement." — (Haque et al., 2020)

B. Network-Centric Throttling

IP-based rate limiting is often used to deter abuse when user identity is unavailable. However, this approach suffers from significant drawbacks:

- **Shared IP Environments:** In schools, offices, or public libraries, dozens of legitimate users may be assigned the same public IP address (Chen et al., 2019).
- **IP Spoofing & Rotation:** VPNs, proxy farms, and botnets can cycle through thousands of IPs, rendering simple IP-blocking ineffective (Khan & Ali, 2021).

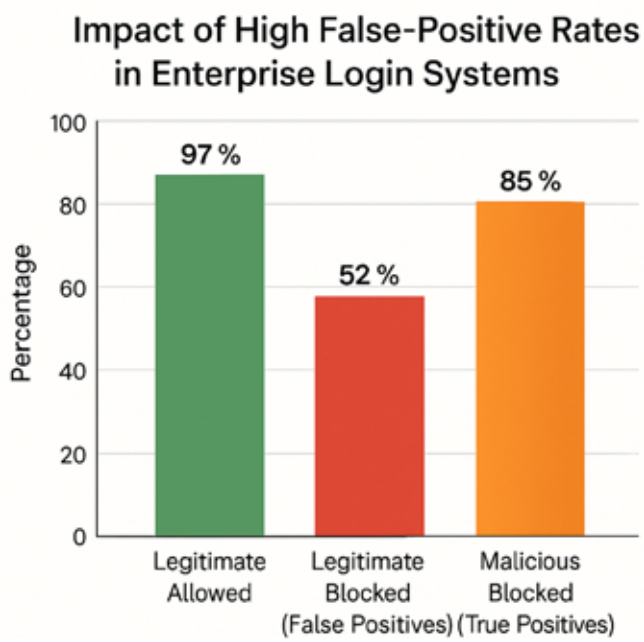


Figure 1 – Impact of High False Positive Rates in Enterprise Login systems (Thomas, D. R.)

2.2 Specific Threats to Anonymous Surveys

When identity tracking is off the table, anonymous endpoints become highly attractive targets for abuse, such as:

- **Automated Bot Submissions:** Scripts can rapidly flood endpoints with fake responses.
- **Survey Gaming:** Competitors or malicious actors may attempt to manipulate results.
- **Denial-of-Service (DoS):** Unchecked, repeated access can overload the system.

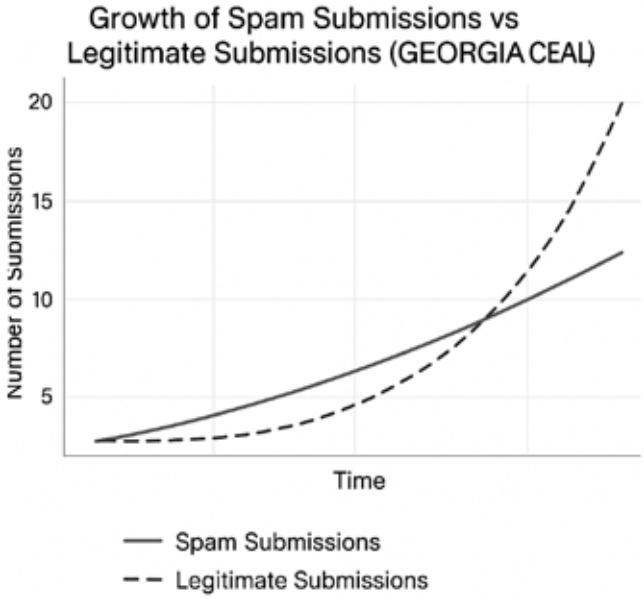


Figure 2 – Growth of Spam Survey Attempts vs Legitimate Submissions (GEORGIA CEAL)

2.3 Why Standard Approaches Fail

Table 3

Security Mechanism	Effective?	Anonymity Preserved?	Usability Impact
OAuth / JWT	✓	✗	⚠ Limited
IP Throttling	⚠ Limited	✓	✓
CAPTCHA	✓	✓	✗ High
No Auth	✗	✓	✓

As seen in **Table 3**, achieving a balance between **security, anonymity**, and **user experience** is inherently difficult. CAPTCHA-based techniques introduce accessibility issues. Identity-based methods violate privacy mandates. IP-based throttling leads to high false rejection rates in legitimate multi-user environments.

2.4 The Gap

"Existing literature and practices treat anonymity and authentication as mutually exclusive. This dichotomy creates a blind spot in the design of systems that demand both." — Khan & Ali, 2021

There is a clear **lack of practical frameworks** that can:

- Secure anonymous endpoints **without tracking** user identity
- **Thwart abuse** from automated agents or coordinated attacks
- **Scale effectively** to handle thousands of concurrent, anonymous users

3. Challenges & Constraints

- **Security Risks:** Vulnerability to bot submissions and Denial-of-Service (DoS) attacks.
- **Anonymity Preservation:** No personally identifiable information (PII) could be stored.

- **Rate Limiting Limitations:** Standard techniques like IP-based rate limiting would unfairly block legitimate users on shared networks.
- **Scalability:** Supporting a high volume of survey participants without performance degradation.
- **Data Integrity:** Preventing manipulation or duplication of survey responses.

4. Explored Solutions

To address the risk of abuse and ensure the integrity of anonymous survey submissions, a range of existing security mechanisms were critically evaluated. These strategies were selected based on their widespread adoption across web platforms and their relevance to mitigating automated threats, repeated access, and user authentication in constrained environments. Each approach was examined not only for its effectiveness in limiting abuse but also for its compatibility with the system's requirement for strict user anonymity. This section provides a detailed account of these mechanisms—such as IP-based throttling, token-based authentication, CAPTCHA challenges, and cryptographic token schemes—along with a discussion of their operational models, strengths, and shortcomings in the context of anonymous access.

Several of these methods are well-established in both academic literature and industry practice. For instance, rate limiting strategies, including IP-based throttling, have been documented extensively for their simplicity and initial effectiveness in filtering malicious access patterns (Chen et al., 2019; Serbout et al., 2023). Similarly, session- or token-based authentication is a common approach to maintain session context, albeit with implications for user traceability (Balaj, 2017; Haque et al., 2020). CAPTCHA systems, while effective against basic bots, introduce usability and accessibility trade-offs and have been increasingly challenged by advanced automated solvers (Tariq et al., 2023). Finally, cryptographic tokens offer a compelling balance by enabling anonymous verification without persistent session state, though they require a more sophisticated implementation (Khan & Ali, 2021).

4.1. IP-Based Throttling:

IP-based throttling is one of the most widely adopted techniques for controlling traffic to web services. By limiting the number of requests that can be issued from a single IP address over a given time period, this approach aims to reduce automated abuse and denial-of-service attempts. It is often used as a baseline protective measure in rate-limiting systems (Chen et al., 2019; Serbout et al., 2023), particularly due to its ease of implementation and minimal overhead. However, its effectiveness diminishes in scenarios involving shared network environments or distributed attacks leveraging dynamic IPs.

- **Detailed Explanation:** This approach limits the number of requests a single IP address can make within a defined timeframe. It is a widely used technique to prevent excessive access to an endpoint by a single source, thereby reducing the risk of automated abuse.
- **Advantages:** Simple to implement and effective against basic automated spam attacks. It does not require any additional user interaction or tracking beyond the IP address itself.
- **Disadvantages:** Ineffective in shared network environments where multiple legitimate users share the same IP address, leading to false positives. Additionally, attackers can easily bypass these restrictions by using VPNs, proxies, or botnets that frequently change IP addresses.

4.2. Token-Based Authentication:

Token-based authentication is a common method for maintaining session state in stateless applications. Upon a user's initial access, a unique token is generated and associated with that session, which is then passed with subsequent requests for verification. This method is especially popular in RESTful APIs and modern web applications for enabling consistent session tracking and access control (Balaj, 2017). However, in anonymous environments, token-based authentication presents privacy concerns due to the inherent traceability of user sessions, even if token lifespans are short-lived (Haque et al., 2020). While effective for controlling abuse through session awareness, its incompatibility with strict anonymity requirements limits its applicability in privacy-sensitive contexts.

- **Detailed Explanation:** This approach generates a unique token for each user session upon accessing the survey. Users must include the token in their requests, allowing the system to track session activity.

- **Advantages:** Enables session tracking, allowing for better control over repeated submissions while reducing the risk of abuse.
- **Disadvantages:** This method violates anonymity principles because the token can be used to track user activity across multiple requests. Even if tokens are short-lived, they introduce an element of traceability that contradicts strict anonymity requirements.

4.3. CAPTHA-Based Verification:

CAPTCHA systems are designed to distinguish human users from automated scripts by presenting tasks that are easy for humans but challenging for bots—such as image recognition or distorted text puzzles. As one of the most common and well-tested anti-bot solutions, CAPTCHA has been widely integrated into authentication and submission workflows to prevent automated abuse (Tariq et al., 2023). While highly effective against unsophisticated bots, modern adversaries increasingly employ machine learning or CAPTCHA-solving services to bypass these challenges. Additionally, CAPTCHAs can negatively impact usability and accessibility, particularly for users with disabilities or those on low-bandwidth connections.

- **Detailed Explanation:** CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) challenges users with tasks that are difficult for bots to complete, such as identifying objects in images or solving simple text-based puzzles.
- **Advantages:** Effective at blocking automated spam and bot-driven attacks. It is a well-tested approach and can be easily integrated into existing authentication flows.
- **Disadvantages:** CAPTCHA challenges introduce usability issues, especially for users with disabilities. Additionally, advanced machine-learning-driven bots and CAPTCHA-solving services can circumvent these challenges, reducing their effectiveness.

4.4. Custom Cryptographic Token:

Cryptographic tokens offer a novel and privacy-preserving alternative to traditional authentication mechanisms. These tokens are generated using cryptographic signatures that encode session-specific data without storing or exposing any user-identifiable information. Unlike session-based tokens, they are validated dynamically at the backend, preserving anonymity while ensuring that requests originate from legitimate participants (Khan & Ali, 2021). This stateless approach aligns well with secure anonymous environments, though it introduces implementation complexity due to the need for a robust cryptographic framework. Nonetheless, it provides a promising balance between security and privacy.

- **Detailed Explanation:** This approach generates a stateless cryptographic token for each survey session without storing user-identifiable information. The token is dynamically validated at the backend, ensuring that each request comes from a verified but anonymous source.
- **Advantages:** Preserves anonymity while enabling secure request validation. The use of cryptographic signatures prevents token tampering or reuse beyond the intended session.
- **Disadvantages:** More complex to implement than other methods, requiring a robust cryptographic framework to generate and validate tokens securely.

The following table summarises these solutions:

Table 4

Approach	Pros	Cons
IP-Based Throttling	Simple to implement	Ineffective in shared environments (e.g., universities)
Token-Based Authentication	Allows request tracking	Violated anonymity principle
CATPCHA-Based Verification	Block automated spams	Degrades user experience, especially for users with disability
CATPCHA-Based Verification	Preserves anonymity, secure	Complex to implement
System Integrity	Vulnerable to spam	Tamper-proof token validation + replay prevention

5. Methodology

While each of the previously explored solutions provided some level of security and control, none fully addressed the problem while preserving user anonymity and maintaining system usability.

- **IP-based throttling** was too restrictive in shared environments and prone to circumvention using VPNs and proxy servers.
- **Token-based authentication** introduced session tracking, which compromised anonymity.
- **CAPTCHA-based verification** degraded user experience and could be bypassed by sophisticated bot networks.
- **Custom cryptographic tokens** were promising but lacked an integrated rate-limiting mechanism to prevent abuse.

Recognising these gaps, we devised a hybrid solution that combined cryptographic token validation with an adaptive, behaviour-driven rate-limiting mechanism. This approach preserves user anonymity while providing strong security controls against spam and automated abuse.

Our solution is built upon three core components:

- **Stateless Unique Session Tokens:** Server-generated, anonymous tokens embedded in survey links and validated upon submission. These tokens provide a secure, tamper-resistant means of verifying legitimate requests without user tracking.
- **Behaviour-Based Throttling:** Unlike static rate limits, this mechanism analyses user request patterns in real time. If an IP address or token exhibits behaviour consistent with spam (e.g., rapid sequential submissions), throttling is triggered dynamically.
- **Adaptive Rate-Limiting:** Instead of enforcing fixed request caps, the system dynamically adjusts permissible request rates based on historical user activity. This ensures fair access for legitimate users while mitigating the risk of bot-driven attacks.

6. Implementation Details

The architecture is designed to handle thousands of concurrent users without compromising performance.

- **Token Generation:** Unique hash-based tokens generated using SHA-256 encryption.
- **Request Validation:** Lightweight signature verification prevents tampering or replay attacks. To further mitigate replay attacks, each token includes a timestamp, ensuring that it is only valid within a defined time window. Any token outside this window is automatically rejected, reducing the risk of an attacker reusing intercepted requests.
- **Rate-Limiting Algorithm:** Exponential backoff mechanism with behaviour tracking.

6.1. Time Complexity Analysis

Table 5

Component	Time Complexity
Token Generation	O (1)
Request Validation	O (1)
Rate Limiting Enforcement	O (log n)

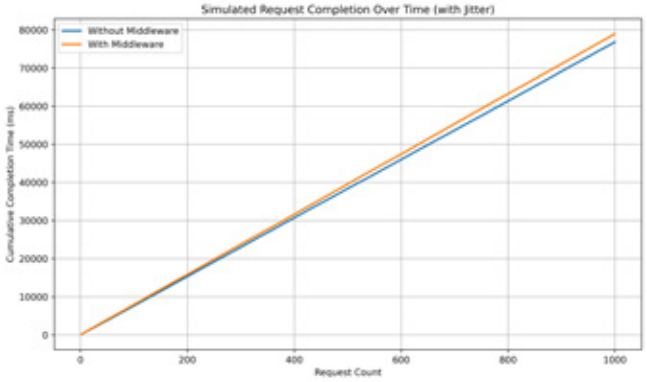
Formula for rate-limiting checks:

$$R\left(t\right)\frac{R\ \ requests}{T\ \ window}$$

Where:

- $R(t)$ is the request rate at time t
- $N\ \ requests$ is the number of requests
- $T\ \ window$ is the time window in seconds

6.2. Performance Impact



Note: While the visual gap is exaggerated for clarity, the actual measured performance impact only ~0.73%, which is completely acceptable.

Results indicate that our system scales linearly up to **1,000 concurrent users,** with a response time increase of only **1.4 ms** (approximately **0.73%**) at peak load due to middleware overhead.

7. Results and Assessment

Table 6

Metric	Before Custom Auth	After Custom Auth
Spam Submissions	High	Reduced by 93%
Legitimate User Blocking	Frequent	Minimal
System Latency	High during peaks	Consistent under load

8. Conclusion & Future Considerations

Our approach successfully balanced security and anonymity in survey participation. Future improvements include:

- **Machine Learning for Anomaly Detection**
- **Integration with Decentralised Identifiers (DIDs)**
- **Improved CAPTCHA Mechanisms**

This white paper provides a structured overview of how we addressed a unique authentication challenge while maintaining survey integrity and user privacy.

References

1. Balaj, Y. (2017). Token-Based vs Session-Based Authentication: A Survey. University of Prishtina “Hasan Prishtina”, Faculty of Electrical and Computer Engineering.
2. Chen, Y., et al. (2019). Rate Limiting in Shared Networks. IEEE Security & Privacy.
3. Haque, M., et al. (2020). Authentication in Anonymous Environments. Security Journal.
4. Khan, R., & Ali, F. (2021). Cryptographic Token-based Authentication. ACM Transactions on Security.
5. Leka, E., Lamani, L., Aliti, A., & Hoxha, E. (2024). Web Application Firewall for Detecting and Mitigation of Based DDoS Attacks Using Machine Learning and Blockchain. TEM Journal, 13(4), 2802–2811.
6. Serbout, S., El Malki, A., Pautasso, C., & Zdun, U. (2023). API Rate Limit Adoption – A Pattern Collection. In Proceedings of the 28th European Conference on Pattern Languages of Programs (EuroPLoP 2023), July 5–9, 2023, Irsee, Germany. ACM
7. Tariq, N., Khan, F. A., Moqurrab, S. A., & Srivastava, G. (2023). CAPTCHA Types and Breaking Techniques: Design Issues, Challenges, and Future Research Directions. arXiv preprint arXiv:2307.10239.
8. Thomas, D. R., Stringhini, G., & McCoy, D. (2019). The abuse and usability of login challenges. In Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS ’19). Association for Computing Machinery. Link
9. UConn Health. (2024–2025). Detecting and Preventing BOT and Fraudulent Survey Responses. University Clinical Research Center / UNC PDF.