Research Paper Proposal: Multi-Expert AI System for Sneaker Bot Detection

Arnav Khinvasara[1]

[1]University of California, San Diego

Abstract

This research proposes a novel mixture-of-experts machine learning architecture to detect and classify automated bot traffic targeting limited-edition sneaker releases. Our approach addresses the technical challenges of identifying increasingly sophisticated bots without relying on extensive labeled training data, using innovative combinations of behavioral analysis, technical fingerprinting, and anomaly detection within a meta-learning framework.

1 Problem Statement

The limited-edition sneaker market faces significant challenges from automated purchasing bots that provide unfair advantages to scalpers, preventing legitimate consumers from accessing products at retail prices. Current detection methods struggle to keep pace with rapidly evolving bot technologies that can mimic human behavior with increasing sophistication. This research aims to develop a robust, adaptive system capable of distinguishing between legitimate human traffic and bot traffic across multiple behavioral and technical dimensions.

2 Current Approaches and Limitations

Current bot detection methods typically fall into several categories:

- **CAPTCHA and Challenge-Based Systems:** These systems present puzzles or challenges designed to be difficult for bots but easy for humans. However, modern bots utilize machine learning, optical character recognition, and human solving services to bypass these controls [1, 2].
- **Rule-Based Systems:** Many e-commerce platforms employ rule-based heuristics examining IP addresses, browsing patterns, and purchasing behavior. These systems lack adaptability to new bot techniques and generate high false positive rates [3].
- **Behavioral Analysis:** Some solutions analyze mouse movements, keystroke dynamics, and session timing. While effective against basic bots, sophisticated bots now incorporate randomized delays and human-like movement patterns to evade detection [4, 5].
- **Machine Learning Approaches:** Recent research has employed supervised learning for bot detection, but these approaches require extensive labeled datasets that are difficult to obtain and quickly become outdated as bot technologies evolve [6].

3 Novelty and Proposed Approach

We propose a novel Multi-Expert AI System for Bot Detection (MEAS-BD) with several innovative components:

3.1 Mixture-of-Experts Architecture

Our system will employ specialized expert models focusing on distinct aspects of web traffic:

- Temporal Pattern Expert (TPE)
- Navigation Sequence Expert (NSE)
- Input Behavior Expert (IBE)
- Technical Fingerprint Expert (TFE)
- Purchase Pattern Expert (PPE)

3.2 Meta-Learning Framework

A meta-learning layer will dynamically weight expert opinions based on context and confidence, allowing the system to adapt to new bot strategies without complete retraining.

3.3 Synthetic Data Generation

To address the lack of labeled training data, we will implement:

- Adversarial synthetic data generation
- Semi-supervised learning techniques
- Active learning procedures to maximize information gain from limited labeled examples

3.4 Adaptive Thresholding

The system will employ contextual risk scoring with dynamic thresholding based on:

- Product release demand levels
- Historical bot activity patterns
- Geographic and temporal anomaly detection

4 Security Applications

The security applications of this research extend beyond the sneaker market:

- E-commerce Fraud Prevention: Techniques developed can be applied to other high-value e-commerce sectors.
- Account Takeover Protection: Methods can help identify automated credential stuffing attacks.
- API Abuse Prevention: The system can identify and mitigate automated API abuse.
- DDoS Attack Mitigation: Bot detection mechanisms may assist in distinguishing malicious from legitimate traffic.
- Ticket Scalping Prevention: Similar techniques can be applied to event ticketing systems.

5 Importance and Impact

This research addresses several critical concerns:

- **Economic Fairness:** Sneaker bots create artificial scarcity, driving prices up by 300-500% on secondary markets [7], limiting access to products based on technical sophistication rather than fair market mechanisms.
- **Consumer Protection:** Legitimate consumers are increasingly excluded from the primary market, facing significantly inflated prices.
- **Brand Reputation:** Brands face consumer frustration when products are unavailable at retail, potentially damaging long-term customer relationships.
- **Technical Innovation:** The proposed techniques will advance the state of the art in behavioral biometrics, anomaly detection, and adversarial machine learning.

6 Risks and Challenges

6.1 Technical Risks

- False positives may block legitimate customers
- Adversarial attacks could manipulate the detection system
- Computational overhead may impact user experience

6.2 Ethical Considerations

- Privacy implications of detailed behavioral monitoring
- Potential for discrimination against certain user groups
- Arms race between detection and evasion technologies

6.3 Implementation Challenges

- Limited availability of labeled training data
- Need for real-time processing capabilities
- Integration with existing e-commerce infrastructure

7 Research Methodology

We will pursue a phased approach:

- Phase 1: Development of individual expert models and synthetic data generation
- Phase 2: Integration into meta-learning framework and baseline performance evaluation
- Phase 3: Simulated deployment testing and refinement
- Phase 4: Limited real-world pilot testing (pending industry partnerships)

8 Conclusion

The proposed MEAS-BD system represents a significant advancement in bot detection technology, specifically tailored to the challenges of the limited-edition sneaker market but with broader applications across e-commerce security. By combining multiple expert models within an adaptive framework, this research aims to create a solution that can evolve alongside bot technologies while minimizing false positives and operational overhead.

References

[1] Sivakorn, S., Polakis, I., & Keromytis, A. D. (2016). I am robot: (deep) learning to break semantic image captchas.

[2] Ye, G., Tang, Z., Fang, D., et al. (2018). Yet another text captcha solver: A generative adversarial network based approach.

[3] Amin, R. M., Ryan, J. J. C. H., & van Dorp, J. R. (2018). Detecting targeted malicious email.

[4] Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2018). Detecting automation of twitter accounts.

[5] Wei, W., Qu, Q., & Lu, J. (2019). Identifying web bot sessions based on mouse behavioral biometrics.

[6] Zhai, Y., Liu, Y., & Wang, L. (2020). Defending against web bot attacks: A machine learning approach.

[7] Thompson, D. (2020). The economics of sneaker reselling.