An Advanced, ML/DL-Driven Autonomous Defense System for Real-Time Detection, Mitigation, and Recovery from DDoS Attacks

Deepesh Patil, Saksham Dharmik, Naman Goyal and Tanishq Ingole

***Abstract*— Distributed Denial-of-Service (DDoS) attacks represent a significant and evolving threat to the stability and availability of cloud infrastructures, leading to severe service disruptions, compromised user experiences, and substantial financial losses [1]. These attacks overwhelm system resources, bypass conventional security measures, and exploit the inherent vulnerabilities of accessible cloud environments [1, 2]. This paper proposes an advanced, autonomous defense system driven by Machine Learning (ML) and Deep Learning (DL) to provide real-time detection, mitigation, and recovery from DDoS attacks on cloud systems [1]. The novelty of our approach lies in the integration of an Incremental Learning Model, which allows the system to adapt to emerging threats seamlessly without the need for complete retraining [4]. The methodology incorporates a multi-layered defense strategy, utilizing ML/DL models to efficiently identify traffic anomalies, dynamically filtering malicious packets via Web Application Firewalls (WAFs), and employing honeypots for the early identification of DDoS patterns [1, 5]. To ensure high availability and resilience, the architecture implements auto-scaling and load-balancing mechanisms for continuous operation during an attack [3]. Experimental results demonstrate the system's effectiveness, with the Incremental Random Forest model achieving a superior detection accuracy of 98.16% and a low False Positive Rate (FPR) of 2.28% (Internal Data). This self-healing, robust, and cost-effective solution significantly strengthens the security and reliability of modern cloud environments [4].**

INDEX TERMS: DDoS Protection, Cloud Security, Incremental Learning, Machine Learning, Honeypots, Anomaly Detection.

I. INTRODUCTION

In the contemporary digital landscape, the proliferation of malicious software presents a formidable threat to security systems globally [1]. The cybersecurity domain is in a state of constant flux, with new and increasingly sophisticated forms of malware emerging regularly [1]. These threats, including viruses, Trojans, and ransomware, are engineered to exploit system vulnerabilities, disrupt operations, and inflict financial damage [1]. As organizations increasingly migrate their critical operations and data storage to cloud infrastructures, this dependency has rendered cloud environments a prime target for Distributed Denial-of-Service (DDoS) attacks [2]. Such attacks can cause widespread service disruptions and significant reputational harm [2].

Traditional DDoS defense mechanisms, such as IP blacklisting and static rate limiting, are often inadequate against modern, multi-vector attack patterns due to their lack of adaptability [2, 6]. This creates a significant research gap in several key areas. First, there is a pressing need for adaptable systems that can intelligently and dynamically differentiate between legitimate user traffic and malicious attack traffic in real-time. Second, while Machine Learning (ML) and Deep Learning (DL) models offer promise, their high computational cost poses a challenge for real-time detection in high-traffic environments. Third, many existing solutions lack scalability and the capacity for incremental learning, requiring complete and resource-intensive retraining to adapt to new attack vectors.

This paper addresses these gaps by proposing a robust, adaptive, and autonomous DDoS defense solution specifically tailored for cloud environments. The novelty of our approach lies in its hybrid architecture that combines ML/DL techniques with an incremental learning framework, enabling the system to evolve its defensive capabilities over time without manual intervention [4]. The key contributions of this work are as follows:

- **Implementation of Hybrid ML and DL Models:** We utilize a combination of Random Forest, Deep Learning, and Incremental Random Forest models to achieve accurate and efficient detection of traffic anomalies.
- **Development of an Adaptive Defense Mechanism:** The system incorporates an incremental learning model, allowing it to adapt to new and evolving DDoS attack patterns without requiring complete retraining, thereby ensuring sustained effectiveness [**?**].
- **Integration of a Multi-Layered Defense Strategy:** We combine Web Application Firewalls (WAFs), honeypots, and automated load-balancing to create a comprehensive, multi-layered defense that enhances system resilience.
- **Design for Real-Time Response and Scalability:** The architecture is designed for minimal latency, ensuring continuous service availability and scalability for high-traffic cloud environments, even during an active attack.

II. RELATED WORK

The body of research on DDoS mitigation is extensive. Previous works can be broadly categorized into several key areas, from analyses of attack evolution to the application of intelligent detection models.

A. EVOLUTION AND IMPACT OF DDOS ATTACKS

DDoS attacks have evolved significantly in complexity, frequency, and scale, posing a persistent threat to cloud infrastructure [1]. The shift from simple volumetric attacks to sophisticated, multi-vector assaults complicates detection and mitigation. Attackers leverage cloud vulnerabilities, botnets, and spoofing techniques to disrupt services. Research by Phan and Park (2019) demonstrated that a successful DDoS attack on a cloud system severely degrades service reliability

and can cause cascading failures across interconnected resources, highlighting the need for adaptive, real-time defense [3].

B. TRADITIONAL MITIGATION TECHNIQUES AND THEIR LIMITATIONS

Conventional DDoS mitigation techniques, including IP blacklisting and rate limiting, have been widely deployed but often fail to distinguish legitimate traffic from malicious traffic, particularly in large-scale cloud environments [2]. Signature-based detection systems are effective against known attack patterns but struggle with novel or polymorphic threats. Traditional methods are insufficient for modern cloud systems, as they cannot adapt to new attack types [7]. This has driven a shift towards dynamic, ML-based approaches that analyze traffic patterns in real-time.

C. MACHINE LEARNING AND DEEP LEARNING APPROACHES

Machine learning has emerged as a powerful tool for DDoS detection. Anomaly detection models can analyze vast datasets to identify unusual patterns indicative of an attack [1]. According to Yin, Zhang, and Yang (2018), ML models show significant potential in distinguishing malicious activity from regular traffic fluctuations [7]. While supervised algorithms like Support Vector Machines (SVM) and Decision Trees are commonly used, their reliance on labeled data is a limitation. Consequently, unsupervised models are increasingly favored for adaptive frameworks. Deep Learning (DL) offers enhanced capabilities, with models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) recognizing intricate patterns that simpler models may miss [2, 6]. Sahi et al. (2017) highlighted that DL can improve detection accuracy in high-volume environments, though computational costs remain a challenge [2].

D. AUTONOMOUS DEFENSE SYSTEMS AND MULTI-LAYERED STRATEGIES

The concept of autonomous defense systems that integrate ML and automation to respond to attacks without manual intervention is gaining traction. These systems often employ self-healing architectures that automatically adjust resources, such as through load balancing and auto-scaling. Self-adapting architectures with incremental learning are critical to maintain effectiveness against evolving threats [4]. This approach is highly scalable and reduces operational costs. To further enhance defense, multi-layered strategies incorporating honeypots are being used for early detection [5]. Honeypots divert malicious traffic, allowing for analysis of attacker behavior to adapt defenses preemptively. Combining honeypots with WAFs and ML models provides a comprehensive and resilient defense against sophisticated DDoS attacks [5, 7].

III. THE PROPOSED AUTONOMOUS DDOS DEFENSE SYSTEM

The proposed system employs a hybrid approach combining traditional machine learning, deep learning, and incremental learning techniques to provide a resilient and adaptive defense against DDoS attacks [1]. The system's architecture is designed for real-time detection and continuous adaptation to new threat patterns. The overall architecture, depicted in Figure 1, consists of a load balancer, an ML/DL classifier, an anomaly detection module, a honeypot, and a primary/backup server configuration.

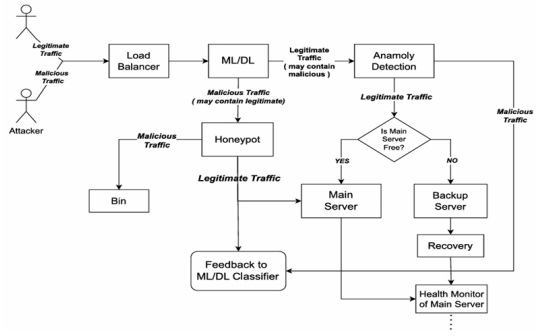

Fig. 1. The overall architecture of the proposed defense system

A. System Components

The architecture is built around several core components working in unison:

- **Main Server:** This server processes all incoming network traffic under normal conditions. It runs the detection algorithms that analyze traffic patterns in real-time to identify potential DDoS attacks.
- **Backup Server:** This server operates in a standby mode, continuously monitoring the health of the main server [8]. If the main server fails or is overwhelmed, the backup server automatically takes over, ensuring uninterrupted service [8]. It utilizes Wake-on-LAN (WoL) technology to remain in a low-power state until needed, optimizing energy consumption [8, 9].
- **ML/DL Classifier:** This is the core detection engine. It receives traffic from the load balancer and uses a combination of ML and DL models to classify it as either legitimate or potentially malicious [1, 2]
- Honeypot: Suspected malicious traffic is diverted to a honeypot, which is a decoy system designed to attract and analyze attacker behavior without risking core infrastructure [5]. Feedback from the honeypot is used to update the ML/DL classifier.

B. Detection and Mitigation Workflow

The operational workflow is as follows:

1) All incoming traffic first passes through a Load Balancer.
2) The traffic is then forwarded to the ML/DL Classifier for initial analysis [**?**].
3) The classifier separates traffic into *Legitimate* and *Malicious (may contain legitimate)* streams.
4) Suspected malicious traffic is redirected to the Honeypot for further analysis and to gather threat intelligence

[**?**]. Legitimate traffic from this stream is forwarded to the main server.

5) Traffic classified as legitimate is sent to the Anomaly Detection module for a final check before being passed to the main server if it is free.
6) If an attack is confirmed, the system initiates mitigation strategies, and the failover mechanism is triggered if the main server becomes unstable [**?**].

C. Server Recovery and Failover Mechanism

The system ensures high availability through an automated recovery and failover process, detailed in Algorithm 1. A health monitor continuously checks the status of the main server [8]. If the main server becomes overwhelmed, traffic is automatically rerouted to the backup server [3]. Once the main server recovers, traffic is gradually re-routed back, starting with a small percentage (5-10%) to ensure stability before a full transition. This process is illustrated in Figure 2.

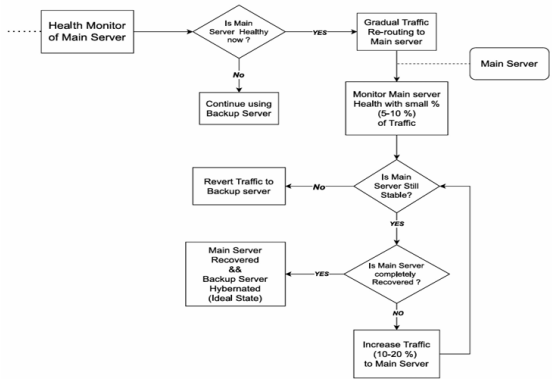

Fig. 2. Continued system architecture showing the server recovery and traffic re-routing logic

Algorithm 1 Network Traffic Monitoring and Mitigation

Require: Server is active
Ensure: Mitigated excessive traffic incidents

while server is active **do**
  Monitor incoming traffic bytes over 1-minute intervals
  if traffic_bytes $>$ threshold **then**
    Log "Excessive traffic detected"
    Identify offending_IP with the highest traffic volume using `tcpdump`
    Block offending_IP using `iptables`
    Log "Blocked IP: {offending_IP}"
    Disconnect network interface temporarily
    Wait for a brief pause
    Reconnect network interface to resume normal operations
  end if
end while

IV. METHODOLOGY AND IMPLEMENTATION

A. DATA COLLECTION AND PRE-PROCESSING

The system processes network traffic data characterized by key features such as packet count (pktcount), byte count (bytecount), network flows (flows), packet rate (pktrate), and transmission/reception rates (`tx_kbps`, `rx_kbps`). The feature set is expanded to include bytes per flow (byteperflow) and total throughput (`tot_kbps`). Data pre-processing involves three crucial steps:

1) **Categorical Encoding:** LabelEncoder is used to convert categorical features into numerical format, specifically for protocol information, source addresses (src), destination addresses (dst), and switch identifiers.
2) **Feature Normalization:** StandardScaler is applied to normalize the feature values, ensuring that all features contribute equally to model training. NaN values are converted to 0.0 to prevent errors during computation.
3) **Data Augmentation:** To address class imbalance in the training data, the Synthetic Minority Oversampling Technique (SMOTE) is used to generate synthetic samples for the minority class (malicious traffic).

B. MODEL ARCHITECTURE AND ENSEMBLE INTEGRATION

The detection system implements a three-tier model architecture to maximize detection performance:

1) **Deep Learning Model (DDoS Detector):** A four-layer neural network is used as the DDoSDetector. It consists of linear layers with decreasing neuron counts (64, 32, 16, 2), ReLU activation functions, batch normalization for training stability, and Dropout (0.3) for regularization to prevent overfitting.
2) **Random Forest Classifier:** An ensemble model comprising 100 decision trees (`n_estimators`=100) with a maximum depth of 10 is implemented. The leaf and split parameters are optimized to enhance performance and prevent overfitting.
3) **Incremental Random Forest:** This model is designed for continuous learning. It uses the `warm_start`=True capability, which allows the model to be updated with new data without retraining from scratch, making it highly efficient for adapting to new threats [4].
4) **Ensemble Integration:** The system achieves its combined results (e.g., Ensemble accuracy of 0.9712) by explicitly combining predictions from all three models using weighted averaging.

C. TRAINING METHODOLOGY AND SYSTEM OPTIMIZATION

The training process follows a multi-phase approach, including an **Initial Training Phase** and an **Incremental Learning Phase**, where models are dynamically updated.

The system incorporates several optimization and monitoring techniques:

1) **Performance Metrics:** The system continuously monitors accuracy metrics, including `rf_accuracy`, `dl_accuracy`, `incremental_accuracy`, and `ensemble_accuracy`. False Positive Rate (FPR) analysis is conducted using the confusion matrix derived from True Negatives and False Positives.
2) **Batch Processing:** Optimization utilizes Batch Processing during operation and testing, dividing data into smaller segments (e.g., processing chunks of size chunk_size = 1000) to enhance efficiency.
3) **Data Augmentation: SMOTE** is applied during the training phase (`smote.fit_resample`) specifically to mitigate the impact of class imbalance in the dataset.

D. PERFORMANCE METRICS

The system's performance is evaluated using standard metrics [1]:

- **Accuracy:** The proportion of total predictions that were correct. It is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

- **False Positive Rate (FPR):** The proportion of negative instances that were incorrectly classified as positive. It is a critical metric for DDoS detection, as a high FPR can lead to legitimate users being blocked. It is calculated as:

$$\text{FPR} = \frac{FP}{FP + TN} \tag{2}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

V. RESULTS AND DISCUSSION

The performance of the proposed system was evaluated based on the accuracy and false positive rates of the different models. The results demonstrate the effectiveness of the hybrid and incremental learning approach.

A. MODEL PERFORMANCE EVALUATION

The final performance metrics for each model are summarized in TABLE I. The Incremental Random Forest model significantly outperformed the others, achieving the highest accuracy of 98.16% and the lowest FPR of 2.28% (Internal Data). The standard Random Forest model also performed well, with an accuracy of 96.89% and an FPR of 4.56%. In contrast, the Deep Learning model exhibited poor performance, with an accuracy of only 72.61% and a very high FPR of 40.49%, making it unsuitable for this application in its current configuration. The high accuracy and low FPR of the Incremental RF validate its suitability for a dynamic defense system where both precision and adaptability are crucial.

TABLE I
FINAL PERFORMANCE METRICS OF CLASSIFICATION MODELS

Model	Accuracy	False Positive Rate (FPR)
Random Forest	0.9689	0.0456
Deep Learning	0.7261	0.4049
Incremental RF	**0.9816**	**0.0228**
Ensemble	0.9712	N/A

B. TEMPORAL PERFORMANCE ANALYSIS

The performance of the models was monitored over time across multiple updates to assess their stability. Figure 3 illustrates the False Positive Rates for the Random Forest, Deep Learning, and Incremental RF models over a series of updates. The Incremental RF consistently maintained the lowest and most stable FPR, while the Deep Learning model showed high volatility and a significantly higher rate of false positives. This demonstrates the robustness of the incremental learning approach in adapting to new data without degrading performance.

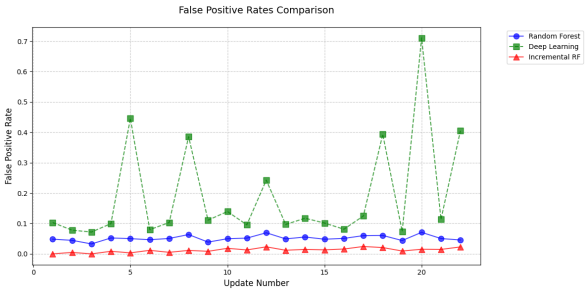

Fig. 3. False Positive Rates Comparison over multiple updates

C. SYSTEM EFFECTIVENESS AND FAILOVER VALIDATION

The overall effectiveness of the system refinements is highlighted in Figure 4, which compares the detection accuracy of the previous model configuration with the improved model. The enhanced model shows a marked improvement, underscoring the value of the incremental adjustments and multi-layered architecture.

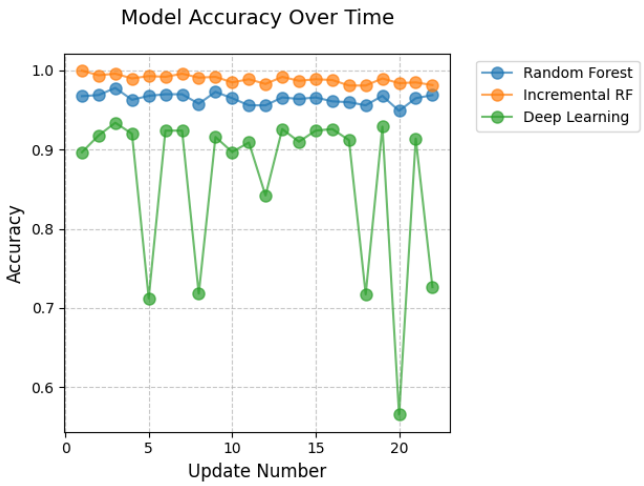

Fig. 4. Detection Accuracy: Previous vs. Improved Model

The system's failover mechanism was also validated. **Figure 5** shows the web service running on the main server during normal operation. **Figure 6** shows the default Apache page served by the backup server after it has taken over due to the main server being compromised by an attack. When the main server is overwhelmed, traffic is automatically rerouted to the backup server, ensuring uninterrupted service. This validation confirms the seamless transition and continuous service availability, demonstrating that the service successfully overtakes while the client's external access point (IP/URL) remains consistent due to the Load Balancer handling the switch.

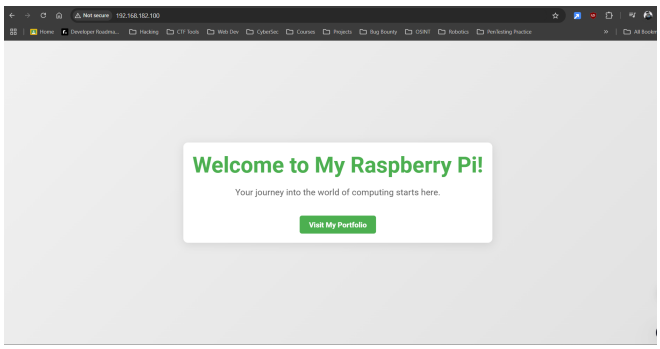

Fig. 5. Service when the main server is active

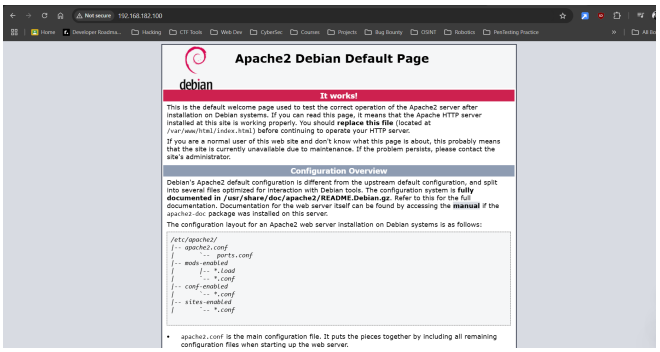

Fig. 6. Service when the main server is compromised

VI. CONCLUSION AND FUTURE WORK

This project successfully developed a robust, autonomous DDoS protection system for cloud infrastructure capable of real-time detection, mitigation, and recovery. By integrating a hybrid of Machine Learning and Deep Learning models with an incremental learning framework, the system demonstrates high adaptability to new and evolving threats. The results confirm the superiority of the Incremental Random Forest model, which achieved 98.16% accuracy with a minimal False Positive Rate of 2.28%. The multi-layered defense strategy, incorporating honeypots and an automated server failover mechanism [5, 8], ensures high resilience and continuous service availability, addressing critical gaps in traditional DDoS defense solutions.

For future work, several directions can be explored to enhance the system's capabilities:

- **Advanced Model Architectures:** Future work could explore more sophisticated architectures, such as Transformer models or novel ensemble methods, to further improve detection accuracy. Implementing Explainable AI (XAI) would also provide transparency into model decisions.
- **Broadened Threat Detection:** The system could be expanded to recognize other attack vectors, such as SQL injection and malware distribution, creating a more comprehensive security solution.
- **Cloud-Native Integration:** Exploring partnerships with cloud security solutions like AWS Shield or Azure DDoS Protection and deploying the system within a serverless architecture would enhance scalability and reduce operational costs.
- **Data Privacy and Ethics:** Implementing privacy-preserving techniques like federated learning or differential privacy would allow the system to learn from sensitive data without compromising privacy, while ensuring compliance with regulations like GDPR.
- **Real-World Deployment:** Pilot programs in live environments are critical to test performance against actual attack conditions and refine the models with real-world data.

REFERENCES

[1] Ajeetha G, Madhu Priya G, "Machine Learning Based DDoS Attack Detection," 2019 Innovations in Power and Advanced Computing Technology (i-PACT), 2019, IEEE.

[2] A. Sahi, D. Lai, Y. Li and M. Diykh, "An Efficient DDoS TCP Flood Attack Detection and Prevention System in a Cloud Environment," in IEEE Access, vol. 5, pp. 6036-6048, 2017, doi: 10.1109/ACCESS.2017.2688460.

[3] T. V. Phan and M. Park, "Efficient Distributed Denial-of-Service Attack Defense in SDN-Based Cloud," in IEEE Access, vol. 7, pp. 18701-18714, 2019, doi: 10.1109/ACCESS.2019.2896783.

[4] D. Yin, L. Zhang and K. Yang, "A DDoS Attack Detection and Mitigation With Software-Defined Internet of Things Framework," in IEEE Access, vol. 6, pp. 24694-24705, 2018, doi: 10.1109/ACCESS.2018.2831284.

[5] M. Zuñiga-Prieto, E. Insfran and S. Abrahão, "Architecture Description Language for Incremental Integration of Cloud Services Architectures," 2016 IEEE 10th International Symposium on the Maintenance and Evolution of Service-Oriented and Cloud-Based Environments (MESOCA), Raleigh, NC, USA, 2016, pp. 16-23, doi: 10.1109/MESOCA.2016.10.

[6] M. H. Rohit, S. M. Fahim and A. H. A. Khan, "Mitigating and Detecting DDoS attack on IoT Environment," 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON), Dhaka, Bangladesh, 2019, pp. 5-8, doi: 10.1109/RAAICON48939.2019.5.

[7] W. H. A. Muragaa, "A hybrid scheme for detecting and preventing single packet Low-rate DDoS and flooding DDoS attacks in SDN," 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA), Benghazi, Libya, 2023, pp. 707-712, doi: 10.1109/MI-STA57575.2023.10169712.

[8] J. Li et al., "Toward Adaptive DDoS-Filtering Rule Generation," 2023 IEEE Conference on Communications and Network Security (CNS), Orlando, FL, USA, 2023, pp. 1-9, doi: 10.1109/CNS59707.2023.10288699.

[9] Power and Energy-efficient VM scheduling in OpenStack Cloud Through Migration and Consolidation using Wake-on-LAN - Krishan Kumar, Kunal Patange, Pushkar Pete, Manjiri Wankhade, Arpitrama Chatterjee & Manish Kurhekar.

[10] M. Popa and T. Slavici, "Embedded server with Wake on LAN function," IEEE EUROCON 2009, St. Petersburg, Russia, 2009, pp. 365-370, doi: 10.1109/EURCON.2009.5167657.