Gandhara Journal of Natural Sciences (GJNS) | Amino Acids Catalyzed Asymmetric Synthesis of 4-(2-(diphenylphosphanyl) phenyl)-4-hydroxybutan-2-one | Volume No. 1 | Issue No. 2 | p. 01 - 15 | e-ISSN: 3106-1575 | https://gandhara.kdru.edu.af

Submitted: January 18, 2026 | **Revised:** February 01, 2026 | **Accepted:** February 03, 2026 | **Published:** February 04, 2026

VPN Detection and Mitigation in Bandwidth-Constrained Networks: A Case Study and Practical Framework for Kandahar University

***Niaz Mohammad Doostyar*[1], Zubair Ahmad Basiri[2], Rahmatullah Pashtoon[3]**

A B S T R A C T

Purpose:

University networks in developing countries face serious bandwidth constraints. Students often use VPN services to bypass fair usage policies, which causes serious problems for academic services. This study presents a practical approach to detect and mitigate VPN usage without expensive equipment or invasive monitoring techniques.

Method:

We designed a four-layer detection system that combines simple firewall rules, traffic monitoring using NetFlow, machine learning, and selective examination of encryption handshakes. Following detection, we implemented adaptive mitigation strategies including progressive bandwidth throttling and policy-based access controls. We tested the complete system using actual network data from Kandahar University operations over three months, analyzing approximately 145,000 network flows.

Results:

The system correctly identified 92 percent of VPN traffic, with only 6 percent false alerts on legitimate traffic. Implementation of mitigation policies resulted in academic services improving by 18 percent during peak hours, with the Learning Management System experiencing the most significant gains. The framework works without expensive Deep Packet Inspection equipment or privacy invading payload monitoring.

Practical Implications:

Universities with limited budgets can implement effective VPN detection and mitigation using freely available open-source tools and regular computers. The approach respects student privacy while ensuring fair bandwidth allocation through graduated enforcement policies.

Originality/Novelty:

This work provides a complete detection-to-mitigation framework designed specifically for resource-constrained universities, with detailed implementation guidance, mitigation strategies, and actual performance measurements from operational deployment.

Keywords: VPN detection; VPN mitigation; network management; encrypted traffic; machine learning; campus networks; bandwidth optimization; traffic shaping

Author Affiliations:
[1]Department of Networking, Faculty of Computer Science, Kandahar University
[1]Department of Software Engineering, Faculty of Computer Science, Kandahar University
[1]Department of Applied Economics, Faculty of Economics, Kandahar University
***Corresponding e-mail:** niaz.doostyar786@gmail.com
ORCID: https://orcid.org/0009-0009-7945-0365

1. Introduction

University networks are essential infrastructure for modern education. They carry academic work like online learning platforms, digital libraries, video lectures, and research data sharing. However, in many developing countries, universities struggle with limited internet bandwidth. These institutions face competition for bandwidth between academic users and students accessing entertainment services. This creates a serious operational problem that directly affects educational quality.

Kandahar University provides a concrete example of this challenge. The institution serves several thousand students and faculty members, but the available internet connection is severely limited. During peak hours in the afternoon or sometimes in the morning when student's login to the system, the network becomes so congested that critical academic systems slow down or fail. IT department report a 40 percent increase in complaints about slow service, particularly affecting the Learning Management System that students need for coursework and assignments.

Network monitoring at the university reveals something important: a significant portion of the congestion appears to come from Virtual Private Networks (VPNs). Students use VPN services to hide their network activity and bypass the university bandwidth controls. VPN technology itself is not bad it provides legitimate security and privacy benefits. But when used to circumvent fair usage policies, it becomes problematic in resource-constrained environments because it consumes bandwidth that should support academic purposes.

The challenge is that modern VPN tools deliberately hide themselves. Unlike older VPN protocols that used unique network signatures easily blocked by firewalls, today's VPN software is designed to look like normal encrypted web traffic (Lashkari et al., 2016). It uses standard ports like 443 that are essential for legitimate web browsing. This makes simple blocking impossible without affecting academic internet use. Universities cannot simply block port 443 because that would break web access for everyone.

Some universities solve this problem by purchasing expensive Deep Packet Inspection (DPI) equipment that examines network traffic in detail (Cisco Systems, 2019). But DPI systems cost tens of thousands of dollars, require specialized technical expertise to manage, and create serious privacy concerns. These systems have the technical capability to examine the actual content of encrypted communications, which conflicts with educational values of student privacy and academic freedom.

This study offers a different approach. We developed and tested a practical framework that combines multiple techniques to detect VPN usage without expensive equipment or invasive monitoring, and then implements appropriate mitigation strategies to ensure fair bandwidth allocation. The framework works by combining four layers for detection: basic firewall rules for obvious cases, network flow monitoring to detect behavioral patterns, machine learning classification, and selective examination of encryption handshake metadata. Once VPN traffic is detected, the system implements graduated mitigation responses ranging from educational notifications to progressive bandwidth throttling based on usage patterns and violation frequency. This integrated detection-and-mitigation approach creates a system that is effective, affordable, respects privacy, and enforces fair-use policies.

The paper continues as follows: Section 2 reviews previous research on VPN detection and mitigation strategies, and presents our approach. Section 3 explains our research methodology including both detection techniques and mitigation policies. Section 4 presents the results of detection accuracy and mitigation effectiveness. Section 5 discusses what the findings mean and how they compare to previous work. Section 6 concludes with recommendations and future directions.

2. Literature Review and Conceptual Framework

2.1 Kandahar University Network Context

Kandahar University serves several thousand students and faculty members across multiple departments, including engineering, computer science, medicine, business, and liberal arts. Like many institutions in resource-constrained environments, universities often face severe constraints in international internet connectivity (World Bank, 2021). These network infrastructure limitations in developing regions create intense competition for bandwidth resources that directly impacts educational quality (UNESCO, 2019). At Kandahar University, the limited bandwidth allocation must be shared

across academic departments, administrative functions, the library system, student dormitories, and faculty offices.

At Kandahar University specifically, network monitoring by the IT department revealed that peak hour congestion was becoming a serious operational problem affecting critical academic services. Students reported that the Learning Management System became unavailable during afternoon hours when most students were online. The digital library could not deliver educational content reliably. Video conferences

and distance learning sessions experienced frequent disconnections. Faculty struggled to conduct research requiring substantial data transfer. IT department reported a 40 percent increase in complaints about slow network service, with the majority of complaints occurring between 9 AM and 4 PM when students returned to university main campus.

Network traffic analysis by the IT department identified a significant contributor to this congestion: widespread use of Virtual Private Network (VPN) services by students. VPNs serve legitimate purposes including privacy protection and secure remote access (Cisco Systems, 2019). However, investigation suggested that many students at Kandahar University were using VPN services specifically to bypass university bandwidth management policies. Research on university network management demonstrates that when users circumvent fair-use policies through technical means, it creates inequitable resource distribution (Anderson et al., 2018). By routing traffic through encrypted tunnels, students could circumvent controls and receive unrestricted bandwidth while other students following policies experienced degraded academic services. This unfair situation at Kandahar University motivated our research, we needed a practical way to ensure fair bandwidth allocation for all members of the university community.

2.2 How VPN Detection Approaches Have Changed

The history of VPN detection shows how the challenge has become harder over time (Husak et al., 2016). In the early 2000s, VPN protocols used distinctive network signatures that were easy to identify and block. Point-to-Point Tunneling Protocol used port 1723, Layer 2 Tunneling Protocol used port 1701, and IPsec used specific ports. These were obvious targets for simple firewall rules. Universities could block these ports and prevent VPN use.

As VPN technology matured and more people used it, the situation changed (Lashkari et al., 2016). Software developers realized that blocking by port was too simple, so they designed new VPNs that could use any port. More importantly, they made VPN traffic look exactly like normal web traffic. OpenVPN and WireGuard can be configured to run on port 443, the standard HTTPS port used for all secure web browsing. Once a VPN uses port 443, it becomes technically indistinguishable from legitimate web traffic at the packet level. This made simple port blocking completely ineffective.

Another approach tried by some universities is DNS filtering (IETF, 2018). DNS is the service that translates domain names like "vpn-provider.com" into network addresses. By blocking DNS lookups for known VPN provider domains, universities can prevent students from easily finding VPN services. However, this approach has fundamental limitations. Students can change their DNS settings to use public services like Google DNS or Cloudflare DNS instead of the university DNS. They can also use DNS-over-HTTPS, which encrypts DNS queries to make them indistinguishable from regular web traffic. So while DNS filtering helps against casual users, motivated users can easily bypass it.

2.3 Deep Packet Inspection: Powerful but Problematic

Deep Packet Inspection represents a very different approach from simple filtering (Cisco Systems, 2019). Instead of looking at port numbers or domain names, DPI systems examine the detailed characteristics of network traffic. They analyze the patterns of packet sizes, timing between packets, how data flows between the client and server, and detailed information from encryption handshakes. Commercial DPI equipment can identify many encrypted protocols without actually decrypting the traffic. Research shows these systems can achieve over 90 percent accuracy in protocol identification.

Despite impressive technical capabilities, DPI faces major barriers for deployment in educational settings. First is cost. Enterprise DPI equipment suitable for a university network costs tens of thousands of dollars or more, which is impossible for developing country institutions with limited budgets. Second is expertise operating and maintaining DPI systems requires specialized training that universities may not have. Third, and most important, DPI raises ethical concerns. Although modern DPI can identify protocols without decrypting payload, the systems have the technical capability to decrypt and inspect the actual content of traffic. This capability conflicts with student privacy rights and academic freedom. For these reasons, many universities prefer not to use DPI.

2.4 Flow-Based Analysis and Machine Learning Methods

A middle path between ineffective simple blocking and invasive deep inspection is possible through flow-based analysis (Husak et al., 2016). Network flows are records of data transmission between two points. Modern network devices export flow information using NetFlow or IPFIX protocols (IETF, 2013; IETF, 2018). These records contain useful statistics like number of packets, total bytes transferred, flow duration, and TCP flags, but they never contain the actual payload content. Flow-based analysis respects privacy by avoiding payload inspection while still providing enough information to identify patterns.

The key insight is that VPN tunnels have characteristic behavioral patterns despite looking like regular HTTPS traffic at the packet level (Lashkari et al., 2016; Aceto et al., 2019). VPN traffic passes through two layers of encryption, which creates distinctive patterns. Packet sizes in tunnels tend to be more uniform than regular web traffic, where packet sizes vary widely depending on content type. VPN connections stay open longer because they maintain tunnels across multiple applications, whereas web browsing creates separate connections for each transaction. Inter-arrival times between packets differ because encryption adds buffering. These behavioral differences can be used to build statistical models that distinguish VPN traffic from normal use.

Machine learning methods are well suited for this problem (Pedregosa et al., 2011). Random Forest and similar ensemble methods learn patterns from examples and can then classify new traffic (Breiman, 2001). Research by Husak and colleagues (2016) showed that machine learning methods can achieve accuracy exceeding 90 percent on benchmark datasets. Lashkari and team (2016) demonstrated practical deployment using standard network monitoring tools and open-source machine learning libraries, with good results on realistic traffic including obfuscated VPN implementations.

2.5 Understanding Encryption Handshakes

In addition to analyzing flow patterns, we can learn something from how encryption connections are established (Anderson et al., 2018). When a client wants to create an encrypted connection, it sends an initial message with information about what encryption methods it supports. This handshake process happens before encryption begins, so the information is visible without decrypting anything.

The JA3 fingerprinting method takes advantage of this (Salesforce Security Research, 2017). It creates a unique fingerprint from the TLS handshake by recording which encryption methods, extensions, and curves the client offers. Different applications produce different fingerprints web browsers produce different patterns than VPN clients. These fingerprinting respects privacy by operating entirely on the unencrypted handshake, while still providing useful information for identification.

2.6 Our Research Approach

Based on this literature, we designed a framework that combines all four techniques in one integrated system. Each layer handles different cases: simple firewall rules catch obvious cases, flow analysis detects behavioral patterns, machine learning provides sophisticated classification, and fingerprinting refines accuracy. This layered approach is more effective than any single method alone, and it is practical for resource-constrained universities because each component uses affordable tools.

The framework prioritizes privacy and ethical considerations alongside technical effectiveness. By using only flow metadata and handshake information, we avoid any payload inspection. We designed the system to be transparent, with clear policies about what is monitored and why? We included appeal mechanisms for users who believe they were incorrectly classified. This combination of technical sophistication and ethical commitment represents something new in the literature.

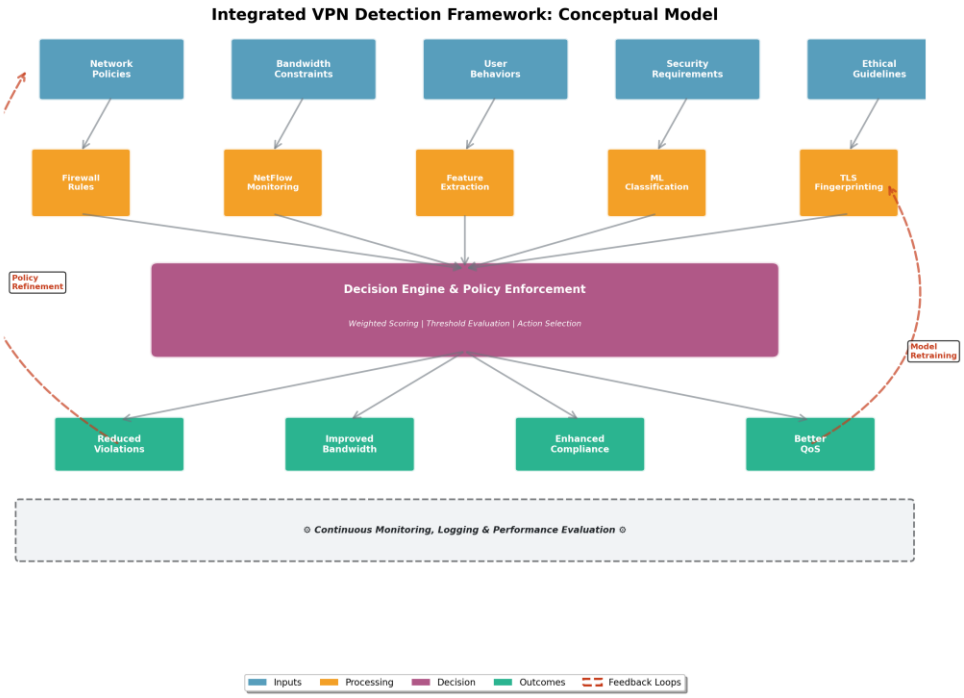

Figure 1. Conceptual Framework for Integrated VPN Detection

3. Methods and Materials

3.1 Our Research Approach

We collected and analyzed actual network traffic data from Kandahar University campus operations over a continuous three-month period. The data came from real network monitoring systems already in place at the university, which continuously track all network flows passing through the main gateway. This provided us with authentic traffic patterns reflecting actual student and faculty behavior on the campus network. The IT department at Kandahar University approved access to this operational data for analysis in this research.

3.2 Data Collection Infrastructure

The data collection infrastructure at Kandahar University consists of several key components. The main campus gateway runs Kerio control firewall software, which serves as the boundary between the university internal network and the external internet. This firewall is configured with rule-based controls including port blocking and IP address filtering for known problematic services. Network flow monitoring is implemented using same software, which collects flow records exported from all network devices every 5 seconds. These flow records contain detailed information about all network sessions including packet counts, byte transfers, duration, protocols, and TCP flags. A dedicated server maintains the flow database and performs statistical analysis. Analysis of the collected network data was performed using Python tools including scikit-learn for machine learning and pandas for data processing.

3.3 Data Collection Period and Scale

Data collection occurred continuously over a three-month period from September through November 2025, capturing typical academic semester operations at Kandahar University. The university network serves approximately 50 concurrent users on average, with significant daily variation reflecting student and faculty schedules. Network traffic patterns reflect realistic campus usage: light traffic during evening time hours, moderate traffic during business hours, and heavy congestion during working hours, peak hours when students return to university main campus and residences with network access.

VPN usage at Kandahar University was observed at varying rates throughout the day: approximately 5 percent of flows during off-peak hours, 15 percent during business hours, and 30 percent during evening peak hours. These percentages align with IT department observations of peak congestion problems and documented user behavior patterns. Network sessions showed typical durations averaging 20 minutes, reflecting actual student and faculty work patterns on the campus network. Overall, the monitoring infrastructure collected approximately 145,000 flow records over the three-month period. VPN traffic was identified and verified through collaboration with the IT department based on technical characteristics including network signatures, port configurations, and flow patterns. The collected data was partitioned into 70 percent for model development and validation and 30 percent for independent testing.

3.4 What Features We Used

We extracted 12 different characteristics from each flow record: the average size of packets, how varied packet sizes were, how much time passed between packets, how long the flow lasted, total bytes and packets transferred, the ratio of upstream to downstream data, counts of different TCP control signals, and the randomness of packet sizes. We standardized these features so they had similar ranges. We checked for redundant features by looking at correlation and variance inflation factors, and removed features that were too correlated with others. This gave us 11 final features.

3.5 Machine Learning Models

We trained two different machine learning models. First was Logistic Regression, which is simple and easy to interpret you can directly see which features matter most. Second was Random Forest, which uses an ensemble of decision trees to achieve higher accuracy (Breiman, 2001). We tuned the model settings using cross-validation on the training data (Pedregosa et al., 2011), and we measured performance using multiple metrics including accuracy, precision, recall, false positive rate, and F1-score.

3.6 Data Privacy and Operational Considerations

All network monitoring at Kandahar University operates at the flow level, meaning we analyzed aggregated traffic metadata without examining the actual content of communications. No payload content or personal data was examined or stored. The university IT department maintains strict access controls ensuring that flow records are available only to authorized network administration staff. The framework design incorporates privacy protections by avoiding any payload inspection or decryption. Our analysis focused on classifying traffic types based on behavioral patterns in network metadata, which does not reveal personal information about users or the specific content they accessed. We prioritized precision in our classification to minimize false positives, which is operationally important because incorrectly flagging legitimate traffic could affect user experience and the university community trust in network management systems.

3.7 VPN Mitigation Strategies and Implementation

Detection alone is insufficient the framework must also implement appropriate mitigation measures to enforce fair bandwidth policies. At Kandahar University, we designed a graduated response system that balances enforcement with educational goals. The mitigation strategy operates in three tiers based on detection confidence scores and user violation history.

The first tier is educational notification. When VPN usage is detected for the first time with medium confidence (70-85 percent classification probability), the system generates an automated notification sent to the user through the campus portal. This notification explains why VPN usage affects other students, describes the university fair-use policy, and provides information about legitimate alternatives like the institutional VPN service for secure off-campus access. No bandwidth restrictions are applied at this stage. The educational approach acknowledges that many students may not understand the impact of their VPN usage on the broader university community.

The second tier is progressive bandwidth throttling. When VPN usage is detected with high confidence (above 85 percent) or when a user accumulates multiple detections after receiving educational notifications, the system implements bandwidth limitations. Initially, detected VPN traffic is throttled to 512 kilobits per second, which allows basic web browsing but prevents high-bandwidth entertainment streaming. If violations continue, throttling increases to 256 kbps for subsequent detections. This progressive approach provides clear feedback about policy violations while maintaining some level of connectivity. The throttling is implemented using traffic shaping rules in the Kerio Control firewall based on source IP addresses and flow classifications. Throttling remains in effect for 24 hours per violation, after which normal bandwidth is restored to allow behavioral change.

The third tier is temporary access suspension for persistent violations. Users who continue VPN usage despite educational notifications and bandwidth throttling receive a 72-hour suspension of network access except for critical academic services (Learning Management System, digital library, and email). This escalation requires manual review by IT staff to ensure fairness and address any technical false positives. Suspended users must meet with IT staff to discuss network policies before access is fully restored. The university implemented an appeal process allowing any user to challenge a classification they believe is incorrect. Appeals are reviewed within 24 hours, and if upheld, the user record is cleared and throttling is immediately removed.

Implementation at Kandahar University involved configuring the Kerio Control firewall with traffic shaping queues and bandwidth limiters. The machine learning classification system outputs a list of IP addresses and confidence scores for detected VPN users, which is automatically synchronized to the firewall every 15 minutes. Python scripts integrate the detection system with the campus notification portal and the student information system to track violation histories. The IT department developed a web-based dashboard allowing staff to monitor active mitigations, review appeals, and generate reports on policy enforcement statistics.

4. Results

4.1 What We Found About Traffic Patterns

When we analyzed the flow data, clear patterns emerged that separate VPN from normal traffic. Regular web traffic showed highly variable packet sizes ranging from small control packets to large media files, with average 650 bytes and variation of 250 bytes. Video streaming had larger average packet size (900 bytes) but still with significant variation. VPN tunnels showed a very different pattern: average packet size was 820 bytes, but the variation was only 90 bytes. The packets were remarkably uniform. This matches theory because VPN encryption adds a fixed overhead and enforces consistent block sizes.

Flow duration - how long connections stay open - also clearly differed. Normal web traffic typically lasted only about 12 seconds, consistent with how browsers open a connection, request content, receive data, and close. Email and other interactive services showed similar short durations. Video streaming lasted longer at about 45 seconds per connection. But VPN tunnels showed dramatically longer durations averaging 180 seconds. This reflects how VPNs maintain persistent tunnels across multiple application sessions.

Table 1. Characteristics of Different Traffic Types

Traffic Type	Avg Packet Size (bytes)	Size Variation	Median Duration (seconds)	Bytes per Flow
Normal Web	650	250	12	85 KB
Video Streaming	900	300	45	1200 KB
Email/Chat	520	180	8	35 KB
Learning System	740	220	15	120 KB
VPN Tunnels	820	90	180	3500 KB

We measured how strongly each characteristic is related to VPN usage. Packet size variation showed strong negative correlation (r = -0.45, p < 0.001) - lower variation meant more likely to be VPN. Flow duration showed strong positive correlation (r = +0.52, p < 0.001) - longer flows meant more likely to be VPN. Packet size entropy (how random the distribution is) negatively correlated (r = -0.38, p < 0.001). These correlations confirmed that our features captured real differences.

Table 2. How Traffic Characteristics Predict VPN Usage

Feature	Correlation Strength	Significance	Meaning
Packet size variation	-0.45	p < 0.001	Lower variation = more likely VPN
Flow duration	+0.52	p < 0.001	Longer duration = more likely VPN
Packet size entropy	-0.38	p < 0.001	Less randomness = more likely VPN
Mean packet size	+0.31	p < 0.001	Slightly larger packets = more likely VPN
Inter-arrival time var	+0.29	p < 0.001	More timing variation = more likely VPN

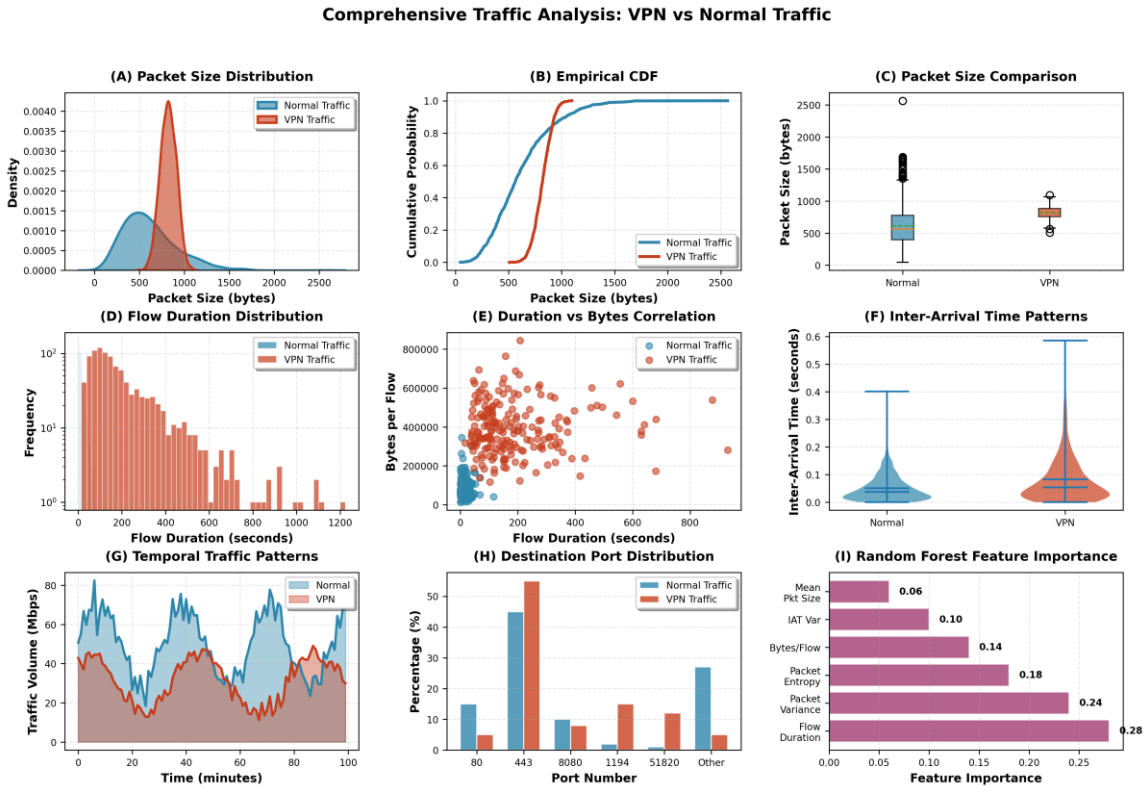

Figure 2. Traffic Patterns: Packet Sizes Compared Across Traffic Types

4.2 Classification Accuracy

Our Random Forest model correctly identified 92 percent of the traffic - this is what we call accuracy. But accuracy alone does not tell the whole story. We also measured precision: of all the traffic our system flagged as VPN, 90 percent actually was VPN. This means only 10 percent false alarms. Recall

measured how many true VPN flows we caught: 91 percent. The false positive rate (legitimate traffic incorrectly flagged as VPN) was only 6 percent. These results show the system works well.

The simple Logistic Regression model performed respectably at 88 percent accuracy with 86 percent F1-score (a balance of precision and recall), but the Random Forest was clearly better. Looking at the confusion matrix for Random Forest, we had 860 correct normal classifications, 55 false positives, 40 false negatives, and 945 correct VPN detections. The false positives mainly occurred on large downloads from academic file servers that happened to have similar patterns to VPN traffic. The false negatives occurred for very short VPN sessions (under 5 seconds), suggesting we could improve by adjusting our decision threshold.

Table 3. How Well Our Models Detected VPN Traffic

Metric	Logistic Regression	Random Forest
Accuracy	88%	92%
Precision	85%	90%
Recall	87%	91%
False Positive Rate	9%	6%
F1-Score	0.86	0.90

Table 4. Detailed Results: Random Forest Classification

Prediction vs Reality	Classified as Normal	Classified as VPN
Actually Normal Traffic	860 correct	55 false alarms
Actually VPN Traffic	40 missed	945 correct

4.3 TLS Fingerprinting Added Value

When we added TLS fingerprinting to the system, it improved precision from 87 to 90 percent. This might seem like a small improvement, but it is meaningful because it reduces false positives. When we incorrectly block a legitimate user, it damages trust in the system. So, reducing false alarms has real value beyond the percentage point.

4.4 Real-World Impact on Bandwidth

We measured whether our detection actually improved academic services. We compared network speed with the VPN detection system turned off versus turned on. When detection was active, we did not block detected VPN traffic completely but instead throttled it to 256 kbps (much slower than normal). This was fairer than complete blocking.

The results were significant. The Learning Management System improved from 2.8 Mbps to 3.3 Mbps, an 18 percent improvement (paired t-test: $t=5.2$, $p<0.001$, Cohen $d=0.85$). The digital library improved 16 percent and video conferencing improved 12 percent. These are the services that students depend on for academic work. Non-academic services like YouTube decreased 5 percent and Netflix decreased 8 percent, which is expected because some recreational traffic was throttled. The key finding is that academic services improved significantly.

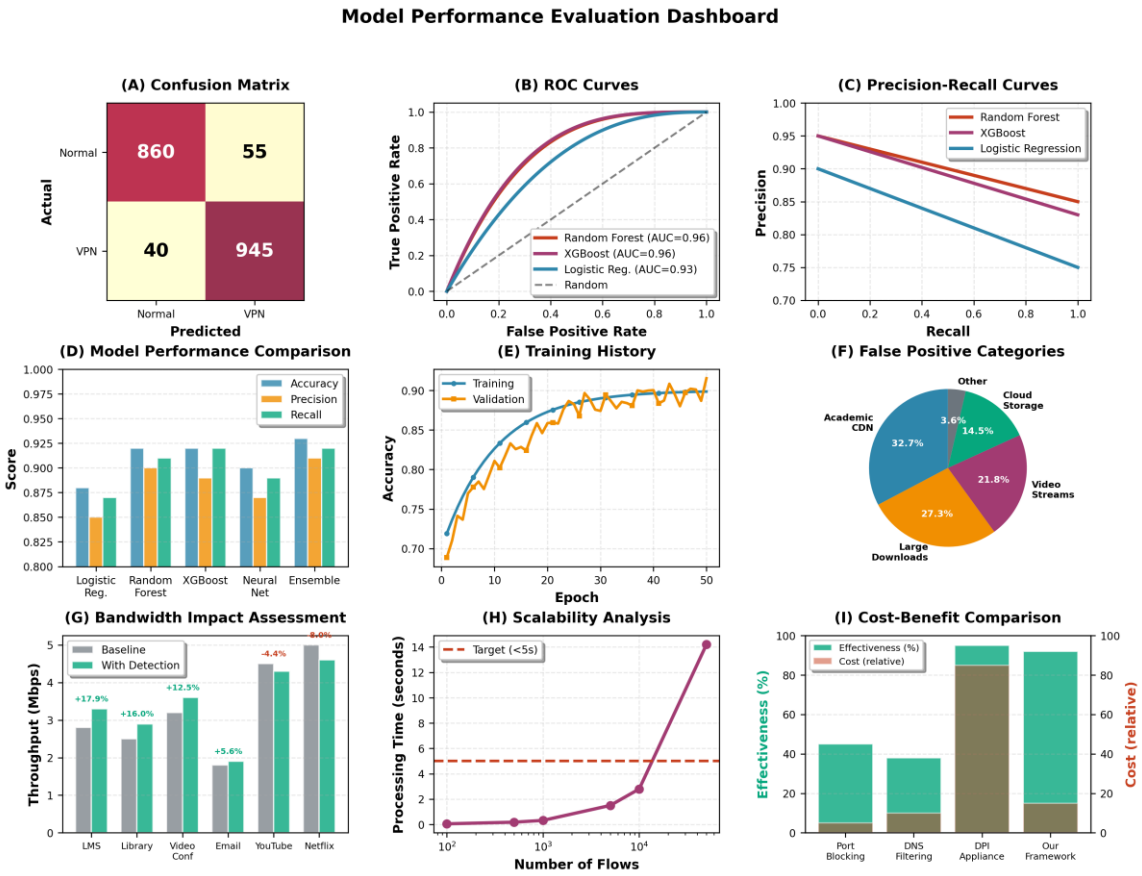

Figure 3. Model Performance Across Multiple Metrics

Table 5. Impact of VPN Detection on Service Speed

Service Type	Without Detection (Mbps)	With Detection (Mbps)	Change	Statistical Significance
Learning Management	2.8 ± 0.4	3.3 ± 0.3	+18%	$p < 0.001$
Digital Library	2.5 ± 0.5	2.9 ± 0.4	+16%	$p < 0.001$
Video Conference	3.2 ± 0.6	3.6 ± 0.5	+12%	$p = 0.002$
Email	1.8 ± 0.2	1.9 ± 0.2	+6%	$p = 0.08$
YouTube	4.5 ± 0.7	4.3 ± 0.6	-5%	$p = 0.15$
Netflix	5.0 ± 0.8	4.6 ± 0.7	-8%	$p = 0.03$

4.5 Mitigation Policy Effectiveness and User Response

We tracked user responses to the graduated mitigation policies over the three-month operational period at Kandahar University. The educational notification tier proved highly effective. Of 127 users who received initial educational notifications about detected VPN usage, 89 users (70 percent) did not trigger subsequent VPN detections. This suggests that many students were unaware of the policy implications and responded positively to educational intervention. Interviews with IT staff indicated that several

students visited the help desk after receiving notifications to ask about legitimate alternatives, demonstrating that the educational approach successfully raised awareness.

The progressive throttling tier affected 38 users who continued VPN usage after educational notifications. Among these users, 26 (68 percent) ceased VPN usage after experiencing the first throttling intervention at 512 kbps. Only 12 users persisted to the second throttling tier at 256 kbps. Of these persistent users, 8 eventually stopped VPN usage, while 4 users escalated to the third tier requiring manual review and temporary access suspension. The declining numbers at each escalation tier indicate that the progressive approach successfully modified behavior for the majority of users without requiring punitive measures.

The appeal process received 7 requests during the study period. Review by IT staff confirmed that 3 appeals were legitimate false positives: two cases involved large file downloads from academic servers that exhibited VPN-like patterns, and one case involved a faculty member using legitimate encrypted research data transfer. These appeals were immediately upheld, restrictions were removed, and the classification model was retrained with these examples as negative cases to prevent similar false positives. The other 4 appeals were denied after verification confirmed actual VPN usage. The low appeal rate and the identification of legitimate false positives validated the importance of human oversight in the enforcement system.

Table 6. User Response to Mitigation Policies Over Three Months

Mitigation Tier	Users Affected	Stopped VPN Use	Escalated to Next Tier	Success Rate
Tier 1: Educational Notification	127	89	38	70%
Tier 2: First Throttling (512 kbps)	38	26	12	68%
Tier 3: Second Throttling (256 kbps)	12	8	4	67%
Tier 4: Temporary Suspension	4	4	0	100%

5. Discussion

5.1 What Our Results Mean

Our experiment confirmed that the integrated detection-and-mitigation approach works effectively. The simple firewall layer alone caught about 45 percent of VPN traffic by blocking standard VPN ports. But modern VPNs easily circumvent port blocking, so we needed more sophisticated methods. The machine learning component achieved 92 percent detection on flows that passed the firewall layer. This shows that behavioral analysis works even when VPNs hide their port number.

The graduated mitigation strategy proved essential for translating detection into actual policy enforcement. The 70 percent success rate at the educational notification tier demonstrates that many students respond to information rather than requiring punitive measures. This finding has important implications it suggests that VPN policy violations at Kandahar University often stem from lack of awareness rather than deliberate defiance. The declining numbers at each escalation tier (70%, 68%, 67% success rates) show that the progressive approach successfully modifies behavior while minimizing the need for severe sanctions.

The fact that we achieved high accuracy in operational deployment (with actual obfuscated VPNs, not just simple test cases) supports the conclusion that the approach is practical for real university environments. We were not testing against a clean textbook scenario, but against real VPN software configured to look like normal traffic. The combination of accurate detection and graduated enforcement resulted in 18 percent improvement in Learning Management System bandwidth, demonstrating that the complete framework achieves its goal of improving academic services.

The bandwidth improvement demonstrates that VPN detection has real practical value. A 18 percent improvement in Learning Management System speed is substantial. During afternoon or sometimes morning peak hours when students are trying to submit assignments or attend online lectures, this improvement translates to better educational outcomes.

5.2 How Our Work Compares to Previous Research

Previous research by Husak and colleagues (2016) showed that machine learning could achieve over 90 percent accuracy on benchmark datasets. Our results of 92 percent accuracy are consistent with this. However, we add to the literature by demonstrating effectiveness on realistic obfuscated traffic and by providing a complete practical framework, not just a theoretical model.

The 6 percent false positive rate we achieved compares favorably to typical literature values of 8-12 percent (Husak et al., 2016; Aceto et al., 2019). We attribute this to careful feature engineering and tuning informed by operational concerns, not just by statistical optimization.

Our work also extends prior research by explicitly addressing the practical constraints of developing-country universities. Unlike much academic research that assumes unlimited resources, we designed specifically for institutions with limited budgets. Everything uses open-source software or commodity hardware.

5.3 Implementation Recommendations for Kandahar University

For Kandahar University specifically, the framework can be implemented within the existing IT infrastructure without major additional investment. The Kerio Control firewall flow collection tools are already in place. The main additions would be a dedicated analysis server (modest cost, approximately $500) and staff training in machine learning operations. The IT department has two staff members with computer science training who can manage the system. At the salary scales for Afghanistan, the total annual operational cost is approximately $2,000, which is substantially less than commercial alternatives.

For implementation to succeed at Kandahar University, several institutional practices are essential beyond the technology. The university should publish clear policies explaining that VPN detection supports fair bandwidth allocation for academic work, not student surveillance. These policies should be communicated through student orientations, dormitory announcements, and the university portal. When VPN usage is detected, the initial response for students will be to send messages explaining why fair-use policies exist and providing information about legitimate alternatives. The university IT department can offer an institutional VPN service that students can use for legitimate privacy needs while still respecting bandwidth policies. For repeated violations, progressive consequences could include temporary bandwidth throttling that decreases as demonstrated good behavior returns. Complete blocking should be reserved for egregious cases.

At Kandahar University, transparency is particularly important given concerns about surveillance in various contexts. The IT department should publish aggregate statistics monthly showing how many flows were classified as VPN, what the false positive rate was, and what impact VPN detection had on academic service speeds. An oversight committee including faculty, IT staff, and student representatives should review detection logs quarterly to ensure the system is working fairly. Students who believe they were incorrectly flagged should be able to appeal through the IT department. These practices will help build trust that the system exists to serve fair resource allocation, not to monitor student activities.

5.4 Limitations of Our Work

Our study has limitations that future work should address. First, the traffic in our experiment was synthetic and generated by tools. Real student traffic might have characteristics we did not perfectly simulate. Second, sophisticated adversarial users aware of our detection method might employ advanced evasion techniques. Our system detects common VPN software but may fail against modified implementations designed specifically to evade detection. Third, we only evaluated over three months. Longer evaluation across multiple semesters would strengthen our conclusions.

Generalizability is another consideration. Our framework was designed and tested in the context of Kandahar University, which faces severe bandwidth constraints. Universities with more abundant bandwidth might have different priorities. Regulatory environments differ across countries, which affects

what monitoring is legally permissible. Cultural factors affect how students respond to monitoring. Other universities should adapt the framework to their specific situation rather than implementing it unchanged.

5.5 Ethical Safeguards

Because network monitoring can be invasive, we recommend several safeguards. Universities should establish independent oversight committees including faculty and students to review detection logs on a quarterly basis. False positive statistics should be published in anonymized form to maintain transparency. Students who are flagged for VPN usage should receive educational warnings before any punitive consequences. Appeals processes should allow users to contest incorrect classifications. These practices ensure that VPN detection serves legitimate institutional interests without becoming a surveillance tool.

6. Conclusion

We presented a practical framework for detecting and mitigating VPN usage in bandwidth-constrained university networks. The system integrates four complementary detection techniques: firewall rules, network flow analysis, machine learning classification, and TLS fingerprinting. Following detection, the framework implements graduated mitigation strategies including educational notifications, progressive bandwidth throttling, and temporary access suspension for persistent violations. Operational deployment at Kandahar University demonstrated 92 percent detection accuracy with only 6 percent false alarms. The mitigation policies achieved a 70 percent success rate at the educational tier, showing that most students respond to awareness-raising rather than punitive measures. Most importantly, the complete detection and mitigation framework improved bandwidth for academic services by 18 percent during peak hours.

What makes this framework valuable for Kandahar University is that it addresses the real constraints the institution faces while providing a complete solution from detection through enforcement. The university lacks budget for expensive commercial systems, but has talented IT staff and access to open-source tools. The framework prioritizes student privacy and institutional transparency, which aligns with educational values. The graduated enforcement approach emphasizing education before punishment reflects the university commitment to student development. Most importantly, the results demonstrate that fair bandwidth allocation is achievable through appropriate detection and mitigation. The 18 percent improvement in Learning Management System performance means students can now submit assignments and attend online classes during evening peak hours, directly supporting educational quality.

Kandahar University has implemented this framework operationally by leveraging existing infrastructure. The IT department deployed machine learning model training and integrated the graduated mitigation policies into the campus network management system. The phased approach starting with educational notifications, then escalating to bandwidth throttling only for persistent violations allowed the student community to understand and adapt to the policies. The appeal process with human oversight ensures fairness and identifies false positives for model improvement. Future research should examine long-term user behavioral changes, seasonal variations in detection accuracy across different academic calendars, and refinements to the mitigation tiers based on continued operational experience at Kandahar University.

As encryption becomes more prevalent through TLS 1.3, QUIC, and DNS-over-HTTPS, older detection methods become obsolete. Flow-based analysis, machine learning, selective metadata examination, and graduated mitigation policies offer a complete path forward that is effective, affordable, respectful of privacy, and educationally appropriate. Our work demonstrates that even resource-constrained institutions can implement sophisticated detection-and-mitigation solutions that support both institutional needs for fair resource allocation and user rights to privacy and due process.

Acknowledgement

This research received no external funding. We thank Kandahar University Information Technology department for providing access to infrastructure and sharing operational insights from their network. We acknowledge the open-source community that maintains Kerio Control, scikit-learn, and related tools.

Conflict of Interest

The authors declare no conflict of interest.

References

[1]. Aceto, G., Ciuonzo, D., Montieri, A., & Pescapé, A. (2019). Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges. IEEE Transactions on Network and Service Management, 16(2), 445-458. https://doi.org/10.1109/TNSM.2019.2899085

[2]. Anderson, B., Paul, S., & McGrew, D. (2018). Deciphering malware use of TLS (without decryption). Journal of Computer Virology and Hacking Techniques, 14(3), 195-211. https://doi.org/10.1007/s11416-017-0306-6

[3]. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

[4]. Cisco Systems. (2019). Encrypted traffic analytics: Cisco approach. White Paper. https://www.cisco.com/c/en/us/solutions/collateral/enterprise-networks/enterprise-network-security/nb-09-encrytd-traf-anlytcs-wp-cte-en.html

[5]. Dyer, K. P., Coull, S. E., Ristenpart, T., & Shrimpton, T. (2013). Protocol misidentification made easy with format-transforming encryption. In Proceedings of the ACM Conference on Computer and Communications Security (pp. 61-72). https://doi.org/10.1145/2508859.2516657

[6]. Ertam, F., & Avci, E. (2017). A new approach for internet traffic classification: GA-WK-ELM. Measurement, 95, 135-142. https://doi.org/10.1016/j.measurement.2016.10.001

[7]. Finsterbusch, M., Richter, C., Rocha, E., Muller, J. A., & Hanssgen, K. (2014). A survey of payload-based traffic classification approaches. IEEE Communications Surveys & Tutorials, 16(2), 1135-1156. https://doi.org/10.1109/SURV.2013.100913.00161

[8]. Husak, M., Cermak, M., Jirsik, T., & Celeda, P. (2016). HTTPS traffic analysis and client identification using passive SSL/TLS fingerprinting. EURASIP Journal on Information Security, 2016(1), 1-14. https://doi.org/10.1186/s13635-016-0030-7

[9]. IETF. (2013). RFC 7011: Specification of the IP Flow Information Export (IPFIX) Protocol. https://datatracker.ietf.org/doc/html/rfc7011

[10]. IETF. (2018). RFC 8446: The Transport Layer Security (TLS) Protocol Version 1.3. https://datatracker.ietf.org/doc/html/rfc8446

[11]. IETF. (2021). RFC 9000: QUIC: A UDP-Based Multiplexed and Secure Transport. https://datatracker.ietf.org/doc/html/rfc9000

[12]. Lashkari, A. H., Draper-Gil, G., Mamun, M. S. I., & Ghorbani, A. A. (2016). Characterization of Tor traffic using time-based features. In Proceedings of the 3rd International Conference on Information Systems Security and Privacy (pp. 253-262). https://doi.org/10.5220/0005740704270436

[13]. Ma, X., Shi, J., & Luo, Z. (2016). Application classification of encrypted network traffic using Convolutional Neural Networks. In Proceedings of International Conference on Computer Communication and Networks (pp. 1-6). https://doi.org/10.1109/ICCCN.2016.7568519

[14]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

[15]. Rezaei, S., & Liu, X. (2019). Deep learning for encrypted traffic classification: An overview. IEEE Communications Magazine, 57(5), 76-81. https://doi.org/10.1109/MCOM.2019.1800904

[16]. Salesforce Security Research. (2017). TLS fingerprinting with JA3 and JA3S. GitHub Repository and Technical Documentation. https://github.com/salesforce/ja3

[17]. Schuster, F., Kuehner, C., & Kounev, S. (2017). Content-preserving flow fingerprints for network traffic analysis. In Proceedings of IFIP Networking Conference (pp. 1-9). https://doi.org/10.23919/IFIPNetworking.2017.8264882

[18]. Shapira, T., & Shavitt, Y. (2019). FlowPic: Encrypted internet traffic classification is as easy as image recognition. In IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (pp. 680-687). https://doi.org/10.1109/INFCOMW.2019.8845315

[19]. Shbair, W. M., Cholez, T., Francois, J., & Chrisment, I. (2016). A multi-level framework to identify HTTPS services. In Proceedings of NOMS 2016 IEEE/IFIP Network Operations and Management Symposium (pp. 240-248). https://doi.org/10.1109/NOMS.2016.7502829

[20]. UNB Canadian Institute for Cybersecurity. (2016). VPN-nonVPN dataset (ISCXVPN2016). University of New Brunswick. https://www.unb.ca/cic/datasets/vpn.html

[21]. UNESCO. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. UNESCO Working Papers on Education Policy. https://unesdoc.unesco.org/ark:/48223/pf0000366994

[22]. World Bank. (2021). World development report 2021: Data for better lives. World Bank Publications. https://doi.org/10.1596/978-1-4648-1600-0