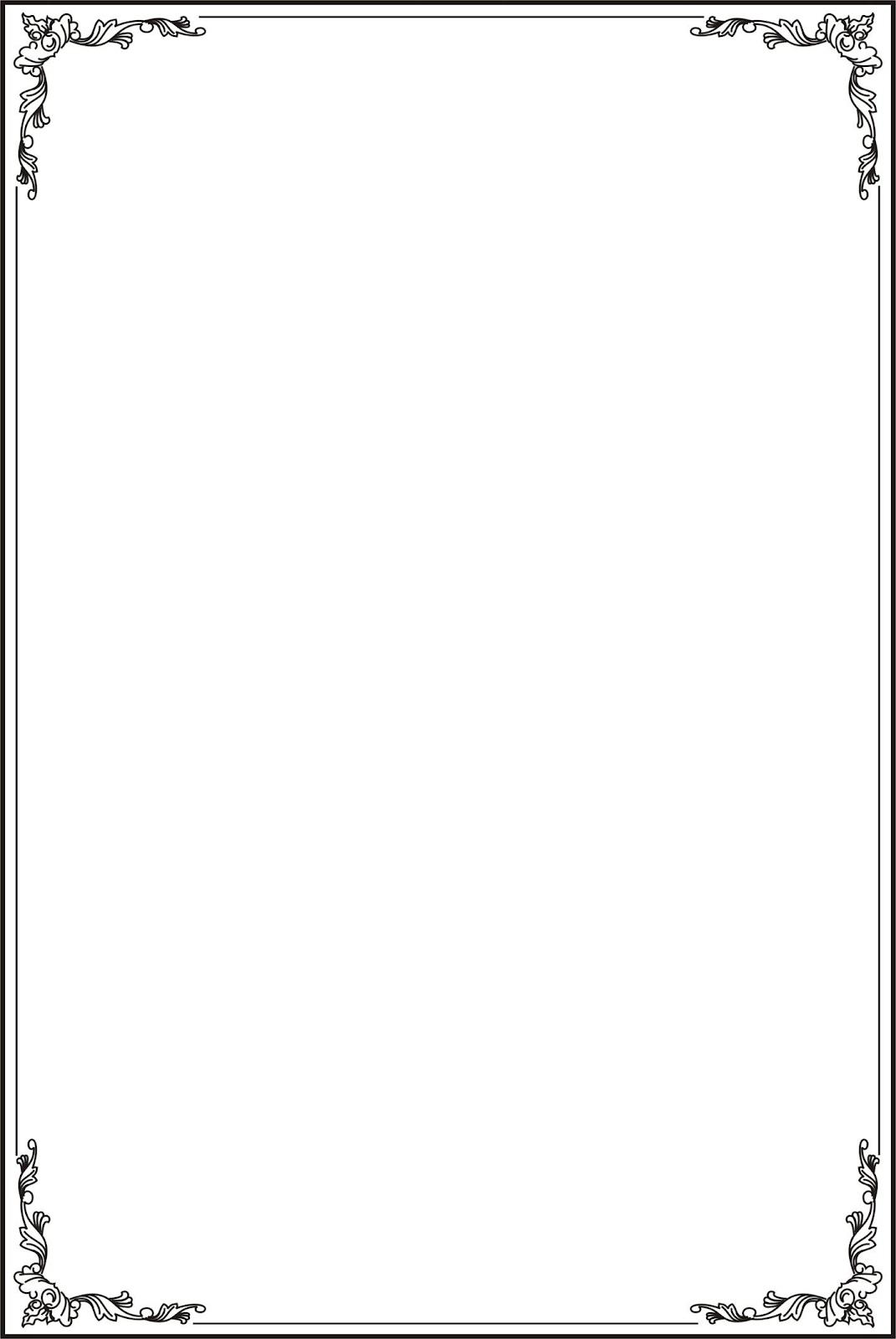
**H****ỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**VIỆN KHOA HỌC KỸ THUẬT BƯU ĐIỆN**

****

**BÁO CÁO PHÂN TÍCH BÀI BÁO**

**Machine Learning based Anomaly Detection for 5G**

|  |  |
| --- | --- |
| ***Giảng viên hướng dẫn*** |  |
| ***Họ và tên sinh viên*** | **Nguyễn Ngọc Duy** |
| ***Mã sinh viên*** | **B23DCCC049** |

***Hà Nội – 2025***

1. **Phần mở đầu ( Introduction )**

* **Bối cảnh:** Sự phát triển mạnh mẽ của 5G và các thách thức bảo mật.
* **Tiêu đề:** *Machine Learning based Anomaly Detection for 5G Netwworks ( Phát hiện bất thường dựa trên Machine Learning cho mạng 5G )*
* **Tác giả:** Robert Abbas và Jordan Lam
* **Ngày công bố:** 06/03/2020
* **Mục tiêu chính:** Phát triển một hệ thống phát hiện bất thường dựa trên máy học cho mạng 5G

1. **Tóm tắt bài báo ( Summary of the Paper )**

Bảo vệ mạng của tương lai được thiết lập để trở thành một lĩnh vực đầy thách thức do các mối đe dọa an ninh mạng ngày càng tăng và bề mặt tấn công mở rộng do Internet vạn vật (IoT) tạo ra, tăng tính không đồng nhất của mạng, tăng cường sử dụng các công nghệ ảo hóa và kiến trúc phân tán. Bài báo này đề xuất SDS (Bảo mật do phần mềm xác định) như một phương tiện để cung cấp một hệ thống phòng thủ mạng tự động, linh hoạt và có thể mở rộng. SDS sẽ khai thác những tiến bộ hiện tại trong học máy để thiết kế CNN (Convolutional Neural Network) bằng cách sử dụng NAS (Neural Architecture Search) để phát hiện lưu lượng mạng bất thường. SDS có thể được áp dụng cho hệ thống phát hiện xâm nhập để tạo ra khả năng phòng thủ chủ động và đầu cuối hơn cho mạng 5G. Để kiểm tra giả định này, các luồng mạng bình thường và bất thường từ một môi trường mô phỏng đã được thu thập và phân tích bằng CNN. Kết quả từ phương pháp này rất hứa hẹn vì mô hình đã xác định được lưu lượng truy cập lành tính với tỷ lệ chính xác 100% và lưu lượng truy cập bất thường với tỷ lệ phát hiện 96,4%. Điều này chứng minh hiệu quả của phân tích luồng mạng đối với nhiều cuộc tấn công độc hại phổ biến và cũng cung cấp một tùy chọn khả thi để phát hiện lưu lượng mạng độc hại được mã hóa.

1. **Phân tích chi tiết (Detailed Anlysis)**
2. *Bài báo trình bày về giải pháp nào trong bảo mật 5G*

Bài báo đề xuất hệ thống **Software Defined Security (SDS)** kết hợp với **Convolutional Neural Network (CNN)** để phát hiện lưu lượng bất thường trong mạng 5G. Giải pháp này sử dụng **Neural Architecture Search (NAS)** để tự động thiết kế CNN tối ưu, nhằm phát hiện các mối đe dọa như zero-day attacks mà không cần giải mã lưu lượng mã hóa. SDS được triển khai end-to-end, từ backhaul đến core network, và có thể tùy chỉnh cho từng network slice. Mục tiêu là tạo ra một hệ thống bảo mật tự động, linh hoạt, và hiệu quả, đáp ứng các thách thức của 5G như IoT, virtualization, và kiến trúc phân tán.

1. *Bài báo tổ chức, tiến hành phương pháp đó như thế nào.*

 **Thu thập dữ liệu:** Sử dụng bộ dữ liệu **CICIDS2018**, mô phỏng mạng thực tế với 50 máy tấn công trên một tổ chức nạn nhân với 5 bộ phận và 420 máy và 30 máy chủ trên AWS.

*Giải thích*

* 50 máy tấn công đề cập đến máy tính được sử dụng trong mô phỏng của tấn công trong cuộc thử nghiệm của tập dữ liệu CICIDS2018 đây không phải là thiết bị vật lý cụ thể mà là các máy tính ảo hoặc vật lý được cấu hình để thực hiện các hành vi độc hại nhằm tạo ra lưu lượng mạng bất thường.
* Các máy này thuộc "attacker network" (mạng của kẻ tấn công) trong mô phỏng.
* Chúng chạy các hệ điều hành như Windows 8/10 và Ubuntu, được trang bị các công cụ tấn công như Metasploit (cho khai thác lỗ hổng), Slowloris, LOIC, HOIC (cho tấn công DoS), hoặc các công cụ brute force.
* Mục đích: Tạo ra các loại tấn công khác nhau (ví dụ: xâm nhập nội bộ, HTTP DoS, tấn công ứng dụng web, brute force, khai thác Heartbleed) để kiểm tra khả năng phát hiện của mô hình học máy.
* 5 bộ phận đề cập đến tổ chức của *victim organisation* là(tổ chức bị tấn công) trong mô phỏng CIDS 2018 đây là 5 phân đoạn mạng đại diện cho các phòng ban khác nhau của một tổ chức giả lập được thiết kế để phản ánh một môi trường doanh nghiệp thực tế.
* Môi trường mạng được chia thành 6 subnets: 5 subnets cho các phòng ban (Dep1 đến Dep5) và 1 subnet cho server.
* **Dep1 đến Dep4**: Các máy tính chạy Windows 8/10, mô phỏng các phòng ban sử dụng hệ điều hành phổ biến.
* **Dep5**: Các máy tính chạy Ubuntu, đại diện cho một phòng ban sử dụng hệ điều hành Linux.
* Mỗi "bộ phận" này chứa các máyđể mô phỏng hành vi người dùng bình thường và làm mục tiêu cho các cuộc tấn công.
* 420 máy tính thuộc về tổ chức nạn nhân đây là số lượng máy tính của người dùng cuối được phân bổ trong 5 bộ phận trên.\
* **Vị trí**: 420 máy này được đặt trong 5 subnets tương ứng với 5 bộ phận (Dep1 đến Dep5).

Dep1 đến Dep4: Chạy Windows 8/10.

Dep5: Chạy Ubuntu.

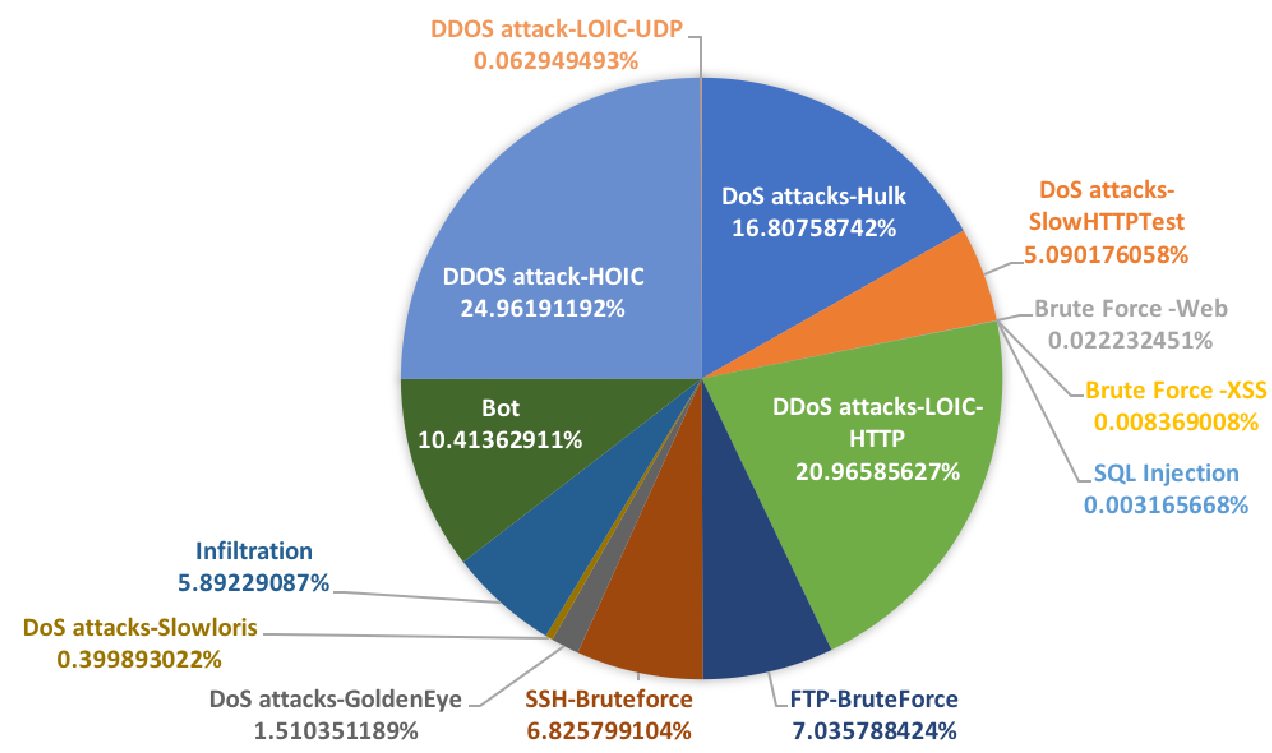
* Chúng đại diện cho các máy trạm (workstations) của nhân viên trong tổ chức, tạo ra lưu lượng bình thường (B-Profile) bằng cách mô phỏng hành vi người dùng (ví dụ: truy cập web, gửi email) thông qua các thuật toán học máy như K-Means, Random Forest, SVM.
* **Mục đích**: Các máy này là mục tiêu của 50 máy tấn công, và lưu lượng từ chúng được thu thập để phân tích sự khác biệt giữa bình thường và bất thường.
* 30 máy chủ làm các máy chủ thuộc tổ chức nạn nhân được sử dụng để cung cấp dịch vụ và lưu trữ dữ liệu đồng thời cũng là mục tiêu của các cuộc tấn công
* **Vị trí**: 30 máy chủ này được đặt trong subnet thứ 6, được gọi là "Servers" trong sơ đồ mạng (xem Hình 7 trong bài báo).
* **Cấu hình**: Chạy các hệ điều hành Microsoft Windows Server (ví dụ: App servers, Active Directory, email servers), cung cấp các dịch vụ như HTTPS, SMTP, POP3, IMAP, SSH, FTP.
* **Vai trò**:

Xử lý lưu lượng mạng từ 420 máy trạm trong 5 bộ phận.

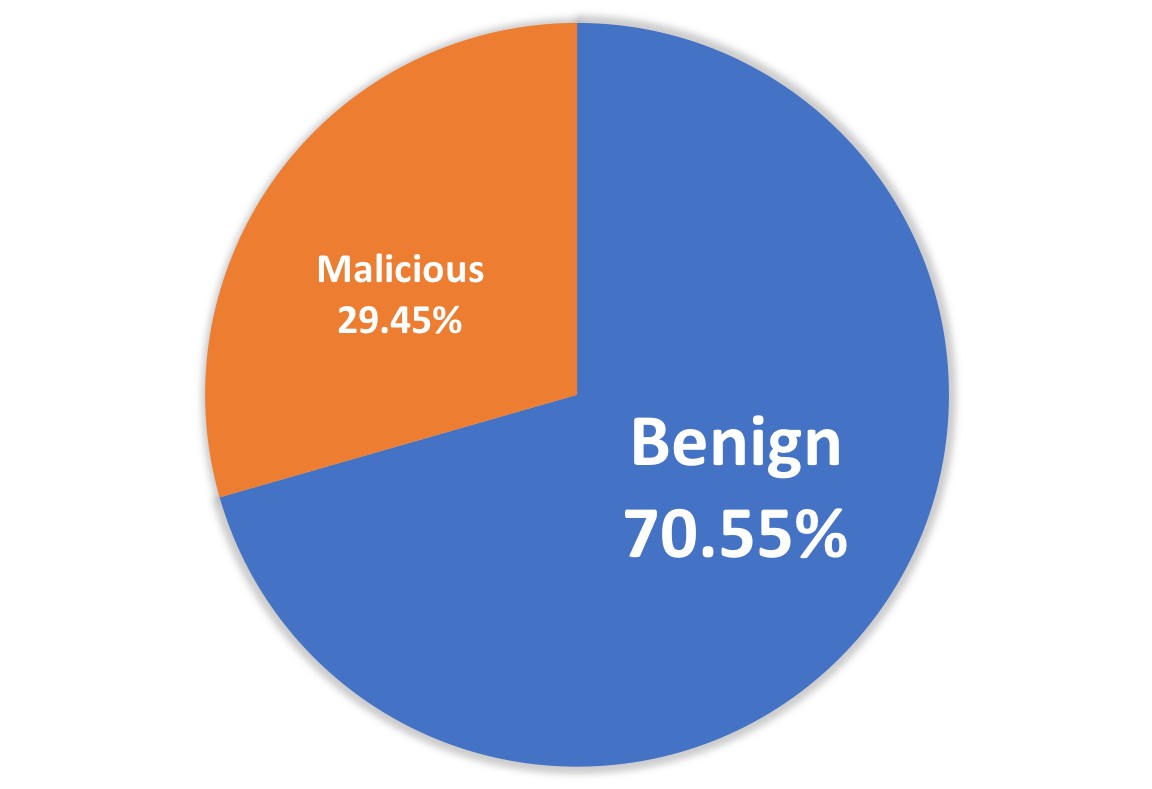
Là mục tiêu chính của các cuộc tấn công như brute force (đối với SSH, MySQL) hoặc khai thác lỗ hổng (Heartbleed).

* Lưu lượng từ các máy chủ này được thu thập để phân tích cùng với lưu lượng từ máy trạm.

Số lượng: Tổng có 9,332,770 luồng (70.55% benign, 29.45% anomaly), với các giao thức như HTTPS, SMTP, SSH.



*Hình 1: Biểu đồ hình tròn về khối lượng loại lưu lượng truy cập độc hại*



*Hình 2: Biểu đồ hình tròn của tất cả các khối lượng lưu lượng truy cập*

* Chọn 20 đặc trưng : Có 6 tính năng cơ bản của Flow ( port, protocol, flow duration, Total Forward Packets, Total Back-ward Packets, Flow Pkts/s) và 14 IAT (Flow IAT Mean, Flow IAT Standard Deviation, Flow IAT Maximum, Flow IAT Minimum, Flow IAT Total, Forward IAT Mean, Forward IAT Standard Deviation, Forward IAT Max,Forward IAT Min, Backward IAT Total, Backward IAT Mean, Backward IAT Standard Deviation, Backward IAT Max, Backward IAT Min)

 **Tiền xử lý:**

* Dữ liệu CSV được chuyển thành hình ảnh RGB với kích thước 100x100x3 để phù hợp với CNN bất kỳ dữ liệu bổ sung nào còn sót lại dưới kích thước này sẽ bị xóa bỏ vì tất cả các hình ảnh cho CNN phải có cùng kích thước đầu vào.
* AutoML Vision tự động tăng kích thước hình ảnh lên 224x224x3 trong quá trình huấn luyện.
* Hình ảnh bất thường có đặc điểm ngẫu nhiên, nhiễu, trong khi đó hình ảnh bình thường có mẫu hình rõ ràng.



Hình ảnh bất thường 1 Hình ảnh bất thường 2



Hình ảnh lành tính 1 Hình ảnh lành tính 2

 **Thiết kế mô hình:**

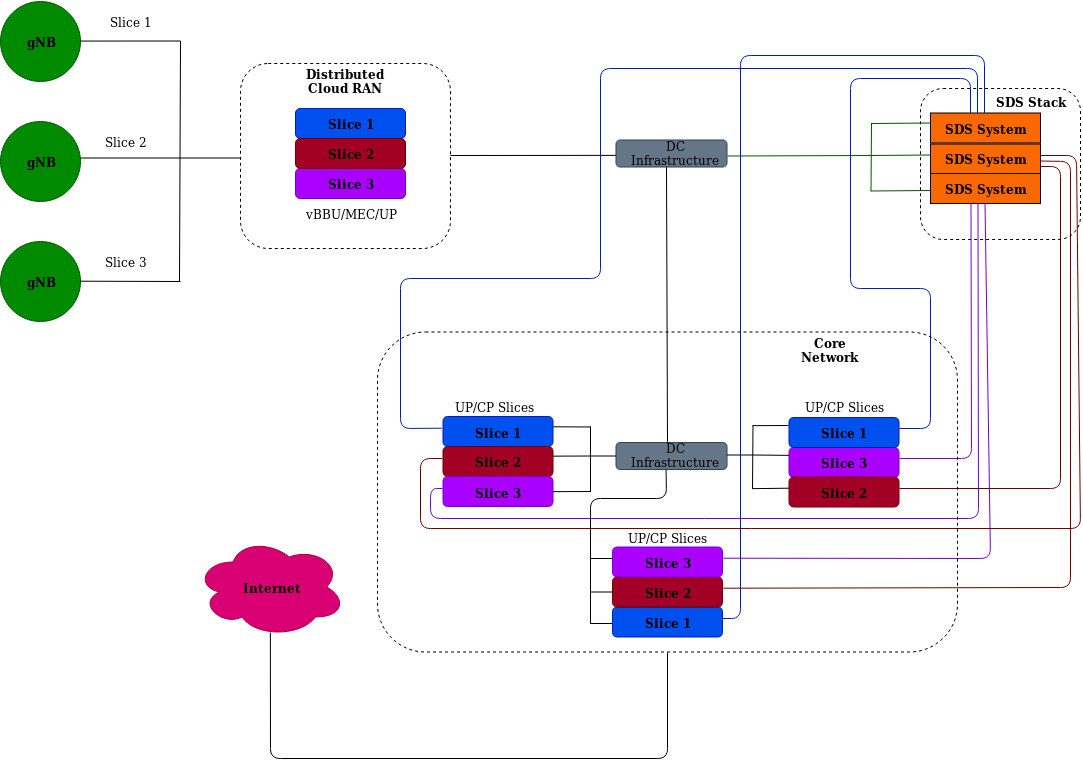
* Dùng **Google AutoML Vision** và Vision Edge với hai kiến trúc:
* **NASNet:** Tối ưu độ chính xác chạy thử nghiệm 24 giờ.
* **MnasNet:** Tối ưu cho tốc độ chạy thử nghiệm trong 3 giờ.
* NAS (Neural Architecture Search) tự động điều chỉnh siêu tham số CNN qua RNN controller để tự động tối ưu siêu tham số (số bộ lọc, kích thước bộ lọc, bước stride)
* Không gian tìm kiếm được chia thành các “cell” với cấu trúc giống nhau nhưng trọng số khác nhau, giảm độ phức tạp và tăng khả năng tổng quát hóa.
* Dữ liệu huấn luyện: 1,433 hình ảnh (925 bình thường, 508 bất thường), chia thành tập huấn luyện, kiểm tra và đánh giá.
* Thời gian xử lý:

MNasNet: 105ms/hình ảnh trên điện thoại Pixel 1.

NASNet: Tốn nhiều tài nguyên hơn nhưng cải thiện độ chính xác nhẹ.

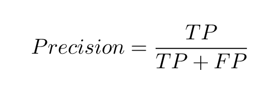
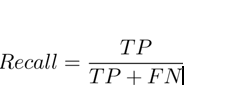
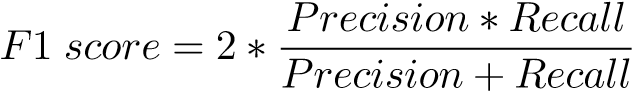
 **Triển khai SDS:**

* Sao chép lưu lượng từ backhaul và core network, phân tích để xây dựng profile, cập nhật chính sách IDS qua VNF manager.



*Hình 4: Ứng dụng SDS cho các lắt cắt mạng*

1. *Bài báo đánh giá và đo thử nghiệm thế nào mà lại cho rằng giải pháp của họ là tốt.*

* Thử nghiệm: Chạy AutoML Vision trên dữ liệu hình ảnh CICIDS2018 với NASNet (24 giờ) và MNasNet (3 giờ).
* Thước đo:
* *Precision:*  🡪 Tỷ lệ dự đoán anomaly đúng.
* *Recall:* **🡪Tỷ lệ nomaly được phát hiện.
* *F1 score :*  🡪 Trung bình hài hòa của precision và recall.
* Kết quả:
* **NASNet:** Precision trung bình 99.2%, recall tối đa 98.58%,
* **MNasNet:** Precision trung bình 97.6%, recall tối đa 98.58%

|  |  |  |  |
| --- | --- | --- | --- |
|  | 24 Hour & 3 Hour Results | |  |
| **Traffic Type** | **Precision** | **Recall** | **F1 Score** |
| Benign | 0.977 | 1 | 0.988 |
| Anomaly | 1 | 0.964 | 0.982 |
| Average | 0.9885 | 0.982 | 0.985 |

*Bảng thống kê hiệu suất*

* Kết luận

Tính thực tiễn cao: Latency thấp, triển khai được trên smartphone, phù hợp với 5G cloud-native.

1. *Bài báo có nêu hạn chế gì không.*

* **Dữ liệu nhỏ:** 1433 ảnh (925 benign, 508 anomaly) chưa đủ lớn để kiểm chứng tính tổng quát, đặc biệt với sự đa dạng của lưu lượng 5G thực tế.
* **Môi trường mô phỏng:** CICIDS2018 không đại diện hoàn toàn cho mạng 5G thực (thiếu network slicing, edge computing thực tế).
* **Miss rate:** 3.6% anomaly bị nhầm thành benign, có thể bỏ sót các tấn công nghiêm trọng như DDoS.
* **Học có giám sát:** Chỉ dùng labeled data, không giải quyết được thực tế lưu lượng mạng phần lớn không nhãn.
* **Chi phí tài nguyên** : Mô hình NASNet cho độ chính xác cao hơn nhưng mất tới 24 giờ huấn luyện, điều này làm tăng chi phí tính toán. Trong đó MNASNet nhanh hơn ( 3 giờ) nhưng đánh đổi một chút về độ chính xác.
* **Chưa được kiểm chứng đầy đủ với các cuộc tấn công chưa từng biết trước:** Các cuộc tấn công độc hại có thể được khái quát hóa thành hai loại tấn công zero-day và tấn công ngày đầu tiên . Các cuộc tấn công zero-day là các mối đe dọa không có dấu vân tay hoặc chữ ký hiện có, các cuộc tấn công ngày là các mối đe dọa có chữ ký hoặc dấu vân tay và có thể được giảm thiểu một cách hiệu quả. Mục tiêu cuối cùng của phát hiện bất thường là cung cấp phản ứng nhanh hơn và chủ động hơn đối với các mối đe dọa chưa từng thấy trước đây( zero day ) và giảm thiểu thích hợp.

**Đề xuất thêm:**

* Tăng dữ liệu (ví dụ: kết hợp thêm dataset như UNSW-NB15).
* Thử nghiệm trên testbed 5G thực tế.
* Kết hợp học không giám sát để xử lý unlabeled data.

1. *Em có thể tiến hành chạy CODE nếu bài báo có công bố đường link để tải về không.*
   * + 1. Bài báo không cung cấp đường link đến mã nguồn hoặc reponsitory Github . Tuy nhiên bài báo có mô tả chi tiết quy trình thực hiện, bao gồm:

* Sử dụng **Google AutoML Vision** và **Vision Edge** để thiết kế và huấn luyện mô hình.
* Dữ liệu được lấy từ bộ **CICIDS2018**, có sẵn công khai từ Canadian Institute of Cybersecurity.

**Đề xuất thêm:** Liên hệ tác giả qua email

([jordan.lam@studentm.mq.edu.au](mailto:jordan.lam@studentm.mq.edu.au), [robert.abbas@mq.edu.au](mailto:robert.abbas@mq.edu.au) ) để xin code.

1. **Kết luận**

Bài báo đề xuất SDS với CNN là giải pháp sáng tạo cho bảo mật 5G, đạt kết quả ấn tượng (100% benign, 96.4% anomaly) và khả năng triển khai thực tế nhờ latency thấp. Phương pháp được tổ chức chặt chẽ, đánh giá minh bạch qua các thước đo tiêu chuẩn, nhưng còn hạn chế về dữ liệu nhỏ, môi trường mô phỏng, và phụ thuộc học có giám sát. Tái hiện thử nghiệm khả thi với AutoML, nhưng cần tài nguyên và mã nguồn bổ sung. Nghiên cứu này đóng góp lớn vào bảo mật 5G, đặc biệt trong bối cảnh năm 2025, dù cần cải thiện để ứng dụng rộng rãi.