مجلة جامعة طرطوس للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (١٠) العدد (٢) ٢٠٢٦
Tartous University Journal for Research and Scientific Studies - engineering Sciences Series Vol. (10) No. (2) 2026

تحسين عملية الاسترداد من اختراق الشبكة في SDN باستخدام التعلم المعزز الاستباقي

م. نيرمين عقول *

(تاريخ الإيداع ٢٠٢٥/١١/٢٣ . قُبِل للنشر في ٢٠٢٦/١/٢٠)

□ ملخّص □

مع التطور السريع للشبكات المعرفة بالبرمجيات (SDN(Software-Defined Networks، تبرز الحاجة إلى استراتيجيات فعالة لاستعادة التدفق بعد التطفل السيبراني لضمان استمرارية الخدمة وتقليل تأثير الهجمات على أداء الشبكة. تهدف هذه الدراسة إلى تحسين عملية استعادة التدفق بعد الاختراقات الأمنية في SDN باستخدام نهج جديد قائم على التعلم المعزز، والذي يُعرف بـ RLNIR (Reinforcement Learning-based Network Intrusion Recovery). يعتمد النهج المقترح على تحليل ديناميكي لحركة المرور داخل الشبكة وتحديد المسارات البديلة بشكل فوري عند اكتشاف التطفل، مما يضمن تقليل زمن استعادة التدفق وتحسين تخصيص عرض النطاق الترددي. تم تقييم أداء النموذج المقترح عبر مقارنته بعدة استراتيجيات تقليدية ومتقدمة، بما في ذلك نهج التعلم الآلي في النموذج الأساسي (ML approach in Baseline)، نهج المسار الأسرع للاسترداد(FRT (Fast Rerouting، النهج الاستباقي (Proactive Approach)، ونهج MLBNIR(Machine Learning Based Network Intrusion Recovery) المدعوم بالتعلم الآلي. أظهرت النتائج أن RLNIR تفوق على جميع المناهج السابقة، حيث تمكن من تقليل زمن استعادة التدفق إلى ٨ مللي ثانية فقط مقارنة بـ ٥٥ مللي ثانية في النهج التقليدي، كما حقق تخصيصاً محسناً لعرض النطاق الترددي بقيمة ٩٠٠ ميجابت في الثانية، وهو أعلى مستوى بين جميع الأساليب المقارنة. تؤكد هذه النتائج أهمية استخدام التعلم المعزز كنهج مستقبلي لاستعادة التطفل في SDN، حيث يسمح بتحسين المرونة الأمنية، الاستجابة السريعة للهجمات، وتحقيق توزيع ديناميكي للموارد الشبكية بفعالية.

الكلمات المفتاحية: الشبكات المعرفة بالبرمجيات، هجمات أمنية، استرداد الشبكة، تعلم الآلة، التعلم المعزز.

* قائم بالأعمال في كلية هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس – سورية.

Improving the recovery process from network attacks in SDN using proactive reinforcement learning

Eng. Nermin Akoul *

(Received 23/11/2025 . Accepted 20/1/2026)

□ **ABSTRACT** □

With the rapid evolution of Software-Defined Networks (SDN), the need for efficient network intrusion recovery strategies has become critical to ensuring service continuity and minimizing the impact of cyberattacks on network performance. This study proposes an innovative Reinforcement Learning-Based Network Intrusion Recovery (RLNIR) approach, which dynamically analyzes network traffic and instantly determines optimal alternative paths upon intrusion detection. This ensures minimal recovery time and improved bandwidth allocation for affected network flows. The proposed model was evaluated against several conventional and advanced strategies, including the ML approach in baseline, Fast Rerouting (FRT), proactive approach, and MLBNIR (Machine Learning-Based Network Intrusion Recovery). Experimental results demonstrate that RLNIR outperforms all existing approaches, reducing the network recovery time to 8 milliseconds, compared to 55 milliseconds in traditional methods. Additionally, it achieves an enhanced bandwidth allocation of 900 Mbps, the highest among all compared techniques. These findings highlight the potential of reinforcement learning as a future-proof solution for network intrusion recovery in SDN, enabling faster response to cyber threats, improved security flexibility, and dynamic resource allocation.

Keywords: Software-defined networks, network attacks, network recovery, machine learning, reinforcement learning.

*Engineer at the Faculty of Information and Communication Technology- Tartous University, Syria.

1- المقدمة

شهد العالم تحولاً رقمياً غير مسبوق في السنوات الأخيرة، حيث تعتمد المؤسسات على نطاق واسع على الأنظمة الرقمية المعتمدة على الإنترنت لتقديم تجارب سلسة لعملائها [1,2]. ومع تسارع الرقمنة في جميع القطاعات، أصبح تأمين العمليات الرقمية أولوية قصوى تتطلب استخدام أحدث تقنيات الأمن السيبراني. يُعد الأمن الشبكي من الركائز الأساسية في هذا التحول، حيث تتزايد الحاجة إلى أنظمة قادرة على اكتشاف التهديدات والتعامل مع التطفل بفعالية، مما يُمكّن من الحد من الأضرار المحتملة وتعزيز استدامة البنية التحتية الرقمية [3,4].

في بيئات الشبكات الحديثة، تمثل استعادة تدفق البيانات بعد التطفل تحدياً بالغ الأهمية، خاصة في سياق الشبكات المعرفة بالبرمجيات Software-Defined Networks (SDN)، التي تعتمد على نموذج تحكم مركزي لفصل طبقة التحكم عن طبقة البيانات [5,6]. هذا الفصل يوفر مرونة في إدارة الشبكة، لكنه في الوقت نفسه يزيد من المخاطر الأمنية، حيث يمكن لاستهداف وحدة التحكم أن يؤدي إلى تعطيل الشبكة بأكملها. لهذا السبب، يُعد تطوير أنظمة استرداد أو تعافي متقدمة أمراً ضرورياً لضمان استمرارية الشبكة ومرونتها في مواجهة الهجمات المتطورة [7].

تعتمد الأساليب التقليدية للاسترداد على استراتيجيات استرجاع استباقية أو تفاعلية. في النهج التفاعلي، تُطلب مسارات بديلة عند حدوث خلل في التدفق، إلا أن ذلك يستدعي تواصلاً إضافياً مع وحدة التحكم، مما قد يؤدي إلى تأخيرات كبيرة تؤثر على أداء الشبكة [8,9]. في المقابل، يعتمد النهج الاستباقي على حساب المسارات البديلة مسبقاً وتثبيتها في مفاتيح الشبكة، مما يقلل من زمن الاستجابة في حالة حدوث اختراق أو فشل في المسار الرئيسي. ومع ذلك، فإن هذه الأساليب غالباً ما تستند إلى حالات شبكة ثابتة ولا تأخذ في الاعتبار التغيرات الديناميكية في أنماط حركة البيانات، مما قد يؤدي إلى ازدحام المسارات الاحتياطية وفقدان كفاءتها [10].

مع التحديات المتزايدة في مجال أمن SDN، تبرز الحاجة إلى حلول ذكية تعتمد على تقنيات التعلم الآلي وتحليل البيانات لاكتشاف الاختراقات واسترداد الشبكة بكفاءة [13-11]. تعتبر أنظمة التعلم الآلي أداة قوية في التعرف على أنماط الهجمات وتصنيفها بناءً على ميزات تدفق البيانات. ومع ذلك، فإن فعالية هذه الأنظمة تعتمد على جودة اختيار الميزات والبيانات المستخدمة في التدريب، فضلاً عن مدى قدرتها على التعامل مع التهديدات المتغيرة والمتطورة. علاوة على ذلك، تحتاج استراتيجيات استرداد الشبكة إلى تجاوز الأساليب التقليدية من خلال التكيف المستمر مع حركة البيانات المتغيرة وتحليل التأثير المستقبلي لمسارات الشبكة المحتملة [10-7].

بناءً على ذلك، تهدف هذه الدراسة إلى تطوير آلية مبتكرة تعتمد على التعلم المعزز لتعزيز استعادة تدفق البيانات في SDN بعد التطفل. يتيح هذا النهج للنظام تعلم أفضل استراتيجيات الاسترداد بناءً على بيانات واقعية، مما يمكنه من التكيف مع الظروف المتغيرة للشبكة واتخاذ قرارات استباقية فعالة دون الحاجة إلى تدخل بشري مستمر. تسهم هذه المقاربة في تقليل زمن الاسترداد وتحسين كفاءة الشبكة من خلال اختيار مسارات احتياطية ديناميكية تتماشى مع تطورات حركة البيانات، مما يعزز من مرونة الشبكة وقدرتها على التصدي للهجمات السيبرانية المتزايدة.

يهدف هذا البحث إلى تطوير نموذج استرداد شبكي يعتمد على التعلم المعزز، قادر على التكيف مع التغيرات الديناميكية في حركة البيانات واختيار أفضل المسارات البديلة تلقائياً بعد حدوث التطفل. يسعى البحث

إلى تحليل الأنماط المختلفة للهجمات الإلكترونية التي تستهدفSDN ، وتطوير آلية استرداد تفاعلية تتعلم من البيانات الفعلية لتحسين كفاءة الاستجابة. بالإضافة إلى ذلك، يهدف البحث إلى تقليل زمن استعادة التدفق وتقليل احتمالية حدوث ازدحام في المسارات الاحتياطية، مما يؤدي إلى تحسين الأداء العام للشبكة. كما يتضمن البحث تقييم أداء النموذج المقترح مقارنةً بالأساليب التقليدية، وقياس مدى تحسينه لاستقرار وأمان الشبكة في ظل سيناريوهات مختلفة من التهديدات السيبرانية.

1-1 أهمية البحث

تبرز أهمية هذا البحث في تقديم حلول مبتكرة لتحسين آليات استرداد الشبكة بعد التطفل في بيئةSDN ، مما يسهم في تعزيز أمن الشبكات الحديثة وضمان استمراريتها في مواجهة التهديدات السيبرانية. يؤدي تحسين استراتيجيات الاسترداد إلى تقليل زمن استجابة الشبكة بعد الاختراق، مما يضمن استقرار تدفق البيانات حتى في الظروف الحرجة. كما أن استخدام تقنيات التعلم الآلي والتعلم المعزز في تحسين استراتيجيات الاسترداد يمثل نقلة نوعية في كيفية تعامل الشبكات مع التهديدات الأمنية. يساعد البحث أيضاً في دعم المؤسسات والشركات التي تعتمد على SDN في تقليل المخاطر التشغيلية الناجمة عن الهجمات الإلكترونية، وبالتالي تعزيز موثوقية الشبكة وتحقيق أقصى قدر من الكفاءة في إدارتها.

2-1 مشكلة البحث

مع التطور السريع في تقنيات الشبكات وزيادة الاعتماد على البنية التحتية الرقمية، أصبحت الشبكات المعرفة بالبرمجيات (SDN) جزءاً أساسياً من بنية الاتصالات الحديثة نظراً لقدرتها على تحسين إدارة الشبكة وتعزيز مرونتها. ومع ذلك، فإن هذه الشبكات معرضة لتهديدات أمنية متزايدة، حيث يمكن أن تؤدي الهجمات الإلكترونية إلى تعطيل التحكم المركزي للشبكة، مما يؤثر على تدفق البيانات واستقرار العمليات. تكمن المشكلة الرئيسية في كيفية استعادة الشبكة بعد حدوث اختراق أو فشل في التدفق دون التأثير على كفاءة التشغيل. تعتمد الأساليب التقليدية على استراتيجيات استرداد ثابتة أو تفاعلية، والتي قد لا تكون فعالة في مواجهة الهجمات الديناميكية أو التغيرات في أنماط حركة البيانات. وبما أن استراتيجيات الاسترداد التقليدية قد تؤدي إلى تأخيرات كبيرة أو فقدان في البيانات، فإن الحاجة إلى حلول ذكية أكثر تكيفاً مع بيئة الشبكة المتغيرة أصبحت ملحة.

2- الدراسات المرجعية

في عام ٢٠٠٦، اقترح Eswaradass وآخرون استخدام الشبكات العصبونية الاصطناعية وتقنيات التعلم المعزز للتنبؤ بعرض النطاق الترددي للروابط الشبكية. وقد كانت هذه واحدة من المحاولات الأولى للاستفادة من التعلم الآلي في تحسين استراتيجيات تخصيص المسارات في الشبكات [14]. لاحقاً، قدم Brun وآخرون (٢٠١٦) دراسة حول استخدام تقنيات الذكاء الاصطناعي لاختيار المسارات المثلى في الشبكات المتداخلة عبر القارات، ما شكل خطوة مهمة نحو توظيف التعلم الآلي في تحسين كفاءة الشبكات [15].

في عام ٢٠١٤، قام Akyildiz وآخرون بتصنيف استراتيجيات استرداد الفشل في الشبكات المعرفة بالبرمجيات (SDN) إلى نوعين رئيسيين: الاستباقي والتفاعلي. وقد سلط البحث الضوء على ضرورة فهم العلاقة بين طبقتي التحكم والبيانات داخل SDN لتحسين كفاءة الاسترداد بعد الفشل [16].

في عام ٢٠١٧، قدم Muthumanikandan & Valliyammai تقنية إعادة التوجيه السريع (FRT) القائمة على أقصر مسار، والتي كانت قادرة على تحقيق استرداد في غضون ٣٠ مللي ثانية، مع استهلاك يصل إلى ٧٠% من عرض النطاق الترددي، مما أظهر مزايا استخدام الذكاء الاصطناعي في تخطيط المسارات

الاحتياطية[17]. في عام ٢٠١٨، قام Xu وآخرون بتوضيح الأثر المتزايد لتوسع الشبكات، مشيرة إلى أن بيانات الإنترنت الخاصة بالمركبات قد تصل إلى ٣٠٠,٠٠٠ إكسابايت، مما يزيد من خطر الأعطال على نطاق واسع [18].

في عام ٢٠١٩، قدم Truong-Huu وآخرون طريقة مبتكرة تستفيد من إمكانيات SDN في إدارة حركة البيانات ومراقبة الشبكة، حيث يقوم المتحكم بمراقبة التغيرات المفاجئة في أنماط المرور وتنبيه المسارات الاحتياطية تلقائياً عند رصد أي تهديد محتمل. ومع ذلك، لاحظت الدراسة أن زيادة عدد العقد في الشبكة يزيد من تعقيد عملية إدارة المسارات الاحتياطية [19]. وفي الوقت نفسه، بحث Srinivasan وآخرون في تطبيقات التعلم الآلي لاكتشاف فشل الروابط استناداً إلى خصائص حركة البيانات، لكن الدراسة لم تقدم حلولاً ديناميكية لاسترداد الشبكة بمجرد حدوث الفشل [20].

في عام ٢٠٢٠، طوّر Ali وآخرون نموذجاً لاسترداد تدفق البيانات في SDN يتميز بزمن استرداد منخفض يصل إلى ٢٢ مللي ثانية، لكنه واجه مشكلة انخفاض كفاءة استغلال المسارات الاحتياطية، التي لم تتجاوز ١٧%. وقد أدى ذلك إلى توجيه المزيد من الأبحاث نحو إيجاد حلول أكثر تكيفاً مع التغيرات الديناميكية في الشبكة [21].

في عام ٢٠٢٣، قدم Hammad وآخرون نهجاً حديثاً تحت مسمى Machine Learning-based Network Intrusion Recovery (MLBNIR)، والذي يعتمد على التعلم الآلي في استرداد تدفق البيانات بعد الهجمات السيبرانية في SDN. حيث تم تدريب نموذج التعلم الآلي باستخدام بيانات حقيقية من الشبكة لفهم الديناميكيات المرورية والتنبؤ بالمسارات الأكثر كفاءة في حالات الطوارئ. وأظهرت نتائج الدراسة أن هذا النهج أدى إلى تقليل وقت الاسترداد بنسبة تصل إلى ٩٠%، وزيادة استهلاك عرض النطاق الترددي بنسبة ٥٧% مقارنةً بالأساليب التقليدية، مما يعكس فعالية الذكاء الاصطناعي في تعزيز استقرار الشبكات المعاصرة وحمايتها من الهجمات المتطورة[22] .

3- الإطار النظري:

أصبحت الشبكات المعرفة بالبرمجيات (SDN – software-defined networking) واحدة من أكثر الحلول ابتكاراً في إدارة الشبكات الحديثة، حيث تعتمد على الفصل بين طبقة التحكم وطبقة البيانات، مما يتيح إمكانية إدارة تدفق البيانات بمرونة عالية عبر وحدة تحكم مركزية. ومع ذلك، فإن هذه الميزة تجعل SDN عرضة لتهديدات سيبرانية متزايدة، حيث يمكن أن تؤدي الهجمات على وحدة التحكم إلى تعطيل الشبكة بأكملها، مما يبرز الحاجة إلى استراتيجيات فعالة لاسترداد الشبكة بعد التطفل [23].

1. مفهوم استرداد الشبكة بعد التطفل فيSDN

يشير استرداد الشبكة بعد التطفل إلى مجموعة من الإجراءات والتقنيات المستخدمة لاستعادة أداء الشبكة الطبيعي بعد اكتشاف حدوث هجوم أو اختراق أمني. يهدف هذا المفهوم إلى تقليل تأثير التهديدات الأمنية، سواءً كانت هجمات رفض الخدمة الموزعة أو التطفل غير المشروع، أو محاولات تعديل تدفق البيانات بطرق خبيثة. في بيئة SDN ، يزداد تعقيد هذه العملية نظراً لاعتماد البنية الشبكية على وحدة تحكم مركزية، والتي إذا تعرضت لهجوم، فإن الشبكة بأكملها قد تتوقف عن العمل.

٢. أنواع استراتيجيات الاسترداد في SDN

يمكن تصنيف تقنيات استرداد الشبكة بعد التطفل إلى فئتين رئيسيتين [24]:

a. **الاسترداد التفاعلي (Reactive Recovery) :**

يعتمد على استراتيجيات الاستجابة بعد وقوع الهجوم، حيث يتم الكشف عن التطفل أولاً من خلال أنظمة كشف التسلل(IDS – Intrusion Detection Systems) ، ثم تُنفّذ إجراءات الاسترداد مثل إعادة توجيه البيانات عبر مسارات بديلة أو إعادة ضبط إعدادات الشبكة. على الرغم من فعالية هذا النهج، إلا أنه قد يؤدي إلى تأخيرات كبيرة نظراً لاعتماده على استجابة الشبكة بعد وقوع الاختراق.

b. **الاسترداد الاستباقي (Proactive Recovery)**

يهدف إلى تقليل التأثير السلبي للهجمات عبر التحضير المسبق لمسارات بديلة لحركة البيانات، بحيث يتم إعادة التوجيه تلقائياً عند اكتشاف التطفل، دون الحاجة إلى تدخل بشري أو انتظار الاستجابة من وحدة التحكم. تتمثل إحدى الطرق الشائعة في تثبيت قواعد إعادة التوجيه مسبقاً في أجهزة التحويل الشبكي لتقليل زمن الاسترداد وزيادة استقرار الشبكة.

٣. أهمية الذكاء الاصطناعي والتعلم الآلي في استرداد الشبكة

مع تزايد حجم البيانات وتعقيد الهجمات السيبرانية، أصبحت تقنيات التعلم الآلي (ML – Machine Learning)والتعلم المعزز (Reinforcement Learning – RL) من الأدوات الفعالة في تعزيز استراتيجيات استرداد الشبكة. يمكن لهذه التقنيات أن تتعلم أنماط حركة البيانات الطبيعية داخل الشبكة، وتكتشف أي انحرافات تدل على وجود اختراق محتمل، ثم تحدد أنسب مسار بديل لإعادة التوجيه [25].

يتيح التعلم الآلي إمكانية التنبؤ بمسارات الفشل المحتملة، وبالتالي تنفيذ عمليات استرداد أسرع وأكثر كفاءة. على سبيل المثال، يمكن لنماذج الشبكات العصبية الاصطناعية (ANN – Artificial Neural Networks)تحليل تدفقات البيانات المرورية في الوقت الحقيقي وتقديم قرارات ديناميكية حول إعادة التوجيه.

أما التعلم المعزز، فهو يتيح لنظام SDN تطوير سياسات استرداد قائمة على التجربة، حيث يقوم الذكاء الاصطناعي بتجربة استراتيجيات استرداد مختلفة، ومن ثم تحسينها بمرور الوقت بناءً على أدائها في تقليل وقت الاسترداد وزيادة كفاءة استخدام الموارد الشبكية.

4. تحديات استرداد الشبكة في SDN

رغم التقدم في استراتيجيات استرداد الشبكة بعد التطفل، لا تزال هناك العديد من التحديات التي تواجه بيئةSDN ، ومنها [26]:

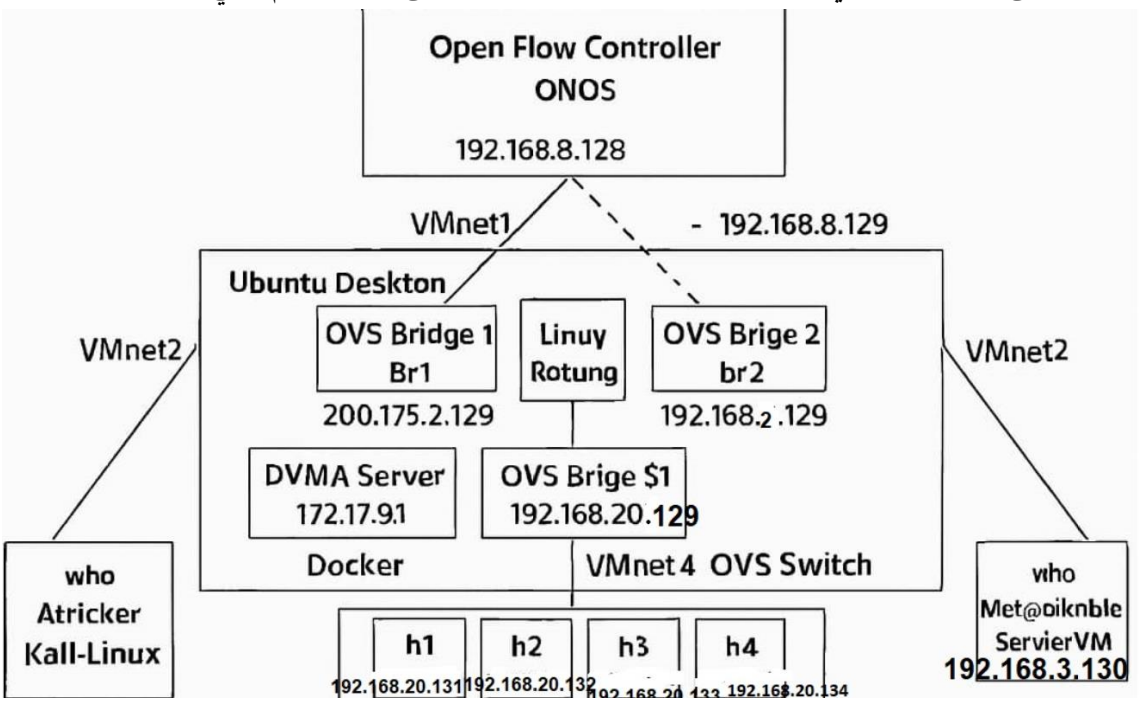
- **الزمن المستغرق في الاسترداد :**بعض الطرق التفاعلية تتطلب وقتاً طويلاً لتحليل الهجوم واتخاذ الإجراء المناسب، مما قد يؤدي إلى تعطيل الشبكة لفترة طويلة.
- **ازدحام المسارات الاحتياطية :**في بعض الحالات، قد تكون المسارات البديلة التي تم إعدادها مسبقاً مزدحمة بسبب مشاركة العديد من التدفقات الشبكية، مما قد يؤثر على جودة الخدمة.
- **تعقيد الإدارة المركزية :**اعتماد SDN على وحدة تحكم مركزية يجعلها نقطة ضعف رئيسية، حيث قد يؤدي أي هجوم على وحدة التحكم إلى انهيار الشبكة بالكامل.

• **صعوبة كشف التهديدات المتقدمة :**تتطور أساليب الهجوم باستمرار، مما يجعل من الصعب اكتشاف بعض أنواع الهجمات المتخفية التي تحاكي حركة البيانات العادية داخل الشبكة.

4- الأدوات والمنهجية

4-1 بناء قاعدة البيانات

في ظل التوسع المتزايد في استخدام الشبكات المعرفة بالبرمجيات، برزت الحاجة إلى تطوير قواعد بيانات متخصصة تسهم في تحسين آليات الكشف عن الهجمات واسترداد الشبكة بعد الاختراق. تم إنشاء قاعدة بيانات InSDN [27,28] (البنية موضحة في الشكل ١) التي توفر بيانات تجريبية متكاملة لاختبار خوارزميات الذكاء الاصطناعي في مجال أمن الشبكات. تهدف هذه القاعدة إلى تمكين الباحثين من تحليل أنماط حركة البيانات، والتعرف على الهجمات، وتطوير استراتيجيات وقائية واستردادية تعتمد على تقنيات التعلم الآلي.



الشكل (١): طوبولوجيا الشبكة حيث تم تكوين جسور Open vSwitch (OVS) المستقلة الفردية باستخدام التبديل L3 وتوصيلها بوحدة تحكم SDN [27].

يعتمد جمع البيانات في InSDN على محاكاة شبكة SDN فعلية، حيث تم تصميم بيئة افتراضية تتكون من عدة وحدات، تشمل وحدة تحكم مركزية، خادم Metasploitable2 المخصص للاختراقات التجريبية، شبكة بيانات SDN تعمل عبر Mininet، بالإضافة إلى خادم Kali Linux المصمم لتنفيذ وتحليل الهجمات السيبرانية. تتميز هذه البنية بقدرتها على توليد سيناريوهات هجومية متعددة تعكس التهديدات الحقيقية التي قد تتعرض لها الشبكات المعرفة بالبرمجيات.

تقوم قاعدة البيانات بتوثيق تدفقات البيانات في كلا الاتجاهين، حيث يتم تحليل أنماط الاتصال لتحديد السلوك الطبيعي والمشبوه لحركة البيانات. وقد تم بناء هذه القاعدة باستخدام أداة CICFlowMeter (أداة مفتوحة المصدر لتوليد وتحليل تدفق حركة المرور في الشبكة، تم تطويرها بواسطة المعهد الكندي للأمن السيبراني في جامعة New Brunswick)، التي توفر بيانات تفصيلية حول خصائص التدفقات الشبكية من خلال أكثر من

80 ميزة تحليلية، تشمل مدة الاتصال، عدد الحزم والبايتات المرسلة والمستقبلة، توقيتات التدفقات، ومجموعة من الأعلام التي تشير إلى طبيعة الاتصال بين الأجهزة.

تم تصنيف الميزات المسجلة ضمن قاعدة البيانات إلى مجموعات رئيسية تغطي مختلف جوانب تدفق البيانات داخل الشبكة، مثل محددات الشبكة التي تتضمن عناوين IP والمنافذ والبروتوكولات، وتحليل الحزم الذي يوفر معلومات عن كمية البيانات المتبادلة، إضافة إلى التحليل الزمني الذي يدرس الفواصل الزمنية بين إرسال واستقبال الحزم، إلى جانب السمات الأمنية التي تكشف عن الأنماط غير الطبيعية في الاتصال. هذه البيانات توفر أرضية صلبة لتطوير خوارزميات قادرة على التعرف على الهجمات بفعالية، وتعزيز أداء الشبكة عبر حلول استباقية لتفادي الانقطاعات. تشمل الميزات ثماني مجموعات رئيسية [27]:

١. **معرفات الشبكة:** تشمل معلومات مثل عنوان **IP**، رقم المنفذ، ونوع البروتوكول، والتي تساعد في تحديد مصدر ومقصد التدفق.
٢. **خصائص الحزم:** تحتوي على بيانات عن عدد الحزم المرسلة والمستقبلة في التدفق الأمامي والخلفي.
٣. **خصائص البيانات (Bytes)** : توفر معلومات عن حجم البيانات المنقولة في الاتجاهين.
٤. **الزمن الفاصل بين الحزم (Interarrival Time)** : تسجل الفترات الزمنية بين الحزم في الاتجاهين.
٥. **مؤقتات التدفق (Flow Timers):** تحدد مدة التدفق، وما إذا كان نشطاً أم غير نشط.
٦. **سمات الأعلام (Flags)** : تشمل بيانات عن أعلام الشبكة مثل **SYN, RST, PUSH**، وغيرها.
٧. **خصائص تدفق البيانات (Flow Descriptors)** : تتضمن معلومات عن العدد الإجمالي للحزم والبايتات في التدفق ثنائي الاتجاه.
٨. **خصائص التدفقات الفرعية (Subflow Descriptors)** : تعرض بيانات عن عدد الحزم والبايتات في التدفقات الفرعية.

تشمل قاعدة بيانات InSDN بيانات ضخمة عن سيناريوهات مختلفة للهجمات، بما في ذلك هجمات حجب الخدمة ، الاختراقات بواسطة Botnets، الهجمات على المواقع الإلكترونية، تقنيات Brute Force، والبرمجيات الخبيثة. وقد بلغ إجمالي السجلات في القاعدة 343,939 إدخالاً، منها 68,424 تدفقات طبيعية، و 275,515 تمثل تدفقات هجومية، مما يجعل هذه القاعدة من أهم الأدوات المستخدمة في تطوير واختبار آليات الحماية الذكية في بيئات SDN توفر هذه البيانات المتنوعة أساساً قوياً للبحث في مجال أمن الشبكات، حيث يمكن استخدامها لتطوير تقنيات كشف التطفل وتحليل استراتيجيات استرداد الشبكة بعد التعرض للهجمات.

٤-٢- المنهجية المقترحة بالاعتماد على التعلم المعزز

تعتمد هذه الدراسة على تطوير نهج متكامل لاسترداد الشبكة بعد التطفل في بيئة الشبكات المعرفة بالبرمجيات باستخدام التعلم المعزز (الموضحة كخطوات في الخورازمية ١) ، بهدف تقليل زمن الاسترداد وتحسين كفاءة استغلال الموارد الشبكية. يعتمد النظام المقترح على وكيل ذكاء اصطناعي يقوم بتحليل تدفق البيانات واكتشاف التغيرات غير الطبيعية التي قد تشير إلى اختراق أمني أو انقطاع في الخدمة. يعمل هذا الوكيل بالاستفادة من خوارزمية التعلم المعزز العميق(Deep Q-Network - DQN) ، حيث يتعلم تدريجياً اتخاذ قرارات مثلى لإعادة توجيه التدفقات المتأثرة.

في بيئة التعلم المعزز، يتم تمثيل مشكلة استرداد الشبكة كعملية ماركوف لاتخاذ القرار (Markov Decision Process - MDP)، والتي تُعرّف بمجموعات (S, A, P, R, γ) حيث [29]:

- **S:** مجموعة الحالات الممكنة للشبكة، مثل حالات التدفق الطبيعي، الاختراق، ومسارات الاسترداد.
- **A:** مجموعة الإجراءات التي يمكن للنظام اتخاذها، مثل إعادة توجيه التدفق، إعادة ضبط القواعد، أو زيادة سعة المسار الاحتياطي.
- **P (s′| s,a):** احتمالية الانتقال من الحالة **s** إلى الحالة **s′** عند تنفيذ الإجراء **a**.
- **R(s,a):** دالة المكافأة التي تحدد مدى جودة الإجراء **a** في الحالة **s** ، بحيث تشجع الإجراءات التي تقلل زمن الاسترداد وتحافظ على جودة الخدمة.
- **γ:** معامل الخصم الذي يحدد مدى تأثير المكافآت المستقبلية على القرارات الحالية، حيث **0≤γ≤10**

يتمثل هدف نموذج DQN في تعلم دالة القيمة المثلى Q∗(s,a) التي تُحدّد قيمة كل إجراء a في الحالة s ، باستخدام تحديثات معادلة Bellman كما يلي:

$$\gamma \max a' Q(s', a')\, R(s, a) = max\, Q(s', a') \qquad (1)$$

يتم تحديث القيم بناءً على التغذية الراجعة التي يحصل عليها النموذج أثناء التدريب. ولتحسين استجابة النظام وتقليل زمن الاسترداد، تم تصميم دالة مكافأة مخصصة بحيث تعاقب أي تأخير في اتخاذ القرار وتعزز المسارات التي تقلل من فقدان الحزم (Packet Loss) وتجنب الازدحام. يمكن تعريف المكافأة بالاستفادة من [30] كما يلي:

$$R(s, a) = -\alpha \cdot T_{rec} + \beta \cdot (1 - P_{loss}) + \lambda \cdot U_{bw} \qquad (2)$$

الخوارزمية (١)

المدخلات:

- **as: عنوان IP المصدر للتدفق الشبكي المتعرّض للهجوم.**
- **ad: عنوان IP الوجهة للتدفق الشبكي المتعرّض للهجوم.**
- **قاعدة بيانات تدفقات الشبكة (InSDN) التي تحتوي على سجلات الاتصال (عناوينIP ، الأحجام، الفواصل الزمنية، إلخ).**

المخرجات:

- **أفضل مسار احتياطي لكل تدفق شبكي متأثر بالهجوم.**

المرحلة الأولى: تحديد التدفقات الشبكية السليمة

١. **تهيئة وكيل التعلم المعزز.**
٢. **تعريف فضاء الحالات S، فضاء الإجراءات A، ودالة المكافأة R.**
٣. **تدريب وكيل التعلم المعزز باستخدام بيانات الشبكة السابقة.**
٤. **إنشاء المصفوفة []BNF لتخزين التدفقات الشبكية السليمة.**
٥. **لكل سجل في قاعدة بيانات InSDN:**
   - **s= عنوان IP المصدر، و d=عنوان IP الوجهة.**
   - **إذا تحقق الشرط (s==as) و(d==ad):**

     أضف الزوج (s,d) إلى []BNF:
٦. **نهاية الحلقة.**

المرحلة الثانية: استخراج الخصائص وبناء شجرة القرار

١. **حساب الخصائص التالية لكل تدفق:**
   - **FIM: الفاصل الزمني بين التدفقات.**
   - **SFP: حجم الحزم ضمن التدفق.**
   - **SFB: الحجم الكلي للتدفق.**
٢. **بدء بناء شجرة القرار BuildTree(N) باستخدام مجموعة البيانات N.**
٣. **إذا كانت البيانات في العقدة N تنتمي إلى فئة واحدة فقط،**

فارجع النتيجة (تصنيف نهائي).

٤. **وإلا:**
   - **اختر نسبة x% من الخصائص القابلة للتقسيم في N .**
   - **حدّد الخاصية F التي تمتلك أعلى قيمة معلومات مكتسبة.(Information Gain)**
   - **أنشئ عقدًا فرعية $N_1, N_2, \dots, N_f$ بناءً على القيم المحتملة للخاصية F .**
   - **لكل عقدة فرعية N_i:**
     - **عيّن القيم المطابقة E_i في X_i.**
     - **استدعِ الإجراء BuildTree(N_i) لبناء الفروع.**
٥. **نهاية الحلقة ونهاية عملية بناء الشجرة.**

المرحلة الثالثة: اختيار المسار الاحتياطي باستخدام التعلم المعزز

١. **حساب قيمة الازدحام في التدفق (FC) لكل تدفق في BNF[i] باستخدام مقاييس الاستخدام، التأخير، وطول الطابور.**
٢. **تحديد التدفق BNF[i] الذي يمتلك أقل قيمة FC.**
٣. **تفعيل المسار الاحتياطي الأفضل المرتبط بهذا التدفق.**

حيث:

- T_{rec}: زمن الاسترداد بعد التطفل.
- P_{loss}: معدل فقدان الحزم.
- U_{bw}: نسبة استخدام عرض النطاق الترددي للمسار البديل.
- **α, β, λ**: معاملات وزن تُحدد أهمية كل عامل في تحسين الأداء.

يتم تنفيذ التدريب باستخدام محاكاة لشبكة SDN داخل بيئة مثل Mininet، حيث يتم تشغيل الهجمات المختلفة واختبار استجابات النموذج. أثناء عملية التعلم، يقوم النموذج بتجربة مختلف المسارات والقرارات، ومن ثم تحسين أدائه باستخدام خوارزمية التدرج العشوائي (Stochastic Gradient Descent - SGD) لتحديث معلمات Q.

بعد مرحلة التدريب، يتم اختبار النموذج عبر مقارنة أدائه مع الأساليب التقليدية للاسترداد، حيث يتم قياس عوامل مثل متوسط زمن الاسترداد Trec، معدل فقدان الحزم Ploss، وكفاءة استخدام الموارد Ubw.

3-4 مقياس الأداء

يتم حساب قيمة الازدحام في التدفق (Flow Congestion - FC) لكل زوج من المصدر والوجهة باستخدام عدة معايير مرتبطة بحركة البيانات داخل الشبكة، وفقاً للمعادلة (3) [22]:

$$FC = \frac{Mean\ IAT}{Max\ IAT} + \frac{Pkts/sec}{Total\ Pkts/sec} + \frac{Bytes/sec}{Total\ Bytes/sec} \qquad (3)$$

تجمع هذه المعادلة بين نسب كل من الفاصل الزمني بين الحزم، معدل إرسال الحزم، واستهلاك عرض النطاق الترددي لكل تدفق، مما يتيح تقييم شامل لمستوى الازدحام في الشبكة المعرفة بالبرمجيات. بناءً على هذه القيم، يتم تحديد أفضل مسار احتياطي، حيث يتم اختيار المسار ذو أقل قيمة FC لضمان استعادة سريعة وكفوءة للتدفق بعد حدوث التطفل. يشير انخفاض قيمة FC إلى كفاءة أعلى في اختيار المسار الاحتياطي، حيث يتم تقييم المسارات بناءً على تدفق البيانات ومدى تأثيرها على الازدحام.

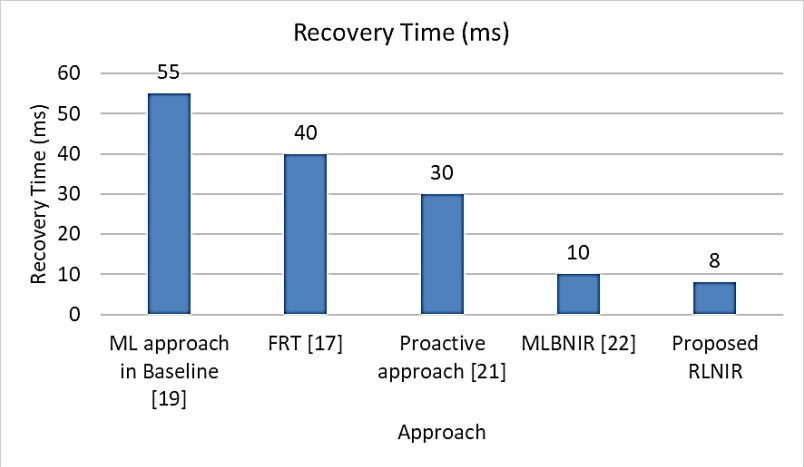
يعتمد الحساب على عدة عوامل رئيسية، أولها الفاصل الزمني بين الحزم (Flow Interarrival Time - IAT)، حيث يتم قياس متوسط الزمن الفاصل بين إرسال حزم البيانات المتتابعة. كلما قل هذا الزمن، دلّ ذلك على كفاءة أعلى في تدفق البيانات، مما يعكس جودة أفضل للمسار المستخدم.

العامل الثاني هو حجم التدفق المقاس بعدد الحزم (Flow Packets per Second - Pkts/s) ، حيث يتم قياس عدد الحزم المرسلة في الثانية. يؤدي ارتفاع عدد الحزم في التدفق إلى استهلاك أكبر لعرض النطاق الترددي، مما يزيد من احتمالية ازدحام المسار الاحتياطي ويقلل من كفاءته.

أما العامل الثالث، فهو حجم التدفق المقاس بعدد البايتات (Flow Bytes per Second - Bytes/s) ، والذي يعكس معدل البيانات المنقولة عبر الشبكة. كلما زاد عدد البايتات في التدفق، ارتفع مستوى استهلاك عرض النطاق الترددي، مما يؤدي إلى زيادة احتمالية حدوث ازدحام في المسارات الاحتياطية.

5-النتائج والمناقشة

كما هو موضح في الشكل ٢، يبين تحليل الأداء أن نهج ML approach in Baseline المقدم في [١٩] سجل زمن استرداد بلغ ٥٥ مللي ثانية، مما يعكس ضعف قدرته على التكيف مع الاختراقات بشكل فعال. من ناحية أخرى، أظهر نهج FRT المُعتمد في [١٧] تحسناً ملحوظاً، حيث تمكن من تقليل زمن الاسترداد إلى ٤٠ مللي ثانية، وهو ما يدل على قدرة هذا النهج على استعادة التدفق بسرعة أكبر عبر إعادة توجيه البيانات إلى مسارات بديلة أكثر كفاءة. أما النهج الاستباقي Proactive Approach، الذي تمت دراسته في [٢١]، فقد استطاع تحسين الاستجابة بشكل أكبر، حيث خفض زمن الاستعادة إلى ٣٠ مللي ثانية، مما يشير إلى أهمية تنفيذ سياسات استباقية لاكتشاف الأعطال قبل وقوعها. عند النظر إلى النهج المدعوم بالتعلم الآلي MLBNIR في [٢٢]،



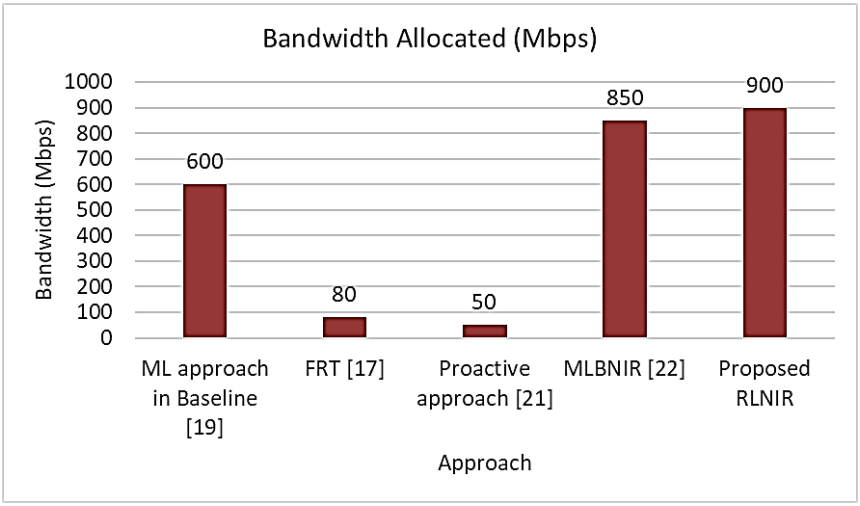
الشكل (٢): مقارنة زمن استعادة التدفق بين النهج المقترح RLNIR والطرق السابقة.

نجد أنه تمكن من تحسين أداء الاستجابة بشكل أكبر من النهج الاستباقي، حيث خفض زمن الاسترداد إلى ١٠ مللي ثانية فقط، مما يدل على تفوق الأساليب القائمة على التعلم الآلي في تحديد المسارات البديلة بذكاء عند حدوث الهجمات السيبرانية. أخيراً، تفوق النهج المقترح RLNIR في هذه الدراسة على جميع الأساليب السابقة، حيث حقق زمن استرداد قدره ٨ مللي ثانية فقط، وهو ما يمثل الحد الأدنى المسجل، مما يشير إلى الاستجابة فائقة السرعة التي يوفرها التعلم المعزز مقارنةً بالنهج التقليدية.

يوضح الشكل ٣ مدى اختلاف كفاءة تخصيص عرض النطاق الترددي بين النهج المختلفة. يظهر أن النهج القائم على التعلم الآلي في النموذج الأساسي (ML approach in Baseline) [١٩] قد خصص ٦٠٠ ميجابت في الثانية، وهو مستوى عالٍ نسبياً ولكنه لا يزال غير كافٍ عند مقارنة بالكفاءة الأعلى التي توفرها التقنيات الأكثر تطوراً. أما نهج FRT [١٧]، فقد حقق ٨٠ ميجابت في الثانية، وهو أقل بكثير من المستوى المطلوب، مما يشير إلى ضعف أداء هذا الأسلوب في إدارة تدفق البيانات عند حدوث اختراقات. بالنسبة للنهج الاستباقي Proactive Approach [٢١]، فقد أظهر قدرة أقل على تخصيص عرض النطاق الترددي بكفاءة، حيث بلغ ٥٠ ميجابت في الثانية فقط، مما يعكس قصور النهج الاستباقي في التعامل مع سيناريوهات الاستخدام الكثيف للبيانات. في المقابل، أظهر نهج MLBNIR [٢٢] المدعوم بالتعلم الآلي تحسناً واضحاً، حيث تمكن

من تحقيق ٨٥٠ ميجابت في الثانية، وهو ما يعكس قدرة أفضل على إدارة الموارد الشبكية وتوجيه حركة المرور بذكاء لتفادي فقدان البيانات.

أخيراً، تفوق النهج المقترح RLNIR على جميع المناهج السابقة، حيث تمكن من تحقيق أقصى تخصيص لعرض النطاق الترددي بقيمة ٩٠٠ ميجابت في الثانية، مما يشير إلى تحسن ملحوظ في كفاءة إدارة تدفق البيانات باستخدام التعلم المعزز. تُظهر هذه النتيجة الدور الحيوي للتعلم المعزز في تحسين استراتيجيات استرداد التطفل، حيث يستطيع النظام التكيف ديناميكياً مع متغيرات الشبكة وتحقيق أفضل استغلال لمواردها في الوقت الفعلي.



الشكل (٣): مقارنة كفاءة تخصيص عرض النطاق الترددي بين النهج المقترح RLNIR والطرق السابقة.

تُثبت هذه النتائج أن استخدام التعلم المعزز يمكن أن يحسن كفاءة تخصيص الموارد الشبكية بشكل كبير، مما يضمن استمرارية تدفق البيانات وتقليل تأثير الاختراقات على جودة الخدمة. إن النهج المقترح يعزز المرونة الشبكية والاستجابة السريعة للتغيرات الأمنية، مما يجعله خياراً مثالياً للشبكات المستقبلية التي تعتمد على الذكاء الاصطناعي في إدارة الأمن والتحكم في حركة البيانات.

5-الاستنتاجات و التوصيات

توصلت هذه الدراسة إلى أن استخدام التعلم المعزز في استعادة التطفل داخل الشبكات المعرفة بالبرمجيات يحقق تحسينات كبيرة في زمن الاستجابة وتخصيص عرض النطاق الترددي مقارنةً بالمناهج التقليدية والمعززة بالتعلم الآلي فقط. أظهرت النتائج أن النهج المقترح RLNIR قد تفوق على جميع الأساليب السابقة من خلال تقليل زمن استعادة التدفق إلى ٨ مللي ثانية (الشكل ٢)، مما يعكس استجابة شبه فورية للأعطال السيبرانية. كما أن النموذج المقترح نجح في تحقيق أقصى تخصيص لعرض النطاق الترددي بقيمة ٩٠٠ ميجابت في الثانية (الشكل ٣)، مما يضمن استمرارية تدفق البيانات بكفاءة عالية دون التأثير على جودة الخدمة.

يؤكد التحليل أن الاستراتيجيات التقليدية مثل المسار الأسرع للاسترداد (FRT) والنهج الاستباقي (Proactive Approach) تعاني من قيود في التكيف مع التغيرات الديناميكية داخل الشبكة، حيث أظهرت هذه الأساليب أزمنة استرداد أطول وتخصيصاً أقل لعرض النطاق الترددي، مما قد يؤدي إلى تدهور أداء الشبكة في سيناريوهات التهديدات المتكررة. وعلى الرغم من أن النهج المدعوم بالتعلم الآلي MLBNIR قد حقق تحسناً

ملحوظاً مقارنةً بالمناهج التقليدية، إلا أنه لا يزال يعاني من تأخير نسبي بالمقارنة مع النهج المقترح القائم على التعلم المعزز.

تُظهر هذه الدراسة أن استخدام التعلم المعزز في SDN لا يقتصر فقط على تحسين زمن الاستجابة، بل يسهم أيضاً في تحسين إدارة الموارد الشبكية بشكل ديناميكي، مما يجعل الشبكات أكثر كفاءة وأماناً وقدرة على التكيف مع التهديدات في الزمن الحقيقي.

بناءً على النتائج التي تم الحصول عليها، يمكن تقديم التوصيات التالية لتحسين استراتيجيات استعادة التطفل داخل الشبكات المعرفة بالبرمجيات:

١. يجب تطوير أنظمة أكثر تقدماً تعتمد على التعلم المعزز العميق (Deep Reinforcement Learning) لضمان استجابة أسرع وتحسين دقة التنبؤ بالمسارات البديلة لاستعادة التدفق.

٢. يمكن تعزيز أمن الشبكات عبر دمج أنظمة كشف التطفل (IDS) المدعومة بالتعلم العميق مع النهج المقترح، مما يسمح بتحديد أنواع الهجمات وتحليلها بذكاء قبل اتخاذ قرار استعادة التدفق.

٣. ينبغي تطوير آليات ذكية لتخصيص عرض النطاق الترددي بناءً على أنماط حركة المرور في الزمن الحقيقي، مما يضمن تحقيق توازن بين الأداء الأمني وكفاءة إدارة الشبكة.

6-المراجع

1. Paul, J., Ueno, A., Dennis, C., Alamanos, E., Curtis, L., Foroudi, P., ... & Wirtz, J. (2024). Digital transformation: A multidisciplinary perspective and future research agenda. *International Journal of Consumer Studies*, *48*(2), e13015.
2. Yaqub, M. Z., & Alsabban, A. (2023). Industry-4.0-Enabled digital transformation: prospects, instruments, challenges, and implications for business strategies. *Sustainability*, *15*(11), 8553.
3. Schmitt, M. (2023). Securing the Digital World: Protecting smart infrastructures and digital industries with Artificial Intelligence (AI)-enabled malware and intrusion detection. *Journal of Industrial Information Integration*, *36*, 100520.
4. Govea, J., Gaibor-Naranjo, W., & Villegas-Ch, W. (2024). Transforming Cybersecurity into Critical Energy Infrastructure: A Study on the Effectiveness of Artificial Intelligence. *Systems*, *12*(5), 165.
5. Etxezarreta, X., Garitano, I., Iturbe, M., & Zurutuza, U. (2023). Software-Defined Networking approaches for intrusion response in Industrial Control Systems: A survey. *International Journal of Critical Infrastructure Protection*, *42*, 100615.
6. Benzekki, K., El Fergougui, A., & Elbelrhiti Elalaoui, A. (2016). Software-defined networking (SDN): a survey. *Security and communication networks*, *9*(18), 5803-5833.

7. Abdi, A. H., Audah, L., Salh, A., Alhartomi, M. A., Rasheed, H., Ahmed, S., & Tahir, A. (2024). Security Control and Data Planes of SDN: A Comprehensive Review of Traditional, AI and MTD Approaches to Security Solutions. *IEEE Access*.

8. Ali, J., Lee, G. M., Roh, B. H., Ryu, D. K., & Park, G. (2020). Software-defined networking approaches for link failure recovery: A survey. *Sustainability*, *12*(10), 4255.

9. Khan, N., Bin Salleh, R., Koubaa, A., Khan, Z., Khan, M. K., & Ali, I. (2023). Data plane failure and its recovery techniques in SDN: A systematic literature review. *Journal of King Saud University-Computer and Information Sciences*, *35*(3), 176-201.

10. Fawcett, L., Scott-Hayward, S., Broadbent, M., Wright, A., & Race, N. (2018). Tennison: A distributed SDN framework for scalable network security. *IEEE Journal on Selected Areas in Communications*, *36*(12), 2805-2818.

11. Jada, I., & Mayayise, T. O. (2024). The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. *Data and Information Management*, *8*(2), 100063.

12. Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, *97*, 101804.

13. Apruzzese, G., Laskov, P., Montes de Oca, E., Mallouli, W., Brdalo Rapa, L., Grammatopoulos, A. V., & Di Franco, F. (2023). The role of machine learning in cybersecurity. *Digital Threats: Research and Practice*, *4*(1), 1-38.

14. Eswaradass, A., Sun, X. H., & Wu, M. (2006, May). Network bandwidth predictor (nbp): A system for online network performance forecasting. In *Sixth IEEE International Symposium on Cluster Computing and the Grid (CCGRID'06)* (Vol. 1, pp. 4-pp).

15. Brun, O., Wang, L., & Gelenbe, E. (2016). Big data for autonomic intercontinental overlays. *IEEE Journal on Selected Areas in Communications*, *34*(3), 575–583.

16. Akyildiz, I. F., Lee, A., Wang, P., Luo, M., & Chou, W. (2014). A roadmap for traffic engineering in SDN-OpenFlow networks. *Computer Networks*, *71*, 1–30.

17. Muthumanikandan, V., & Valliyammai, C. (2017). Link failure recovery using shortest path fast rerouting technique inSDN. *Wireless Personal Communications*, *97*(2), 2475–2495.

18. Xu, W., Zhou, H., Cheng, N., Lyu, F., Shi, W., Chen, J., & Shen, X. (2018). Internet of vehicles in big data era. *IEEE/CAA Journal of Automatica Sinica*, *5*(1), 19–35.

19. Truong-Huu, T., Prathap, P., Mohan, P. M., & Gurusamy, M. (2019, May) Fast and adaptive failure recovery using machine learning in software defined networks. In *2019 IEEE international conference on communications workshops (ICC workshops)*.

20. Srinivasan, S. M., Truong-Huu, T., & Gurusamy, M. (2019). Machine learning-based link fault identification and localization in complex networks. *IEEE Internet of Things Journal*, *6*(4), 6556–6566.

21. Ali, J., Min Lee, G., Hee Roh, B., Ryu, D. K., & Park, G. (2020). Software-defined networking approaches for link failure recovery: A survey. *Sustainability*, *12*(10), 4255.

22. Hammad, M., Hewahi, N., & Elmedany, W. (2023). Enhancing Network Intrusion Recovery in SDN with machine learning: an innovative approach. *Arab Journal of Basic and Applied Sciences*, *30*(1), 561-572.

23. Urrea, C., & Benítez, D. (2021). Software-defined networking solutions, architecture and controllers for the industrial internet of things: A review. *Sensors*, *21*(19), 6585.

24. Sousa, P., Bessani, A. N., Correia, M., Neves, N. F., & Verissimo, P. (2009). Highly available intrusion-tolerant services with proactive-reactive recovery. *IEEE Transactions on Parallel and Distributed Systems*, *21*(4), 452-465.

25. Feltus, C. (2020). Reinforcement Learning's Contribution to the Cyber Security of Distributed Systems: Systematization of Knowledge. *International Journal of Distributed Artificial Intelligence (IJDAI)*, *12*(2), 35-55.

26. Li, Z., Hu, Y., Wu, J., & Lu, J. (2022). P4Resilience: Scalable resilience for multi-failure recovery in SDN with programmable data plane. *Computer Networks*, *208*, 108896.

27. Elsayed, M. S., Le-Khac, N. A., & Jurcut, A. D. (2020). InSDN: A novel SDN intrusion dataset. *IEEE access*, *8*, 165263-165284.

28. Kaggle. *InSDN dataset* [Dataset]. Kaggle. Retrieved [2024/10/15], from https://kaggle.com/datasets/badcodebuilder/insdn-dataset

29. Wang, X., Yang, Z., Chen, G., & Liu, Y. (2023). A reinforcement learning method of solving Markov decision processes: an adaptive exploration model based on temporal difference error. *Electronics*, *12*(19), 4176.

30. Naqvi, H. A., Hilman, M. H., & Anggorojati, B. (2023). Implementability improvement of deep reinforcement learning based congestion control in cellular network. *Computer Networks*, *233*, 109874.