Hybrid Intrusion Detection System for Internet of Things (IoT)

Abstract:

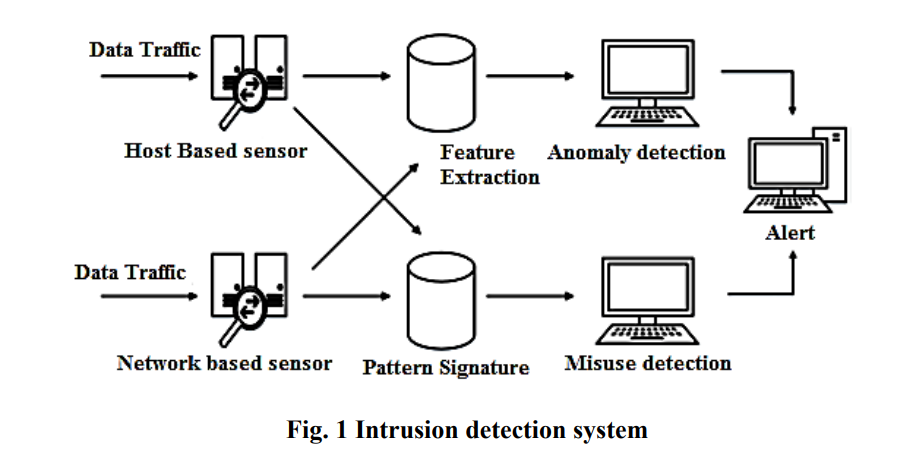
* Đoạn văn này nói về lợi ích và thách thức của Internet of Things (**IoT**), một **công nghệ kết nối các thiết bị khác nhau thông qua internet**.
* Mặc dù IoT **cung cấp** một loạt các **dịch vụ và ứng dụng**, nhưng nó cũng đối mặt với **nhiều vấn đề bảo mật** và **dễ bị tấn công**, chẳng hạn như tấn công sinkhole, nghe trộm và tấn công từ chối dịch vụ.
* **Để giải quyết** vấn đề này, tác giả đề xuất một **hệ thống phát hiện xâm nhập** cho các mạng IoT dựa trên **hybrid convolutional neural network model**.
* Mô hình đc đề xuất này phù hợp cho một loạt các ứng dụng IoT và **được so sánh** với **các mô hình học máy và học sâu** thông thường.
* Kết quả thử nghiệm cho thấy **hybrid model** nhạy hơn đối với các cuộc tấn công trong mạng IoT.

**INTRODUCTION**

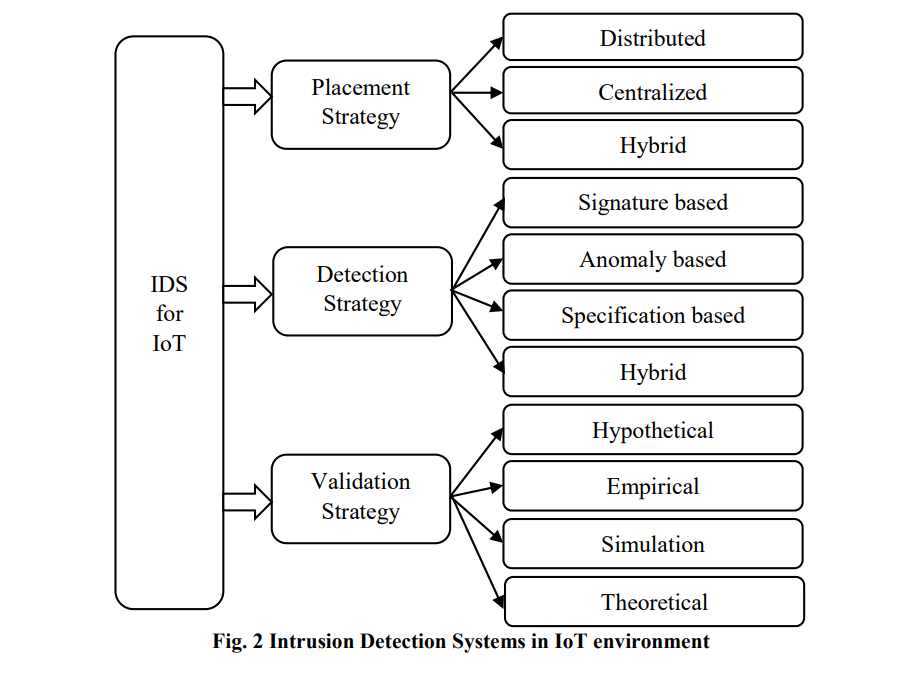
* Internet of Things (**IoT**) là một ứng dụng mới và **hỗ trợ** cho **nhiều lĩnh vực** như quy trình công nghiệp, chăm sóc sức khỏe, tự động hóa, môi trường thông minh,...
* Tuy nhiên, IoT đang phải **đối mặt với nhiều vấn đề bảo mật** và **các cuộc tấn công** do tính đa dạng và **không tương thích** với các p**hương pháp bảo mật thông thường**.
* Để đảm bảo an ninh cho mạng IoT, hệ thống phát hiện xâm nhập (**IDS**) được đề xuất nhằm **nâng cao** tính **bảo mật** và giảm thiểu các lỗ hổng bảo mật.
* Trước khi phát triển hệ thống IDS, cần có kiến thức về môi trường IoT và các vấn đề bảo mật của nó. Các tính chất bảo mật chính của IoT được tóm tắt như sau:
* **Data Confidentiality:** Các yêu cầu bảo vệ căn bản của IoT bao gồm việc **không cho phép người dùng được ủy quyền** thực hiện **sửa đổi** hoặc **thay đổi dữ liệu trên các thiết bị IoT.**

**Quyền riêng tư** **dữ liệu** là cần thiết cho các ứng dụng công nghiệp và cá nhân, và sửa đổi dữ liệu sức khỏe sẽ dẫn đến các vấn đề nghiêm trọng. Do đó, **tính bảo mật của dữ liệu** là một **yếu tố quan trọng** trong môi trường IoT.

* **Data Integrity** : **Bảo mật tính toàn vẹn** dữ liệu là **yếu tố cần thiết** trong môi trường IoT. Vì IoT hoạt động trong môi trường đa dạng và **dữ liệu được di chuyển từ những nơi xa về trung tâm**. Độ tin cậy và **dịch vụ đảm bảo** để **chuyển dữ liệu từ các hệ thống xa là cần thiết trong IoT**. Vì vậy, việc xác định **tính toàn vẹn** của dữ liệu bằng cách **xác minh nguồn dữ liệu** và xác định các cuộc tấn công độc hại **trong các thiết bị , một bộ phận của hệ thống Iot** [2] là cần thiết.
* **Data Availability :** Để hoạt động hiệu quả, **việc truy cập dữ liệu** trong môi trường IoT là rất quan trọng và **cần được đảm bảo liên tục**. Tuy nhiên, **các thiết bị** được sử dụng bởi người dùng **có thể gặp phải các lỗ hổng** bảo mật dẫn đến **những vấn đề** liên quan đến an ninh mạng. Do đó, **việc cung cấp truy cập dữ liệu an toàn cho người dùng là rất cần thiết trong môi trường IoT.**
* **Authentication :** Nhìn chung, **quá trình xác minh trong IoT khác nhau từ hệ thống này sang hệ thống khác** [4] và việc **phân loại đối tượng** cần được thực hiện ở **giai đoạn ban đầu [5]** là một quá trình **cần thiết**. IoT **yêu cầu** một quá trình **xác minh** **tốt** để cung cấp **cho phép truy cập** và **truy xuất dữ liệu** kèm theo **khả năng thích ứng** tốt hơn. Cho nên ta phải **cân bằng giữa khả năng thích ứng và bảo mật dữ liệu trong quá trình thiết kế và triển khai một môi trường IoT.**. [3]
* **Authorization** : Việc **xác định quyền** của người dùng đối với việc truy cập dữ liệu trong môi trường IoT cần được định nghĩa bằng **quy trình ủy quyền**. Bằng cách đo lường các thiết bị và tính bảo mật thông tin, **quy trình ủy quyền cần được định nghĩa trong một mạng [6]**
* Đoạn văn trình bày về nhu cầu **phát triển cơ chế bảo mật cho môi trường Internet of Things (IoT).** Nó nhấn mạnh rằng **Data oriented security mechanism** [7] là cần thiết để **ngăn chặn truy cập trái phép** đến nguồn dữ liệu từ các người dùng độc hại. Đoạn văn lưu ý rằng các **cơ chế bảo mật thông thường** dựa trên kỹ thuật mật mã **không được sử dụng rộng rãi** trong môi trường **IoT** do lượng dữ liệu lớn.
* Tác giả cho rằng **xác định các mối đe dọa** trong **thời** **gian tối thiểu** là **quan trọng** để giảm thiểu các vấn đề trong mạng. Do đó, **hệ thống Phát hiện Xâm nhập (IDS)** là **cần thiết** để xác định kẻ xâm nhập trong mạng IoT và ngăn chặn truy cập trái phép vào dữ liệu.
* Tuy nhiên, do **giới hạn** về **tài nguyên**, việc **triển khai IDS** có thể **phức tạp**. Để đơn giản hóa quá trình, một hệ thống **IDS trung tâm** có thể **được sử dụng** **để giám sát mạng** và các nút **từ xa** và **kích hoạt cảnh báo** cho quản trị mạng trong trường hợp có sự cố bảo mật. Đoạn văn bao gồm một hình ảnh minh họa việc triển khai hệ thống IDS.



* Các **hệ thống phát hiện xâm nhập** được chia thành **ba giai đoạn**.
  + **Giám sát** là giai đoạn ban đầu trong IDS dựa trên cảm biến mạng hoặc máy chủ.
  + **Phân tích** là giai đoạn thứ hai của IDSs, thực hiện trích xuất tính năng và quá trình nhận dạng mẫu dựa trên đó.
  + **Phát hiện** là giai đoạn cuối cùng của IDSs, phát hiện bất thường hoặc xâm nhập trong một mạng.
* IDS giúp **theo dõi và phân tích thông tin, dịch vụ và mạng**, phân tích **lưu lượng** thông qua quản lý mạng hiệu quả và **xác định các lỗ hổng** trong thời gian ngắn nhất.
* Nó **bảo vệ mạng** khỏi các cuộc **tấn công** và cải thiện tính bảo mật và toàn vẹn dữ liệu mạng. IDS tóm tắt lưu lượng dữ liệu của hệ thống và phân tích nó để phát hiện các hoạt động độc hại.
* **Kiến trúc hệ thống phát hiện xâm nhập truyền thống** chủ yếu tập trung vào **cung cấp bảo mật** cho các **đặc tính quản lý internet** và **thiếu bảo mật** dòng **dữ liệu lớn thời gian thực. [8]**
* Cơ bản, **IDS truyền thống** được phân loại thành **ba loại** như **placement strategy, detection strategy** và **validation strategy**
* Hình 2 mô tả các loại **hệ thống phát hiện xâm nhập khác nhau** trong môi trường **IoT**. Trong số ba danh mục này, **chiến lược phát hiện** thu hút nhiều sự chú ý nhất và **hầu hết các hệ thống** **được phát triển** dựa **trên chiến lược phát hiện**.
* **Signature based IDS, Anomaly based IDS, specification based IDS and Hybrid IDS** là các **phân loại con** trong **chiến lược phát hiện.**



* **Signature based IDS**: hệ thống phát hiện xâm nhập **dựa trên chữ ký.** Nó **mô tả các cuộc tấn công** và **các mẫu tấn công**, và **nhận dạng các cuộc tấn công** bằng cách **so sánh các hoạt động trên mạng** **với** **các chữ ký tấn công đã biết trước**. **Khi phát hiện** một cuộc tấn công trong mạng, hệ thống phát hiện dựa trên chữ ký sẽ **tạo ra cảnh báo** về các hoạt động đáng ngờ và **thực hiện phân tích các mẫu tấn công**. Dựa trên **sự tương tự và khác biệt**, hệ thống **cung cấp truy cập hoặc cảnh báo cho người dùn**g và phát hiện các cuộc tấn công một cách hiệu quả.
* **Anomaly based IDS** là hệ thống **phát hiện xâm nhập** **dựa trên sự bất thường**. Nó **thu thập dữ liệu** và **nhận dạng các sự bất thường** trong hệ thống **dựa trên giá trị ngưỡng**.. Anomaly based IDS **phát hiện** các cuộc **tấn công không xác định** một cách **hiệu quả**, tuy nhiên nó **đòi hỏi bộ nhớ lớn để xử lý**  và **chi phí tính toán** là hạn chế của hệ thống phát hiện xâm nhập dựa trên sự bất thường.

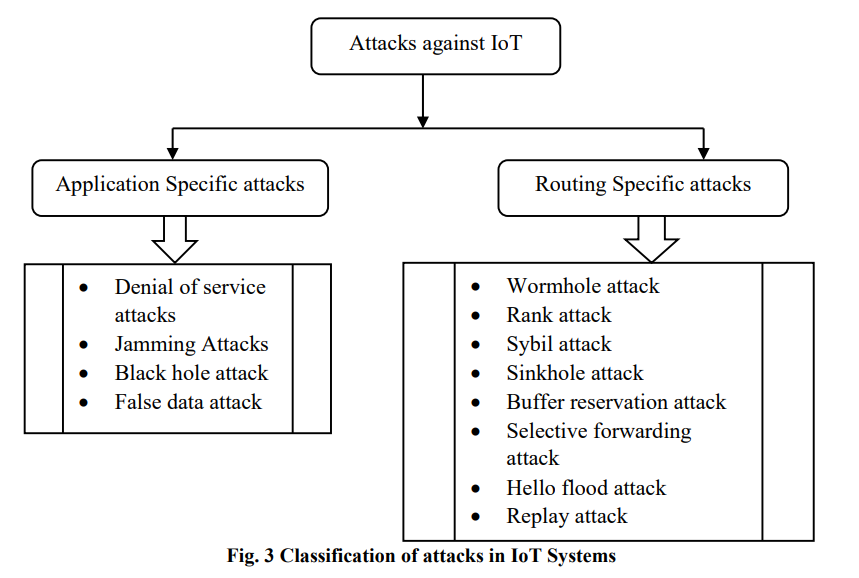
( lưu trữ data và so sánh data với mô hình dữ liệu lớn )

* **Specification-based IDS** là hệ thống phát hiện xâm nhập **dựa trên thông số kỹ thuật**. Hệ thống này liên tục đánh giá các hoạt động trong hệ thống dựa trên các thông số kỹ thuật **cụ thể do quản trị mạng định nghĩa** và **liên tục giám sát** quá trình để xác nhận các hoạt động đó. Nếu phát hiện bất thường đối với các hoạt động đã định nghĩa, hệ thống sẽ gửi cảnh báo đến quản trị mạng.
* **Hybrid IDS** là mô hình kết hợp giữa phát hiện xâm nhập dựa trên **Anomaly based IDS** và **Signature based IDS**. Các mô hình kết hợp này được xem là tốt hơn trong việc đưa ra sự cân đối giữa chi phí lưu trữ và tính toán với **ít cảnh báo** dương tính **sai**. Hiện nay, hầu hết các hệ thống dựa trên Hybrid IDS do khả năng phát hiện hiệu quả và hoạt động đơn giản của nó.

**2. Related Works**

**-** Trong hệ thống phát hiện xâm nhập, **chú trọng** chính vào việc **phát hiện các cuộc tấn công** và điều quan trọng là **xác định các loại tấn công khác nhau trong môi trường IoT**.

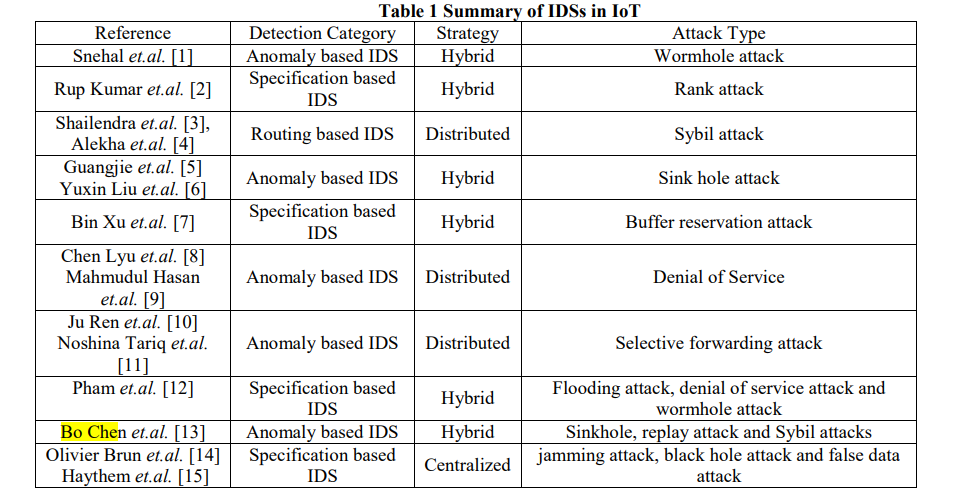
Một số tấn công chính được phân loại trong môi trường IoT bao gồm: **Sybil attack, selective forwarding attack, service attack, wormhole attack, replay attack, sinkhole attack, false data attack, black hole attack, jamming attack .**Hình 4 cung cấp một cái nhìn chi tiết về các loại tấn công khác nhau trong môi trường IoT.



* Nghiên cứu của **Snehal et.al**. [1] tập trung vào cuộc tấn công **wormhole** trong môi trường **IoT**, trong đó nút mục tiêu **bị tấn công** hoặc **xâm nhập từ hai hướng khác nhau.** Nghiên cứu tập trung vào việc **xác định vị trí kẻ xâm nhập trong mạng** và **cảnh báo cho quản trị mạng**.

**Rup Kumar et.al**. [2] nói về các cuộc **rank attacks** trong môi trường IoT. Đó là một cuộc **xâm nhập dựa trên định tuyến** xảy ra **trong low power networks [9]**. Dựa trên nguyên tắc xếp hạng, một topology hiệu quả được tạo ra trong nghiên cứu bằng cách tránh quá trình định hình vòng lặp truyền thống để giảm thiểu chi phí.

* Các nghiên cứu liên quan tập trung vào những **tác động của các cuộc tấn công** như **Sybil attack** và **sinkhole attack** đến mạng IoT.
  + **Shailendra** et.al. [3], tác động của **Sybil attack** đến mạng IoT được đưa ra.
  + **Alekha** et.al. [4] tập trung vào các **cuộc tấn công Sybil** dựa trên **social graph based** và phân loại dựa trên đồ thị xã hội là một lĩnh vực quan trọng trong bảo mật IoT.
  + Nghiên cứu của **Guangjie** et.al. [5] tập trung vào **cuộc tấn công sinkhole** và ảnh hưởng của nó đến mạng
  + **Yuxin Liu et.al**. [6] đề xuất một **hệ thống phát hiện xâm nhập** để **phát hiện** cuộc **tấn công sinkhole** bằng cách **sử dụng giao thức định tuyến RPL** và **đánh giá tỷ lệ gói nhận và gói truyền để xác định tỷ lệ xâm nhập**. Khi phát hiện các nút độc hại, hệ thống sẽ tạo ra cảnh báo để giảm thiểu ảnh hưởng của xâm nhập.
* Trong nghiên cứu của **Bin Xu et.al**. [7], tác giả báo cáo về (**buffer reservation attack**) trong môi trường **IoT**, tấn công này **xảy ra do sự trùng lặp các nút trong mạng**. Kẻ tấn công **phân mảnh các nút** và **tạo các nút giả** để tấn công và **thực hiện các hoạt động độc hại trong mạng**.
* Trong khi đó, **Chen Lyu et.al**. [8] báo cáo về vấn đề quan trọng trong bảo mật IoT như **tấn công từ chối dịch vụ (DoS).** Trong quá trình này, kẻ tấn công **tấn công vào nút và từ chối các yêu cầu dữ liệu từ các nút khác**. Trên cơ sở đó, các tác giả **đề xuất mô hình phát hiện xâm nhập để xác định người dùng độc hại trong mạng**.
* Nghiên cứu của **Mahmudul Hasan et.al**. [9] tổng hợp các chiến lược khác nhau của các cuộc tấn công từ chối dịch vụ như một cuộc khảo sát. Tóm tắt nghiên cứu cung cấp một cái nhìn tổng quan về các cơ chế phòng chống và kỹ thuật phát hiện khác nhau chống lại các cuộc tấn công DoS và DDoS.
* Bài báo của **Ju Ren et.al**. [10] tập trung vào vấn đề của **selective forwarding attack** trong mạng IoT. Trong tấn công này, **một nút độc hại giả mạo là một nút chính** cho **việc truyền tải và ảnh hưởng đến định tuyến mạng và hiệu suất truyền tải.** Trong quá trình này, chỉ **những tin nhắn được lựa chọn** sẽ **được truyền** và các **tin nhắn khác sẽ bị kẹt lại** trong chính nút đó, **gây ảnh hưởng đến hoạt động định tuyến.**
* **Noshina Tariq** et.al. [11] đã đề xuất **mô hình phát hiện xâm nhập** s**ử dụng mạng nơ-ron nhân tạo** để **nhận diện** **các** **cuộc tấn công** như vậy trong một mạng **IoT**. Công trình nghiên cứu được xác minh với một **hệ thống** truyền tải dữ liệu **thời gian thực**, trong đó các **thông điệp điều khiển được chuyển tiếp** và **các thông điệp khác bị chặn**. Với **tỷ lệ phát hiện cao và chi phí tính toán thấp**, mô hình đề xuất có khả năng nhận diện người xâm nhập trong mạng
* **Pham và đồng nghiệp** [12] đã đề xuất một **hệ thống phát hiện xâm nhập** để **xác định cuộc tấn công** "**hello flood**" trong một mạng. Hệ thống **này sử dụng mô hình mạng nơ-ron truyền thẳng** và **mô hình mạng nơ-ron truyền ngược** để **phát hiện các cuộc tấn công của đối thủ trong mạng IoT.**
* Các tác giả **Bo Chen và đồng nghiệp** [13] đã **tập trun**g vào việc **phát hiện** các cuộc tấn công **sinkhole**, **replay** và **Sybil** trong môi trường **IoT**. Đề xuất của họ **hiệu quả** trong việc phát hiện các cuộc tấn công này **thông qua hệ thống phát hiện xâm nhập dựa trên fuzzy**.
  + Trong cuộc tấn công "**replay**", kẻ xâm nhập t**hực hiện cuộc tấn công vào các thời điểm khác nhau để thu thập dữ liệu cần thiế**t, sau đó **phát lại dữ liệu đó trong mạng**. Điều này d**ẫn đến các vấn đề không mong muốn cho cộng đồng người dùng**, gây ra nh**ững vấn đề nghiêm trọng cho các truyền thông quan trọng**. Mô hình đề xuất **hiệu quả** trong việc xác định các cuộc tấn công này thông qua hệ thống phát hiện xâm nhập **dựa trên "fuzzy"** (mờ) thực hiện quá trình phát hiện thông qua một tập các quy tắc.
* **Olivier Brun** et.al. [14] nghiên cứu về **tấn công jamming** trong môi trường **IoT**, trong đó **kênh truyền được giám sát bởi kẻ tấn công**. Họ đề xuất một **hệ thống phát hiện xâm nhập** sử dụng hệ thống suy luận **fuzzy neuro** thích nghi để xác định các cuộc tấn công **jamming**, **Sybil** và **từ chối dịch vụ**. Tuy nhiên, **chi phí tính toán** của hệ thống này là **rất cao.**
* **Haythem** et.al. [15] đề xuất một hệ thống phát hiện xâm nhập sử dụng **mạng nơ-ron sâu** để **xác định black hole attack** và **false data attack**. Hệ thống đề xuất **sử dụng phương pháp học má**y kết hợp để **phát hiện các cuộc tấn công** trong mạng. Trong bảng 1, cung cấp một bản tóm tắt về các cuộc tấn công và tính chất của chúng.



**3. Proposed Work (important as hell)**

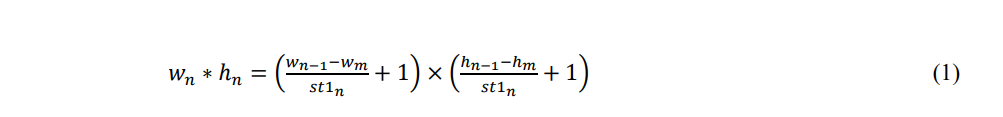
* Bài báo **đề xuất một hệ thống phát hiện xâm nhập (IDS)** mới cho mạng IoT **dựa trên mô hình mạng nơ-ron tích chập (*CNN*)** và mô hình mạng nơ-ron tái phát sinh dài ngắn hạn (***LSTM***).
  + Điều này nhằm **giải quyết các vấn đề của các IDS hiện có**, chúng thường chỉ tập trung vào phát hiện duy nhất một loại tấn công, hoặc đối **với phát hiện nhiều loại tấn công thì tính toán rất tốn kém**.
  + Mô hình **CNN** [10] là một phương pháp **deep learning** phổ biến trong **nhận dạng ảnh**, đặc biệt là **phân loại dữ liệu**, trong khi mô hình **LSTM** [11] là một mô hình mạng nơ-ron tái phát sinh **dùng để phân tích dữ liệu chuỗi** và có khả năng giữ trạng thái trước đó để áp dụng cho dữ liệu hiện tại, từ đó giúp cải thiện hiệu suất phát hiện của hệ thống
* Mô hình được đề xuất bao gồm bốn giai đoạn như sau: **thu thập dữ liệu**, **tiền xử lý dữ liệu**, **trainning the network** và **identify attack**.
  + Trong đó, việc **thu thập dữ liệu quan trọng để tạo dữ liệu đầu vào cho mô hình**.
  + Dữ liệu là **system logs** và **features** của nó và sau đó được **tiền xử lý** để **loại bỏ các nhiễu** không cần thiết.
  + **Dữ liệu đã được tiền xử lý** được **đưa vào mô hình đào tạo** với các thông số được định nghĩa như lớp tích chập, kích thước của cửa sổ trượt, trọng số liên kết neuron và đầu ra.
  + Cuối cùng, trong **giai đoạn xác định tấn công**, **dữ liệu đã được đào tạo** và **dữ liệu thực tế được xử lý cùng nhau để tính toán các trọng số và xác định các cuộc tấn công.**

**===============**

* **Mô hình mạng nơ-ron** đề xuất bao gồm **một lớp đầu** **vào** có tập **ma trận** kích thước **m0 x n0** và lớp **đầu ra** có **tập neuron cho mỗi nhãn**. [14]
  + **Các lớp ẩn** [15] được sử dụng để **tổng quát hóa (trích xuất đặc trưng)** các **tính năng** bao gồm nhiều **convolutional matrixes [12]** và **filter matrices [13]**.
  + Một **tập các tham số** **(𝑠, 𝑤𝑛 ∗ ℎ𝑛, 𝑠𝑡1 )** and **(𝑠, 𝑤𝑚 ∗ ℎ𝑚, 𝑠𝑡2 )** được sử dụng để tính tích của tích chập và tính năng bộ lọc, bao gồm **độ sâu trọng số ma trận chia sẻ (s) [18]** và **các bước trượt** (sliding steps) được biểu diễn là **st1**, **st2**.
  + **Kích thước cửa sổ trượt** (**The sliding window size**) [19] là **wn x hn**

**kích thước cửa sổ bộ lọc [20]** được định nghĩa là **wm x hm** .

* + **Lớp ẩn** là lớp cuối cùng **được sử dụng để kết nối lớp đầu ra** **và thu được kết quả phân loại.**
  + Các kích thước ma trận kết quả được tính dựa trên kích thước cửa sổ trượt [19] ( **sliding window size** ) và **tích chập [17] của matrix size** , tương tự với **filter size**.



Trong đoạn văn này, kích thước của **resulting matrix [21]** được đại diện bởi **𝑤𝑛 ∗ ℎ𝑛** và **𝑤𝑛−1 ∗ ℎ𝑛−1**, kích thước của **sliding window** được biểu diễn bởi **𝑤𝑚 ∗ ℎ𝑚** và **sliding step size [22]** được cho là **𝑠𝑡1𝑛**.

Quá trình này **sử dụng Long short term memory (LSTM)** **để** **học nội dung trên toàn mạng**.

LSTM giúp **trích xuất các đặc trưng quan trọng từ các nút** và **giúp xác định các nút độc hại** và **sự tấn công của nó**.

LSTM được **sử dụng cùng với mạng neural tích chập** để **giảm độ phức tạp của hệ thống**.

LSTM là một mô hình mạng nơ-ron dựa trên RNN được sử dụng để đưa ra đầu ra dựa trên các time stamps cho đầu vào. Tuy nhiên, **việc sử dụng LSTM đơn lẻ** **không** thể **hiệu quả phát hiện xâm nhập trong mạng** vì **LSTM** **có vấn đề về gradient vanishing**, **không thể học thông tin trong thời gian dài**.

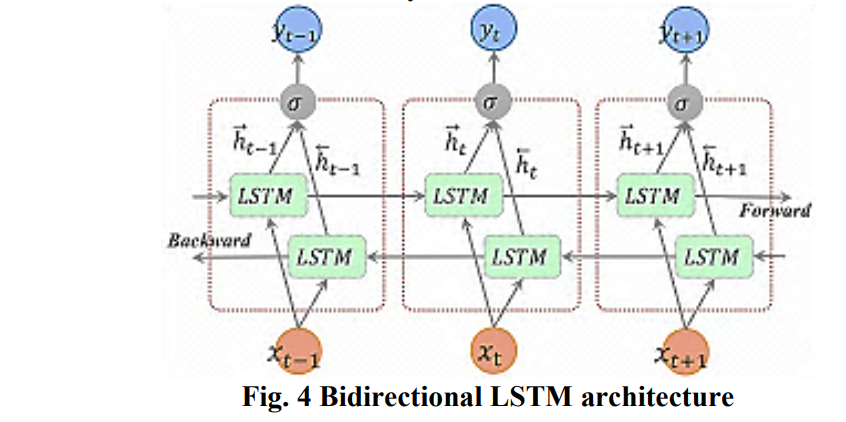
Tuy nhiên, **với thời gian ngắn** thì **hiệu suất của LSTM tốt hơn** và giúp **giảm độ phức tạp của hệ thống.**

Do đó, LSTM **được sử dụng cùng với mạng nơ-ron tích chập** trong hệ thống phát hiện xâm nhập đề xuất.

============

Mô hình Bidirectional LSTM **được mô tả trong hình 4**, trong đó một **chuỗi đầu vào** được **chuyển tiếp đến các lớp ẩn** của mạng và **tạo ra đầu ra** trong lớp đầu ra tương ứng.

Sau đó, các mô hình **LSTM hai chiều** **được phát triển để xử lý chuỗi đầu vào** theo **hướng thuận** và **ngược** bằng **hai lớp ẩn**. Những tính năng này liên quan đến truyền dữ liệu trong mạng và thu được dữ liệu cần thiết



**Trong giai đoạn huấn luyện**, các **đặc trưng đầu vào** được **đánh nhãn** và mỗi phần tử được **gán trong ma trận đặc trưng dựa trên các neuron đầu vào**.

**Sau khi xử lý đầu vào**, mạng tính toán **trọng số** **liên kết neural** và **ma trận trọng số** cho lớp tích chập.

**Hệ thống** có **x lớp tích chập** và **các lớp vector được biểu diễn dưới dạng f**, các **neuron đầu ra** được biểu diễn dưới dạng **y** và hàm trọng số được thu được là :



Wy : tổng trọng số của mạng

Xn \* fn : kích thước window filter

Mn  : số lượng filter ở lớp tích chập

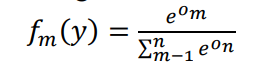
* **Trong quá trình huấn luyện** , mỗi đầu vào và **trọng số** [25] của nó tạo ra **labels** dựa trên **hàm mất mát [24]** .
  + Hàm mất mát thu được dựa trên **compactness** và **descriptiveness [26]** của dữ liệu xác thực và đào tạo mô hình và nó được đưa ra dưới dạng



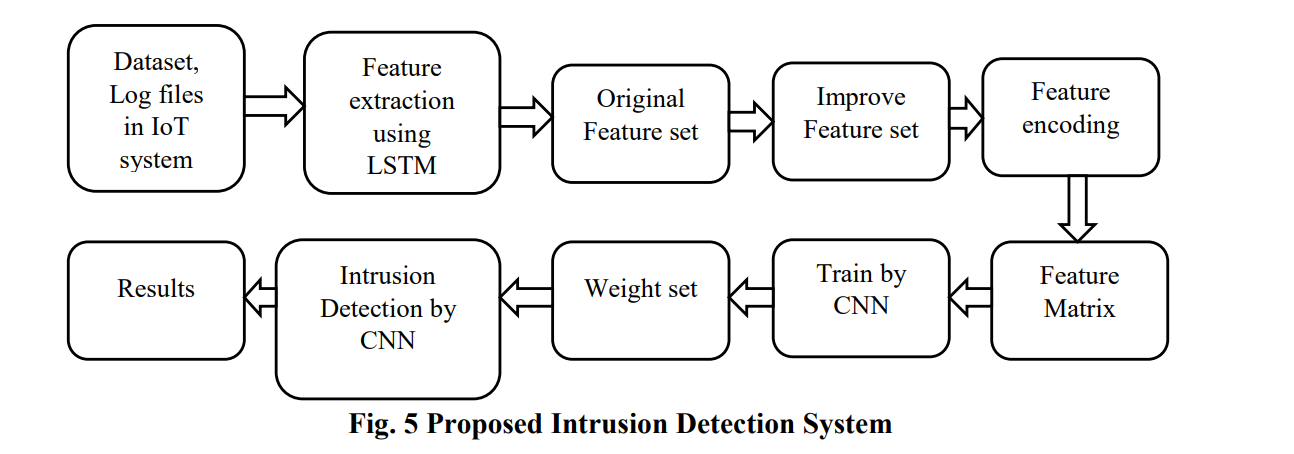
𝜑 : hệ số tỷ lệ được sử dụng để đánh giá mức độ ưu tiên của nút

𝐿𝑐 , 𝐿D : đại diện cho **compactness** **loss** và **descriptiveness** **loss**

* **Trong giai đoạn phát hiện**, các **tính năng** **được trích xuất từ tập dữ liệu** và sau đó **được chuyển đổi thành feature matrix [27]**.
  + Nó giúp **lấy tập trọng số từ quá trình huấn luyện** và **output layers defines neural label [28]**. Sử dụng chức năng kích hoạt, các kết quả đầu ra được quan sát và nó được đưa ra dưới dạng
  + Sử dụng **activation function** [29], các **kết quả đầu ra** được quan sát và nó được đưa ra dưới dạng :



* **Activation function** lấy **đầu ra xác suất của mỗi neuron** và **trọng số** của các **fully connected layer weigh**. [30]
  + Thông thường, **giá trị của hàm kích hoạt** [29] nằm trong **khoảng từ 0 đến 1** và giá trị tối đa đại diện cho nhãn đầu ra. **Hình 5** miêu tả quy trình tổng thể của hệ thống phát hiện xâm nhập IoT được đề xuất



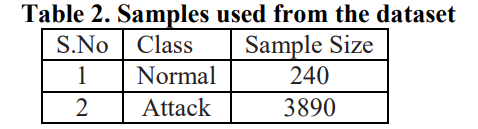
* **Trong giai đoạn huấn luyện**, **dataset** và các **logs** của nó **được thu thập** **để trích xuất các đặc trưng cần thiết**, và **so sánh với bộ đặc trưng ban đầu** để **cải thiện các đặc trưng của dữ liệu** hiện có **bằng cách sử dụng LSTM**.
  + Sau khi xác định **label function [31]** , **dữ liệu** được lựa chọn và **huấn luyện bằng CNN**.
  + Dựa trên **weight function [32]** được thu được **sau quá trình huấn luyện CNN**, sự xâm nhập **được phân loại thành các kết quả**

**4. Result and discussion**

- Đoạn văn miêu tả một cuộc thí nghiệm **so sánh** hệ thống **phát hiện xâm nhập đề xuất** với hệ thống phát hiện xâm nhập dựa trên mạng nơ-ron hồi quy (**RNN**).

Tập dữ liệu **UNSW NB15** được sử dụng với tỷ lệ xác thực **70%** cho việc huấn luyện và **30%** cho việc kiểm tra.

Hệ thống đề xuất trích xuất đặc trưng từ tập dữ liệu và xác định các điều kiện tấn công và bình thường. Mô hình được thực nghiệm bằng **Tensorflow** được cài đặt trên bộ vi xử lý **Intel i5 2.4GHz** với **8GB RAM**. Bảng 2 mô tả chi tiết về các mẫu được sử dụng cho thí nghiệm

****

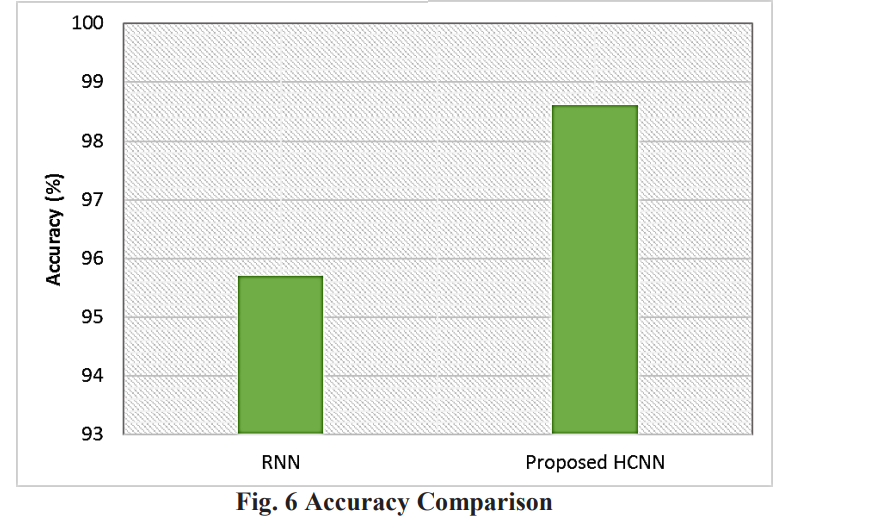
* Bài văn này trình bày **các chỉ số đánh giá được sử dụng để kiểm tra hiệu suất** của **hệ thống** phát hiện xâm nhập đề xuất, bao gồm **true positive**, **false positive**, **accuracy**, **precision**, **recall**, **f-score** và **error function**.
  + Bảng 2 mô tả **giá trị trung bình** của **mô hình đề xuất** và mô hình **RNN** dựa trên các hệ thống phát hiện xâm nhập.
  + Từ giá trị này, quan sát được rằng mô hình đề xuất đạt được **hiệu suất** phát hiện **tốt hơn** so với mô hình **RNN**.
  + Một số tham số như **recall** và **precision** tương tự như **RNN**, tuy nhiên tỷ lệ **true positive** và **false positive** của mô hình đề xuất **tốt hơn rất nhiều** so với RNN.

Table

Description automatically generated

Bài viết **so sánh độ chính xác phát hiện của mô hình đề xuất với mô hình RNN** và được **miêu tả trong** **hình 6.**

Từ hình, quan sát được rằng mô hình **Hybrid convolutional neural network** đề xuất đạt được **hiệu suất tốt hơn với độ chính xác 98**%, **cao hơn 3%** so với mô hình **RNN** truyền thống.

****

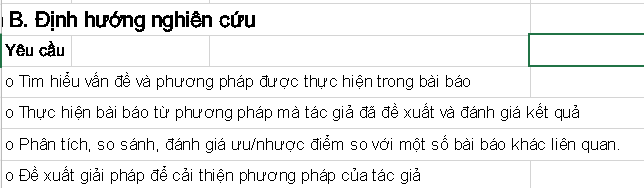
**5. Conclusion**

Bài viết **giới thiệu** về **hệ thống phát hiện xâm nhập**, một đơn vị xử lý không thể thiếu trong các mạng không dây hiện nay do thiếu an ninh và số lượng kẻ xâm nhập tăng lên.

**IoT** là một mạng không đồng nhất mà **đối mặt** nghiêm trọng với các **mối đe dọa** an ninh tương tự như các mạng không dây, và việc **phát triển một hệ thống phát hiện xâm nhập** **là cần thiết** để tránh giảm hiệu suất trong các mạng IoT.

Công trình nghiên cứu đề xuất **phân tích các loại tấn công khác nhau** trong IoT và **đề xuất một mô-đun mạng nơ-ron tích chập kết hợp với quá trình bộ nhớ ngắn hạn**.

Mô hình đề xuất được **thực nghiệm và so sánh** với m**ạng nơ-ron hồi quy truyền thống** và đạt được độ chính xác phát hiện tốt hơn là **98%**, điều này làm cho ứng dụng phù hợp với các môi trường IoT khác nhau.



-Đang giải quyết vấn đề gì

-PP chung

-Thực nghiệm ntn

-4,5 slide nội dung

**1.Vấn đề và phương pháp**

- Trong bài báo "Hybrid Intrusion Detection System for Internet of Things (IoT)", tác giả **thực hiện phương pháp phát triển một hệ thống phát hiện xâm nhập dựa trên mạng nơ-ron tích chập (CNN)** kết hợp **với mạng nơ-ron dài ngắn hạn** (hybrid convolutional neural network with long short term memory). **LSTM**

Phương pháp này giúp **trích xuất đặc trưng và phân loại dữ liệu**, nhằm giúp **xác định các trạng thái tấn công** và bình thường

Tác giả cũng sử dụng bộ dữ **liệu UNSW-NB15** để **thực hiện thí nghiệm** và **so sánh hiệu quả của phương pháp đề xuất** với **phương pháp phát hiện xâm nhập dựa trên mạng nơ-ron hồi quy (**recurrent neural network-based intrusion detection system**). RNN-based IDS**

* Kết quả cho thấy phương pháp đề xuất đạt hiệu quả phát hiện tốt hơn so với phương pháp truyền thống.

**CNN** được sử dụng để **trích xuất đặc trưng** từ các gói tin dữ liệu trong mạng

**LSTM** được sử dụng để **giải quyết vấn đề** của **time series [33]** và hỗ trợ **việc học tập từ các trạng thái trước đó**. Sự kết hợp của hai loại mạng nơ-ron này trong mô hình phát hiện xâm nhập đề xuất có thể giúp cải thiện độ chính xác của hệ thống phát hiện xâm nhập và giảm số lượng false positive (cảnh báo sai).

**Vd** : CNN được sử dụng để trích xuất các đặc trưng tần số của âm thanh từ mỗi khung thời gian, sau đó RNN được sử dụng để học các mối quan hệ giữa các khung thời gian liên tiếp trong dữ liệu.

**3.Phân tích , so sánh, đánh giá ưu nhược điểm với bài báo khác**

Chọn hai bài báo trong phần này để so sánh với bài báo "Hybrid Intrusion Detection System for Internet of Things (IoT)":

1."Intrusion Detection for IoT Networks Using Deep Learning" của nhóm tác giả Hasan Ali Khattak, Ali Kashif Bashir và Muhammad Husnain Malik.

2."Machine Learning Based Intrusion Detection System for Internet of Things in Healthcare" của nhóm tác giả Chhagan Lal, Arshdeep Kaur và Vishal Bhatnagar.

**So sánh về phương pháp**:

* Bài báo "Intrusion Detection for IoT Networks Using Deep Learning" sử dụng các kiến trúc mạng nơ-ron tích chập (CNN) và mạng nơ-ron hồi quy (RNN) để phát hiện xâm nhập trong mạng IoT.
  + Trong khi đó, bài báo "Hybrid Intrusion Detection System for Internet of Things (IoT)" sử dụng mô hình mạng nơ-ron tích chập hồi quy (CRNN) để phát hiện xâm nhập trong mạng IoT.

(

* Với **CNN**, nó có **khả năng trích xuất các đặc trưng** tần số và không gian của dữ liệu, đặc biệt là trong ảnh và âm thanh
  + Tuy nhiên, CNN **không** đáp ứng được cho các tác vụ **xử lý dữ liệu thời gian** vì nó không có khả năng xem xét quá khứ hoặc tương lai của dữ liệu.
* Với **RNN**, nó có **khả năng xử lý dữ liệu** chuỗi và bối cảnh dữ liệu, do đó có thể được sử dụng để xử lý dữ liệu thời gian hoặc ngôn ngữ tự nhiên.
  + Tuy nhiên, vấn đề của RNN là khi **đầu vào quá dài** thì các thông tin cần thiết có thể bị mất đi hoặc **không hiệu quả trong quá trình huấn luyện**.
* **CRNN** kết hợp sức mạnh của CNN và RNN để giải quyết những vấn đề này.
  + Có thể xử lý dữ liệu chuỗi dài và giữ nguyên thông tin quan trọng của dữ liệu trong quá trình huấn luyện.

)

* Bài báo "Machine Learning Based Intrusion Detection System for Internet of Things in Healthcare" sử dụng một hệ thống phát hiện xâm nhập dựa trên các thuật toán học máy như Random Forest, Naïve Bayes và Support Vector Machine (SVM) để phát hiện xâm nhập trong mạng IoT trong lĩnh vực chăm sóc sức khỏe.

**So sánh về kết quả thực nghiệm:**

* Bài báo "**Intrusion Detection for IoT Networks Using Deep Learning**" đạt được độ chính xác **96,5%** khi sử dụng mô hình CNN và RNN. Trong khi đó, bài báo "**Hybrid Intrusion Detection System for Internet of Things (IoT)"** đạt được độ chính xác **98%** khi sử dụng mô hình CRNN.
* Bài báo "**Machine Learning Based Intrusion Detection System for Internet of Things in Healthcare**" đạt được độ chính xác **89,3%** khi sử dụng các thuật toán học máy như **Random Forest, Naïve Bayes và SVM**.

**Ưu điểm của bài báo "Hybrid Intrusion Detection System for Internet of Things (IoT)":**

* Sử dụng mô hình **CRNN** mới để phát hiện xâm nhập trong mạng IoT, **kết hợp giữa tính linh hoạt của CNN** và **khả năng học tập dài hạn của RNN**, giúp **tăng độ chính xác của hệ thống phát hiện xâm nhập.**
* Sử dụng tập dữ liệu UNSW-NB15 để thực nghiệm, đây là tập dữ liệu phổ biến

**Đề xuất cải thiện phương pháp của tác giả**

* Ngoài ra, để cải thiện phương pháp đề xuất trong bài báo "Hybrid Intrusion Detection System for Internet of Things (IoT)", ta có thể **sử dụng các kỹ thuật tăng cường dữ liệu (data augmentation) [34]** để **tăng độ đa dạng của dữ liệu đầu vào**. [35]
  + Việc này có thể giúp cho mô hình đào tạo được **hiệu quả hơn** và **giảm thiểu** hiện tượng **overfitting**.
* Sử dụng kiến trúc mạng nơ-ron phức tạp hơn như **Transformer** hay một số kiến trúc mạng nơ-ron sử dụng attention mechanism để giải quyết bài toán phát hiện xâm nhập.
* Tối ưu hóa **siêu tham số [36]** của mô hình để giảm thiểu **overfitting** và **tăng độ chính xác của mô hình.**
* Sử dụng **một số phương pháp tiền xử lý dữ liệu [37]** mới để tăng tính phân loại của mô hình.
* Áp dụng một số **kỹ thuật ensemble [38]** để **kết hợp** các **mô hình khác nhau** và **giảm thiểu bias** và **variance [38]** của mô hình.
* Tuy nhiên, việc cải tiến phương pháp cần phải được thực hiện dựa trên mục tiêu và bản chất của bài toán cũng như đặc điểm của tập dữ liệu.