

# مجلة جامعة البعث

سلسلة العلوم الهندسية الميكانيكية  
والكهربائية والمعلوماتية



مجلة علمية محكمة دورية

المجلد 44 . العدد 5

1443 هـ - 2022 م

الأستاذ الدكتور عبد الباسط الخطيب

رئيس جامعة البعث

المدير المسؤول عن المجلة

رئيس هيئة التحرير

أ. د. ناصر سعد الدين

رئيس التحرير

أ. د. درغام سلوم

مديرة مكتب مجلة جامعة البعث

بشرى مصطفى

عضو هيئة التحرير	د. محمد هلال
عضو هيئة التحرير	د. فهد شريباتي
عضو هيئة التحرير	د. معن سلامة
عضو هيئة التحرير	د. جمال العلي
عضو هيئة التحرير	د. عباد كاسوحة
عضو هيئة التحرير	د. محمود عامر
عضو هيئة التحرير	د. أحمد الحسن
عضو هيئة التحرير	د. سونيا عطية
عضو هيئة التحرير	د. ريم ديب
عضو هيئة التحرير	د. حسن مشرقي
عضو هيئة التحرير	د. هيثم حسن
عضو هيئة التحرير	د. نزار عبشي

تهدف المجلة إلى نشر البحوث العلمية الأصيلة، ويمكن للراغبين في طلبها

الاتصال بالعنوان التالي:

رئيس تحرير مجلة جامعة البعث

سورية . حمص . جامعة البعث . الإدارة المركزية . ص . ب (77)

. هاتف / فاكس : 963 31 2138071 ++

. موقع الإنترنت : [www.albaath-univ.edu.sy](http://www.albaath-univ.edu.sy)

. البريد الإلكتروني : [magazine@ albaath-univ.edu.sy](mailto:magazine@albaath-univ.edu.sy)

ISSN: 1022-467X

قيمة العدد الواحد : 100 ل.س داخل القطر العربي السوري

25 دولاراً أمريكياً خارج القطر العربي السوري

قيمة الاشتراك السنوي : 1000 ل.س للعموم

500 ل.س لأعضاء الهيئة التدريسية والطلاب

250 دولاراً أمريكياً خارج القطر العربي السوري

توجه الطلبات الخاصة بالاشتراك في المجلة إلى العنوان المبين أعلاه.

يرسل المبلغ المطلوب من خارج القطر بالدولارات الأمريكية بموجب شيكات

باسم جامعة البعث.

تضاف نسبة 50% إذا كان الاشتراك أكثر من نسخة.

## شروط النشر في مجلة جامعة البعث

### الأوراق المطلوبة:

- 2 نسخة ورقية من البحث بدون اسم الباحث / الكلية / الجامعة) + CD / word من البحث منسق حسب شروط المجلة.
- طابع بحث علمي + طابع نقابة معلمين.
- إذا كان الباحث طالب دراسات عليا:  
يجب إرفاق قرار تسجيل الدكتوراه / ماجستير + كتاب من الدكتور المشرف بموافقة على النشر في المجلة.
- إذا كان الباحث عضو هيئة تدريسية:  
يجب إرفاق قرار المجلس المختص بإنجاز البحث أو قرار قسم بالموافقة على اعتماده حسب الحال.
- إذا كان الباحث عضو هيئة تدريسية من خارج جامعة البعث :  
يجب إحضار كتاب من عمادة كليته تثبت أنه عضو بالهيئة التدريسية و على رأس عمله حتى تاريخه.
- إذا كان الباحث عضواً في الهيئة الفنية :  
يجب إرفاق كتاب يحدد فيه مكان و زمان إجراء البحث , وما يثبت صفته وأنه على رأس عمله.
- يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (العلوم الطبية والهندسية والأساسية والتطبيقية):  
عنوان البحث .. ملخص عربي و إنكليزي ( كلمات مفتاحية في نهاية الملخصين).
- 1- مقدمة
- 2- هدف البحث
- 3- مواد وطرق البحث
- 4- النتائج ومناقشتها .
- 5- الاستنتاجات والتوصيات .
- 6- المراجع.

- يتم ترتيب البحث على النحو الآتي بالنسبة لكليات ( الآداب - الاقتصاد - التربية - الحقوق - السياحة - التربية الموسيقية وجميع العلوم الإنسانية):
- عنوان البحث .. ملخص عربي و إنكليزي ( كلمات مفتاحية في نهاية الملخصين).
- 1. مقدمة.
- 2. مشكلة البحث وأهميته والجديد فيه.
- 3. أهداف البحث و أسئلته.
- 4. فرضيات البحث و حدوده.
- 5. مصطلحات البحث و تعريفاته الإجرائية.
- 6. الإطار النظري و الدراسات السابقة.
- 7. منهج البحث و إجراءاته.
- 8. عرض البحث و المناقشة والتحليل
- 9. نتائج البحث.
- 10. مقترحات البحث إن وجدت.
- 11. قائمة المصادر والمراجع.
- 7- يجب اعتماد الإعدادات الآتية أثناء طباعة البحث على الكمبيوتر:
  - أ- قياس الورق 25×17.5 B5.
  - ب- هوامش الصفحة: أعلى 2.54- أسفل 2.54 - يمين 2.5- يسار 2.5 سم
  - ت- رأس الصفحة 1.6 / تذييل الصفحة 1.8
  - ث- نوع الخط وقياسه: العنوان . Monotype Koufi قياس 20
- . كتابة النص Simplified Arabic قياس 13 عادي . العناوين الفرعية Simplified Arabic قياس 13 عريض.
- ج . يجب مراعاة أن يكون قياس الصور والجداول المدرجة في البحث لا يتعدى 12سم.
- 8- في حال عدم إجراء البحث وفقاً لما ورد أعلاه من إشارات فإن البحث سيهمل ولا يرد البحث إلى صاحبه.
- 9- تقديم أي بحث للنشر في المجلة يدل ضمناً على عدم نشره في أي مكان آخر، وفي حال قبول البحث للنشر في مجلة جامعة البعث يجب عدم نشره في أي مجلة أخرى.
- 10- الناشر غير مسؤول عن محتوى ما ينشر من مادة الموضوعات التي تنشر في المجلة

11- تكتب المراجع ضمن النص على الشكل التالي: [1] ثم رقم الصفحة ويفضل استخدام التهميش الإلكتروني المعمول به في نظام وورد WORD حيث يشير الرقم إلى رقم المرجع الوارد في قائمة المراجع.

تكتب جميع المراجع باللغة الانكليزية (الأحرف الرومانية) وفق التالي:  
آ . إذا كان المرجع أجنبياً:

الكنية بالأحرف الكبيرة . الحرف الأول من الاسم تتبعه فاصلة . سنة النشر . وتتبعها معترضة ( - ) عنوان الكتاب ويوضع تحته خط وتتبعه نقطة . دار النشر وتتبعها فاصلة . الطبعة ( ثانية . ثالثة ) . بلد النشر وتتبعها فاصلة . عدد صفحات الكتاب وتتبعها نقطة .  
وفيما يلي مثال على ذلك:

-MAVRODEANUS, R1986- Flame Spectroscopy. Willy, New York, 373p.

ب . إذا كان المرجع بحثاً منشوراً في مجلة باللغة الأجنبية:

. بعد الكنية والاسم وسنة النشر يضاف عنوان البحث وتتبعه فاصلة, اسم المجلد ويوضع تحته خط وتتبعه فاصلة . المجلد والعدد ( كتابية مختزلة ) وبعدها فاصلة . أرقام الصفحات الخاصة بالبحث ضمن المجلة.  
مثال على ذلك:

BUSSE,E 1980 Organic Brain Diseases Clinical Psychiatry News ,  
Vol. 4. 20 – 60

ج . إذا كان المرجع أو البحث منشوراً باللغة العربية فيجب تحويله إلى اللغة الإنكليزية و  
التقيد

بالبنود (أ و ب) ويكتب في نهاية المراجع العربية: ( المراجع In Arabic )

## رسوم النشر في مجلة جامعة البعث

1. دفع رسم نشر (20000) ل.س عشرون ألف ليرة سورية عن كل بحث لكل باحث يريد نشره في مجلة جامعة البعث.
2. دفع رسم نشر (50000) ل.س خمسون ألف ليرة سورية عن كل بحث للباحثين من الجامعة الخاصة والافتراضية .
3. دفع رسم نشر (200) مئتا دولار أمريكي فقط للباحثين من خارج القطر العربي السوري .
4. دفع مبلغ (3000) ل.س ثلاثة آلاف ليرة سورية رسم موافقة على النشر من كافة الباحثين.

## المحتوى

الصفحة	اسم الباحث	اسم البحث
48-11	علي ياسين د. كمال السلوم د. وسيم رمضان	تحديد التوليفة الأمثل من ضبط البارامترات الفائقة واختيار الميزات لتحسين أداء أنظمة كشف الشذوذ
88-49	د. أحمد صقر أحمد حازم ديب	تحسين أداء الشبكات المعرفة برمجياً ذات المتحكمات المتعددة من خلال خوارزمية موازنة الحمل
102-89	د. أسامة عجمي	العوامل المؤثرة على تحسين الطلي الكيميائي بالفضة
120-103	د. أسامة عجمي	دراسة تأثير العوامل المختلفة على الطلي الكيميائي للألياف الكربونية بطبقة النيكل- بور





## تحديد التوليفة الأمثل من ضبط البارامترات الفائقة واختيار الميزات لتحسين أداء أنظمة كشف الشذوذ

المهندس علي ياسين<sup>1</sup> - إشراف: الدكتور كمال السلوم<sup>2</sup> - الدكتور وسيم رمضان<sup>3</sup>

### المخلص

لازال اكتشاف الشذوذ محط اهتمام الكثير من الباحثين لما له من دور في حل مشكلات العالم الحقيقية مثل الكشف عن الاحتيال واكتشاف الحالات غير الطبيعية...الخ. يتأثر أداء خوارزميات اكتشاف الشذوذ بشكل كبير بكيفية ضبط البارامترات الفائقة من جهة، واختيار الميزات الأكثر أهمية ضمن مجموعات البيانات من جهة أخرى. على الرغم من ذلك فإن الأدبيات السابقة المتعلقة باكتشاف الشذوذ لا تتطرق في معظمها لضبط البارامترات الفائقة واختيار الميزات عند المقارنة التجريبية للخوارزميات المختلفة. حيث تقارن معظم الأوراق الأداء باستخدام القيم الافتراضية للبارامترات الفائقة وجميع الميزات ضمن مجموعات البيانات.

يهدف البحث إلى تحليل أداء مصنفات تعلم الآلة المستخدمة في كشف الشذوذ عند ضبط البارامترات الفائقة الخاصة بها واختيار الميزات الأكثر أهمية، بالإضافة إلى تحديد الاستراتيجية الأفضل لضبط البارامترات الفائقة واختيار الميزات، حيث يؤدي اختيار مجموعة مختلفة من الميزات إلى اختلاف مجموعة البارامترات الفائقة والعكس صحيح.

أكدت نتائج الدراسة أن اختيار الميزات الأكثر أهمية ضمن مجموعة البيانات متنوعاً بضبط البارامترات الفائقة يساهم في تحسين أداء خوارزميات الكشف عن الشذوذ بشكل كبير، حيث تم اختزال الزمن من 50% إلى 53% وزيادة دقة الكشف من 5% إلى 6% وتحسين كل مقاييس الأداء المستخدمة.

**الكلمات المفتاحية:** اكتشاف الشذوذ، ضبط البارامترات الفائقة، اختيار الميزات، تعلم الآلة

<sup>1</sup> طالب دكتوراه - قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث - حمص - سوريا

<sup>2</sup> أستاذ دكتور - قسم هندسة البرمجيات - كلية الهندسة المعلوماتية - جامعة البعث - حمص - سوريا

<sup>3</sup> دكتور مشارك - قسم الاقتصاد الزراعي - كلية الزراعة - جامعة البعث - حمص - سوريا

# Selecting the Optimal Combination of Hyperparameter Tuning and Feature Selection to Improve the Performance of Anomaly Detection Systems

Eng. Ali Yassin, Dr. Kamal Al-Salloum, Dr. Wassim Ramadan

## Abstract

Anomaly detection is still an important research subject for many researchers. It has a crucial role in solving real-world problems such as fraud detection...etc.

The performance of anomaly detection algorithms is greatly affected by how the hyperparameters are tuned on one hand, and the selection of the most important features within datasets on the other. However, previous literature on anomaly detection mostly does not deal with hyperparameter tuning and feature selection when comparing different algorithms using experiments. Most papers compare performance using default values for hyperparameters and all features within datasets.

This research aims to analyze the performance of machine learning classifiers used in anomaly detection when tuning their hyperparameters and selecting the most important features, in addition to defining the best strategy for hyperparameters tuning and feature selection. Because selecting a different set of features causes a different set of hyperparameters and vice versa.

The results of the study confirm that selecting the most important features within the dataset followed by tuning the hyperparameters contributes significantly in improving the performance of anomaly detection algorithms, as the time was reduced by 50% to 53%, the accuracy of detection increased by 5% to 6%, and all performance measures used were improved.

**Keywords:** Anomaly Detection, Tuning Hyperparameters, Feature selection, Machine Learning

---

## 1- مقدمة

نتيجة ازدياد كم البيانات المتوافر في كل مكان بشكل كبير في الآونة الأخيرة (من المتوقع أن يصل إنتاج البيانات إلى 175 Zettabytes بحلول عام 2025) [1] ، أصبح هناك حاجة ضرورية لتحليل مجموعات البيانات في العالم الحقيقي بكافة أشكالها، واستخراج أنماط البيانات الخفية والمعقدة، لاستخدامها في العديد من المجالات المختلفة. يتأثر تحليل البيانات في مختلف المجالات التطبيقية بالبيانات الشاذة التي يمكن أن تؤثر بشكل أو بآخر عن شيء خارج نطاق البيانات الطبيعية. وبالتالي فالتحليل السليم للبيانات للحصول على المعلومات الصحيحة والدقيقة والمعبرة عن الحالة الطبيعية يجب أن يميز بين البيانات الطبيعية وغير الطبيعية. وهذا ما يدعو للحاجة إلى الاهتمام بشكل أو بآخر بالكشف عن الشذوذ.

تهدف عملية الكشف عن الحالات الشاذة [1] إلى تحديد كل نقاط البيانات التي تسلك سلوكاً مختلفاً عن باقي نقاط المجموعة. يُمكن أن تنتج الحالات الشاذة عن خطأ في البيانات؛ ولكنها تدل أيضاً على عمليات أساسية جديدة لم تكن معروفة مسبقاً وغالباً ما تكون حرجة في مجموعة واسعة من التطبيقات. يوجد إسقاط كبير لمفهوم الشذوذ في تطبيقات العالم الحقيقي كتطبيقات اكتشاف الاحتيال المالي والاختراقات الشبكية والأمراض النادرة.

توجهت العديد من شركات القطاع المالي مؤخراً لتشغيل خدماتها عبر الإنترنت كعملية الدفع الإلكتروني، وبالتالي تتزايد عمليات الاحتيال المالي من حيث الشكل والعدد حول العالم. يُعد كل من الوصول غير المصرح به والهجمات غير المنتظمة أمثلة على التهديدات التي يجب اكتشافها باستخدام أنظمة الكشف عن الاحتيال المالي. مما يؤدي إلى خسائر مالية هائلة تجعل الاحتيال المالي مشكلة كبيرة. دفع كل ذلك كبرى الشركات ومواقع التجارة الإلكترونية إلى استخدام مفهوم تعلم الآلة لبناء أنظمة قادرة على كشف تلك الحالات الشاذة (الاحتمالية). كما أصبح هذا الأمر مهماً في الجمهورية العربية السورية حيث هناك توجه كبير لاستخدام خدمات الدفع الإلكتروني بدءاً من عام 2020 مع انطلاق منظومة الدفع الإلكتروني.

يُمكن تعريف أنظمة كشف الشذوذ على أنها أجهزة أو برامج تقوم بمراقبة البيانات المتعلقة بمجال معين لتحديد الحالات الشاذة والمختلفة عن السلوك الطبيعي باستخدام طرائق معينة مثل تعلم الآلة.

تَسْتَطِيع هذه الأنظمة والتي تعتمد على خوارزميات تعلم الآلة التعامل مع البيانات بأشكالها المختلفة كالبيانات الرقمية والنصية بنوعها الهيكلية وغير الهيكلية بالإضافة لقدرتها على استخراج أنماط البيانات المعقدة. بالمقابل يتأثر أدائها بشكل كبير بكيفية ضبط البارامترات الفائقة (Hyperparameters Tuning) من جهة، واختيار الميزات (Features Selection) الأكثر أهمية ضمن مجموعات البيانات من جهة أخرى.

حيث تُعبر البارامترات الفائقة (Hyperparameters) عن مجموعة من البارامترات الرياضية والتي يتم ضبطها بشكل مختلف عن البارامترات العادية، حيث لن يقوم النموذج بتحديثها وفقاً لاستراتيجية التحسين ولازالت الخوارزميات حالياً تحتاج إلى تدخل يدوي لضبطها في بداية التنفيذ. يوجد عدة عوامل تحفز على ضبط هذه البارامترات الفائقة بشكل أوتوماتيكي [3] ومنها:

- الأبعاد العالية للبارامترات الفائقة (High Dimensions of Hyperparameters): أصبحت نماذج تعلم الآلة بشكل عام أكثر تعقيداً مع تزايد عدد البارامترات الفائقة. مما يضطر المختصون في هذا المجال إلى تقييم النماذج مع مجموعات مختلفة من قيم البارامترات الفائقة الخاصة بنماذجهم للحصول على أفضل النتائج.
- زيادة وقت التدريب: بتزايد حجم مجموعات البيانات يصبح تدريب النماذج أكثر تكلفة من الناحية الحسابية بشكل كبير، وغالباً ما يستغرق من ساعات إلى أيام على أجهزة مختصة عالية الأداء. يعد ذلك مرهقاً بشكل خاص في سياق ضبط البارامترات الفائقة، حيث يجب تدريب نموذج جديد في كل مرة لتقييم المجموعات المختلفة من البارامترات الفائقة المرشحة.

يوجد العديد من الطرائق والأساليب المختلفة [4] لأتمتة ضبط البارامترات الفائقة (Hyperparameters Tuning) أهمها البحث الشبكي (Grid Search) وطريقة البحث العشوائي (Random Search).

من جهة أخرى، تعرف هندسة الميزات (Feature engineering) بأنها عملية إنشاء مجموعة ميزات باستخدام خصائص البيانات التي تعزز أداء خوارزميات تعلم الآلة. يمكن أن تكون هذه الميزات ذات أبعاد عالية (High Dimensions) ويصعب تدريبها. يعد تقليل الأبعاد (Dimensionality reduction) أحد أكثر الطرق شيوعاً لتحويل (Mapping) الميزات من فضاء ذو أبعاد عالية إلى فضاء بعدد أقل من الأبعاد والتي لها معنى [5] يوجد مجموعة من تقنيات تقليل الأبعاد وأهمها استخراج الميزات ( Feature Extraction) واختيار الميزات (Feature Selection) [6]. تنشئ تقنية استخراج الميزات مجموعة ميزات جديدة باستخدام مجموعة من الميزات الأصلية وإسقاطها إلى فضاء بأبعاد أقل. بينما يهدف اختيار الميزات إلى تحديد مجموعة فرعية من الميزات وثيقة الصلة باستخدام مقياس معياري. يوجد مجموعة من الطرائق لاختيار الميزات أهمها التصفية (Filter) والتغليف (Wrapper) وتقنيات التضمين (Embedded Methods).

يعد كلاً من تحديد الميزات وضبط البارامترات الفائقة مهمتين أساسيتين في تعلم الآلة. يؤثر تنفيذ أحدهما قبل الآخر على أداء النماذج، حيث أن لاختيار ميزات النموذج تأثير كبير على ضبط البارامترات الفائقة والعكس صحيح. سوف نعتمد في هذا البحث على استخدام تقنيات التضمين لتحديد الميزات الأكثر أهمية وطريقة البحث العشوائي لضبط البارامترات الفائقة، وتحديد الاستراتيجية الأفضل لترتيب تحقيقهما ضمن أنظمة كشف الشذوذ المعتمدة على خوارزميات تعلم الآلة.

## 2- أهمية وأهداف البحث

تتأثر أنظمة كشف الشذوذ المعتمدة على خوارزميات تعلم الآلة بتحديد الميزات الأكثر أهمية بالإضافة لاختيار أفضل البارامترات الفائقة. لكن بالمقابل فإن لترتيب تحقيق هذه العمليات دور هام في تحسين أداء هذه الأنظمة من حيث وقت ودقة الكشف. يهدف البحث بشكل أساسي إلى إيجاد التوليفة الأمثل بين اختيار الميزات الأكثر أهمية وضبط البارامترات الفائقة لتحقيق أفضل أداء في أنظمة الكشف عن الشذوذ، وذلك من خلال تحقيق ما يلي:

- 1- إيجاد أفضل قيم البارامترات الفائقة لنماذج تعلم الآلة المقترحة في سياق التعلم الخاضع للإشراف (Supervised Learning) والتعلم غير الخاضع (Unsupervised Learning).
- 2- إيجاد الميزات الأكثر أهمية ضمن مجموعة البيانات.
- 3- تحديد تأثير ترتيب الخطوتين السابقتين على أداء نظام كشف الشذوذ بحيث يتم اختيار الترتيب الأمثل للوصول إلى أفضل دقة كشف.

### 3- مواد وطرائق البحث

#### 3-1 البيانات البحثية

تمّ الاعتماد على مجموعة بيانات (Dataset) لمعاملات بطاقات ائتمان أوروبية تمت على مدار يومين في أيلول عام 2013، وتُعدّ من أكثر مجموعات البيانات الحقيقية المتاحة في الدراسات حتى الآن. تحتوي مجموعة البيانات على 284807 معاملة (سجل)، منها 492 معاملة احتيالية (حوالي 0.17%)، مما يجعل مجموعة البيانات هذه غير متوازنة (Unbalanced Dataset) إلى حد كبير.

تمّ كل معاملة ب 30 ميزة (Feature) جميعها رقمية. تمّ تحويل القيم الأصلية لـ 28 من هذه الميزات باستخدام تحليل المكونات الرئيسية (Principal Component Analysis) (PCA) وسميت بأسماء المتغيرات من V1 وحتى V28، ولم يتم الكشف عن معلومات حول هذه الميزات لأسباب تتعلق بالسرية. كما لا يتوفر أي معلومات عن معرف حامل البطاقة ID حيث يتم اعتبار كل معاملة مستقلة عن المعاملات الأخرى. تُعبر ميزة الوقت (Time) عن الثواني المنقضية بين كل معاملة والمعاملة الأولى، وتحتوي ميزة الكمية (Amount) على مبلغ المعاملة. يُعدّ المتغير Class متغير الاستجابة ويأخذ القيمة 1 في حالة المعاملات الاحتيالية و 0 على خلاف ذلك [7].

#### 3-2 الحزم والمكتبات المستخدمة Used Packages and Libraries

تمّ بناء نظام الكشف عن الشذوذ في مجموعة البيانات السابقة باستخدام لغة بايثون Python وبالاعتماد على مجموعة من المكتبات البرمجية [8] وأهمها:

- **Numpy**: هي اختصار لعبارة (Numerical Python Library) وتستخدم هذه الحزمة من أجل المصفوفات متعدد الأبعاد والعمليات الجبرية الخطية.
- **Pandas**: توفر هذه الحزمة أداة لتحليل ومعالجة البيانات. تم استخدامه لقراءة مجموعة البيانات وتحميلها.
- **Scikitlearn**: تستخدم هذه الحزمة من أجل الأساليب الإحصائية وتعلم الآلة.
- **Keras**: توفر هذه الحزمة واجهات برمجية متناسقة وبسيطة من أجل التخاطب مع المستخدم النهائي وليس الآلة ، وتحتوي على مجموعة من النماذج (Models) مثل الشبكات العصبونية وأشجار القرار وتوابع التنشيط ، كما تتميز بقابليتها للتوسع أي القدرة على إضافة نماذج جديدة. تكون المهمة الأساسية للمكتبة جعل التطبيق أكثر استجابة مع إمكانية إعطاء المستخدم المزيد من القدرة على التحكم به.
- **TensorFlow** : تُطبق هذه الحزمة في العديد من المجالات كحساب المشتقات والمصفوفات الضخمة بالإضافة إلى استخدامها في توزيع العمليات الحاسوبية على وحدات المعالجة المركزية CPU وكذلك على شبكة موزعة مكونة من مجموعة أجهزة بعيدة تتضمن هذه المكتبة. يُستخدم TensorFlow بشكل أساسي في تعلم الآلة في الوقت الحالي.

### 3-3 مقاييس الأداء Performance Metrics

يعتبر مقياس الدقة Accuracy المقياس الأكثر استخداماً والأكثر منطقية في تقييم أداء خوارزميات التصنيف وذلك عندما تكون مجموعات البيانات متوازنة. فهو يمثل النسبة بين عدد العينات المصنفة بشكل صحيح وعدد العينات الكلية. ولكن بالمقابل لا يمكن اعتباره مقياساً جيداً في البيانات غير المتوازنة (كما في حالة اكتشاف الشذوذ) لأنه يسبب ملائمة زائدة (Overfitting) لجهة صف الأغلبية.

يوجد مجموعة واسعة من المقاييس التي يمكن استخدامها لتقييم أداء أنظمة الكشف عن الشذوذ ونذكر منها هنا فقط ما تم استخدامه لتقييم الأداء كونها الأكثر استخداماً في تقييم خوارزميات اكتشاف الشذوذ:

## 1. مصفوفة الارتباك Confusion Matrix

- يَتَّبَعُ النموذج بصف عينات البيانات وينسب إلى كل عينة تسميتها (Label) المتوقعة (إيجابية أو سلبية)، لتقع كل عينة في نهاية المطاف ضمن أحد الحالات التالية:
- (1) الإيجابيات الحقيقية (TP) True Positives تدل على العينات الإيجابية المتوقعة بشكل صحيح.
  - (2) السلبيات الحقيقية (TN) True Negatives تدل على العينات السلبية المتوقعة بشكل صحيح.
  - (3) الإيجابيات الخاطئة (FP) False Positives والسلبيات الخاطئة (FN) False Negatives عندما يتعارض الصف الحقيقي مع الصف المتوقع.
- يُمكن تلخيص الحالات السابقة ضمن مصفوفة  $2 \times 2$  تسمى مصفوفة الارتباك
- $$M = \begin{pmatrix} TP & FN \\ FP & TN \end{pmatrix}$$

## 2. مقياس F1 Score

يُعدّ F1 Score في التحليل الإحصائي مقياساً لدقة الاختبار، فهو المتوسط التوافقي للدقة (Precision) والاستنكار (Recall). تحسب قيمة F1 بالاعتماد على مصفوفة الارتباك بالشكل التالي:

$$F1 \text{ score} = \frac{2 * TP}{2 * TP + FP + FN} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

تمثل القيمة 1 أفضل نتيجة، بينما 0 أسوأها.

## 3. Matthews Correlation Coefficient (MCC)

يُعدّ مقياس Matthews Correlation Coefficient (MCC) والذي يعتبر حالة خاصة من معامل phi حلاً للتغلب على مشكلة عدم توازن البيانات. فهو يعبر عن الارتباط بين القيم المتوقعة والحقيقية. تعتمد قيمة MCC على جميع عناصر مصفوفة الارتباك بالشكل التالي:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}}$$

تمثل القيمة 1 أفضل نتيجة، بينما 1- أسوأها، أما القيمة 0 فهي للدلالة على عشوائية النموذج. يعتبر MCC المقياس الوحيد [9][9] الذي يستخدم لمعرفة فيما إذا كان المصنف الثنائي قادر على توقع معظم الحالات الإيجابية والسلبية بشكل صحيح.

#### 4. منحنيات الدقة والاستدكار (Precision-Recall Curves (PR

يوصى عادةً بمنحنيات الدقة والاستدكار [9] عند التعامل مع مجموعات بيانات غير متوازنة، حيث توفر منحنيات Receiver Operating Characteristic (ROC) رؤية مضللة ومنحازة لصف الأغلبية.

يوفر منحنى PR رؤية بصرية لقيم كل من الدقة والاستدكار وذلك من أجل قيم عتبات مختلفة بدلاً من قيمة واحدة. تُقاس منحنيات PR درجة الفصل بين الصفوف بالاعتماد على مساحة السطح تحت المنحنى (Area Under the Curve (AUC ويُعتبر النموذج أكثر دقة مع ازدياد قيمة AUC حيث يزداد الفصل والتمايز بين الصفوف، تتراوح قيم AUC ضمن المجال [0,1].

#### 5. معدل الفشل Failure Rate

قد يكون في بعض الحالات للمقاييس المستخدمة نفس القيم، نحتاج عندئذٍ إلى تقاضل بين هذه الحالات. اقترح البحث لتحقيق ذلك ما يسمى معدل الفشل (Failure Rate). يعطي معدل الفشل نسبة كل من السلبيات الخاطئة (تصنيف الحالات الشاذة على أنها طبيعية) والإيجابيات الخاطئة (تصنيف الحالات الطبيعية على أنها شاذة). يجب أن تكون قيمة هذه النسبة أقل ما يمكن لتحقيق أفضل أداء لنماذج كشف الشذوذ.

$$failure\ rate = \frac{FP + FN}{TP + FP + TN + FN}$$

اعتمد البحث الحالي على كل من منحنى PR و MCC لتحليل أداء مصنفات تعلم الآلة لما توفره من قدرة على إعطاء تصور واضح لأداء هذه المصنفات مع مسائل كشف الشذوذ.

#### 4- الدراسة المرجعية

اشتملت أدبيات الدراسات على العديد من الوسائل والطرائق والخوارزميات لمعالجة الحالات الشاذة ومحاولة تصنيفها واكتشافها ومن أهمها خوارزميات تعلم الآلة. بالمقابل

تقتصر معظم هذه الأدبيات إلى استخدام آليات ضبط البارامترات الفائقة واختيار الميزات الأكثر أهمية من جهة، ومن جهة أخرى إلى تكامل هذه الآليات معاً.

تُشير الدراسات السابقة [10] لوجود حوالي 49 أسلوب تعلم خاضع للإشراف في اكتشاف الحالات الشاذة [10]، تتفوق فيها كل من والشبكات العصبونية (Neural Network)، والانحدار اللوجستي (Logistic Regression)، الغابات العشوائية (Random Forests)، وآلة شعاع الدعم (Support Vector Machines (SVM)).

تمّ تطبيق مجموعة من خوارزميات تعلم الآلة مثل Naïve Bayes وأقرب جار (K-Nearest Neighbor) وآلة شعاع الدعم والغابات العشوائية بهدف كشف الاحتيال المالي ضمن مجموعة بيانات غير متوازنة تحتوي على نسبة 3% فقط من المعاملات الاحتمالية [11]. أظهرت النتائج النهائية أن نموذج الغابات العشوائية لديه أفضل معدل للإيجابيات الحقيقية (True Positive Rate) مقارنةً مع النماذج الأخرى المستخدمة. لكن بالمقابل لم تذكر الدراسة آلية ضبط البارامترات الفائقة الخاصة بالنماذج المقترحة.

استخدمت خوارزمية الغابات العشوائية (Random Forest) بغرض كشف الاحتيال المالي ضمن أحد الشركات الصينية [12]. اعتمدت الدراسة على بناء نموذجين من الغابات العشوائية يختلفان عن بعضهما بآلية اختيار الميزات (Features) ضمن العقد (Nodes). تعتمد الآلية الأولى على حساب المسافة ما بين نقاط البيانات وصفي البيانات (الطبيعية والشاذة)، بينما تعتمد الآلية الثانية على حساب قيمة سائبة جيني (Gini Impurity) لكل ميزة واختيار تلك التي تحقق أقل قيمة. أظهرت النتائج تفوق الآلية الثانية من حيث الاستدكار (Recall) والدقة (Precision)، حيث بلغت قيمتهما 95% و 89% على التوالي. بالمقابل نوهت الدراسة أنها تتعامل مع مجموعة بيانات غير متوازنة ولذلك يمكن أن تكون دقة هذه النماذج مضللة بعض الشيء.

تمّ مقارنة عدد من تقنيات تعلم الآلة غير خاضعة للإشراف والأكثر شيوعاً في مجال اكتشاف الشذوذ، وهي غابة العزل (Isolation Forest) وخوارزمية K-Means وخوارزمية المعامل الخارجي المحلي (Local Outlier Factor) لتحديد أفضل تنقية في اكتشاف الاحتيال المالي ضمن مجموعة البيانات الأوربية[13]]. أظهرت نتائج الدراسة تفوق خوارزمية غابة العزل بالاعتماد على منحنى ROC حيث بلغت مساحة السطح تحت المنحني (AUC) 91%. أحد أهم قيود الدراسة هو الاعتماد على منحنى ROC فقط لتقييم نماذج الكشف وهذا أمر مضر بسبب طبيعة البيانات غير المتوازنة، كما اعتمدت أيضاً على القيم الافتراضية للبارامترات الفائقة.

تمّ مقارنة ثماني خوارزميات تعلم آلة (Machine Learning) في مسائل كشف الاحتيال المالي ضمن بطاقات الائتمان[14]]. أظهرت النتائج **Error! Reference source not found.** تفوق كل من خوارزميات آلة شعاع الدعم (SVM) والشبكات العصبونية والغابات العشوائية وفقاً لثلاث مقاييس أداء وهي الاستدكار (Recall) ومساحة السطح تحت منحنى الدقة والاستدكار (AUCPR) والدقة (Accuracy). لكن من جهة أخرى، نوه الباحثون لضرورة إجراء ضبط البارامترات الفائقة لتحقيق نتائج أفضل لهذه الخوارزميات.

تم استخدام أداة Features Selector لتحديد الميزات الأكثر أهمية ضمن مجموعة البيانات الأوربية الخاصة بالاحتيال المالي حيث تم اختيار 27 ميزة لهذه التجربة. ليتم في المرحلة التالية تطبيق مجموعة من خوارزميات تعلم الآلة وهي الانحدار اللوجستي (Logistic Regression) والغابات العشوائية (Random Forests) ومصنف بايز (Naive Bayes) لتصنيف الحالات الاحتمالية ضمن مجموعة البيانات المستخدمة [15][15]. حيث أظهرت نتائج هذه الدراسة **Error! Reference source not found.**

تفوق الغابات العشوائية بنسبة استذكار (Recall) حوالي 81%. لكن بالمقابل لم تطرُق الدراسة لضبط البارامترات الفائقة الخاصة بالنماذج المقترحة.

استخدمت خوارزميات الانحدار اللوجستي والغابات العشوائية وأشجار القرار (Decision tree) لاكتشاف العمليات الاحتمالية ضمن مجموعة البيانات الأوربية [16]. واختبرت الدراسة مدى فعالية هذه الخوارزميات باستخدام جميع ميزات مجموعة البيانات وعند اختيار مجموعات جزئية من الميزات مكونة من (5-10 ميزات) حيث أظهرت النتائج تفوق الغابات العشوائية بدقة (Accuracy) حوالي 90% عند استخدام جميع الميزات. لم تذكر الدراسة الطريقة التي تم اختيار الميزات بها من جهة، ومن جهة أخرى لا يمكن تبني هذه النتائج بشكل كبير بسبب اعتمادها على مقياس الدقة (Accuracy).

تم تطبيق البحث الشبكي (Grid Search) لضبط البارامترات الفائقة الخاصة بخوارزميتي آلة شعاع الدعم (SVM) والغابات العشوائية من أجل بناء نماذج لاكتشاف الاحتيال المالي ضمن عدة مجموعات غير متوازنة [17][17] أظهرت نتائج الدراسة **Error!** **Reference source not found.** تفوق شعاع آلة الدعم على الغابات العشوائية من أجل جميع مجموعات البيانات المستخدمة. حيث تم تقييم النتائج باستخدام مقياس (MCC) Matthews Correlation Coefficient، وبلغت قيمته في خوارزمية SVM حوالي 81% من أجل مجموعة البيانات الأوربية. إن أحد قيود هذه الدراسة أنها تتعامل فقط مع اكتشاف الاحتيال في سياق التعلم الخاضع للإشراف، بالإضافة لعدم تحديد الميزات الأكثر أهمية.

تم ملاحظة التفوق الكبير من خلال الدراسات السابقة للغابات العشوائية وآلة شعاع الدعم وغابات العزل في مجال اكتشاف الشذوذ على باقي الخوارزميات. لكن وبنفس الوقت تم ملاحظة القصر الحاصل فيما يتعلق بضبط البارامترات الفائقة لهذه الخوارزميات، مما يستدعي المزيد من البحث للحصول على أفضل البارامترات الفائقة. بالمقابل تم اعتماد أكثر

من طريقة لاختيار الميزات الأكثر أهمية ضمن الأدبيات، لكنها لازالت تعاني من الدقة من جهة ومن عدم تكاملها مع ضبط البارامترات الفائقة من جهة أخرى. استوجب كل ذلك المزيد من البحث لتحديد الآلية الأفضل لضبط البارامترات الفائقة، واختيار الميزات الأكثر أهمية وتكاملها معاً لتحقيق أفضل دقة كشف عن الحالات الشاذة باستخدام هذه الخوارزميات.

### 5- الخوارزميات المقترحة للدراسة

بناء على ما تم ملاحظته من الدراسات السابقة فقد تم التركيز في هذا البحث على الخوارزميات التي أعطت أفضل أداء في اكتشاف الشذوذ. وعليه تم تفصيل كل منها.

### 5-1 خوارزمية الغابات العشوائية Random Forests Algorithm

تتدرج خوارزمية الغابات العشوائية [18]] تحت خوارزميات التعلّم الخاضع للإشراف. تتشكل الغابة (Forest) من مجموعة من أشجار القرار ويتم تدريبها باستخدام مفهوم التعبئة (Bagging). تتمثل إحدى الميزات الكبيرة للغابات العشوائية في أنه يمكن استخدامها لكل من مسائل التصنيف والانحدار، والتي تشكل غالبية أنظمة تعلّم الآلة الحالية. باختصار: تقوم الغابات العشوائية ببناء مجموعة من أشجار القرار ودمجها معاً باستخدام تقنيات التجميع (Ensemble Technique) للحصول على تنبؤ أكثر دقة.

### 5-1-1 مصنفات الغابات العشوائية Random Forests Classifiers

يستخدم التصنيف في الغابات العشوائية تقنية التعبئة المعروفة أيضاً باسم "Bootstrap Aggregation" والتي تقوم باختيار مجموعة جزئية عشوائية مع الاستبدال من بيانات التدريب (مما يعني أنه يمكن اختيار نفس العينة عدة مرات، سحب مع إعادة). تُعرف الخطوة السابقة باسم Bootstrap. يتم في كل مرة اختيار عينة جديدة Bootstrap Sample ليتم تدريب النموذج عليها لنحصل في نهاية الأمر على مجموعة من النماذج المدربة، ويعتمد الناتج النهائي على تصويت الأغلبية (Majority Voting) بعد دمج نتائج جميع النماذج (بمعنى آخر يصبح الناتج الذي تختاره غالبية أشجار القرار هو الناتج النهائي لمصنف الغابة العشوائية). تُعرف الخطوة الأخيرة باسم التجميع (Aggregation).

## 2-1-5 البارامترات الفائقة Hyperparameters

يوجد مجموعة من البارامترات الفائقة [18] ضمن نموذج الغابات العشوائية وأهمها:

### 1- عدد نماذج التنبؤ Count of Estimators

تمثل قيمة هذا المتحول عدد أشجار القرار التي تبنيها الخوارزمية.

### 2- الحد الأعظمي لعدد الميزات Maximum Number of Features

تمثل الحد الأعظمي لعدد الميزات التي تأخذها الخوارزمية بعين الاعتبار عند

تقسيم العقدة.

### 3- الحد الأدنى لعدد العينات Minimum Number of Samples

تمثل الحد الأدنى لعدد العينات المطلوبة لتقسيم عقدة داخلية.

### 4- عمق الشجرة Depth of The Tree

يحدد أقصى عمق يمكن أن تصل إليه الشجرة. في حال لم يتم تحديد قيمة ضمن

هذا البارامتر فسوف يتم التوسع في الشجرة حتى تصبح جميع الأوراق نقية، أو حتى تحتوي

على عينات أقل من الحد الأدنى لعدد العينات.

## 3-1-5 اكتشاف الشذوذ باستخدام الغابات العشوائية

تعتمد الفكرة الرئيسية في استخدام الغابات العشوائية في اكتشاف الشذوذ على دمج

نتائج مجموعة من أشجار القرار المتنوعة، نظراً لأن كل شجرة تأخذ بعين الاعتبار مجموعة

مختلفة من الميزات (Features). يؤدي تنوع أشجار القرار إلى اختزال فضاء أبعاد الميزات

من خلال تقسيمها بين هذه الأشجار، ويحقق ذلك القدرة على التعامل مع بيانات عالية

الأبعاد ولعل ذلك من أهم ما تتطلبه أنظمة كشف الشذوذ ضمن البيانات.

## 5-2 خوارزمية غابات العزل Isolation Forests

تشبه غابات العزل (Isolation Forests) الغابات العشوائية لحد ما حيث يتم بناء كل منهما بالاعتماد على تقنيات التجميع لأشجار القرار. تتدرج غابة العزل تحت خوارزميات التعلم غير الخاضع للإشراف

يشير مصطلح العزل (Isolation) إلى فصل نقطة بيانات (حالة/Instance) عن بقية النقاط [23][19]]، ونظراً لأن الحالات الشاذة قليلة ومختلفة بالتالي فهي أكثر عرضة للعزل. يحدث في شجرة القرار انقسام لنقاط البيانات بشكل متكرر حتى يتم عزل جميع النقاط. ينتج عن هذه الانقسامات العشوائية مسارات أقصر ملحوظة للحالات الشاذة.

### 5-2-1 آلية عمل غابة العزل

كما ذكرنا سابقاً أن غابات العزل ليست إلا مجموعة من أشجار القرار الثنائية، وتسمى كل شجرة في غابة العزل بشجرة العزل. تبدأ الخوارزمية بالتدريب عن طريق توليد مجموعة من أشجار العزل باستخدام مفهوم التعبئة (Bagging).

تتلخص مراحل عمل الخوارزمية [20] بالخطوات التالية:

1) يتم تحديد عينة جزئية عشوائية من بيانات التدريب باستخدام مفهوم التعبئة، وتسمى هذه العينة Bootstrap Sample.

2) يبدأ تقسيم شجرة العزل باختيار ميزة عشوائية من مجموعة الميزات، ثم يتم إجراء الانقسامات المتتالية باختيار قيمة عتبة (Threshold) عشوائية تقع بين القيم القصوى والدنيا للميزة التي تم اختيارها أولاً.

3) نختبر قيمة نقطة البيانات فإذا كانت أصغر من قيمة العتبة فإنها تصبح ابن يساري للعقدة الأب وإلا تصبح ابن يميني.

4) يتم تكرار الخطوات 2 و3 بشكل متكرر حتى يتم عزل كل نقطة بيانات تماماً أو حتى يتم الوصول إلى أقصى عمق (إذا تم تحديده).

5) يتم تكرار الخطوات المذكورة أعلاه لإنشاء أشجار ثنائية عشوائية.

نحصل في نهاية عملية تدريب الخوارزمية على مجموعة من أشجار العزل (غابة عزل). تجتاز نقطة البيانات جميع الأشجار التي تم تدريبها مسبقاً، ويتم تعيين درجة الانحراف (Anomaly Score) لكل نقطة من نقاط البيانات بناءً على طول المسار

المطلوب للوصول إلى عزل تلك النقطة. تسحب القيمة النهائية لدرجة الانحراف (الشذوذ) لنقطة ما من خلال حساب متوسط أطوال المسارات التي تم الحصول عليها من جميع أشجار العزل.

## 2-2-5 البارامترات الفائقة Hyperparameters

يوجد مجموعة من البارامترات الفائقة ضمن نموذج غابة العزل [23][21] وأهمها

هي:

### 1. الحد الأعظمي لعدد العينات Maximum Number of Samples

تمثل قيمة هذا البارامتر الحد الأعظمي لعدد العينات التي يمكن سحبها لتدريب كل شجرة.

### 2. التلوث Contamination

تمثل النسبة المتوقعة من القيم الشاذة في مجموعة البيانات وهي حساسة للغاية.

### 3. عدد نماذج التنبؤ Count of Estimators

يشير إلى عدد الأشجار التي سيتم بناؤها في الغابة. القيمة الافتراضية 100.

## 3-2-5 اكتشاف الشذوذ باستخدام غابات العزل

يتم الكشف عن الشذوذ باستخدام غابات العزل على مرحلتين: الأولى هي مرحلة التدريب لبناء شجرة عزل باستخدام عينات فرعية من مجموعة بيانات التدريب. توفر المرحلة الثانية اختبار كل نقطة بيانات باستخدام أشجار العزل الناتجة من المرحلة الأولى للحصول على قيمة درجة الشذوذ لكل حالة.

## 3-5 خوارزمية آلة شعاع الدعم Support Vector Machine Algorithm

تعد خوارزمية آلة شعاع الدعم [23] (SVM) واحدة من أكثر خوارزميات تعلم الآلة شيوعاً [22] والمستخدم على نطاق واسع. وذلك بسبب بساطة فكرة عملها وأناقته رياضياً من جهة، وسهولة استخدامها من جهة أخرى.

تقوم فكرة الخوارزمية على الفصل بين الصفوف داخل فضاء البيانات بمستوى يسمى المستوى الفائق (Hyperplane). لكي تتمكن SVM من إيجاد ذلك المستوى، لا بد من أن تكون البيانات قابلة للفصل الخطي (Linearly Separable).

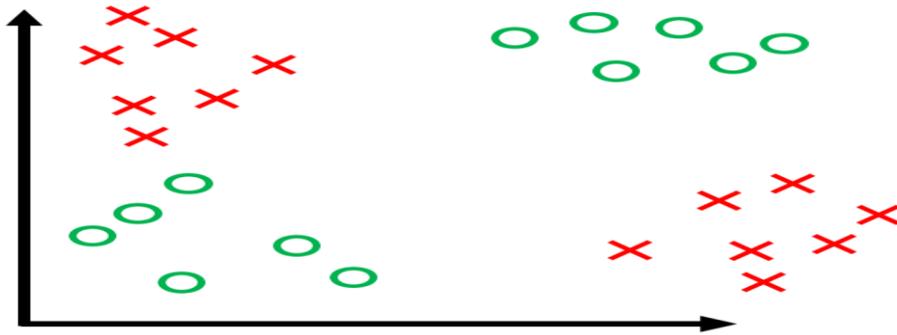
### 5-3-1 آلية عمل آلة شعاع الدعم

يتم تدريب المصنف بإيجاد قيم  $w$  و  $b$  التي تفصل مجموعة البيانات بأكبر هامش (Margin) ممكن. تسعى الخوارزمية إلى إيجاد أفضل مستوى فائق قادر على فصل أكبر عدد ممكن من نقاط البيانات، حيث تقوم بمعاقبة المصنف من أجل جميع نقاط البيانات التي تقع في الاتجاه الخاطئ من المستوى.

ماذا لو كانت البيانات غير قابلة للفصل خطياً؟ حيث إننا نرى في **Error!**

**Reference source not found.** مثلاً على هذه الحالة. تعجز خوارزمية SVM

التقليدية عن إيجاد مستوى فائق يفصل هذه البيانات مما يؤدي بدوره إلى عدم القدرة على تصنيف البيانات بشكل صحيح.



الشكل 1 بيانات غير قابلة للفصل خطياً

يتم حل المشكلة السابقة [22]. **Error! Reference source not found.** باستخدام

تعديل على الخوارزمية يسمى "نوى SVM" (Kernel SVM) وهي عبارة عن تعميم لخوارزمية آلة شعاع الدعم تسمح لها بتصنيف البيانات التي لا يمكن فصلها خطياً.

### 5-3-2 البارامترات الفائقة Hyperparameters

#### 1. التنظيم Regularization:

تحدد قيمة هذه البارامتر مدى السماح لخوارزمية SVM في تخطي تصنيف كل نقطة تدريب. يطلق على هذه البارامتر في لغة بايثون C.

### متحول Gamma:

مدى تأثير نقاط بيانات التدريب في إيجاد المستوى الفائق الأمثل.

## 3-3-5 اكتشاف الشذوذ باستخدام خوارزمية SVM

جعل كل ما سبق من SVM خوارزمية فعالة في مسألة اكتشاف الشذوذ نظراً لإمكانية استخدام نواة مناسبة في حال كانت البيانات غير قابلة للفصل خطياً من جهة، وقدرتها على التعامل مع البيانات ذات الأبعاد العالية (وهي سمة عامة لأغلب مسائل كشف الشذوذ) من جهة أخرى. فهي تحاول إيجاد المستوى الفائق الأفضل بناءً على عامل تعظيم الهوامش الفاصلة بين صفي البيانات.

## 8- النتائج والمناقشة Results and Discussion

يستعرض في هذا القسم بدايةً وصفاً تجريبياً لتصميم التجارب التي أجريناها لبناء أنظمة كشف الشذوذ (اكتشاف الاحتيال المالي) باستخدام خوارزميات تعلم الآلة المقترحة وأساليب ضبط البارامترات الفائقة واختيار الميزات. يتبع التوصيف استعراض النتائج وتقييم الأداء.

## 8-1 ضبط البارامترات الفائقة Hyperparameters Tuning

تُعني عملية ضبط (أو تحسين) البارامترات الفائقة إيجاد مجموعة قيم البارامترات التي تحقق أفضل أداء لنموذج تعلم الآلة. يركز هذا البحث على استخدام طريقة البحث العشوائي لضبط البارامترات الفائقة، حيث أن البحث العشوائي [23][23] أكثر كفاءة في تحسين البارامترات من البحث الشبكي من الناحية التجريبية والنظرية.

يعمل البحث العشوائي بشكل أفضل عندما يكون فضاء البحث عالي الأبعاد ويحتوي على عدد كبير من التركيبات المختلفة للبارامترات الفائقة. حيث تبحث تقنية البحث العشوائي ضمن مجموعة عشوائية من هذه التركيبات لاختيار أفضل القيم للبارامترات الفائقة

الخاصة بالنموذج المقترح، وبالتالي فإن الوقت المستغرق للعثور على المجموعة الصحيحة يكون أقل مع عدد أقل من التكرارات.

## 2-8 اختيار الميزات Features Selection

يعتمد البحث على الغابات العشوائية لاختيار الميزات، حيث تتمتع بقدرتها على قياس الأهمية النسبية (Relative Importance) لكل ميزة في التنبؤ [24]. يتم قياس أهمية الميزة (Feature Importance) من خلال النظر بعدد العقد النقية في نهاية جميع الأشجار التي تستخدم هذه الميزة (يحدث الانقسام عندها باتجاه واحد). بمعنى آخر تكون العقد الشائبة (Impurity Nodes) في بداية الشجرة، بينما تحدث الملاحظات (Observations) التي تسبب نقصاً في شوائب العقد في نهاية الشجرة. بتقليم الأشجار أسفل العقد النقية يمكننا إنشاء مجموعة فرعية من أهم الميزات. يمكن من خلال النظر إلى أهمية الميزة تحديد الميزات التي من المحتمل حذفها لأنها لا تساهم بشكل كافٍ (أو أحياناً لا تساهم على الإطلاق) في عملية التنبؤ.

يبين الجدول 1 ترتيب ميزات مجموعة البيانات المقترحة باستخدام خوارزمية الغابات العشوائية ترتيب ميزات مجموعة البيانات المستخدمة وفقاً لدرجة أهمية كل منها باستخدام خوارزمية الغابات العشوائية. يتم اختيار الميزات الأكثر أهمية التي تحقق أفضل أداء للنموذج من خلال تصفية جميع الميزات التي لها درجة أهمية منخفضة، أي لها تأثير ضعيف في التصنيف.

الجدول 1 ترتيب ميزات مجموعة البيانات المقترحة باستخدام خوارزمية الغابات العشوائية

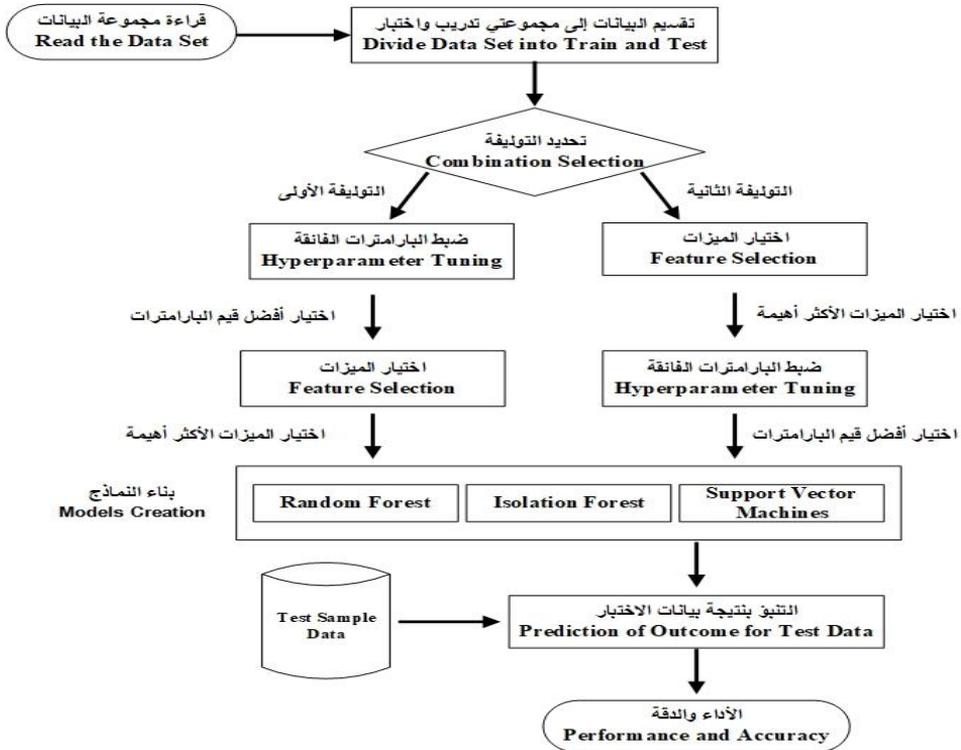
Feature	Score	Feature	Score	Feature	Score
V14	0.184	V9	0.022	V15	0.009
V4	0.113	V21	0.020	V1	0.007
V12	0.109	V19	0.017	V13	0.007
V10	0.101	V27	0.012	V6	0.007
V17	0.088	V5	0.012	V22	0.006
V3	0.058	V18	0.012	V23	0.006
V11	0.046	Amount	0.012	V25	0.006

## تحديد التوليفة الأمثل من ضبط البارامترات الفائقة واختيار الميزات لتحسين أداء أنظمة كشف الشذوذ

V16	0.044	V8	0.011	V24	0.005
V2	0.036	V26	0.009	V28	0.005
V7	0.023	V20	0.009	Time	0.005

### 3-8 بناء النموذج Model Creation

يوضح المخطط التالي الاستراتيجية المقدمه ضمن هذا البحث لاستخدام الأساليب المقترحة لبناء أنظمة كشف الشذوذ (كشف الاحتيال المالي كمثال تطبيقي على النموذج).



الشكل 2 النظام المقترح لاكتشاف الشذوذ

#### 1- اختيار التوليفة

استخدمت الخوارزميات المقترحة لبناء هذه الأنظمة وفق سيناريوهين يختلفان بترتيب التوليفة بين ضبط البارامترات الفائقة واختيار الميزات الأكثر أهمية، وذلك لتحديد التوليفة التي تحقق أفضل أداء لأنظمة كشف الشذوذ من حيث الدقة والوقت.

يكون ترتيب التوليفة من أجل السيناريو الأول بالشكل التالي: (1) ضبط البارامترات الفائقة للخوارزميات المقترحة، (2) اختيار الميزات الأكثر أهمية التي تحقق أفضل أداء لهذه الخوارزميات، باستخدام البارامترات الفائقة المحددة من الخطوة السابقة. بينما يكون ترتيبها من أجل السيناريو الثاني: (1) اختيار الميزات الأكثر أهمية، (2) ضبط البارامترات الفائقة للخوارزميات، بالاعتماد على الميزات المحددة من الخطوة السابقة.

## 2- بناء النموذج

## 3- اختبار النموذج

يجب التنويه أنه تمّ تقسيم مجموعة بيانات الدراسة إلى مجموعات تدريب واختبار (Training and Test Sets)، حيث تحتوي مجموعة التدريب على 70% من إجمالي عدد العينات.

## 8-4 تقييم الأداء Performance Evaluation

### 8-4-1 خوارزمية الغابات العشوائية

#### • السيناريو الأول (scenario 1)

تظهر النتائج المبينة في الجدول 2 أداء نموذج الغابة العشوائية في المرحلة الأولى عند ضبط البارامترات الفائقة. بينما يُبين

الجدول 3 أداء النموذج في المرحلة الثانية عند اختيار الميزات الأكثر أهمية باستخدام أفضل البارامترات الفائقة المحددة من المرحلة الأولى.

الجدول 2 نتائج ضبط البارامترات الفائقة لنموذج الغابة العشوائية (السيناريو 1)

Precision	Recall	F1	MCC	AUCPR	Failure Rate	Time
0.87	0.85	0.86	0.85	0.84	0.046	31.23M

الجدول 3 نتائج اختيار الميزات لنموذج الغابة العشوائية (السيناريو 1)

#Features	Precision	Recall	F1	MCC	AUCPR	Failure Rate
-----------	-----------	--------	----	-----	-------	--------------

## تحديد التوليفة الأمثل من ضبط البارامترات الفائقة واختيار الميزات لتحسين أداء أنظمة كشف الشذوذ

1	0.27	0.74	0.39	0.44	0.53	0.364
2	0.61	0.76	0.67	0.68	0.72	0.117
3	0.82	0.83	0.83	0.83	0.82	0.055
4	0.82	0.83	0.82	0.82	0.84	0.056
<b>5</b>	<b>0.87</b>	<b>0.88</b>	<b>0.87</b>	<b>0.87</b>	<b>0.86</b>	<b>0.041</b>
6	0.87	0.86	0.86	0.86	0.86	0.043
7	0.86	0.85	0.86	0.85	0.86	0.046
8	0.85	0.87	0.86	0.86	0.86	0.046
9	0.86	0.88	0.87	0.87	0.86	0.043
10	0.85	0.85	0.85	0.85	0.85	0.042

### 1. ضبط البارامترات الفائقة

نلاحظ من الجدول 2 أن أفضل أداء يمكن أن يصل إليه نموذج الغابات العشوائية لاكتشاف الاحتيال المالي بالاعتماد على جميع ميزات مجموعة البيانات المقترحة يكون عند ضبط البارامترات الفائقة بالشكل التالي: عدد الأشجار 100، عمق الشجرة 150، الحد الأعظمي للميزات  $\sqrt{30}$  (عدد الميزات=30)، الحد الأدنى للعينات 5. حيث يحقق النموذج أداء عالي في تصنيف الحالات الاحتمالية وذلك بالنظر إلى قيم كل من AUCPR و MCC حيث تصل قيمة كل منهما إلى 84% و 85% على الترتيب.

### 2. اختيار الميزات الأكثر أهمية

تظهر نتائج المرحلة الحالية من خلال

الجدول 3 أن نموذج الغابة العشوائية يحقق أفضل أداء له في اكتشاف العمليات الاحتمالية عند اختيار أهم 5 ميزات وهي: (V14, V4, V10, V12, V17) وفق ترتيب الجدول 1، وباستخدام البارامترات الفائقة المحددة من المرحلة السابقة. حيث تصل قيمة AUCPR إلى 86% وقيمة MCC إلى 87% وذلك عندما تكون قيمة معدل الفشل أقل ما يمكن  $\text{Failure rate} = 0.041$ ، وهي أفضل نتيجة تم الحصول عليها مع قيم عالية

أيضاً لباقي المقاييس، وبالتالي فإن النموذج يصنف معظم الحالات الشاذة (الاحتمالية) بشكل صحيح.

• السيناريو الثاني (scenario 2)

تظهر النتائج المبينة في الجدول 4 أداء نموذج الغابة العشوائية في المرحلة الأولى عند اختيار الميزات الأكثر أهمية. بينما يُبين

الجدول 5 أداء النموذج في المرحلة الثانية عند ضبط البارامترات الفائقة باستخدام أهم الميزات المحددة من المرحلة الأولى.

الجدول 4 نتائج اختيار الميزات لنموذج الغابة العشوائية (السيناريو 2)

#Features	Precision	Recall	F1	MCC	AUCPR	Failure Rate
1	0.46	0.52	0.49	0.49	0.52	0.172
2	0.86	0.65	0.74	0.75	0.73	0.071
3	0.92	0.80	0.86	0.86	0.84	0.042
4	0.93	0.83	0.86	0.86	0.85	0.041
5	0.94	0.79	0.86	0.86	0.85	0.041
<b>6</b>	<b>0.96</b>	<b>0.80</b>	<b>0.87</b>	<b>0.88</b>	<b>0.85</b>	<b>0.037</b>
7	0.95	0.78	0.85	0.86	0.85	0.042
8	0.96	0.79	0.87	0.87	0.85	0.039
9	0.96	0.79	0.86	0.87	0.85	0.040
10	0.96	0.79	0.87	0.87	0.85	0.039

الجدول 5 نتائج ضبط البارامترات الفائقة لنموذج الغابة العشوائية (السيناريو 2)

Precision	Recall	F1	MCC	AUCPR	Failure Rate	Time
0.96	0.81	0.88	0.88	0.86	0.036	15.12M

### 1. اختيار الميزات الأكثر أهمية

أظهرت نتائج المرحلة الحالية المبينة في الجدول 4 أن أفضل أداء لنموذج الغابة العشوائية باستخدام قيم البارامترات الافتراضية يكون عند اختيار أهم 6 ميزات ضمن مجموعة البيانات وفق لترتيبها في الجدول 1 وهي: (V14, V4, V10, V12, V17, V3)، حيث تبلغ قيمة AUCPR 85% وقيمة MCC 88% وذلك عندما تكون قيمة معدل الفشل أقل ما يمكن Failure rate=0.037، وهذا مؤشر جيد في تصنيف الحالات الشاذة (الاحتمالية).

### 2. ضبط البارامترات الفائقة

بالاعتماد على الميزات الأكثر أهمية التي تم اختيارها من المرحلة السابقة، أظهرت نتائج المرحلة الحالية والمبينة في

الجدول 5 تحسين بسيط في أداء نموذج الغابة العشوائية بالنسبة لقيمة AUCPR

حيث تبلغ قيمته 86% عند ضبط البارامترات الفائقة بالشكل التالي: عدد الأشجار: 250، عمق الشجرة: 150، الحد الأعظمي للميزات:  $\sqrt{6}$ ، الحد الأدنى للعينات: 1.

يمكن مقارنة نتائج السيناريوهين السابقين من خلال الجدول 6 نتائج نموذج الغابات

العشوائية التالي.

الجدول 6 نتائج نموذج الغابات العشوائية

Metrics / Hyperparameters	Scenario 1	Scenario 2
F1	0.87	<b>0.88</b>
MCC	0.87	<b>0.88</b>
AUCPR	<b>0.86</b>	<b>0.86</b>
Failure Rate	0.041	<b>0.036</b>
Run Time	31.23M	<b>15.12M</b>

No. Features	5	6
No. Tree	100	250
Tree Depth	150	150
Maximum Features	$RoundUp(\sqrt{30})$	$RoundUp(\sqrt{6})$

حيث نجد أن أفضل سيناريو لعمل مصنف الغابة العشوائية في اكتشاف الاحتيال المالي هو السيناريو الثاني باستخدام التوليفة "اختيار الميزات الأكثر أهمية ثم ضبط البارامترات الفائقة".

نلاحظ بدايةً تقليص الزمن المستغرق لضبط البارامترات الفائقة حوالي نصف الزمن المستهلك في السيناريو الأول (من 31د إلى 15د)، وهذا مؤشر مهم جداً في أنظمة الكشف عن الشذوذ بشكل عام. كما نجد انخفاض في عدد كل من الإيجابيات الخاطئة FP (المعاملات المشروعة التي صنفت على أنها احتيال) والسلبيات الخاطئة FN (المعاملات الاحتمالية التي صنفت على أنها مشروعة) وهذا ما يظهر من خلال قيمة معدل الفشل (ينخفض من 0.041 إلى 0.036)، حيث يعد ذلك من أهم الأهداف التي يسعى إليها أي نظام كشف عن الاحتيال المالي.

#### 8-4-2 خوارزمية غابات العزل

##### • السيناريو الأول (scenario 1)

تظهر النتائج المبينة في الجدول 7 أداء نموذج غابة العزل في المرحلة الأولى عند ضبط البارامترات الفائقة. بينما يبيّن الجدول 8 أداء النموذج في المرحلة الثانية عند اختيار الميزات الأكثر أهمية باستخدام أفضل البارامترات الفائقة المحددة من المرحلة الأولى.

الجدول 7 نتائج ضبط البارامترات الفائقة لنموذج لغابة العزل (السيناريو 1)

Precision	Recall	F1	MCC	AUCPR	Failure Rate	Time
0.49	0.29	0.37	0.38	0.24	0.16	40M

الجدول 8 نتائج اختيار الميزات لنموذج غابة العزل (السيناريو 1)

#Features	Precision	Recall	F1	MCC	AUCPR	Failure Rate
1	0.75	0.41	0.53	0.55	0.55	0.116
2	0.76	0.38	0.51	0.54	0.61	0.117
3	0.84	0.47	0.60	0.63	0.68	0.098
4	0.83	0.48	0.61	0.63	0.65	0.098
5	0.83	0.51	0.64	0.65	0.66	0.094
6	0.81	0.49	0.61	0.63	0.65	0.099
7	0.80	0.51	0.62	0.64	0.64	0.098
<b>8</b>	<b>0.83</b>	<b>0.53</b>	<b>0.65</b>	<b>0.66</b>	<b>0.66</b>	<b>0.092</b>
9	0.81	0.49	0.61	0.63	0.62	0.099
10	0.76	0.46	0.58	0.59	0.59	0.109

### 1. ضبط البارامترات الفائقة

نلاحظ من الجدول 7 أن أفضل أداء يمكن أن يصل إليه نموذج غابة العزل لاكتشاف الاحتيال المالي بالاعتماد على جميع ميزات مجموعة البيانات المقترحة يكون عند ضبط البارامترات الفائقة بالشكل التالي: عدد الأشجار: 250، العدد الأعظمي للعينات: 550، التلوث: 0.001. لكن بالمقابل نلاحظ أن قيمة AUCPR تبلغ 24% وقيمة MCC 38% وهذا مؤشر على سوء تصنيف الحالات الاحتمالية.

### 2. اختيار الميزات الأكثر أهمية

تظهر نتائج المرحلة الحالية من خلال الجدول 8 تحسن كبير في أداء نموذج غابة العزل عند اختيار أهم 8 ميزات ضمن مجموعة البيانات وهي: ( V14, V4, V10, V12, V17, V3, V11, V16 ) وفق ترتيبها في الجدول 1، وباستخدام البارامترات الفائقة المحددة من المرحلة السابقة، حيث نلاحظ ازدياد قيمة AUCPR من 24% في المرحلة السابقة إلى 66% وقيمة MCC من 38% إلى 66%، كما نلاحظ انخفاض قيمة Failure Rate من 0.16 إلى 0.092.

### • السيناريو الثاني (scenario 2)

تظهر النتائج المبينة في الجدول 9 أداء نموذج غابة العزل في المرحلة الأولى عند اختيار الميزات الأكثر أهمية. بينما يُبين الجدول 10 أداء النموذج في المرحلة الثانية عند ضبط البارامترات الفائقة باستخدام أهم الميزات المحددة من المرحلة الأولى.

الجدول 9 نتائج اختيار الميزات لنموذج غابة العزل (السيناريو 2)

#Features	Precision	Recall	F1	MCC	AUCPR	Failure Rate
1	0.62	0.64	0.63	0.63	0.61	0.119
2	0.58	0.58	0.58	0.58	0.61	0.113
3	0.68	0.67	0.66	0.66	0.66	0.110
<b>4</b>	<b>0.68</b>	<b>0.72</b>	<b>0.70</b>	<b>0.70</b>	<b>0.68</b>	<b>0.099</b>
5	0.68	0.66	0.67	0.67	0.60	0.104
6	0.70	0.65	0.67	0.67	0.62	0.101
7	0.57	0.60	0.58	0.58	0.51	0.137
8	0.67	0.69	0.68	0.68	0.65	0.103
9	0.67	0.68	0.67	0.67	0.64	0.104
10	0.55	0.64	0.59	0.59	0.60	0.142

الجدول 10 نتائج ضبط البارامترات الفائقة لنموذج لغابة العزل (السيناريو 2)

Precision	Recall	F1	MCC	AUCPR	Failure Rate	Time
0.80	0.71	0.75	0.74	0.71	0.083	18M

### 1. اختيار الميزات الأكثر أهمية

أظهرت نتائج المرحلة الحالية المبينة في الجدول 9 أن أفضل أداء لنموذج غابة العزل باستخدام قيم البارامترات الافتراضية يكون عند اختيار أهم 4 ميزات ضمن مجموعة البيانات المستخدمة وفق ترتيب الجدول 1 وهي: (V14, V4, V10, V12). حيث تبلغ قيمة AUCPR 68% وقيمة MCC 70% وذلك عندما تكون قيمة معدل الفشل أقل ما يمكن في هذه المرحلة Failure rate=0.099، أي أن النموذج يحقق أداء مقبول في تصنيف الحالات الاحتمالية.

## 2. ضبط البارامترات الفائقة

بالاعتماد على الميزات الأكثر أهمية التي تم اختيارها من المرحلة السابقة، أظهرت نتائج المرحلة الحالية والمبينة في الجدول 10 تحسن في أداء نموذج غابة العزل عند ضبط البارامترات الفائقة بالشكل التالي: عدد الأشجار: 100، العدد الأعظمي للعينات: 300، التلوث: 0.00178. حيث تصل قيمة AUCPR إلى 71% وقيمة MCC إلى 74% كما تنخفض قيمة Failure rate إلى 0.083.

يمكن مقارنة نتائج السيناريوهين السابقين من خلال الجدول 11/الجدول 6 نتائج نموذج الغابات العشوائية التالي. حيث نجد أن أفضل سيناريو لعمل نموذج غابة العزل في اكتشاف الاحتيال المالي هو السيناريو الثاني باستخدام التوليفة "اختيار الميزات الأكثر أهمية ثم ضبط البارامترات الفائقة". حيث أظهرت النتائج أن غابة العزل لا يمكن أن تحقق أداء جيد في اكتشاف الحالات الاحتمالية دون اختيار الميزات الأكثر أهمية رغم ضبط البارامترات الفائقة الخاصة بها. نلاحظ أيضاً تقليص الزمن المستغرق لضبط البارامترات الفائقة أكثر من نصف الزمن المستهلك في السيناريو الأول (من 40د إلى 18د)، كما نجد انخفاض في عدد كل من الإيجابيات الخاطئة FP والسلبيات الخاطئة FN وهذا ما يظهر من خلال قيمة معدل الفشل (ينخفض من 0.092 إلى 0.083).

الجدول 11 نتائج نموذج غابات العزل

Metrics / Hyperparameters	Scenario 1	Scenario 2
F1	0.65	<b>0.75</b>
MCC	0.66	<b>0.74</b>
AUCPR	0.66	<b>0.71</b>
Failure Rate	0.092	<b>0.083</b>
Run Time	40M	<b>18M</b>
No. Features	8	4
No. Tree	250	100

Sup Sample	550	300
Contamination	0.001	0.00178

### 3-4-8 خوارزمية آلة شعاع الدعم

#### • السيناريو الأول (scenario 1)

تظهر النتائج المبينة في الجدول 12 أداء نموذج آلة شعاع الدعم في المرحلة الأولى عند ضبط البارامترات الفائقة. بينما يُبين الجدول 13 أداء النموذج في المرحلة الثانية عند اختيار الميزات الأكثر أهمية باستخدام أفضل البارامترات الفائقة المحددة من المرحلة الأولى.

الجدول 12 نتائج ضبط البارامترات الفائقة لنموذج آلة شعاع الدعم (السيناريو 1)

Precision	Recall	F1	MCC	AUCPR	Failure Rate	Time
0.79	0.82	0.81	0.81	0.77	0.063	15.74M

الجدول 13 نتائج اختيار الميزات لنموذج آلة شعاع الدعم (السيناريو 1)

#Features	Precision	Recall	F1	MCC	AUCPR	Failure Rate
1	0.79	0.36	0.49	0.53	0.61	0.117
2	0.74	0.40	0.52	0.55	0.67	0.117
3	0.84	0.57	0.68	0.69	0.73	0.087
4	0.85	0.62	0.71	0.72	0.76	0.078
5	0.85	0.76	0.81	0.81	0.76	0.059
6	0.84	0.76	0.80	0.80	0.75	0.061

7	0.86	0.79	0.82	0.82	0.76	0.055
8	0.84	0.80	0.82	0.82	0.76	0.055
9	0.84	0.80	0.82	0.82	0.77	0.056
10	0.82	0.82	0.82	0.82	0.77	0.057

## 1. ضبط البارامترات الفائقة

نلاحظ من الجدول 12 أن أفضل أداء يمكن أن يصل إليه نموذج آلة شعاع الدعم لاكتشاف الاحتيال المالي بالاعتماد على جميع ميزات مجموعة البيانات المقترحة يكون عند ضبط البارامترات بالشكل التالي: C: 3000، Gamma: 0.001، kernel: 'rbf'. حيث تبلغ قيمة كل من AUCPR و MCC 77% و 81% على الترتيب، أي أن للنموذج أداء جيد في تصنيف الحالات الاحتمالية.

## 2. اختيار الميزات الأكثر أهمية

تظهر نتائج المرحلة الحالية من خلال الجدول 13 إن النموذج يحافظ على أدائه من المرحلة الأولى عند اختيار أهم 7 ميزات وهي: (V14, V4, V10, V12, V17, V3, V11) ضمن مجموعة البيانات وفق الجدول 1. **Error! Reference source not found.** وباستخدام البارامترات الفائقة المحددة من المرحلة السابقة. لكن من جهة أخرى نلاحظ انخفاض عدد كل من الإيجابيات الخاطئة (FP) والسلبيات الخاطئة (FN) بناء على قيمة معدل الفشل (ينخفض من 0.063 إلى 0.055) وهذا مؤشر هام في أنظمة كشف الشذوذ عموماً.

### • السيناريو الثاني (scenario 2)

تظهر النتائج المبينة في الجدول 14 أداء نموذج آلة شعاع الدعم في المرحلة الأولى عند اختيار الميزات الأكثر أهمية. بينما يُبين الجدول 15 أداء النموذج في المرحلة الثانية عند ضبط البارامترات الفائقة باستخدام أهم الميزات المحددة من المرحلة الأولى.

الجدول 14 نتائج اختيار الميزات لنموذج آلة شعاع الدعم (السيناريو 2)

#Features	Precision	Recall	F1	MCC	AUCPR	Failure Rate
1	0.71	0.53	0.61	0.61	0.51	0.109
2	0.78	0.53	0.63	0.67	0.62	0.098
3	0.85	0.63	0.73	0.73	0.66	0.076
4	0.85	0.68	0.76	0.76	0.71	0.069
5	0.85	0.80	0.83	0.83	0.75	0.055
<b>6</b>	<b>0.85</b>	<b>0.81</b>	<b>0.83</b>	<b>0.83</b>	<b>0.77</b>	<b>0.054</b>
7	0.85	0.82	0.83	0.83	0.73	0.054
8	0.82	0.82	0.82	0.82	0.77	0.057
9	0.81	0.83	0.82	0.82	0.77	0.059
10	0.79	0.83	0.81	0.81	0.76	0.062

الجدول 15 نتائج ضبط البارامترات الفائقة لنموذج آلة شعاع الدعم (السيناريو 2)

Precision	Recall	F1	MCC	AUCPR	Failure Rate	Time
0.87	0.8	0.83	0.83	0.82	0.051	7.59

### 1. اختيار الميزات الأكثر أهمية

أظهرت نتائج المرحلة الحالية المبينة في الجدول 14 أن أفضل أداء لنموذج آلة شعاع الدعم باستخدام قيم البارامترات الافتراضية يكون عند اختيار أهم 6 ميزات ضمن مجموعة البيانات المستخدمة وفق ترتيب الجدول 1 وهي: (V14, V4, V10, V12, V17, V3). حيث تبلغ قيمة AUCPR 77% وقيمة MCC 83% وذلك عندما تكون قيمة معدل الفشل أقل ما يمكن في هذه المرحلة Failure rate=0.054، أي أن النموذج يحقق أداء جيد في تصنيف الحالات الاحتمالية.

### 2. ضبط البارامترات الفائقة

## تحديد التوليفة الأمثل من ضبط البارامترات الفائقة واختيار الميزات لتحسين أداء أنظمة كشف الشذوذ

بالاعتماد على الميزات الأكثر أهمية التي تم اختيارها من المرحلة السابقة، أظهرت نتائج المرحلة الحالية والمبينة في الجدول 15 تحسن ملحوظ في أداء نموذج آلة شعاع الدعم عند ضبط البارامترات بالشكل التالي: C:1000، Gamma: 0.01، kernel: "rbf". حيث يحقق النموذج أداء جيد للغاية في تصنيف الحالات الاحتمالية لتصل قيمة AUCPR إلى 82% وقيمة MCC إلى 83% كما تنخفض قيمة Failure rate إلى 0.051.

يمكن مقارنة نتائج السيناريوهين السابقين من خلال الجدول 16/الجدول 6 نتائج نموذج الغابات العشوائية التالي.

الجدول 16 نتائج نموذج آلة شعاع الدعم

Metrics / Hyperparameters	Scenario 1	Scenario 2
F1	0.82	<b>0.83</b>
MCC	0.82	<b>0.83</b>
AUCPR	0.76	<b>0.82</b>
Failure Rate	0.055	<b>0.051</b>
Run Time	15.74M	<b>7.59M</b>
No. Features	7	6
C	3000	1000
Gamma	0.001	0.01

حيث نجد أن أفضل سيناريو لعمل مصنف آلة شعاع الدعم في اكتشاف الاحتيال المالي هو السيناريو الثاني باستخدام التوليفة "اختيار الميزات الأكثر أهمية ثم ضبط البارامترات الفائقة". حيث أظهرت النتائج تقليص الزمن المستغرق لضبط البارامترات الفائقة حوالي نصف الزمن المستهلك في السيناريو الأول (من 15د إلى 7.5د). كما نجد انخفاض في عدد كل من الإيجابيات الخاطئة FP والسلبيات الخاطئة FN وهذا ما يظهر من خلال قيمة معدل الفشل (ينخفض من 0.055 إلى 0.051).

أكدت نتائج التجارب العملية لهذا البحث، أن اختيار الميزات الأكثر أهمية ضمن مجموعة البيانات متنوعاً بضبط البارامترات الفائقة يساهم في تحسين أداء جميع الخوارزميات المقترحة بشكل كبير. حيث أظهرت النتائج أن خوارزمية الغابات العشوائية لها أفضل أداء في كشف الحالات الشاذة  $MCC=88\%$ ، تليها خوارزمية آلة شعاع الدعم  $MCC=83\%$ ، ومن ثم خوارزمية غابة العزل  $MCC=74\%$ .

كما ساهمت النتائج التي توصل إليها البحث في اختزال الزمن من 50% إلى 53%، وزيادة دقة الكشف من 5% إلى 6%، وتحسين كل مقاييس الأداء المستخدمة لجميع الخوارزميات المقترحة.

### 9-الخاتمة والاستنتاجات والتوصيات

تمّ في هذا البحث دراسة أهم العوامل التي تؤثر على أداء أنظمة الكشف عن الشذوذ بشكل عام، وهي تحديد التوليفة الأمثل من ضبط البارامترات الفائقة واختيار الميزات الأكثر أهمية؛ وذلك من خلال تحليل أنظمة كشف الاحتيال المالي ضمن بطاقات الائتمان الأوروبية بالاعتماد على عدة خوارزميات تعلم آلة وهي الغابات العشوائية وآلة شعاع الدعم وغابة العزل.

توصلت نتائج البحث إلى أن أفضل سياق لبناء أنظمة كشف الشذوذ هو اختيار الميزات الأكثر أهمية ضمن مجموعة البيانات ومن ثم ضبط البارامترات الفائقة الخاصة بها. حيث أظهرت التجارب العملية تقليص الزمن المستغرق لبناء هذه الأنظمة عند استخدام السياق المحدد إلى النصف أو أكثر، بالإضافة إلى زيادة في كشف الحالات الشاذة وذلك بالنظر إلى قيم AUCPR بحدود 5% لخوارزمية غابة العزل و6% لخوارزمية آلة شعاع الدعم ومن أجل جميع الخوارزميات المقترحة بشكل عام، مما يؤكد على التحسين الحاصل عند اعتماد الترتيب المذكور.

تفوقت خوارزمية الغابات العشوائية على الخوارزميات الأخرى المستخدمة من حيث دقة كشف الشذوذ وبالتالي كشف العمليات الاحتيالية، وذلك باختيار أهم 6 ميزات ضمن مجموعة البيانات وضبط البارامترات الفائقة بالشكل التالي: عدد الأشجار: 250، عمق الشجرة: 150، الحد الأعظمي للميزات:  $\sqrt{6}$ ، الحد الأدنى للعينات: 1. حيث كانت أفضل

النتائج لهذه الخوارزمية AUCPR: 0.86-F1: 0.88-MCC: 0.88، بينما كان الزمن المستغرق 15.12 دقيقة. بالمقابل حققت خوارزمية آلة شعاع الدعم أفضل أداء من حيث زمن التنفيذ بمقدار 7.59 دقيقة، وذلك عند اختيار أهم 6 ميزات وضبط البارامترات الفائقة بالشكل التالي: C: 1000، Gamma: 0.01، kernel: "rbf". أن أفضل النتائج لهذه الخوارزمية AUCPR: 0.82-F1: 0.83-MCC: 0.83. أما بالنسبة لخوارزمية غابة العزل فقد حققت أداء جيد عن اختيار أهم 4 ميزات وضبط البارامترات الفائقة بالشكل التالي: عدد الأشجار: 100، العدد الأعظمي للعينات: 300، التلوث: 0.00178. أن أفضل النتائج لهذه الخوارزمية AUCPR: 0.71-F1: 0.75-MCC: 0.74 وبزمن تنفيذ 18 دقيقة.

وبناء على النتائج التي تم التوصل إليها يمكن أن نقدم بعض المقترحات:

- توسيع نطاق العمل ليشمل خوارزميات التعلم العميق ومجالات أخرى للكشف.
- محاولة التوسع في الدراسة لتشمل مجالات أخرى في كشف الشذوذ.
- السعي لتطبيق الدراسة على بيانات الشركات والمؤسسات المالية السورية.

## المراجع

- [1] How Much Data Is on the Internet? [Online] Available: [seedsscientific.com](https://seedsscientific.com) [Accessed 1 Jul. 2021]
- [2] Zenati, H., Romain, M., Foo, C. S., Lecouat, B., & Chandrasekhar, V. (2018, November). Adversarially learned anomaly detection. In 2018 IEEE International conference on data mining (ICDM) (pp. 727-736).
- [3] Li, L., Jamieson, K., Rostamizadeh, A., Gonina, E., Hardt, M., Recht, B., & Talwalkar, A. (2018). Massively parallel hyperparameter tuning.
- [4] Florea, A. C., & Andonie, R. (2020). Weighted random search for hyperparameter optimization. arXiv preprint arXiv:2004.01628.
- [5] Van Der Maaten, L., Postma, E. and Van den Herik, J., (2009). Dimensionality reduction: a comparative. J Mach Learn Res, 10(66-71).
- [6] Tang, J., Alelyani, S. and Liu, H., 2014. Feature selection for classification: A review. Data classification: Algorithms and applications, p.37.
- [7] Credit card fraud detection anonymized credit card transaction labeled as fraudulent or genuine [Online] Available: <https://www.kaggle.com/mlg-ulb/creditcardfraud>. [Accessed 1 Jul. 2021]
- [8] Python 3.10.0 documentation [Online] Available: <https://docs.python.org/3/> [Accessed 1 May. 2021]
- [9] Chicco, D. (2017). Ten quick tips for machine learning in computational biology. BioData mining, 10(1), 1-17.
- [10] Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision support systems, 50(3), 559-569.
- [11] Zareapoor, M., & Shamsolmoali, P. (2015). Application of credit card fraud detection: Based on bagging ensemble classifier. Procedia computer science, 48(2015), 679-685.
- [12] Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., & Jiang, C. (2018, March). Random forest for credit card fraud detection. In 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC) (pp. 1-6). IEEE.
- [13] Ounacer, S., El Bour, H. A., Oubrahim, Y., Ghomari, M. Y., & Azzouazi, M. (2018). Using Isolation Forest in anomaly detection: the case of credit card transactions. Periodicals of Engineering and Natural Sciences (PEN), 6(2), 394-400.
- [14] Makki, S., Assaghir, Z., Taher, Y., Haque, R., Hacid, M. S., & Zeineddine, H. (2019). An experimental study with imbalanced classification approaches for credit card fraud detection. IEEE Access, 7, 93010-93022.
- [15] Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., & Anderla, A. (2019, March). Credit card fraud detection-machine learning methods. In 2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-5). IEEE.

- [16] Lakshmi, S. V. S. S., & Kavilla, S. D. (2018). Machine learning for credit card fraud detection system. *International Journal of Applied Engineering Research*, 13(24 Pt. 1), 16819-16824.
- [17] Raghavan, P., & El Gayar, N. (2019, December). Fraud detection using machine learning and deep learning. In *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)* (pp. 334-339). IEEE.
- [18] Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In *Ensemble machine learning* (pp. 157-175). Springer, Boston, MA.
- [19] Staerman, G., Mozharovskiy, P., Cl  men  on, S., & d'Alch  -Buc, F. (2019, October). Functional isolation forest. In *Asian Conference on Machine Learning* (pp. 332-347). PMLR.
- [20] "Anomaly detection using Isolation Forest – A Complete Guide" [online]. <https://www.analyticsvidhya.com>. [Accessed 26 July 2021].
- [21] "Isolation Forest Algorithm for Anomaly Detection" [online]. <https://heartbeat.comet.ml>. [Accessed Oct 2021]
- [22] F. Cady (2017)-*The Data Science Handbook*, 1st edition, NJ: John Wiley and Sons Ltd
- [23] Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).
- [24] Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*.



## تحسين أداء الشبكات المعرفة برمجياً ذات

### المتحكمات المتعددة من خلال خوارزمية

#### موازنة الحمل

د. أحمد صقر أحمد \* حازم ديب \*\*

##### ملخص البحث:

تم اقتراح الشبكات المعرفة برمجياً (SDN) Software Defined Network كحل لمعالجة المشاكل التي تعاني منها الشبكات التقليدية، ولكن النموذج الأولي المقترح لشبكات SDN، والذي يعتمد على وجود متحكم مركزي وحيد، عانى من مشكلة رئيسية هي مشكلة نقطة الفشل الوحيدة، هذا الأمر دفعنا إلى الانتقال إلى بنية شبكات SDN بمتحكمات متعددة، وقد واجهت هذه البنية الجديدة تحديات عديدة أهمها هي عملية موازنة الحمل بين المتحكمات المتعددة الموجودة في الشبكة، بحيث لا يتعرض أحد هذه المتحكمات إلى حمل زائد بشكل يجعله يتوقف عن العمل، لذا فقد كان لعملية اختيار خوارزمية موازنة حمل مناسبة دور كبير في أداء الشبكة.

تم التطرق في هذا البحث إلى الحاجة إلى الشبكات المعرفة برمجياً وبنيتها، والحاجة إلى الانتقال من البنية المركزية إلى البنية الموزعة، مع ذكر فوائد البنية الموزعة والتحديات التي تتعرض لها، والنماذج المختلفة منها، كما تم أيضاً تعريف عملية موازنة الحمل، والتعرف على آليتها، والعوامل التي تلعب دوراً في هذه العملية، بالإضافة إلى أنه تم ذكر أهم خوارزميات موازنة الحمل المستخدمة في وقتنا الحالي.

وفي الجانب العملي، تم اقتراح خوارزمية جديدة لموازنة الحمل وقد تمت تسميتها بخوارزمية الترحيل باستخدام العتبات الثلاث، حيث تمت مقارنة هذه الخوارزمية مع مجموعة من خوارزميات موازنة الحمل المستخدمة سابقاً، من خلال ثلاثة عوامل هي: إنتاجية الشبكة، معدل استخدام الموارد، واحتمال فشل المتحكم، وقد بينت نتائج التجارب التي تم إجراؤها أن الخوارزمية المقترحة تعطي نتائج أفضل من الخوارزميات السابقة،

وتقدم تحسناً مقبولاً على جميع الخوارزميات الأخرى وفقاً لجميع العوامل والسيناريوهات التي تمت دراستها في هذا البحث.

**الكلمات المفتاحية:** الشبكات المعرفة برمجياً، موازنة الحمل، إنتاجية الشبكة، معدل استخدام الموارد.

---

\*أستاذ-قسم النظم والشبكات الحاسوبية-كلية الهندسة المعلوماتية-جامعة تشرين-اللاذقية-سوريا.

Email: [ahmad.s.ahmad@tishreen.edu.sy](mailto:ahmad.s.ahmad@tishreen.edu.sy)

\*\*طالب دكتوراه-قسم النظم والشبكات الحاسوبية-كلية الهندسة المعلوماتية-جامعة تشرين-اللاذقية-

سوريا. Email: [hazem.deeb@tishreen.edu.sy](mailto:hazem.deeb@tishreen.edu.sy)

# Improvement of Multiple Software Defined Networks Controllers with Load Balancing Algorithm

\*Dr. Ahmad Saker Ahmad

\*\*Hazem Deeb

## Abstract

Software Defined Networks were proposed as a solution for the problems that traditional networks suffer from. But the first proposed SDN prototype which depends on having a single main controller suffered from a main problem of having single point of failure, which is what force us to go with multiple controllers SDN, this new architecture faced multiple challenges such as load balancing between the different controllers in the network, so that no controller gets overloaded to the point which makes it stop working. So choosing the correct load balancing algorithm had a major rule in the performance of the network.

This research discussed the need for software defined networks, their infrastructures and the need to move from the central infrastructure to the distributed one with mentioning the benefits of distributed infrastructures, the challenges that it's facing and the different types that it takes. The load balancing process, its mechanism and the factors that play an important role in this process were introduced. In addition to that, the most common load balancing algorithms used nowadays were also mentioned in this research.

A new algorithm was proposed to balance the load, which was called "Three thresholds migration algorithm", this algorithm was compared to multiple previously used load balancing

algorithms by using three factors: network throughput, resource utilization, and controller failure probability. The experiments results show decent improvement over the other algorithms based on the different factors and scenarios studied in this research.

**Keywords:** SDN, Load balancing, Throughput, Resource Utilization.

---

**\*Professor, Department of Computer Networks & Systems, Faculty of Informatics Engineering Faculty, Tishreen University, Lattakia, Syria.**

**Email:** [ahmad.s.ahmad@tishreen.edu.sy](mailto:ahmad.s.ahmad@tishreen.edu.sy)

**\*\* PhD Student, Department of Computer Networks & Systems, Faculty of Informatics Engineering Faculty, Tishreen University, Lattakia, Syria.**

**Email:** [hazem.deeb@tishreen.edu.sy](mailto:hazem.deeb@tishreen.edu.sy)

## 1- مقدمة:

كانت الفكرة الأساسية من شبكات SDN هي معالجة المشكلات التي تعاني منها الشبكات التقليدية، فقد اتسمت الشبكات التقليدية بقلّة المرونة وعدم القدرة على التكيف عند ازدياد حجم الشبكة بشكل كبير، وذلك لأنها تعتمد على عملية إدخال الإعدادات يدوياً لكل جهاز شبكي موجود في الشبكة، وعند حصول تعديل على بنية الشبكة مثل إضافة جهاز جديد وجب إدخال كم كبير من الإعدادات، وهذا يعتبر عملية معقدة للغاية وخاصةً في الشبكات الكبيرة، لذا تم اقتراح الانتقال إلى شبكات SDN والتي تهدف إلى جعل الشبكات أكثر مرونة، وتسهيل عملية الإدارة بفصل مستوي التحكم عن مستوي البيانات، وجعل مستوي التحكم موجود في جهاز واحد هو متحكم SDN المركزي الموجود في الشبكة. بدأت بنية شبكات SDN بوجود عدد من الأجهزة الشبكية مع متحكم SDN مركزي وحيد للشبكة ككل، وتتم جميع عمليات التحكم واتخاذ القرارات في هذا المتحكم، ولكن مع مرور الوقت وزيادة حجم الشبكات أصبح هذا المتحكم الوحيد يتعرض لحمل زائد، مما قد يسبب في توقفه عن العمل، وبما أنه لا يوجد بديل لهذا المتحكم سوف تظهر مشكلة نقطة الفشل الوحيدة في هذا النوع من الشبكات، وقد كان هذا الدافع الرئيسي للانتقال إلى بنية شبكات SDN موزعة بمتحكمات متعددة، حيث أتاحت هذه البنية تجنب المشاكل الحاصلة في الشبكات ذات المتحكم الوحيد.

## 2- أهمية البحث وأهدافه:

لقد برزت الحاجة الملحة للانتقال من بنية شبكات SDN ذات المتحكم المركزي إلى بنية ذات متحكمات متعددة، وذلك لمعالجة المشاكل التي واجهتها شبكات النوع الأول، ولكن هذا الانتقال لم يكن سهلاً، فعند محاولة الانتقال إلى البنية ذات المتحكمات المتعددة برزت مجموعة من المشاكل والتحديات من أهمها: عدد المتحكمات اللازمة لإعطاء أحسن أداء في الشبكة، وأماكن توضع هذه المتحكمات، بالإضافة إلى ضرورة وجود آلية فعالة لتوزيع الحمل بين هذه المتحكمات، أي طريقة تخصيص المهام الواردة إلى الموارد الموجودة بالشكل الذي يضمن أكبر فعالية في آلية عمل الشبكة، لذا حاولنا التطرق في هذا البحث إلى خوارزميات موازنة الحمل المستخدمة حالياً في بيئات SDN ذات

المتحكمات المتعددة، مع محاولة دراسة إيجابيات وسلبيات كل من الخوارزميات السابقة، والاستفادة من هذه الدراسة لاقتراح خوارزمية موازنة حمل جديدة، ومقارنتها مع الخوارزميات المستخدمة سابقاً، بهدف تحسين أداء الشبكة من ناحية ثلاثة عوامل رئيسية تمت دراستها في هذا البحث هي: إنتاجية الشبكة، معدل استخدام الموارد، واحتمال فشل المتحكم.

### 3- طرائق البحث ومواده:

يتطرق هذا البحث إلى الحاجة إلى الانتقال من الشبكات التقليدية إلى شبكات SDN، وبنية شبكات SDN، وعدم مواكبة البنية التقليدية المركزية لشبكات SDN للتطور الحاصل في عالم الشبكات، والحاجة إلى الانتقال إلى شبكات SDN بمتحكمات متعددة، مع ذكر الفوائد التي تحققها هذه البنية الموزعة، بالإضافة إلى أهم التحديات التي تتعرض لها، والنماذج المختلفة لشبكات SDN ذات المتحكمات المتعددة، كما تم التطرق إلى تعريف آلية موازنة الحمل، والعوامل التي تؤثر على هذه الآلية، وذكر أهم الخوارزميات المستخدمة في موازنة الحمل في الوقت الراهن، وأخيراً تم اقتراح آلية موازنة حمل جديدة ومقارنتها مع الخوارزميات المستخدمة سابقاً، وتم التطرق إلى التجارب التي أجريناها، والنتائج التي حصلنا عليها، وتوضيح التحسين الذي قدمته هذه الخوارزمية المقترحة على أداء الشبكة .

### 4- الحاجة إلى الشبكات المعرفة برمجياً:

برزت الحاجة في السنوات الأخيرة إلى مفهوم جديد للشبكات التقليدية بهدف معالجة المشاكل التي تعاني منها، كان الهدف الأساسي لشبكات SDN هو تبسيط عملية إدارة الشبكة وإضافة المزيد من المرونة إليها بالمقارنة مع الشبكات التقليدية. وقد تمت ملاحظة أربع مشاكل أساسية تعاني منها الشبكات التقليدية يمكن تلخيصها في:

1. **التعقيد وإدخال الإعدادات الشبكية بشكل يدوي:** تعتبر عملية إدخال الإعدادات الشبكية هي عملية معقدة في الشبكات التقليدية، حيث إن كل جهاز شبكي يتم إدخال الإعدادات إليه بشكل يدوي وبالطريقة التي تناسب الشركة المصنعة لهذا الجهاز، وقد تختلف هذه العملية بين جهاز وآخر على نفس الشبكة تبعاً للشركة المصنعة له،

بالإضافة إلى أن بنية الشبكة تتغير بشكل متكرر، فمن الممكن أن تتعرض إحدى الوصلات للفشل، أو أن نضطر إلى إضافة أو إزالة بعض الأجهزة الشبكية، كل هذه الحالات تضطرننا إلى إدخال الكثير من الإعدادات الشبكية بشكل يدوي وهو ما يعتبر عملية معقدة ستقودنا إلى حدوث مزيد من الأخطاء في إعدادات الشبكة [1].

**2. الحالة الديناميكية للشبكة:** يزداد حجم الشبكات بشكل مطرد يوماً بعد يوم، مما يزيد من تعقيد عمليات إدارة الشبكات، علاوةً على ذلك فإن ظهور تقنيات حوسبة جديدة مثل الحوسبة السحابية والضبابية وغيرها، بالإضافة إلى تقنيات الشبكات الافتراضية جعل الشبكات تصبح أكثر ديناميكية لأن المضيفين يتحركون باستمرار ويتصلون ويغادرون بشكل متكرر، هذا جعل حركة المرور وظروف الشبكة تتغير بشكل أسرع وبطريقة ملحوظة [2].

**3. اختلاف مستويات التعقيد:** تتطلب الشبكات الكبيرة في يومنا الحالي القيام بالعديد من المهام ذات مستوى تعقيد عالٍ، وهذه المهام يتم تنفيذها في الشبكات التقليدية عن طريق واجهات ذات مستوى تعقيد منخفض وميزات أقل، وهذا بسبب الترابط الوثيق بين طبقة الإدارة وطبقة البيانات، حيث تكون العديد من مزايا الإدارة متضمنة في عتاديات الأجهزة الشبكية.

**4. عدم التجانس في أجهزة الشبكة:** تتكون الشبكات الحالية من عدد كبير من الأجهزة غير المتجانسة مثل الموجهات والمبدلات وغيرها، ويملك كل جهاز من هذه الأجهزة تكوينها الخاص وإعداداتها الخاصة وكل منها يعمل وفقاً لبروتوكولات محددة قد تختلف من جهاز إلى آخر، مما يزيد من تعقيد إدارة الشبكة ويظهر عدم كفاءة هذه العملية.

أوضحت هذه الأسباب جميعها أن إدارة الشبكة عملية تزداد تعقيداً يوماً بعد آخر، وخاصةً في الشبكات التقليدية التي لا تتسم بالمرونة الكافية للتعامل مع الحاجات المتزايدة للشبكات في الوقت الحالي والتي يجب أن تتصف بالديناميكية العالية. هذا الأمر دفع بمنظمة ONF لاقتراح نموذج جديد هو نموذج الشبكات القابلة للبرمجة SDN [3].

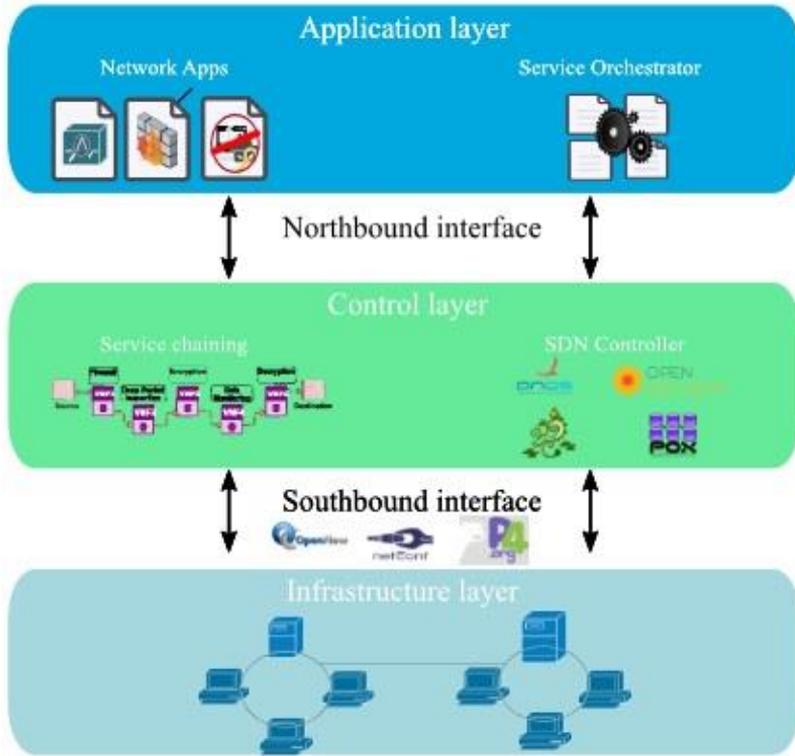
## 5- بنية الشبكات المعرفة برمجياً:

تعتمد شبكات SDN بشكل رئيسي على فصل طبقة التحكم عن طبقة البيانات، أي بشكل أساسي نقل عمليات اتخاذ القرار من العناصر الشبكية والتي هي عتاديات وتجميعها في عنصر مركزي وحيد هو المتحكم المبني على أساس برمجي، يتم اتخاذ كافة القرارات في المتحكم المركزي مما يجعل عملية إدارة الشبكة أكثر سهولة.

تتألف بنية شبكات SDN من ثلاث طبقات رئيسية و3 مكونات فرعية:

- **طبقة التطبيقات:** تحوي العديد من البرامج المسؤولة عن عمليات إدارة الشبكة وتهدف هذه الطبقة إلى تطوير العديد من التطبيقات وذلك من خلال الاستفادة من المعلومات المتوافرة كحالة الشبكة والمعلومات والإحصائيات المتوافرة حولها.
- **طبقة التحكم:** هي الطبقة الأذكى والأكثر أهمية في شبكات SDN والتي تحوي المتحكم الوحيد أو مجموعة المتحكمات الموجودة في شبكات SDN وهي التي تقوم باتخاذ القرارات في الشبكة.
- **طبقة البيانات أو طبقة البنية التحتية:** هي التي تحوي عدة أنواع من الأجهزة الشبكية من موجّهات ومبدلات وغيرها وظيفتها الأساسية هي المساعدة في تمرير التدفقات وتنفيذ القرارات المتخذة في طبقة التحكم [4].
- ✓ **الجسر الشمالي:** هو صلة الاتصال بين طبقة التحكم وطبقة الإدارة والذي تكون عادة مجموعة من واجهات برمجة التطبيقات API.
- ✓ **الجسر الجنوبي:** هو صلة الاتصال بين طبقة التحكم وطبقة البيانات والذي يكون عادةً عبارة عن بروتوكول يتم به التخاطب وهو في أغلب الأحيان بروتوكول OpenFlow
- ✓ **الجسرين الشرقي والغربي:** هما صلة الاتصال التي يتم فيها التخاطب بين المتحكمات المتعددة في طبقة التحكم وعادة تكون عبارة عن بروتوكولات شبكية لتبادل الرسائل مثل Border Gateway Protocol (BGP) و Open Shortest Path First (OSPF) [5].

حيث يبين الشكل (1) بنية الشبكات المعرفة برمجياً:



الشكل (1) بنية الشبكات المعرفة برمجياً.

## 6- الحاجة إلى البنية الموزعة بمتحكمات متعددة:

تطبق هيكلية شبكات SDN مبدأ التحكم المركزي من الأعلى إلى الأسفل، في هذه الحالة فإن متحكم SDN مع تطبيقات SDN المثبتة تعمل معاً لتتحكم وتنظم عمل المبدلات في الشبكة. عند تطبيق التحكم المركزي فإن هذا التصميم يعطينا العديد من الفوائد مثل إمكانية إدارة الشبكة بكفاءة والتعامل مع التغيرات الديناميكية في بنية الشبكة بشكل سهل وسريع. هناك تحقيقان ممكنان للوصول إلى تصميم التحكم المركزي، أولهما من خلال متحكم وحيد في بنية مركزية، والثاني من خلال عدة متحكمات في بنية موزعة. عند وجود متحكم وحيد في بنية مركزية يوجد متحكم SDN وحيد في الشبكة ككل، يقوم هذا المتحكم بجميع عمليات اتخاذ القرارات في الشبكة ويتحكم بطريقة مركزية بجميع المبدلات

في الشبكة بما أنها جميعها متصلة به. العديد من تطبيقات SDN سابقاً استخدمت هذا النموذج ولكنه للأسف عانى من مشكلتين أساسيتين لا يمكن إغفالهما وهما:

❖ **قابلية التوسع:** تعرف قابلية التوسع بأنها قدرة الشبكة على النمو بشكل يليبي الازدياد المطرد في أعداد المستخدمين وطلباتهم، وتعرف قابلية التوسع في شبكات SDN بأنها قدرة المتحكم على معالجة الطلبات الواردة من المبدلات المتصلة فيه. تملك متحكمات SDN موارد محدودة عند معالجة أعداد كبيرة من الطلبات مما يمنع من معالجة جميع الطلبات الواردة. حاول الباحثون حل هذه المشكلة من خلال تقييد عدد الطلبات الواردة إلى المتحكم، ولكن هذه الاستراتيجية لم تعمل بشكل جيد وخاصةً أنها تعتمد على منع خروج الطلبات من المبدلات، مما يعني أن جزءاً من عملية اتخاذ القرار عاد إلى المبدلات، وهذا الأمر كان ضد الفكرة الأساسية من شبكات SDN والتي تقوم على نقل جميع عمليات اتخاذ القرارات إلى المتحكم، لذا كان لا بد من وجود حل آخر وقد كان الحل الوحيد هو إضافة عدد من المتحكمات إلى الشبكة مما دفعنا بالحاجة للانتقال إلى البنية الموزعة بمتحكمات متعددة.

❖ **نقطة الفشل الوحيدة:** عند وجود متحكم SDN مركزي وحيد سوف يعاني من نقطة الفشل الوحيدة، وفي حال فشل المتحكم المركزي فإن جميع المبدلات سوف تفقد القدرة على تمرير الحزم الجديدة الواردة، وفي النهاية سوف تتعطل الشبكة بشكل كامل. وقد تم اقتراح حلين، كان الحل الأول يعتمد على أنه في حال حصول نقطة الفشل الوحيدة يعود اتخاذ القرارات إلى المبدلات، وتنتقل الشبكة من شبكة SDN لتعود إلى شبكة تقليدية، ولكن هذا الحل كان مكلفاً لأنه في حال العودة إلى الشبكة التقليدية فإن المشاكل المتعلقة بمثل هذا النوع من الشبكات، والتي ذكرناها سابقاً سوف تعود إلى الواجهة. بينما يتمثل الحل الثاني باعتماد شبكات SDN ذات بنية موزعة مع متحكمات متعددة وقد كان هذا الحل هو الحل الأفضل [6].

## 7- فوائد البنية الموزعة:

في شبكات SDN الكبيرة يمكن أن نحتاج إلى تقسيم الشبكة إلى مجالات مختلفة أصغر، عندها يحتاج كل مجال من هذه المجالات إلى متحكم واحد على الأقل ليشرف عليه، في هذا السيناريو فإننا نحتاج إلى عدة متحكمات لإدارة الشبكة بشكل كامل. إن وجود عدة متحكمات في الشبكة يساهم في زيادة استقرار الشبكة حيث إن المبدلات من الممكن أن تستمر بالعمل حتى في حال فشل أحد المتحكمات أو إحدى الوصلات إلى هذه المتحكمات، يمكن تحقيق البنية الموزعة بحيث يكون التحكم مركزي بشكل منطقي ولكنه موزع بشكل فيزيائي [7]، وهناك مجموعة من الفوائد الإضافية تتلخص في:

أ . **سهولة الإدارة:** تعتبر الإدارة الفعالة عاملاً مهماً لاستخدام متحكمات متعددة، حيث إن قدرات المتحكم الوحيد تعتبر محدودة بالمقارنة مع وجود متحكمات متعددة، حيث إنه من الصعب على متحكم وحيد معالجة جميع الطلبات الواردة في الشبكات الكبيرة ذات المجالات المختلفة، بالنظر إلى الميزات المختلفة الموجودة في عدة متحكمات، فإن وجود عدد من المتحكمات يمكن أن يطور من كفاءة إدارة الشبكة وخاصة في الشبكات الكبيرة ذات المجالات المتعددة [8].

ب. **قابلية التوسع:** إن نهج OpenFlow الحالي سوف يؤدي إلى زيادة الحمل بشكل كبير في حال وجود متحكم وحيد، مما يقلل من قابلية التوسع في الشبكة، بينما البنية الموزعة في حال وجود متحكمات متعددة يمكن أن تكون حلاً مثالياً، حيث يمكن إزالة وإضافة المتحكمات من وإلى الشبكة بسهولة كبيرة وبشكل ديناميكي، بالإضافة إلى أنه يمكن أن نقوم بتوزيع الحمل بين المتحكمات المختلفة لتجنب مشكلة عنق الزجاجة الموجودة في المتحكم المركزي الوحيد. كما أن الهيكلية الموزعة يمكن أن تكون تصميمًا مناسباً لدعم عدد أكبر من المستخدمين.

ج. **تقليل زمن الاستجابة:** هناك العديد من التطبيقات الحساسة لزمن الاستجابة، فإذا تمت معالجة جميع الطلبات من قبل متحكم وحيد، هذا سوف يؤدي إلى زيادة الحمل

على هذا المتحكم وسوف تتم معالجة الطلبات بشكل متسلسل الواحد تلو الآخر، مما سيؤدي إلى زيادة زمن استجابة هذه الطلبات بشكل كبير، وعند وجود متحكمات متعددة يمكن توزيع الحمل بينها أو تمرير الطلبات إلى المتحكم الأكثر ملاءمة مما يقلل من زمن الاستجابة للطلبات بشكل ملحوظ.

د . **المثانة:** إن بنية شبكات SDN بمتحكمات متعددة توفر العديد من نقاط الاستعادة والنسخ الاحتياطي من أجل تجنب مشكلة نقطة الفشل الوحيدة الموجودة في شبكات SDN ذات المتحكم الوحيد، وهذا سوف يزيد بالضرورة من مثانة الشبكة، عند فشل أحد المتحكمات يمكن للمتحكمات الباقية معالجة هذه المشكلة بتحمل العبء الناتج عن فشل هذا المتحكم من خلال توزيع حمله على باقي المتحكمات، حيث تتم معالجة هذه المشكلة بزيادة طفيفة على أحمال باقي المتحكمات، بالشكل الذي يسمح بالحفاظ على استقرار الشبكة. إن المتحكمات المتعددة يمكن أن تقلل بشكل كبير من احتمالية فشل الشبكة وخاصة في حال الأخطاء الناتجة عن فشل المتحكمات [9].

#### 8- تحديات البنية الموزعة:

في بيئة SDN تعمل المتحكمات مع المبدلات جنباً إلى جنب لرفع مستوى أداء الشبكة، وخاصة الشبكات الكبيرة، ولرفع هذا المستوى كان من الأهمية بمكان تحديد تصميم مناسب لأماكن توضع المتحكمات المتعددة في الشبكة، والذي يعتبر من أهم القضايا الشائكة في شبكات SDN ذات المتحكمات المتعددة إلى يومنا هذا، بالإضافة إلى أماكن توضع المتحكمات يوجد قضايا أخرى مثل: المرونة وقابلية التوسع والأمن والثبات وتقليل زمن الاستجابة، وتعتبر هذه القضايا هي الأخرى قضايا مهمة لرفع مستوى أداء الشبكة [10].

يمكن تلخيص أهم التحديات بالنقاط التالية:

- أ . الإدارة المركزية: تم تطوير المتحكمات المتعددة في شبكات SDN الكبيرة لموازنة الحمل في الشبكات، وعند زيادة عدد المتحكمات في الشبكة سوف يختلف مفهوم الإدارة المركزية، يجب استخدام متحكمات متعددة لتحقيق جودة خدمة عالية ، وذلك لأن سعة المتحكمات المتعددة تكون أكبر من سعة متحكم واحد، ولكن بالمقابل فإن المتحكمات المتعددة تملك مزايا مختلفة وبالتالي تتطلب طرق متعددة للتعامل معها، الأمر الذي يعتبر تحدياً بالنسبة لمطوري الشبكات.
- ب . قابلية التوسع: بالرغم من أن البنية الموزعة تسمح بإضافة المتحكمات وزيادة حجم الشبكة بسهولة، إلا أن تحقيق عملية موازنة الحمل بين هذه المتحكمات تعتبر تحدياً مهماً لتحقيق جودة خدمة عالية.
- ج . الاختلاف بين المتحكمات المتعددة: بسبب وجود عدد من المتحكمات فإن المشكلة الرئيسية هي عملية المزامنة بين هذه المتحكمات، ولذلك فإن عملية تصميم المتحكمات ومزامنتها تعتبر عملية معقدة [11].
- د . توضع المتحكمات: إن وجود متحكم واحد في الشبكة له العديد من الميزات كسهولة الإدارة والتحكم ومراقبة الشبكة بشكل كامل من قبل نقطة وحيدة هي المتحكم المركزي. ولكن في نفس الوقت توجد مشاكل بالمقابل تتعلق بالثبات وقابلية التوسع، وبالتالي كلما زاد حجم الشبكة كلما أصبح الأداء أسوأ. ولكن عند حل هذه المشكلة باعتماد بنية SDN موزعة بوجود متحكمات متعددة تبرز مشكلة رئيسية هي اختيار عدد المتحكمات المطلوب وضعها في الشبكة، بالإضافة إلى اختيار أماكن توضع هذه المتحكمات والمسافة بين كل منها، ويمكن اعتبار المشكلتين الرئيسيتين في شبكات SDN ذات المتحكمات المتعددة تتلخص في سؤالين: الأول هو كم عدد المتحكمات المطلوبة في الشبكة لتحقيق أفضل أداء؟ والثاني هو في أي موضع من

الشبكة يجب وضع هذه المتحكمات؟ هذه المشاكل تعتبر معقدة جداً ومن الصعب الإجابة عليها، وهي تأخذ الحيز الأكبر من وقت الباحثين في وقتنا الراهن [12].

هـ . بروتوكولات الاتصال المختلفة: إن وجود عدة متحكمات مطلوب في البنية الموزعة ويؤثر بشكل رئيسي على جودة الخدمة في شبكات SDN، لحل مشكلة الاتصال بين هذه المتحكمات تم تطوير واجهات تدعى بالواجهات الشرقية والغربية والتي تتضمن مجموعة من البروتوكولات المستخدمة في عملية التخابط بين المتحكمات، لذا فإن إيجاد لغة برمجة وبروتوكولات مشتركة بين المتحكمات يعتبر تحدياً إضافياً في شبكات SDN.

و . عملية الجدولة بين المتحكمات المختلفة: تتحسن جودة الخدمة بوجود متحكمات متعددة، ولكن التحدي هنا هو عملية جدولة الطلبات بين هذه المتحكمات بالشكل الذي يعطي أفضل أداء للشبكة، ودون وجود حمل زائد على أي من المتحكمات في الشبكة. وكلما كانت عملية اتخاذ قرار الجدولة أسرع زادت جودة الخدمة [13].

### 9- نماذج الشبكات المعرفة برمجياً ذات المتحكمات المتعددة:

من الناحية النظرية في حال وجود متحكم وحيد في شبكة SDN يمكن لهذا المتحكم أن يتحكم بجميع المبدلات الموجودة في الشبكة فجميعها متصلة به، كما أنه يملك رؤية عامة للشبكة ككل، لذا في حالة المتحكمات المتعددة وللحصول على تحكم مركزي منطقياً لا بد من وجود متحكم واحد على الأقل يملك رؤية عامة للشبكة ككل في كل مجال من مجالات الشبكة، لذا فإن تبادل معلومات الشبكة المحلية بين متحكمات SDN المختلفة يعتبر أمراً لا مفر منه. يوجد نموذجان مستخدمان لتبادل معلومات الشبكة بين متحكمات SDN، وكل نموذج يقوم بتبادل هذه المعلومات بطريقة مختلفة عن الآخر.

❖ **النموذج الهرمي:** في هذا النموذج يكون أحد أو بعض المتحكمات (ولكن بالضرورة ليس جميع المتحكمات) في المجال تملك رؤية عامة للشبكة ككل. ويعرف هذا

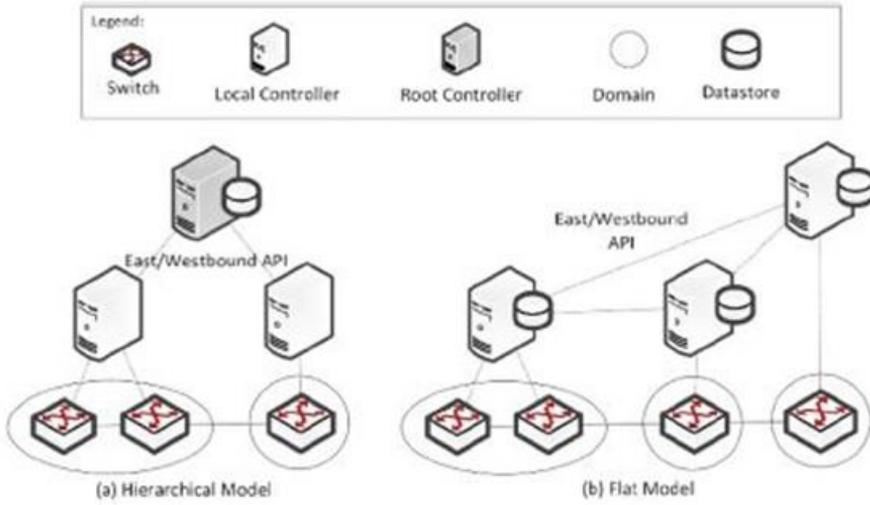
النموذج أيضاً بالنموذج العمودي لأن عمليات التحكم تتم من أعلى إلى أسفل عبر متحكمات SDN، يوضح الشكل (2) هذا النموذج، حيث يوجد متحكمي SDN محلين في كل مجال، بينما يوجد متحكم جذر هو المسؤول عن عمليات التنسيق بين المتحكمات المحلية. تكون معلومات الشبكة الكاملة مخزنة في المتحكم الجذر، لذا يجب على المتحكمات المحلية أن تطلب استعلام من المتحكم الجذر للحصول على المعلومات اللازمة قبل أي عملية اتخاذ قرار داخل مجالها المحلي، بسبب هذه الفكرة يمكن أن يدعى هذا النموذج بنموذج مخدم-زبون، حيث يعمل المتحكم الجذر كمخدم، وتعمل المتحكمات المحلية كزبائن. يمكن وضع أكثر من متحكم جذر وذلك لتجنب نقطة الفشل الوحيدة في المتحكم الجذر. لا تملك المتحكمات المحلية في هذا النموذج جسراً شرقياً وغربياً، أي أنها لا تملك وصلات مع باقي المتحكمات المحلية، ولكنها تملك وصلة واحدة فقط مع المتحكم الجذر الذي يقوم بعمليات التنسيق فيما بينها، ويحتفظ المتحكم الجذر بنسخة أو رؤية عامة عن بيانات الشبكة بشكل كامل، بينما تملك المتحكمات المحلية رؤية محلية للشبكة من منظورها الخاص فقط.

❖ **النموذج المسطح:** في هذا النموذج تملك جميع المتحكمات في المجال رؤية عامة عن حالة الشبكة، يعرف هذا النموذج أيضاً بالنموذج المسطح بما أنه لا يوجد معمارية هرمية بين متحكمات SDN، يوضح الشكل (2) هذا المبدأ حيث نلاحظ وجود ثلاث متحكمات SDN في المجال، هذه المتحكمات الثلاث تتحكم بالمبدلات، وتملك جميعها رؤية عامة عن حالة الشبكة، كما تملك هذه المتحكمات جسرين شرقي وغربي مما يسمح بعملية الاتصال فيما بينها بشكل مباشر. يجب على أي متحكم قبل البدء بعملية إنشاء حالة الشبكة من استقبال المعلومات عن حالة الشبكة من جميع المتحكمات الموجودة في المجال، وعند حصول أي تغيير في مجال أحد المتحكمات يجب أن تتم مشاركة هذا التغيير إلى جميع المتحكمات الباقية بالشكل

الذي يسمح لهذه المتحكمات بتحديث معلومات الشبكة لديهم وملاحظة هذه التغييرات، إذا تمت هذه العملية بشكل صحيح وبدون حدوث مشاكل هذا سوف يؤدي بالضرورة إلى أن جميع المتحكمات في الشبكة تملك نفس الرؤية العامة عن حالة الشبكة. بسبب هذه الفكرة يمكن أن يدعى هذا النموذج أيضاً بنموذج الند للند حيث يمكن لكل متحكم الوصول إلى المتحكمات الأخرى بشكل مباشر أثناء مشاركة معلومات الشبكة.

هناك طريقتان لتبادل معلومات الشبكة بين متحكمات SDN المختلفة في الشبكة، وهاتان الطريقتان هما:

- **الطلب الدوري:** في هذه الحالة يقوم كل متحكم بطلب المعلومات بشكل دوري من باقي المتحكمات في الشبكة، على سبيل المثال يقوم المتحكم بطلب معلومات جديدة عن حالة الشبكة من باقي المتحكمات كل 10 دقائق، سوف تتم عملية الطلب هذه بشكل دوري حتى لو لم يطرأ أي تغيير على الشبكة، هذا الأمر قد يؤدي إلى استلام نفس المعلومات أكثر من مرة مما يجعل هذه الطريقة غير فعالة بشكل كافي.
- **النشر والاشتراك:** حيث يقوم كل متحكم SDN بعملية نشر المعلومات الموجودة لديه إلى المشتركين، وعملية الاشتراك بالحصول على المعلومات من باقي المتحكمات في الشبكة، على سبيل المثال إذا كان أحد المتحكمات وليكن C1 يحتاج إلى معلومات من متحكم جار هو C2، عندها يجب على المتحكم C1 الاشتراك بنشرة المعلومات الصادرة عن C2، في هذه الحالة يعتبر C1 كمشارك و يعتبر C2 كناشر، وفي وقت لاحق سوف يقوم C2 بإعلام C1 عند حصول تغييرات في مجال C2، هذا يعني أن C2 يقوم بإعلام C1 فقط في حال حدوث تغييرات وليس بشكل دوري، مما يوضح أن هذه الطريقة أكثر فعالية من الطريقة الأولى [14].



الشكل (2) النموذج الهرمي و النموذج المسطح

### 10- موازنة الحمل:

تعرف موازنة الحمل بأنها العملية التي يتم توزيع مجموعة من المهام على مجموعة من الموارد وذلك بهدف جعل عملية معالجة هذه المهام أكثر كفاءة، حيث يمكن لهذه العملية أن تقلل من زمن الاستجابة للطلبات الواردة وتجنب الحالات التي تتعرض فيها بعض العقد في الشبكة إلى حمل زائد، بينما تكون هناك عقد أخرى في حالة سكون ( Idle State).

### 10-1 آلية موازنة الحمل:

كما ذكرنا سابقاً يمكن تعريف موازنة الحمل بأنها العملية التي يتم توزيع وتخصيص الطلبات والمهام الواردة من الزبائن على الموارد المتاحة في الشبكة بفعالية وبالشكل الذي يضمن تخديم هذه الطلبات بأسرع طريقة ممكنة. تهدف عملية موازنة الحمل إلى تحسين نسبة استخدام الموارد لتجنب حدوث مشاكل الازدحام وفقدان البيانات. يمكن تطبيق عملية موازنة الحمل على الأجهزة العتادية، ويمكن تطبيقها أيضاً على البرمجيات. كما يجب على الجهاز الذي يقوم بموازنة الحمل أن يكون له بديل لتجنب نقطة الفشل الوحيدة. يوجد العديد من خوارزميات موازنة الحمل، قد يكون بعضها ستاتيكي والبعض

الآخر ديناميكي، النوع الأول يناسب البيئات التي تتسم بالثبات والاستقرار بما أنه يفقد المرونة وسرعة التكيف في حال حصول تغيرات ديناميكية في بنية الشبكة، في جميع الأحوال وبسبب أن سلوك المستخدمين لا يمكن التنبؤ به تتجه الخوارزميات الستاتيكية في توزيع الحمل بالفشل في أغلب الحالات. إن خوارزميات موازنة الحمل الديناميكية أكثر فعالية، لأن الطلبات يتم توزيعها على الموارد المتاحة تبعاً لحالة الشبكة في الوقت الراهن، وتبعاً للقواعد البرمجية المدخلة والتي تكون ديناميكية في هذه الحالة [15].

### 10-2- موازنة الحمل في شبكات SDN:

من ميزات شبكات SDN هي قدرتها على المحافظة على رؤية عامة للشبكة ككل، ولكن من جانب آخر فإن جميع عمليات اتخاذ القرارات تتم في المتحكم، مما سيسبب في زيادة الحمل على المتحكمات وزيادة تأخير زمن الاستجابة، لهذه الأسباب ظهرت الحاجة الماسة إلى تطبيق تكنولوجيا موازنة الحمل في شبكات SDN، مما سيسمح لعناصر الشبكة من تجنب مشاكل الازدحام ومعالجتها في حال حدوث حمل زائد على الشبكة. يجب على متحكم شبكة SDN تحسين معدل استخدام الموارد من خلال برمجيات معينة، هذا سيؤدي بالضرورة إلى زيادة الإتاحة في الشبكة وتحسن أدائها. إنتاجية الشبكة أيضاً سوف تتحسن عندما نقوم بتقليل زمن استجابة العناصر المشكلة للشبكة بالشكل الذي سوف يزيد من رضا الزبائن.

بروتوكول OpenFlow الذي يضمن عمليات إدارة الشبكة يعتبر العامل الأساسي في عمليات موازنة الحمل في شبكات SDN، حيث يعتبر هذا البروتوكول هو المسؤول عن تخزين جدول التدفقات، وكل مدخل في هذا الجدول يتألف من مجموعة من الحقول والعدادات، ويقابلها اتخاذ مجموعة من التعليمات والإجراءات الواجب تنفيذها في حال حدثت المطابقة مع أحد هذه المدخل في الجدول. باستخدام بروتوكول OpenFlow يمكن التحكم بالمدخل في جدول التدفق والحصول على الإحصائيات الخاصة بالشبكة وبكل مبدل وفقاً للقواعد والسياسات المطبقة في المتحكم الذي يستخدم هذا البروتوكول [16].

## 10-3- عوامل موازنة الحمل:

يمكن أخذ العديد من العوامل لتقييم نجاح آلية موازنة الحمل، ولكن أهم هذه العوامل التي أجمع عليها الباحثون يتلخص في أربع نقاط رئيسية هي:

1. **قابلية التوسع:** لا يمكن للآلية المركزية أن تتجح في الشبكات كبيرة الحجم والتي تنمو باستمرار لأنها تفتقر إلى المرونة، ولذلك كان لابد من استخدام بنية موزعة بوجود عدة مجالات وعدة متحكمات في الشبكة، وذلك من أجل الحفاظ على الإنتاجية في الشبكة. أحد أهم المعايير المستخدمة في التقييم هي إنتاجية الشبكة والتي يمكن حسابه بالقانون [17]:

$$F(N) = \emptyset(N) * T(N)/C(N) \quad (1)$$

حيث  $F(N)$  تمثل إنتاجية الشبكة،  $\emptyset(N)$  تمثل إنتاجية طبقة التحكم،  $T(N)$  تمثل متوسط زمن الاستجابة للطلبات في الشبكة،  $C(N)$  هي كلفة التطوير لطبقة التحكم.

2. **المرونة:** يقصد بهذا المعيار هو حالة الفشل التي يتعرض لها المتحكم، وهو من أهم المعايير في تقييم أداء آلية موازنة الحمل، حيث لا بد من ضمان عملية الانتقال إلى متحكم احتياطي في حالات فشل المتحكم الرئيسي، كما من الممكن أن تتأثر الشبكة في حال فشل أحد العقد أو الوصلات، يجب أن تضمن آلية موازنة الحمل تخصيص عادل للعقد على المتحكمات بالشكل الذي تقوم فيه هذه المتحكمات بمعالجة الأخطاء الناتجة عن الفشل في أحد الوصلات أو العقد بشكل سهل، وبدون أن تسبب زيادة في الحمل على أحد المتحكمات، كما يجب أن يتم أخذ المسافة والكلفة إلى المتحكم الاحتياطي في حال حدوث فشل في المتحكم الرئيسي بعين الاعتبار [18].

3. **معدل استخدام الموارد:** تم اقتراح عدة معايير من أجل قياس تأثير عملية موازنة الحمل على معدل استخدام الموارد ومن أهمها: معدل استخدام عرض الحزمة، حيث

يقوم هذا المقياس بقياس معدل النقل في الشبكة، ويعكس الحمل على الوصلات. يقوم المتحكم بحساب عدد البايتات المرسلة خلال فترتين متتاليتين وحساب المعدل في هاتين الفترتين، ثم يتم تقسيم هذا المعدل على العرض الأعظمي للحزمة ، وبالتالي نحصل على معدل استخدام عرض الحزمة، أو يمكن كفكرة عامة حساب معدل استخدام الموارد مثل الذاكرة وعرض الحزمة أو موارد أخرى مستخدمة في نقل المعلومات في الشبكة، من خلال حساب نسبة استخدام هذا النوع من الموارد خلال فترة معينة وتقسيمه على السعة الأعظمية لهذا المورد، وكلما زادت هذه النسبة زاد معدل استخدام المورد، وبالتالي فإن الشبكة تصبح تعمل بكفاءة أعلى [19].

**4. جودة الخدمة:** يوجد العديد من العوامل المهمة في حساب عامل جودة الخدمة، ولكن يمكن تلخيص أهم هذه العوامل في النقاط التالية:

✓ **زمن التأخير:** وهو الزمن الذي تحتاجه الحزمة للوصول من المصدر إلى الهدف عبر الشبكة، وهو يعتمد على الزمن الذي يستغرقه المبدل في تمرير الحزمة الواردة إليه بالإضافة إلى حالة الازدحام في الشبكة، وحجم البيانات المرسلة. لحساب زمن التأخير في مسار ما يجب حساب مجموع التأخيرات في كل وصلة من الوصلات الموجودة على طول هذا المسار، ويعكس هذا الزمن حالة الشبكة، ويحدد إذا كانت مزدحمة ام لا، كما يوجد زمن آخر يدعى زمن الإنهاء وهو الزمن المستغرق في عملية اتخاذ القرارات في المتحكم بالشبكة، والذي قد يختلف تبعاً لخوارزمية موازنة الحمل المتبعة.

✓ **زمن الاستجابة:** يعرف بأنه الزمن المستغرق بين الطلب وحدث الاستجابة، عندما تكون الشبكة مزدحمة، أو عند وجود آلية ضعيفة لموازنة الحمل سيتجاوز معدل إنشاء الحزم السعة العظمى لموارد الشبكة، مما سيزيد من أرتال الانتظار وبالتالي سوف يزيد حكماً زمن الاستجابة. يوجد عامل مهم يؤثر على زمن

الاستجابة وهو درجة انشغال المتحكم والتي يمكن حسابها من خلال حساب الفرق بين زمن استجابة المتحكم في الوقت الراهن مع معدل استجابته العام. إن حساب زمن الاستجابة يعطي انطباعاً واضحاً عن مدى فعالية آلية موازنة الحمل المطبقة.

✓ **معدل فقدان الحزم:** يشير هذا العامل إلى عدد الحزم المفقودة خلال فترة الإرسال، ويعكس مقدار ازدحام المبدلات والوصلات في الشبكة، حيث إن المبدلات يمكن أن تكون مشغولة بشكل كبير وأن تعاني من حمل زائد، الأمر الذي يضطرها إلى تجاهل ورمي بعض الحزم الواردة. يمكن حساب عدد الحزم المفقودة في مسار ما من خلال طرح عدد الحزم المرسل من عدد الحزم المستقبل، كما يمكن حساب معدل فقدان الحزم لمسار ما من خلال تقسيم عدد الحزم المفقودة على عدد الحزم الكلي المستقبل في هذا المسار.

✓ **الخرج:** يعرف بأنه سرعة معالجة البيانات في عقد الشبكة، ويمكن حسابه من خلال حساب كمية أو حجم البيانات المرسل من المصدر إلى الهدف خلال واحدة مختارة من الزمن، وإن اختيار آلية موازنة حمل ناجحة سوف تؤدي إلى زيادة هذا العامل إلى أكبر رقم ممكن [20].

#### 10-4- خوارزميات موازنة الحمل:

تلعب خوارزمية موازنة الحمل المستخدمة دوراً رئيسياً في أداء الشبكة، حيث إن اختيار خوارزمية موازنة حمل مناسبة وفعالة سوف يؤدي بالضرورة إلى تحسين أداء الشبكة، من خلال تقليل زمن الاستجابة للطلبات الواردة، وزيادة عدد الطلبات التي تتم معالجتها في الشبكة دون فشل، وتخفيض نسبة احتمال فشل أحد المتحكمات نتيجةً لتعرضه إلى حمل زائد، وبعض من أهم خوارزميات موازنة الحمل المستخدمة في وقتنا الحالي هي:

➤ **الخوارزمية العشوائية:** تقوم هذه الخوارزمية بتوزيع الحمل على المتحكمات بغض النظر عن حالة الشبكة، حيث تقوم بتوزيع الحمل بشكل عشوائي على الموارد المتوفرة بشكل عشوائي. تسند هذه الخوارزمية الطلبات القادمة إلى أحد الموارد وليكن  $i$  بشكل عشوائي و باحتمال  $1/N$ ، حيث  $N$  هو عدد الموارد الموجودة في الشبكة، خلال هذه العملية لا يملك المتحكم المسؤول عن عملية موازنة الحمل أي معلومات عن حالة الشبكة والحمل الموجود عليها، وقد بينت الدراسات أن هذه الخوارزمية تعطي النتيجة الأسوأ من بين خوارزميات موازنة الحمل الأخرى وخاصة من ناحية زمن الاستجابة.

➤ **خوارزمية القادم أولاً يخدم أولاً:** تقوم هذه الخوارزمية على مبدأ أن الطلب الوارد أولاً يعطى إلى المورد الأكبر سعة، حيث يتم تخزين رتل للطلبات الواردة ريثما يتم إسنادها إلى مورد الشبكة في حال عدم توافر موارد متاحة. بالرغم من سرعة وبساطة خوارزمية موازنة الحمل هذه، إلا أنها أعطت نتائج سيئة بالنسبة لزمن الاستجابة، وذلك لأنها لا تأخذ بعين الاعتبار حجم المهمة الواردة ولا أولويتها [21].

➤ **الخوارزمية الدائرية:** تقوم هذه الخوارزمية بتوزيع الطلبات الواردة إلى الموارد المتاحة بشكل متساوٍ ووفقاً لخوارزمية جدولة دائرية. عند وصول أول طلب يقوم المتحكم المسؤول عن عملية موازنة الحمل بإسناد هذا الطلب إلى المورد الأول في قائمته، ويتم إسناد الطلب الثاني إلى المورد الثاني وهكذا حتى يصل إلى المورد  $N$ ، حيث  $N$  تمثل عدد الموارد المتاحة، في حال ورود طلب جديد  $N+1$  يقوم بإسناده إلى المورد الأول، وتتكرر العملية إلى النهاية. تملك جميع الموارد في هذه الخوارزمية أولوية متساوية بغض النظر عن الحزم والطلبات الواردة،

وبغض النظر عن الزمن الذي تستغرقه الموارد في الاستجابة، هذا ما يفسر زمن الانتظار الكبير الذي تعاني منه المهمات الواردة.

➤ **الخوارزمية الدائرية مع أوزان:** تعتبر هذه الخوارزمية هي تطوير للخوارزمية الدائرية، حيث تختلف عنها بأن يوجد لكل مورد وزن في البداية، ويحدد هذا الوزن حجم الحمل الذي يجب تخصيصه لكل مورد، أي عدد المداخل التي سوف يملكها كل مورد في رتل الانتظار الدائري، أي أنه إذا كان وزن المورد الأول 2 ووزن المورد الثاني 3 ، فإنه سوف يتم تخصيص أول طلبين إلى المورد الأول، والطلبات الثلاث التالية إلى المورد الثاني وهكذا. تعتبر هذه الخوارزمية خوارزمية ستاتيكية، فهي لا تأخذ بعين الاعتبار تطور الأحمال في الشبكة أو حدوث أي تغييرات على بنيتها. لا تراعي هذه الخوارزمية عمليات البعد الجغرافي حيث من الممكن إسناد طلب وارد من مبدل ما إلى أبعد مورد عن هذا المبدل مما قد يسبب بمرور هذا الطلب بعدد كبير من العقد للوصول إلى الهدف وهذا يؤثر بشكل كبير على زمن الاستجابة.

➤ **خوارزمية الحمل الأقل:** تعتبر هذه الخوارزمية هي خوارزمية ديناميكية، حيث يقوم المتحكم المسؤول عن عملية موازنة الحمل بحساب عدد الطلبات في كل مورد، حيث يقوم بإسناد الطلب الوارد إلى المورد الأقل حملاً، أي المورد الذي يملك أقل عدد من الطلبات، قد يتغير الحمل في كل مورد عند تخصيص طلب جديد إليه، أو عند الانتهاء من تنفيذ طلب أو تجاوزه للمهلة المحددة. تعتبر هذه الخوارزمية هي خوارزمية ديناميكية، إذ أنها توزع الطلبات بناءً على أحمال الموارد في الزمن الحقيقي، ولكن من سلبياتها أنها لا تأخذ بعين الاعتبار الاختلاف في السعة بين الموارد، مما يؤثر على أداء الشبكة بشكل عام. تعتبر

هذه الخوارزمية مناسبة للشبكات المؤلفة من مجموعة موارد متساوية السعة  
[22].

### 11-النتائج والمناقشة:

تم اقتراح خوارزمية جديدة لموازنة الحمل وقد اقترحنا تسميتها بخوارزمية الترحيل باستخدام العتبات الثلاث، وتمت مقارنة هذه الخوارزمية الجديدة مع خمس خوارزميات لموازنة الحمل والمستخدمة في الوقت الراهن، تمت عمليات مقارنة هذه الخوارزميات وفقاً لثلاثة عوامل وهي: إنتاجية الشبكة، معدل استخدام الموارد، واحتمال فشل المتحكم، وقد أثبتت التجارب إن الخوارزمية المقترحة تقدم تحسناً ملحوظاً على الخوارزميات السابقة من ناحية العوامل الثلاثة المدروسة.

### 11-1- السيناريو المقترح:

تمت مقارنة ست خوارزميات لموازنة الحمل وهي: الخوارزمية العشوائية، خوارزمية القدام أولاً يخدم أولاً، الخوارزمية الدائرية، الخوارزمية الدائرية مع أوزان، خوارزمية الحمل الأقل، والخوارزمية الأخيرة التي اقترحناها والتي اقترحنا تسميتها بخوارزمية الترحيل باستخدام العتبات الثلاث، وقد تم استخدام متحكم OpenDayLight المكتوب بلغة Java. تمت عمليات المقارنة وفقاً لثلاثة عوامل وهي: إنتاجية الشبكة، معدل استخدام الموارد، واحتمال فشل المتحكم. الأداة التي قمنا باختبار التجارب عليها تمتلك معالج Intel® Core™ i9 [8-core@2.3](#) GHz 9th generation Processor و ذاكر GB 16 من نوع 2666MHz DDR4، و قرص صلب بسعة 1 TB SSD بنظام تشغيل MacOS مع أداة افتراضية على برنامج Virtual Box تم عليها تنصيب نظام تشغيل Linux 22.04 – 64 bit. تمت عمليات المقارنة باستخدام المحاكي الشبكي Mininet والذي تم تنصيبه على الأداة الافتراضية المذكورة سابقاً.

### 11-2- الخوارزمية المقترحة:

تم اقتراح الخوارزمية التي تمت تسميتها خوارزمية الترحيل باستخدام العتبات الثلاث، حيث تحدد هذه الخوارزمية ثلاث عتبات لتحديد حالة المتحكم، هذه العتبات هي:  $T_{min}$ ,  $T_{med}$ ,  $T_{max}$  حيث يتم إدخالها يدوياً وفقاً لسعة المتحكم، وتكون  $T_{med} = T_{max} + T_{min} / 2$ ، باستخدام هذه العتبات الثلاث يصبح لدينا أربع حالات لحمل المتحكم أ وهي:

إذا كان حمل المتحكم  $i$ :  $L(i) < T_{min}$  يكون المتحكم ذا حمل قليل.

أما إذا كان حمل المتحكم  $i$ :  $T_{min} < L(i) < T_{med}$  يكون المتحكم ذا حمل عادي.

أما إذا كان حمل المتحكم  $i$ :  $T_{med} < L(i) < T_{max}$  يكون المتحكم ذا حمل عالٍ.

أما إذا كان حمل المتحكم  $i$ :  $T_{max} < L(i)$  يكون المتحكم ذا حمل مفرط.

يمكن حساب حمل المتحكم  $i$  من خلال العلاقة:  $L(i) = L_c(i) + L_m(i)$

حيث:  $L_c(i)$  تمثل مقدار استهلاك وحدة التحكم المركزية في المتحكم  $i$ ,

$L_m(i)$  تمثل مقدار استهلاك الذواكر في المتحكم  $i$ .

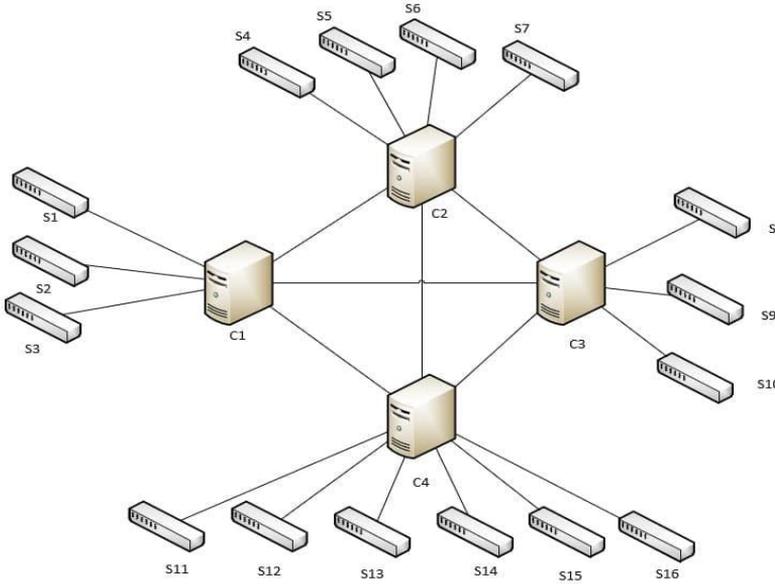
تتم عملية حساب حمل المتحكمات بشكل دوري، وإذا ظهرت حالة حمل مفرط أو عالٍ في أحد المتحكمات، يتم إطلاق عملية الترحيل، حيث يجب أن يتم ترحيل جزء من حمل المتحكم المفرط الحمل إلى المتحكم الأقل حملاً بين المتحكمات المتوفرة، عن طريق اختيار أحد المبدلات المتصلة بالمتحكم المفرط الحمل ونقل حمل هذا المبدل إلى المتحكم الذي تم اختياره، يتم اختيار المبدل الأقل حملاً في المتحكم المفرط الحمل، ليتم نقله إلى المتحكم المختار، وبعدها يتم تكرار العملية حتى نصل إلى حالة الاستقرار في الشبكة، بحيث تصبح جميع المتحكمات في حالة أحمال متقاربة. وفي حالة وجود أكثر من متحكم يعاني من نفس حالة الحمل المفرط أو العالي تكون الأولوية للمتحكم ذي الحمل الأعلى.

### 11-3- تحليل النتائج:

تم اقتراح بنية مؤلفة من 4 متحكمات هي المتحكمات C1,C2,C3,C4 متصلة مع 16 مبدل هي : S1,S2,...,S16.

في خوارزمية الترحيل باستخدام العتبات تم اقتراح حالة ابتدائية، حيث المتحكم C1 يتصل بالمبدلات S1,S2,S3 والمتحكم C2 يتصل بالمبدلات S4,S5,S6,S7 والمتحكم C3 يتصل بالمبدلات S8,S9,S10 بينما المتحكم الأخير C4 يتصل بالمبدلات S11,S12,S13,S14,S15,S16. تم توضيح البنية في الشكل (3):

باختيار عتبة  $T_{min}=10\%$  و  $T_{max}=90\%$  و بحساب  $T_{med}$  تنتج  $T_{med}=50\%$ .

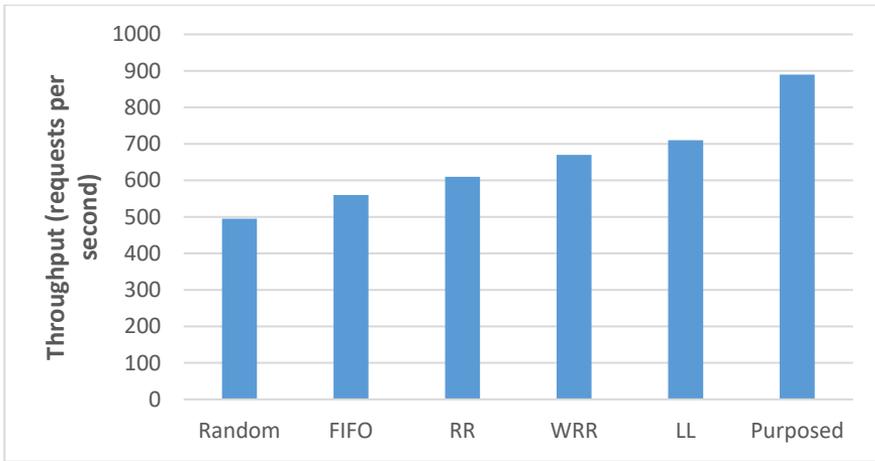


الشكل (3) البنية المقترحة.

### 11-3-1- السيناريو الأول: دراسة تأثير الخوارزميات المختلفة على إنتاجية الشبكة:

تعرف إنتاجية الشبكة بأنها عدد الطلبات التي تمت معالجتها خلال واحدة الزمن وقد اخترنا أن تكون ثانية واحدة، وكلما زادت إنتاجية الشبكة، كلما زاد عدد الطلبات التي تمت

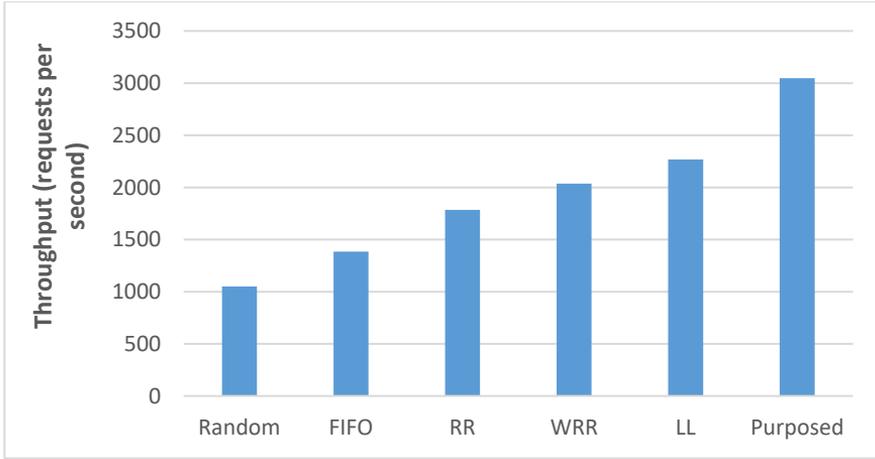
معالجتها في الشبكة دون فشل، وبالتالي تحسن أداء الشبكة بشكل عام، وإن استخدام خوارزمية موازنة حمل فعالة سوف يؤدي بالضرورة إلى رفع إنتاجية الشبكة. تمت دراسة كل من الخوارزميات الستة في أربع حالات هي حالة ورود 1000,4000,8000,16000 طلب في كل ثانية على التوالي وتم حساب إنتاجية الشبكة في كل من الحالات الأربعة كما بينت الأشكال (4,5,6,7). حيث يوضح الشكل (4) إنتاجية الشبكة في حالة ورود 1000 طلب كل ثانية:



الشكل (4) إنتاجية الشبكة في حالة ورود 1000 طلب كل ثانية.

بالنظر إلى الشكل (4) وهي حالة ورود 1000 طلب كل ثانية نجد أن الخوارزميتين العشوائية والقادم أولاً يخدم أولاً تعالج حوالي نصف الطلبات الواردة، أي أنها تعطي خرجاً يتراوح بين 50-60%، بينما تعالج الخوارزميات الدائرية، والدائرية بأوزان، وخوارزمية الحمل الأقل حوالي 60-70% من الطلبات الواردة، بينما الخوارزمية المقترحة تعالج حوالي 90% من الطلبات التي ترد إلى الشبكة.

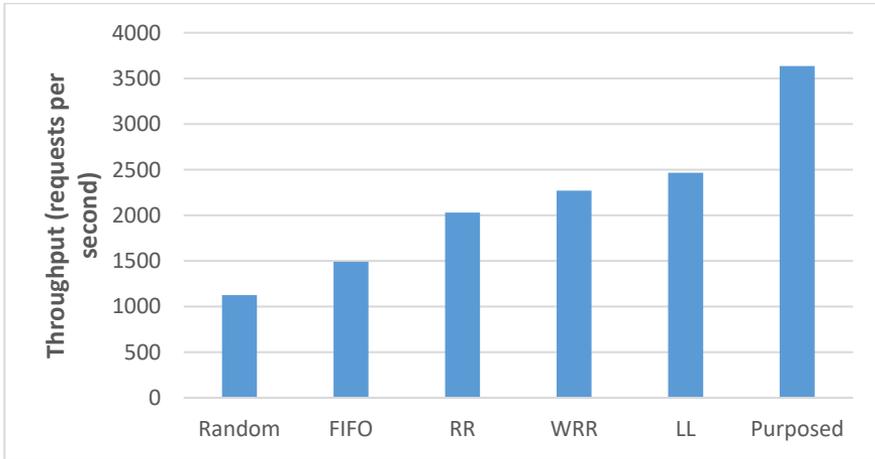
ومع زيادة عدد الطلبات الواردة خلال واحدة الزمن، والتي يوضحها الشكل (5) والذي يعبر عن إنتاجية الشبكة في حالة ورود 4000 طلب كل ثانية:



الشكل (5) إنتاجية الشبكة في حالة ورود 4000 طلب كل ثانية.

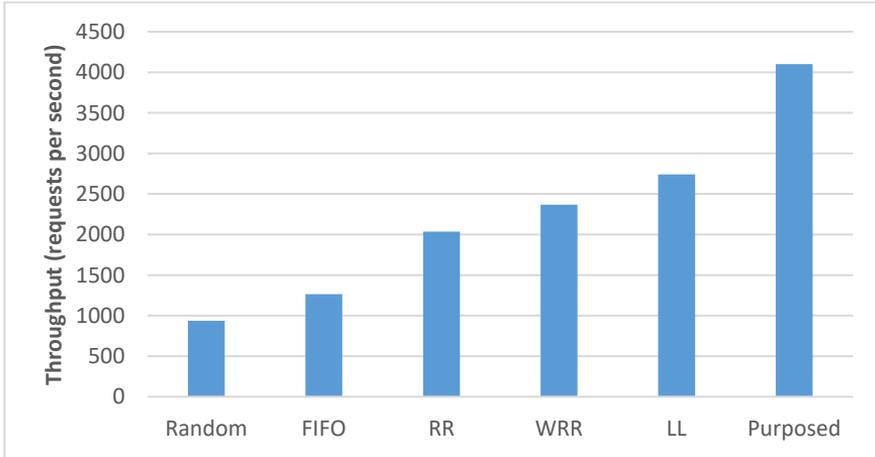
نستنتج من الشكل (5) وهي حالة ورود 4000 طلب كل ثانية نلاحظ انخفاض أداء الخوارزميات بشكل عام، حيث إن الخوارزمية العشوائية تعالج 25% من الطلبات الواردة، وخوارزمية لقدام أولاً يخدم أولاً تعالج حوالي 35% من الطلبات الواردة، بينما تعالج الخوارزميتان الدائرية والدائرية مع أوزان وخوارزمية الحمل الأقل حوالي 45% و 51% و 56% على التوالي، وتبقى الخوارزمية المقترحة هي الأفضل بنسبة معالجة 76% من الطلبات التي ترد إلى الشبكة.

و باستمرار زيادة الطلبات الواردة، تصبح إنتاجية الشبكة في حالة ورود 8000 طلب كل ثانية كما هو موضح في الشكل (6):



الشكل (6) إنتاجية الشبكة في حالة ورود 8000 طلب كل ثانية.

بالنظر إلى الشكل (6) وهي حالة ورود 8000 طلب كل ثانية، يستمر أداء الخوارزميات بالانخفاض، حيث إن الخوارزمية العشوائية تعالج 14% من الطلبات الواردة، وينخفض أداء خوارزمية القادم أولاً يخدم أولاً بشكل أكبر حيث تعالج حوالي 18% من الطلبات الواردة، بينما تعالج الخوارزميتان الدائرية والدائرية مع أوزان وخوارزمية الحمل الأقل حوالي 25% و 28% و 30% على التوالي، وتبقى الخوارزمية المقترحة هي الأفضل ولكن مع انخفاض بالإنتاجية، حيث تعالج بنسبة معالجة 45% من الطلبات الواردة. ومع الوصول إلى حالة 16000 طلب في الثانية تصبح إنتاجية الشبكة كما هي موضحة في الشكل (7):



الشكل (7) إنتاجية الشبكة في حالة ورود 16000 طلب كل ثانية.

من الشكل (7) نجد أن الخوارزميتين العشوائية والقادم أولاً يخدم أولاً تخدم عدداً أقل من الطلبات التي كانت تخدمها عند معدل ورود 8000 طلب في كل ثانية، وينخفض أداؤها حيث تعطي نسبة معالجة 6% و 8% على التوالي، وتحافظ الخوارزميتان الدائرية والدائرية مع أوزان على عدد الطلبات التي كانت تعالجها في حال ورود 8000 طلب، ولكن النسبة تتخفف بسبب أنها تعالج نفس العدد رغم زيادة المعدل من 8000 طلب إلى 16000 طلب كل ثانية، وهذا يعني انخفاض الإنتاجية إلى النصف، وتعطي نسبة

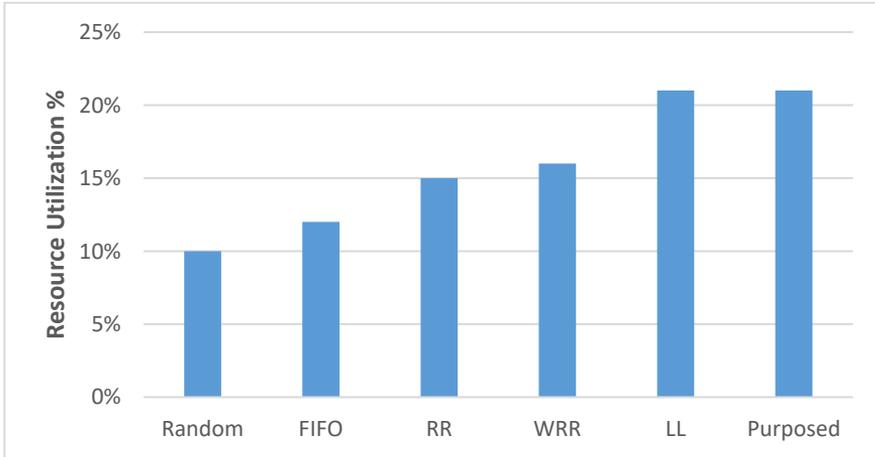
معالجة قدرها 12% و 14% على التوالي، كما تعطي خوارزمية الحمل الأقل 17% من الطلبات الواردة، بينما تعطي الخوارزمية المقترحة زيادة طفيفة في عدد الطلبات المعالجة، وتعطي نسبة معالجة قدرها 25%.

### 11-3-2- السيناريو الثاني: دراسة تأثير الخوارزميات المختلفة على معدل استخدام الموارد:

يعرف معدل استخدام الموارد بأنه نسبة انشغال المورد بالنسبة إلى سعته الكلية، وكلما زادت هذه النسبة كلما تحسن أداء الشبكة بشكل عام، لأنه يعني زيادة استغلال الموارد في معالجة الطلبات الواردة، وهذا الأمر يؤثر بشكل طردي على زيادة عدد الطلبات التي تمت معالجتها بنجاح، وبالتالي على أداء الشبكة.

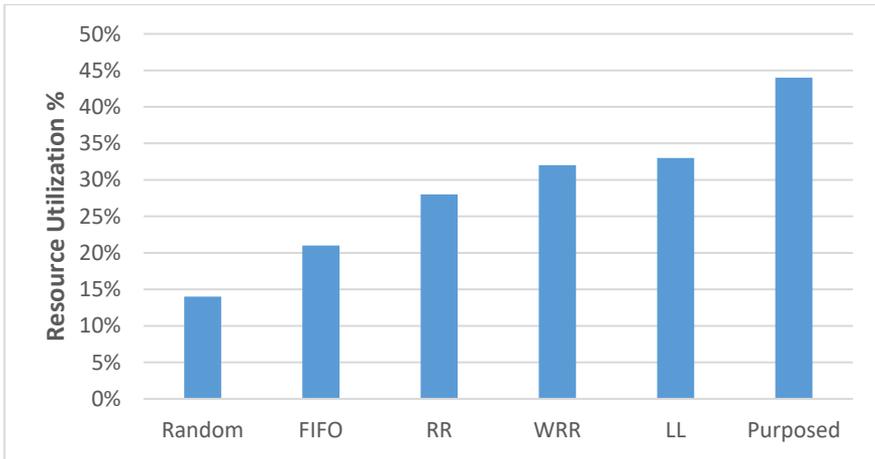
تمت دراسة كل من الخوارزميات الستة في أربع حالات هي حالة ورود 1000,4000,8000,16000 طلب في كل ثانية على التوالي وتم حساب معدل استخدام الموارد في كل من الحالات الأربعة بعد مرور 60 ثانية كما بينت الأشكال (8,9,10,11).

يوضح الشكل (8) معدل استخدام الموارد في حالة ورود 1000 طلب كل ثانية:



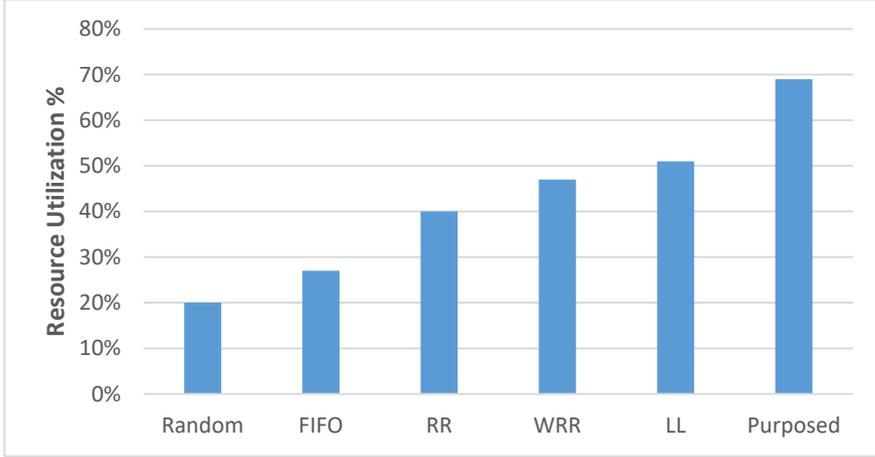
الشكل (8) معدل استخدام الموارد في حالة ورود 1000 طلب كل ثانية.

بالنظر إلى الشكل (8) وفي حال ورود 1000 طلب كل ثانية، نجد أن الخوارزميات العشوائية والقادم أولاً يخدم أولاً والدائرية والدائرية مع أوزان يكون معدل استخدام الموارد فيها متقارباً ويتراوح بين 10-15%، بينما خوارزميتي الحمل الأقل والخوارزمية المقترحة تعطي معدل استخدام متماثل بنسبة 21% من الموارد المتاحة، ونلاحظ انخفاض هذا المعدل بسبب أن عدد الطلبات وحجمها يعتبر قليلاً بالنسبة لسعة الموارد المتوفرة. ومع زيادة عدد الطلبات الواردة خلال وحدة الزمن والتي يوضحها الشكل (9) والذي يعبر عن معدل استخدام الموارد في حالة ورود 4000 طلب في الثانية الواحدة:



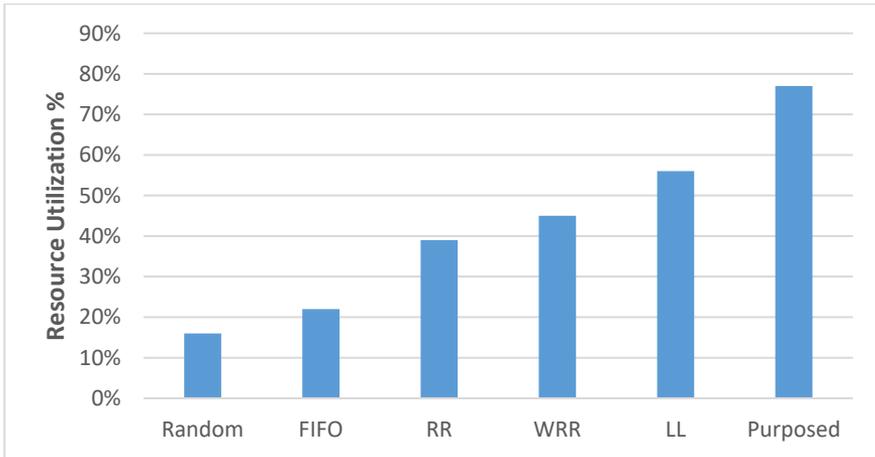
الشكل (9) معدل استخدام الموارد في حالة ورود 4000 طلب كل ثانية.

من الشكل (9) الذي يمثل حالة ورود 4000 طلب كل ثانية، نجد ارتفاع نسبة انشغال الموارد حيث تسجل الخوارزمية العشوائية، والقادم أولاً يخدم أولاً معدل استخدام قدره 14% و 21% على التوالي، بينما يتقارب أداء الخوارزمية الدائرية مع الخوارزمية الدائرية مع أوزان، وخوارزمية الحمل الأقل وتعطي نسب تتراوح بين 28-33% كمعدل استخدام، وتكون الخوارزمية المقترحة هي الأفضل مع معدل استخدام يبلغ 44%. وباستمرار زيادة الطلبات الواردة إلى 8000 طلب في الثانية، يصبح معدل استخدام الموارد كما هو موضح في الشكل (10):



الشكل (10) معدل استخدام الموارد في حالة ورود 8000 طلب كل ثانية.

بملاحظة الشكل (10) عند ورود 8000 طلب كل ثانية، نلاحظ ازدياد المعدل بشكل عام في جميع الخوارزميات المدروسة، حيث يبلغ 20% في الخوارزمية العشوائية، و27% في خوارزمية القادم أولاً يخدم أولاً، و40% في الخوارزمية الدائرية، كما تتشابه خوارزمتا الدائرية مع أوزان، والحمل الأقل في النتائج، حيث يعطيان معدل استخدام 47% للأولى و51% للثانية، ويكون أداء الخوارزمية المقترحة أفضل من الخوارزميات الباقية من حيث هذا العامل بمعدل 69%. ومع الوصول إلى حالة 16000 طلب في الثانية يصبح معدل استخدام الموارد كما هو موضح في الشكل (11):



الشكل (11) معدل استخدام الموارد في حالة ورود 16000 طلب كل ثانية.

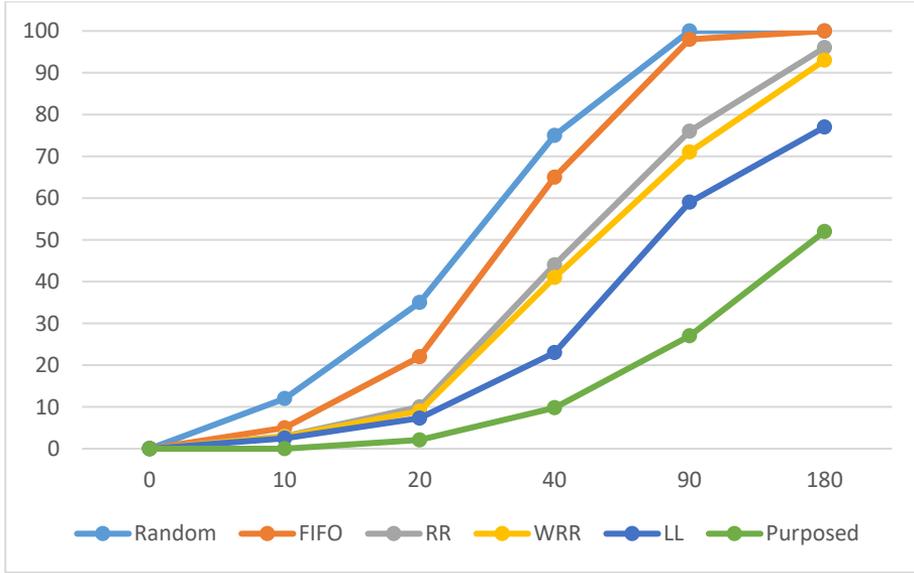
بقراءة النتائج من الشكل (11) في حال ورود 16000 طلب كل ثانية، نلاحظ انخفاض في معدل استخدام الموارد للخوارزميتين العشوائية والقادم أولاً يخدم أولاً، رغم زيادة عدد الطلبات الواردة، ويعود ذلك إلى حالة الفشل التي قد تصيب بعض المتحكمات وخروجها عن الخدمة، وتعطيان معدلات قدرها 16% و 22% للخوارزميتين السابقتين، كما نلاحظ أن الخوارزميتين الدائرية، والدائرية مع أوزان تحافظان على نفس المعدل تقريباً وتعطيان نتائج 39% و 45% على التوالي، بينما تظهر خوارزمية الحمل الأقل تحسناً طفيفاً بمعدل استخدام يبلغ 56%، و تظهر الخوارزمية المقترحة تحسناً أفضل وتعطي معدلاً قدره 77% من الموارد المتاحة.

### 11-3-3- السيناريو الثالث: دراسة تأثير الخوارزميات المختلفة على احتمال فشل

#### المتحكم:

يعرف هذا العامل بأنه النسبة المئوية المتوقعة لاحتمال فشل أحد المتحكمات، وكلما زادت هذه النسبة زادت احتمالات فشل المتحكمات وبالتالي زادت احتمالات انهيار الشبكة ككل، ومن الضروري مراعاة جعل أحمال المتحكمات متفاربة قدر الإمكان، وذلك لتجنب حالات فشل أحد المتحكمات بسبب الحمل الزائد، في حين وجود متحكمات أخرى بأحمال منخفضة.

تمت دراسة كل من الخوارزميات الستة في حال ورود 2000 طلب في كل ثانية، وتمت عملية الدراسة حتى 180 ثانية، وتم حساب احتمال فشل المتحكم تجريبياً، حيث تم إعادة التجربة 3 مرات لكل خوارزمية، وكانت النتائج مبينة في الشكل (12):



الشكل (12) احتمال فشل المتحكمات في الخوارزميات المختلفة.

من الشكل (12) نجد أن الخوارزميتين العشوائيتين، والقادم أولاً يخدم أولاً تعطيان أسوأ نتائج حيث إنه بعد مرور 20 ثانية يعطيان احتمالات فشل تتراوح بين 20-35%، وبعد مرور 40 ثانية ترتفع احتمالات الفشل إلى حوالي 60-70%، وبعد مرور 90 ثانية تصبح نسبة الفشل حوالي 100%، وذلك لأن كلتا الخوارزميتين لا تراعيان الأحمال الحالية للمتحكمات عند إسناد طلبات إليها، بينما تعطي الخوارزميتان الدائرية، والدائرية مع أوزان نتائج متقاربة، مع أفضلية طفيفة للخوارزمية الدائرية مع أوزان، وباحتمالات فشل تتراوح بين 40-44% في الثانية 40، وبين 70-75% في الثانية 90، ولكن كليهما تبلغ احتمالات الفشل فوق 90% عند الوصول إلى الثانية 180، ونلاحظ من الشكل أيضاً أن خوارزمية الحمل الأقل تعطي نتائج أفضل من الخوارزميات المذكورة السابقة وذلك لأنها تراعي الأحمال الحالية للمتحكمات، وتعطي احتمالات فشل لا تتجاوز 8% في الثانية 20، ولا تتجاوز 25% في الثانية 40، ولكن احتمال الفشل يرتفع إلى حوالي 59% في الثانية 90، ويبلغ 77% بعد مرور 180 ثانية، وفي النهاية تبقى الخوارزمية المقترحة هي الأفضل حيث إن بعد مرور 90 ثانية يبقى احتمال الفشل حوالي 27%، ويبلغ حوالي 50% عند وصولنا للثانية 180، وذلك لأنها تعتمد على عملية توزيع الحمل بين

المتحكمات في حال وصول أحد هذه المتحكمات إلى حمل زائد على عكس الخوارزميات الخمس التي تسبقها.

#### الخاتمة:

قمنا في هذا البحث بتوضيح الحاجة للانتقال من الشبكات التقليدية إلى شبكات SDN، وأيضاً الحاجة إلى استبدال نموذج شبكات SDN بمتحكم مركزي إلى شبكات SDN بمتحكمات متعددة، والفوائد المرجوة من استخدام البنية ذات المتحكمات المتعددة، كما تعرضنا إلى التحديات التي تواجهها هذه البنية الجديدة وكان أهمها: أماكن توضع المتحكمات، وخوارزميات موازنة الحمل المستخدمة لتحسين أداء الشبكة بشكل عام. كما قمنا في هذا البحث بالاهتمام بخوارزميات موازنة الحمل، وذلك عن طريق اقتراح خوارزمية جديدة لموازنة الحمل في شبكات SDN ذات المتحكمات المتعددة، وقمنا بمقارنة هذه الخوارزمية مع مجموعة من الخوارزميات المستخدمة سابقاً. تمت عملية الدراسة من خلال ثلاثة عوامل هي: إنتاجية الشبكة، ومعدل استخدام الموارد، واحتمال فشل المتحكم، وقد أثبتت الخوارزمية المقترحة تفوقها على الخوارزميات المستخدمة سابقاً من ناحية العوامل المدروسة.

فمن ناحية إنتاجية الشبكة، عالجت الخوارزمية المقترحة حوالي 90% من الطلبات الواردة عند ورود 1000 طلب في الثانية، وأعطت تحسناً قدره 20% عن أفضل الخوارزميات الأخرى في حالتها و 1000 و 4000 طلب في الثانية الواحدة، كما أعطت تحسناً قدره 15% في حالة ورود 8000 طلب في الثانية، وتحسناً قدره 8% في حالة ورود 16000 طلب في الثانية، وذلك بالمقارنة مع أفضل الخوارزميات الأخرى المدروسة. أما بالنسبة لمعدل استخدام الموارد، وفي حالة ورود 1000 طلب في الثانية تعطي الخوارزمية المقترحة نتائج متقاربة مع خوارزمية الحمل الأقل، ولكنها أفضل من جميع الخوارزميات الباقية، أما في حالات ورود 4000 و 8000 و 16000 طلب في الثانية

الواحدة، فقد أعطت الخوارزمية المقترحة تحسیناً ملحوظاً بنتائج تحسين تقدر بـ 11%، 18%، 21% على التوالي، مقارنةً مع أفضل الخوارزميات الأخرى. أما بالنسبة للعامل الأخير وهو احتمال فشل المتحكم، نجد أنه بعد مرور 180 ثانية فإن الخوارزمية المقترحة تعطي تحسیناً قدره حوالي 27% عن خوارزمية الحمل الأقل، وتحسیناً يقترب من 50% بالنسبة للخوارزميات الباقية. من جهة أخرى لا نستطيع إغفال أهمية عامل توازن المتحكمات في تحسين أداء الشبكة، و يلعب دوراً متكاملاً مع خوارزمية موازنة الحمل المستخدمة، لذا كان لابد من التوجه لدراسة هذا العامل مستقبلاً، وهذا هو موضع بحثنا في الوقت الحالي بعد الانتهاء من دراسة خوارزميات موازنة الحمل.

### المراجع

- [1] HYOJOON KIM, NICK FEAMSTER; *Improving Network Management with software Defined Networking*; IEEE Communications Magazine, February 2013.
- [2] M. ABU SHARKH, M. JAMMAL, A. SHAMI, and A. OUDA; *Resource Allocation in a Network Based Cloud Computing Environment: Design Challenges*; IEEE Communications Magazine, September 2013, DOI: 10.1109/MCOM.2013.6658651.
- [3] FETIA BANNOUR, SAMI SOUIHI, and ABDELHAMID MELLOUK; *Distributed SDN Control: Survey, Taxonomy and Challenges*; *IEEE Communications Surveys & Tutorials*, December 2017, DOI: 10.1109/COMST.2017.2782482.
- [4] WENJUAN LIAB, YU WANGA, ZHIPING JIN, KEPING YUD, JIN LIA, and YANG XIANGA; *Challenge-based Collaborative Intrusion Detection in Software Defined Networking: An Evaluation*; *Digital Communications and Networks*, 9 September 2020, <https://doi.org/10.1016/j.dcan.2020.09.003>.
- [5] JUAN CAMILO CORREA CHICA, JENNY CUATINDIOY IMBACHI, and JUAN FELIPE BOTERO VEGA; *Security in SDN: A comprehensive survey*; *Journal of Network and Computer Applications*, 19 March 2020, <https://doi.org/10.1016/j.jnca.2020.102595>.
- [6] ESMAEIL AMIRI, EMAD ALIZADEH, and KHALILOLLAH RAEISI; *An Efficient Hierarchical Distributed SDN Controller Model*; 5th Conference on Knowledge-Based Engineering and Innovation, September 2019.
- [7] STEFAN SCHMID, JUKKA SUOMELA; *Exploiting Locality in Distributed SDN Control*; *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*, ACM, 2013, pp. 121–126.
- [8] T. FRATCZAK, M. BROADBENT, P. GEORGOPOULOS, and N. RACE; *Homevisor: Adapting home network environments*, in: *Software Defined Networks*, 2013 Second European Workshop on, IEEE, 2013, pp. 32–37.
- [9] YUAN ZHANGA, LIN CUIA,B, WEIWANGC, and YUXIANG ZHANGA; *A Survey on Software Defined Networking with Multiple Controllers*; *Journal of Network and Computer Applications*, December 7, 2017, <https://doi.org/10.1016/j.jnca.2017.11.015>.

- [10] RAHIM MASOUDI, ALI GHAFARI; *Software defined networks: A survey*; Journal of Network and Computer Applications 67 (2016) 1–25, 26 March 2016, <http://dx.doi.org/10.1016/j.jnca.2016.03.016>.
- [11] MURAT KARAKUSA, ARJAN DURRESIA; *A Survey: Control Plane Scalability Issues and Approaches in Software-Defined Networking (SDN)*; Computer Networks November 10, 2016.
- [12] ASHUTOSH KUMAR SINGH, SHASHANK SRIVASTAVA; *A survey and classification of controller placement problem in SDN*; International Journal of Network Management · March 2018, DOI: 10.1002/nem.2018
- [13] SURENDRA KUMAR KESHARI, VINEET KANSAL, and SUMIT KUMAR; *A Systematic Review of Quality of Services (QoS) in Software Defined Networking (SDN)*; Springer Science+Business Media, 24 September 2020, <https://doi.org/10.1007/s11277-020-07812-2>.
- [14] YUSTUS EKO OKTIAN, SANGGON LEE, HOONJAE LEE, and JUNHUY LAM; *Distributed SDN controller system: A survey on design choice*; Computer Networks 121 (2017) 100–111, 11 April 2017, <http://dx.doi.org/10.1016/j.comnet.2017.04.038>.
- [15] AMIT S HANAMAKKANAVA, VIDYA S.HANDUR; *Load Balancing in Distributed Systems: A survey*; International Journal of Emerging Technology in Computer Science & Electronics(IJETCSE) ISSN: 0976-1353 Volume 14 Issue 2 –APRIL 2015.
- [16] JACOB H. COX, JR.1, JOAQUIN CHUNG, SEAN DONOVAN, JARED IVEY, RUSSELL J. CLARK, GEORGE RILEY, and HENRY L. OWEN; *Advancing Software-Defined Networks: A Survey*; IEEE Access, October 12, 2017, DOI: 10.1109/ACCESS.2017.276229.
- [17] JIE HU, CHUANG LIN, XIANGYANG LI, JIWEI HUANG; *Scalability of control planes for Software defined networks: modeling and evaluation*; IEEE 22nd international symposium of quality of service (IWQoS), Hong Kong. 2014. p. 147–52.  
<https://doi.org/10.1109/IWQoS.2014.6914314>
- [18] SOUKAINA BOUZGHIBA, HAMZA DAHMOUNI, ANOUAR RACHDI, AND JEAN-MARIE GARCIA; *Towards an Autonomic Approach for Software Defined Networks: An Overview*; <https://www.researchgate.net/publication/309693700>, November 2017, DOI: 10.1007/978-981-10-1627-1\_12.
- [19] YANING ZHOU, YING WANG, JINKE YU, JUNHUA BA, and SHILEI ZHANG; *Load Balancing for Multiple Controllers in SDN Based on Switches Group*; APNOM2017. IEEE. 2017.

[20] YING-DAR LIN, CHIH CHIANG WANG, YI-JEN LU, YUAN-CHENG LAI, and HSICHANG YANG; *Two-tier dynamic load balancing in SDN-enabled Wi-Fi networks*; Springer Science+Business Media New York 2017, 9 April 2017, DOI 10.1007/s11276-017-1504-3.

[21] GUDU BEKAMA HAILE, JIANXUN ZHANG; *Dynamic Load Balancing Algorithm in SDN-based Data Center Networks*; International Journal of Engineering Research & Technology (IJERT); ISSN: 2278-0181, Vol. 10 Issue 03, March-2021.

[22] VIVEK SRIVASTAVA, RAVI SHANKAR PANDEY; *Load balancing for software-defined network: a review*; International Journal of Computers and Applications, 27 April 2021, <https://doi.org/10.1080/1206212X.2021.1919835>.



## العوامل المؤثرة على تحسين الطلي

### الكيميائي بالفضة

الدكتور: أسامة عجمي

مدرس، قسم العلوم الأساسية، كلية الهندسة الميكانيكية و الكهربائية، جامعة دمشق،

#### ملخص

في بحثنا هذا درسنا تأثير تركيب عناصر المحلول الكيميائي و المؤلف من هيدروكسيد الصوديوم كوسط قلوي و هيدروكسيد الأمونيوم كمادة مخليية و طرطرات الصوديوم و البوتاسيوم كمادة مرجعة على عملية الطلي الكيميائي بمعدن الفضة علي سطح صقيل من الزجاج. لم تتجاوز سماكة طبقة التلبيس 0.3 ميكرو متر كما أن زمن ثبات هذه المحاليل لم يتجاوز 20 دقيقة  
تم زيادة كل من زمن الثبات و سماكة طبقة التلبيس عن طريق إضافة بعض المواد مثل السيسيتين كما تم دراسة تأثير زيادة درجة الحرارة بوجود المادة المضافة المختارة.

#### كلمات مفتاحية:

الطلي الكيميائي بالفضة ، ثباتية المحاليل، محاليل قلوية نشادرية ، سيسيتين، سيسيتيين، ميثيونين، ثنائي ايتيل ثنائي كربامات الصوديوم، ثيو سلفات الصوديوم، تأثير درجة الحرارة.

# Factors effect to improvement of the electroless silver plating

**Dr. Osama Ajami**

**Doctor, Department of Basic Sciences, Faculty of mechanical and electrical Engineering, Damascus university, Syria**

## Abstract

in this research, we studied the effect of the composition of the chemical solution, on electroless silver plating, on a smooth surface of glass. using sodium hydroxide as an alkaline medium, ammonium hydroxide as a chelating substance, and sodium and potassium tartrate as a reducing agent.

The thickness of the coating layer did not exceed 0.3  $\mu\text{m}$ , and the stability time of these solutions did not exceed 20 minutes

Both settling time and coating thickness were increased by adding some substances such as cysteine, and the effect of increasing the temperature was studied in the presence of the selected additives.

## Keywords:

electroless silver plating, solution stability, ammonium alkaline solutions, cystine cysteine, sodium thiosulfate, methionine, Diethyldithiocarbamate.Na, temperature effect.

## دراسة مرجعية:

تترسب الفضة بتأثير المرجمات الضعيفة نسبياً مثل الغلوكوز و أملاح الطرطرات من محاليل معقدات الفضة النشارية. و على الرغم من أن هذه المحاليل معروفة من زمن طويل و تطبق في الصناعة بشكل واسع الا أن جميع تطبيقاتها تتصف بقلّة الثباتية و ببطء شديد في السرعة.

يمكن اجراء عملية التفضييز على الأسطح المختلفة سواء كانت صقيله أو خشنة مثل الزجاج و الكوارتز و السيراميك و المواد البلاستيكية و الألياف الكربونية و كذلك النسيجية، [8,7] وخصوصاً من المحاليل الثابتة نسبياً لذلك كان من المناسب البحث عن التركيب الكيميائي الأنسب لمثل هذه المحاليل. غير أن تأثير إضافة العوامل المثبتة لعملية الطلي الكيميائي بالفضة من المحاليل القلوية النشارية غير مدروس بشكل كاف و المعلومات المتوافرة حول ذلك قليلة.

تمت دراسة تأثير تراكيز أيونات الفضة و الغلوكوز و هيدروكسيد الصوديوم و النشار و ارتفاع درجة حرارة المحلول على عملية الطلي [1]. , تم ترسيب الفضة من المحاليل القلوية النشارية على سطح الألياف النسيجية عند زمن قصير نسبياً [2]. كما تمت دراسة تأثير إضافة الميثيونين على ثباتية عملية الطلي الكيميائي بالفضة في المحاليل السابقة [6]. أشارت بعض الدراسات السابقة إلى إمكانية تحسين الطلي الكيميائي بالنحاس باستخدام إضافات مثل أو ثنائي ايتيل ثنائي ثيوكاربامات الصوديوم أو  $K_4Fe(CN)_6$  أو الميثيونين و الغلايسين [3,4] كما أشارت دراسات أخرى على دور الأحماض الأمينية الحاوية على الكبريت ( السيستين و الميثيونين ) ك مثبتات في عملية الطلي الكيميائي بالنيكل [5].

## هدف البحث

هدف هذه الدراسة إيضاح العوامل المختلفة على الطلي الكيميائي بالفضة بشكل عام في درجة حرارة الغرفة و عند درجات حرارة أعلى.

### طريقة العمل و شروط التجارب:

1. جميع المواد الكيميائية المستخدمة كانت درجة نقاوتها خاصة بالتحليل (Analytical grade) كما استخدمنا الماء المقطر لتحضير المحاليل.
2. تمت عملية التفضييز على سطح صقيل من الزجاج المنظف بواسطة المزيج الكرومي و منشط بمحلول من كلوريد القصدير المائي ( $\text{SnCl}_2 \cdot \text{H}_2\text{O}$ ) تركيزه 1% (w/v).
3. تم تعيين كتلة الطبقة المترسبة من الفضة بوزن القطعة الجافة قبل و بعد الترسيب باستخدام ميزان حساس، درجة حساسيته 0.1 mg.
4. تم تقدير ثباتية المحاليل بقياس الزمن اللازم لإرجاع الفضة ذاتياً ( أي الزمن الفاصل بين تحضير المحلول وحتى لحظة تعكيره). وعلى هذا الأساس تم حساب السرعة الوسطية لعملية الطلي الكيميائي.
5. تم اجراء التجارب بين مرتين و أربع مرات و سجلت النتائج الوسطية.
6. تمت التجارب عند درجة الحرارة  $(25 \pm 2)^\circ\text{C}$  ما عدا تجارب دراسة تأثير درجة الحرارة.
7. كانت درجة التحميل في كافة التجارب 4 دسم<sup>2</sup>/ليتر.
8. تم حساب عامل الإستفادة A من أجل تقييم تأثير إضافة المواد على عملية الطلي الكيميائي بالفضة من خلال العلاقة:

$$A = \frac{v \cdot t - v_0 \cdot t_0}{v_0 \cdot t_0}$$

حيث:

- v: السرعة الوسطية لعملية التلييس بوجود المادة المضافة.
- t: فترة ثباتية المحاليل المستخدمة في عملية الطلي بوجود المادة المضافة.
- v.t=h: سماكة طبقة التلييس بوجود المادة المضافة.
- v<sub>0</sub>: السرعة الوسطية لعملية التلييس بدون وجود المادة المضافة.
- t<sub>0</sub>: فترة ثباتية المحاليل المستخدمة في عملية الطلي بدون وجود المادة المضافة.
- v<sub>0</sub>.t<sub>0</sub>=h<sub>0</sub>: سماكة طبقة التلييس بوجود المادة المضافة.

A: عامل الاستفادة و كلما كان أكبر، كانت الفائدة المرجوة من إضافة المادة المحددة أكبر.

### أهم نتائج البحث

تتعلق سرعة عملية ارجاع شوارد الفضة من محاليلها النشارية بوجود المادة المرجعة (طرطرات) و بتركيز كل من شوارد الفضة و هيدروكسيد الأمونيوم. تم اختيار تركيز شوارد الأمونيوم أكبر من تركيز شوارد الفضة بحيث تكون النسبة بينهما من 10-40 مرة. تم اختيار التركيب الكيميائي التالي لدراسة تأثير تركيز شوارد الفضة على زمن الثباتية و سماكة طبقة التلييس

المادة	[NH <sub>3</sub> ]	NaOH	Tartrate.Na.K
التركيز	600, mM	100 mM	45, mM

تم عرض النتائج في الجدول (1)

جدول (1) تأثير تركيز شوارد الفضة على زمن الثباتية و سماكة طبقة التلييس

[Ag <sup>+</sup> ], mM	15	22	30	45
T <sub>0</sub> , min	20	17	15	15
h <sub>0</sub> , μm	0.2	0.22	0.20	0.20

ارتفاع تركيز شوارد الفضة حتى ثلاثة أضعاف لا يؤثر بشكل ملحوظ على سرعة العملية في حين تتناقص فترة الثباتية بشكل ضئيل.

تم اختيار التركيب الكيميائي التالي لدراسة تأثير النشار على زمن الثباتية سماكة طبقة التلييس

المادة	[Ag <sup>+</sup> ]	NaOH	Tartrate.Na.K
التركيز	45, mM	100 mM	45, mM

تم عرض النتائج في الجدول (2).

جدول (2) تأثير تركيز النشار على زمن الثباتية و سماكة طبقة التلييس

[NH <sub>3</sub> ], mM	200	400	600
t <sub>0</sub> , min	6	8	15
h <sub>0</sub> , μm	0.12	0.15	0.20

عند زيادة تركيز النشادر في المحلول من 200 mM و حتى 600 mM تزداد فترة الثباتية و يرافق ذلك زيادة سماكة طبقة التلييس.

تم اختيار التركيب الكيميائي التالي لدراسة تأثير المادة المرجعة ( طرطرات الصوديوم و البوتاسيوم) على زمن الثباتية و سماكة طبقة التلييس

المادة	[Ag <sup>+</sup> ]	NaOH	[NH <sub>3</sub> ]
التركيز	45, mM	100 mM	600, mM

تم عرض النتائج في الجدول (3).

جدول (3) تأثير تركيز طرطرات الصوديوم و البوتاسيوم على زمن الثباتية و سماكة طبقة

التلييس

[Tartrate], mM	15	30	45
t <sub>0</sub> , min	25	19	15
h <sub>0</sub> , μm	0.3	0.25	0.20

تؤدي زيادة تراكيز طرطرات الصوديوم و البوتاسيوم إلى تناقص ثباتية المحلول بشكل كبير و يرافق ذلك تناقص في سماكة طبقة التلييس.

من النتائج السابقة نخلص إلى أن السماكة العظمى لطبقة التلييس تكون بالمحافظة على ثباتية المحلول و تتزايد عند انقاص تركيز الطرطرات أو زيادة تركيز النشادر في المجالات المدروسة من التراكيز.

من أجل زيادة ثباتية المحاليل تم إضافة مواد مختلفة معروفة بقدرتها على زيادة ثباتية المحاليل و المستخدمة في الطلي الكيميائي بالنحاس أو النيكل و يمكن تقسيم النتائج إلى مجموعتين:

1. مواد مثل KCN و KI و  $K_3Fe(CN)_6$  ساعدت على ظهور العكر بشكل سريع و لذلك استبعدت من الدراسات اللاحقة.

2. مواد أعطت نتائج مقبولة مثل السيستئين و السيتين و الميثيونين و ثيو سلفات الصوديوم و دي ايتيل ثيو كربامات الصوديوم ولذلك تمت دراسة عملية الطلي بالفضة من المحاليل النشادرية باستخدام إضافات مختلفة التركيز من هذه المواد.

نتائج هذه الدراسة معروضة في الجدول (4) حيث تم استخدام المحلول ذي التركيب

الكيميائي التالي:

المادة	[Ag <sup>+</sup> ]	NaOH	[NH <sub>3</sub> ]	Tartrate.Na.K
التركيز	45, mM	100 mM	600, mM	15, mM

الجدول (4) تأثير الإضافات على عملية الطلي الكيميائي بالفضة.

No	المادة المضافة	Concentration (mg/l)	V (μM/h)	t, h	(A)
1	-	-	0.75	0.42	-
2	Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub>	0.79	0.23	3	2.19
3	Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub>	3.16	0.07	6	0.3
4	Cystine	1	0.4	1.5	1.86
5	Cystine	2	0.75	2.1	4
6	Cystine	4	0.2	3.6	1.3
7	Diethyldithiocarbamate.Na	2	1.1	2	6.0
8	Diethyldithiocarbamate.Na	4	0.45	5	6.1
9	Diethyldithiocarbamate.Na	20	0.03	20	0.9
10	Methionine	1	0.45	1.3	0.86
11	Methionine	2	0.6	2	2.8
12	Methionine	4	0.15	3.2	0.52
13	Cysteine	10	0.1	3	-0.05
14	Cysteine	20	0.12	3	0.14

من الجدول (4) يتضح أن أفضل المحاليل من حيث عامل الإستفادة هي المحاليل 2 و 4 و 5 و 7 و 8 و 11 ولكن مظهر و تجانس طبقة الطلي كان أفضل ما يمكن في المحلول رقم 5 الحاوي على السيستين كمادة مضافة لذلك تم استخدام هذه المادة في اختباراتنا اللاحقة.

في الجدول (5) توجد تأثير الإضافات المختلفة من السيستين على محاليل التلبس الكيميائي بالفضة عند تراكيز مختلفة من شوارد الفضة و النشادر و و بتركيز لمحلول هيدروكسيد الصوديوم يساوي 0.1M.

جدول (5) الصفات العامة للطلي بالفضة بوجود السيستين

No	التركيز، mmol/l			Cystine mg/l	t, hour	سماكة طبقة التلبس عند الزمن t (µm)
	Ag <sup>+</sup>	NH <sub>3</sub>	Tartrate			
1	45	200	15	2	1.5	1.7
2	45	200	15	4	1.5	1.9
3	45	200	15	8	3	1.06
4	90	400	15	2	1.0	2.5
5	90	400	15	4	1.5	3.2
6	90	400	15	8	2	1.1

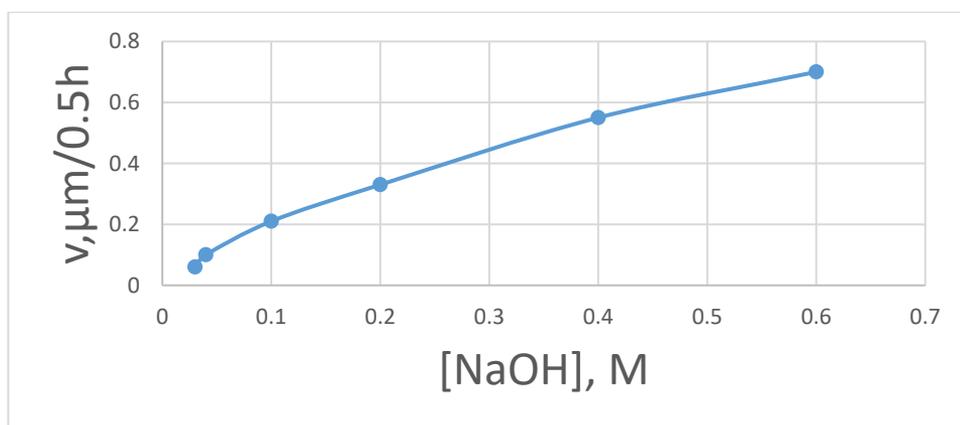
يتضح من الجدول (5) أنه يمكن الحصول على سماكة طبقة التلبس من 2 و حتى 3 ميكرون تقريباً في أغلب المحاليل المدروسة، و أنه بزيادة تركيز السيستين تتناقص سرعة التلبس و يتزايد زمن الثباتية t كما أن الكمية العظمى المترسبة تكون في المجال من 2-4 mg/l.

عند زيادة تركيز شوارد الفضة تزداد سرعة التلبس و بالاختيار المناسب للكميات يمكن الوصول إلى كتلة مترسبة كبيرة.

تزداد سرعة تفضيض بشكل كبير عند زيادة قلوية المحلول في المحاليل الحاوية على السيستين فعند زيادة تركيز هيدروكسيد الصوديوم من 0.03M و حتى 0.6M (الشكل 1).

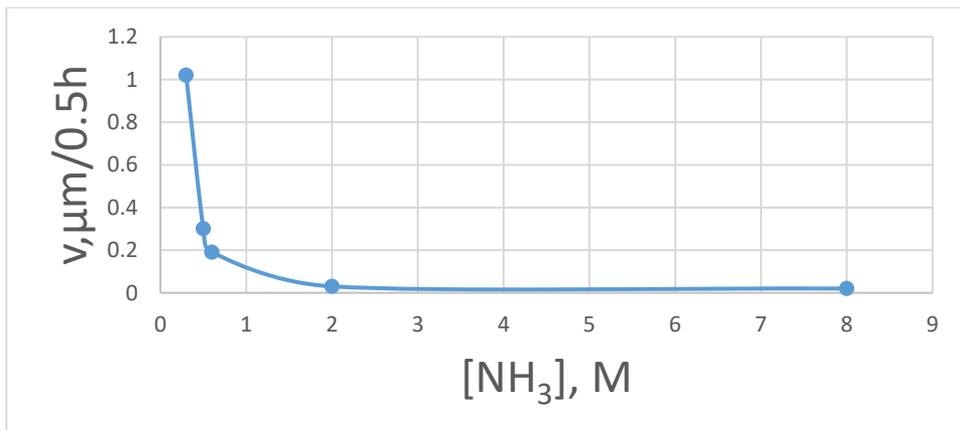
أي في مجال الـ pH من (12.5-13.5) تزداد السرعة أكثر من 10 مرات ولكن يلاحظ أنه في درجات الـ pH العالية يصبح مظهر طبقة التلبس داكناً و لذلك تم اختيار التركيز الأمثل لهيدروكسيد الصوديوم بين 0.1M و حتى 0.2M الذي يجمع بين السرعة المقبولة و المظهر الجيد.

يوضح الشكل (1) العلاقة ما بين سرعة الطلي و تركيز هيدروكسيد الصوديوم في المحلول الحاوي على السيستين بتركيز 2mg/l و و المؤلف من شوارد الفضة (45 mM) و النشادر (600 mM) و طرطرات الصوديوم و البوتاسيوم (15 mM).



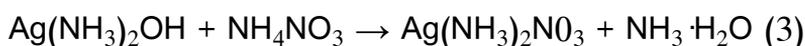
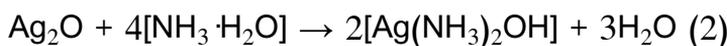
الشكل (1) علاقة سرعة الطلي بتركيز هيدروكسيد الصوديوم

من أجل إيضاح العلاقة بين تركيز النشادر وسرعة التفضيض تمت الدراسة في نفس المحلول المستخدم في الشكل (1) و لكن مع تغيير تركيز النشادر من 0.3M و حتى 8M و عند تركيز لهيدروكسيد الصوديوم يساوي 0.2M.



الشكل (2) علاقة سرعة الطلي بتركيز النشادر

من الشكل (2) يتضح أن سرعة التفضيض تتناقص بزيادة تركيز النشادر و لكن العملية لا تتوقف حتى في التراكيز العالية جداً. يمكن تفسير هذا التأثير بتناقص تركيز شوارد الفضة نتيجة لتشكيل المعقد  $Ag(NH_3)_2^+$  [2].



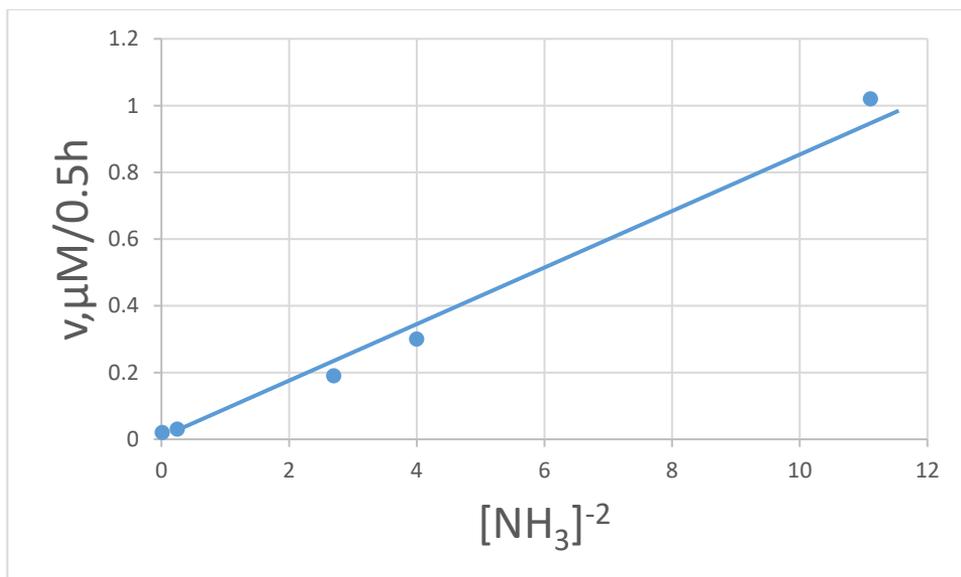
يتفكك المعقد  $Ag(NH_3)_2^+$  ببطء لإعطاء شوارد الفضة و النشادر من جديد:



$$K = \frac{[Ag^+] \cdot [NH_3]^2}{[Ag(NH_3)_2^+]}$$

$$[Ag^+] = K \frac{[Ag(NH_3)_2^+]}{[NH_3]^2} = K' \cdot [NH_3]^{-2}$$

من الشكل (3) يتضح أن سرعة التلييس تتناسب عكساً مع مربع تركيز النشادر و أن رتبة التفاعل تساوي -2 أي يتناسب تركيز شوارد الفضة التوازنية عكساً مع مربع تركيز النشادر.



الشكل (3) العلاقة ما بين سرعة الطلي و  $[NH_3]^{-2}$

يتعلق الاستخدام الأمثل لكمية الفضة ( $Ag^+$ ) المستخدمة في الطلي بدرجة التحميل فعند استخدام المحلول (5) وعند نفس الشروط الواردة في الجدول (5) يترسب حوالي 50% من كمية الفضة أما عند مضاعفة درجة التحميل فيترسب حوالي 75% من كمية شوارد الفضة في حين تتناقص سماكة الطبقة المترسبة من 2.5 ميكرومتر إلى 1.9 ميكرومتر بسبب زيادة مساحة السطح المطلي بالفضة.

يمكن تنشيط الطلي الكيميائي عادة برفع درجة الحرارة و يترافق ذلك مع انخفاض في ثباتية المحلول و ارتفاع في سرعة الترسيب لذلك تمت دراسة العملية عند درجة الحرارة  $(50 \pm 2)^\circ C$  و بتركيز لهيدروكسيد الصوديوم يساوي 0.1M و بدرجة تحميل 2 دسم<sup>2</sup> لكل ليتر.

الجدول (4) التفضيض بوجود السيستين عند الدرجة  $50^\circ C$

No	التركيز ، mM/l			Cystine mg/l	T, h	t.v, μm	v, μm/h
	Ag <sup>+</sup>	NH <sub>3</sub>	Tartrate				
1	20	150	5	2	0.5	1.4	2.8
2	20	300	5	2	0.75	2.1	2.8

3	20	600	5	2	1.5	2.7	1.8
4	20	150	5	2	0.25	1.5	6.0
5	20	300	5	2	0.5	1.5	3.0
6	20	600	5	2	1	2.5	2.5
7	20	1200	5	2	1	2.2	2.2
8	90	400	15	4	0.25	1.8	7.2
9	90	600	15	4	0.5	3.9	7.8
10	90	800	15	4	0.75	3.7	4.8
11	90	400	15	4	0.17	2.6	15
12	90	600	15	4	0.5	3.8	7.6
13	90	800	15	4	0.5	5.0	10
14	90	1600	15	4	0.5	3.8	7.5

يتضح أنه بالرغم من نقصان ثباتية المحلول إلا أن سماكة طبقة التلييس يمكن أن تصل إلى 5 ميكرومتر ( المحلول رقم 13) و خلال زمن قصير نسبياً، كما أن مظهر طبقة التلييس أفضل من تلك الناتجة عند درجة حرارة الغرفة.

#### الاستنتاجات:

- (1) الطلي الكيميائي بالفضة من المحاليل القلوية يعاني من انخفاض في زمن ثبات المحاليل و سماكة طبقة التلييس.
- (2) تؤدي إضافة بعض المركبات الحاوية على الكبريت في المحاليل القلوية النشادرية للتلييس الكيميائي بالفضة إلى زيادة ثباتية هذه المحاليل عشرات المرات كما زادت طبقة الفضة المترسبة حتى 10 مرات.
- (3) الإضافة المثلى في شروط التجربة للسيستين كانت بتركيز من 4-2 mg/l.
- (4) للحصول على سماكات جديدة لطبقة التلييس يجب إضافة السيستين بالمقادير الموصوفة مع رفع درجة الحرارة حتى 50°C.
- (5) رغم المحاولات العديدة لم نستطع الحصول على محاليل قلوية نشادرية للتلييس الكيميائي ثابتة بشكل كامل.

## References

1. Bo.Ban, et al 2021, Preparation of Nano–Porous Silver by Electroless Plating, J. Phys.: Conf. Ser. 1838 012030.
2. Chen.Ming, 2020, A Method of Electroless Silver Plating on the Surface of PA6 Fiber, IOP Conf. Ser.: Mater. Sci. Eng. 774 012045.
3. F. Pearstein, and R. Weightman, 1974, Electroless Deposition of Silver, Plating, 61, N2, 174–157.
4. VASKALIS.A, JACIAUSKIENE.J,2007 Accelerating effect of ammonia on electroless copper deposition in alkaline formaldehyde–containing solution, Journal of Electroanalytical Chemistry. 600 (2007) 6-12
5. Wang. Ke, Hong. Liang,2008, Investigation into the Roles of Sulfur–Containing Amino Acids in Electroless Nickel Plating Bath, Ind. Eng. Chem. Res, 47, 6517–6524.
6. Wang. Xu, Zhang.Cheng, Zhou.Hongqiang, 2016, Effect of additives on electroless silver plating, 6th International Conference on Machinery, Materials, Environment, Biotechnology and Computer (MMEBC).

7. Xu Chanju, Zhou Ruihua, Chen Huiyu, 2014, Silver-coated glass fibers prepared by simple electroless plating technique, Journal of Material in Electronics, 25(10), 4638-4642.
8. Yi. Feng, Hailong Yuan, 2004, Electroless plating of carbon nanotubes with silver, Journal of material science 39(9), 3241-3243.

## دراسة تأثير العوامل المختلفة على الطلي الكيميائي للألياف الكربونية بطبقة النيكل - بور

الدكتور: أسامة عجمي

مدرس، قسم العلوم الأساسية، كلية الهندسة الميكانيكية والكهربائية، جامعة دمشق

### ملخص

في بحثنا هذا نقدم تأثير دراسة تأثير بعض العوامل المعروفة بفعاليتها على سرعة الطلي الكيميائي بالنيكل في الأوساط القلوية الحاوية على هيدريد الصوديوم و البور كمادة مرجعة. تم استخدام كلوريد القصدير كمادة مثبتة غير ضارة بالبيئة عوضاً عن أملاح الرصاص الأوسع انتشاراً وتم تحديد النسبة الملائمة للحصول على أعلى سماكة ضمن شروط التجربة. تمت دراسة تأثير كل من الـ pH و تركيز كل من المادة المرجعة و المادة المخليبية و درجة الحرارة ودرجة التحميل على سرعة ترسيب طبقة نيكل-بور في الأوساط القلوية.

### كلمات مفتاحية:

الطلي بدون تيار كهربائي، الطلي الكيميائي ، طبقة نيكل-بور، ألياف كربونية، وسط قلوي، كلوريد القصدير، هيدريد الصوديوم و البور، مادة مخليبية، مادة مثبتة.

# Study of the effect of different factors on the chemical coating of carbon fibers with a layer of nickel-bor

**Dr. Osama Ajami**

**Doctor, Department of Basic Sciences, Faculty of mechanical and electrical Engineering, Damascus university, Syria**

## Abstract

In this paper, we present the effect of studying the effect of some factors known to be effective on the speed of chemical nickel plating in alkaline media containing  $\text{NaBH}_4$  as a reducing agent. Tin chloride was used as a stabilizer that is not harmful to the environment instead of the more widespread lead salts, and the appropriate ratio was determined to obtain the highest thickness within the conditions of the experiment.

The effect of pH, concentration of the reducing agent, chelating agent, temperature and loading degree on the deposition rate of Nickel-Bor layer in alkaline solutions was studied.

## Keywords:

Electroless plating, Chemical coating, graphite fibers nickel-borate layer, alkaline medium, tin chloride, sodium hydride and boron, chelating agent, stabilizer.

## دراسة مرجعية:

تشغل الألياف الكربونية المطلية بطبقة معدنية اهتماماً متزايداً في التطبيقات الصناعية المختلفة مثل صناعة السيارات و المركبات المختلفة ولذلك فإن تحسين مواصفات هذه الألياف يعتبر هدفاً لكثير من الدراسات.

توجد عدة طرق لإنشاء طبقة معدنية على الأسطح المختلفة مثل الطلي الكيميائي و الطلي الالكتروكيميائي و الترسيب من الطور الغازي. و لكن الطريق الأهم هي الطلي الكيميائي بسبب تساوي سماكة طبقة التلييس على طول الليف الكربوني.

يكتسب طلي المواد المختلفة مثل البلاستيك و الزجاج و الألياف الكربونية أهمية خاصة بسبب الصلابة و القساوة ومقاومة التآكل التي تبديها القطعة المطلية.

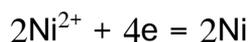
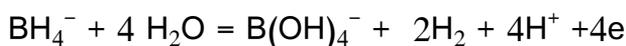
يتم ترسيب النيكل كيميائياً بدون استخدام التيار الكهربائي بطريقتين أساسين: في وسط حمضي بترسيب طبقة نيكول-فوسفور أو بوسط قلوي و أهمها ترسيب طبقة نيكول-بور في هيدريد الصوديوم و البور [3].

اكتشفت طريقة التلييس الكيميائي بالنيكل بوجود هيدريد الصوديوم و البور في منتصف القرن العشرين و لكن ما يحد من استخدامها حالياً متطلبات حماية البيئة بسبب استخدامها لأملاح المعادن الثقيلة مثل الرصاص كمثبتات لمحلول الطلي الكيميائي القلوي [1,2]. و لذلك حاولت بعض الأبحاث إيجاد الشروط الملائمة للاستغناء عن هذه المثبتات [5] مما تطلب تخفيض تركيز  $\text{NaBH}_4$  مما سبب انخفاضاً كبيراً في سرعة الترسيب.

تم التركيز في أبحاث أخرى على استبدال هذه المثبتات بمواد أقل ضرراً مثل شوارد البزموت [2] أو كلوريد القصدير أو أكسيد التنغستين [1].

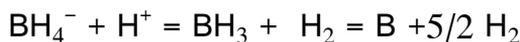
يمكن تقسيم تفاعلات تكون طبقة Ni-B الى المراحل التالية [6]:

- تتألف المرحلة الأولى من تفاعل ارجاع النيكل:

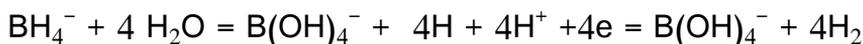




• تتألف المرحلة الثانية من تفاعل ارجاع هيدريد البور:



• في المرحلة الثالثة تتحلله هيدريدات البور:



### هدف البحث

هدف هذه الدراسة ترسيب طبقة من النيكل على السطح الخارجي للألياف الكربونية في الأوساط القلوية باستخدام هيدريد الصوديوم و البور و كلوريد القصدير كمادة مرجعة غير ضارة بالبيئة ودراسة تأثير العوامل المختلفة سرعة العملية وسماكة الطبقة المعدنية المترسبة.

### المواد المستخدمة و طريقة العمل:

1. تحضير الألياف الكربونية: تم غلي الألياف الكربونية في الماء المقطر لمدة 30 دقيقة للتخلص من الدهون و الأوساخ المحتملة ثم نقعها في محلول مؤلف من كلوريد القصدير (100 غ/ل) وحمض كلور الماء المركز (50 مل/ل) و أخيرا تمت عملية التنشيط في محلول كلوريد البالاديوم (4 غ/ل) و ايتلين دي أمين تترا حمض الخل (12 غ/ل) و هيدروكسيد الأمونيوم (130 مل/ل).
2. تم استخدام المواد التالية كمحلول للطلي الكيميائي: كلوريد النيكل سداسي الماء كمصدر للنـيكل - هيدروكسيد الصوديوم لتثبيت قلوية مكونات الحوض - ايتلين دي أمين كمادة مخليبة - كلوريد القصدير كمادة مثبتة-هيدريد الصوديوم و البور كمادة مرجعة. جميع المواد الكيميائية المستخدمة كانت درجة نقاوتها خاصة بالتحليل (Analytical grade) كما استخدمنا الماء المقطر لتحضير المحاليل.

3. طريقة العمل: تم حساب سرعة ترسيب النيكل R من خلال حساب فرق الوزن بين العينة المجففة عند الدرجة 90°C قبل التلييس و بعده باستخدام ميزان حساس درجة حساسيته 0.1 mg من خلال العلاقة [6]:

$$R = \frac{w \cdot 10^4}{A \cdot \rho \cdot t}$$

حيث:

W الفرق في الوزن (غ)

A=LX2.6 مساحة السطح الجانبي لليف الكربوني(سم<sup>2</sup>)

L طول الليف الكربوني و 2.6 محيط مقطعه.

$\rho$  الكثافة الوسطية لطبقة Ni-B المترسبة مع الأخذ بعين الاعتبار أن النسبة التقريبية للبور في هذه الطبقة 6%.

4. تم اجراء التجارب بين مرتين و أربع مرات و سجلت النتائج الوسطية.

5. حجم محلول الترسيب الكيميائي لطبقة نيكول-بور في كل التجارب كان 250 مل.

أهم نتائج البحث

1. تأثير تركيز المادة المثبتة:

تم اختيار التركيب الكيميائي التالي لدراسة تأثير تركيز كلوريد القصدير على سماكة الطبقة المترسبة و سرعة عملية الطلي:

المادة	NiCl <sub>2</sub> .6H <sub>2</sub> O	NaOH	Ethylenediamine	NaBH <sub>4</sub>
التركيز	30 g/l	40 g/l	90 ml/l	0.7 g/l

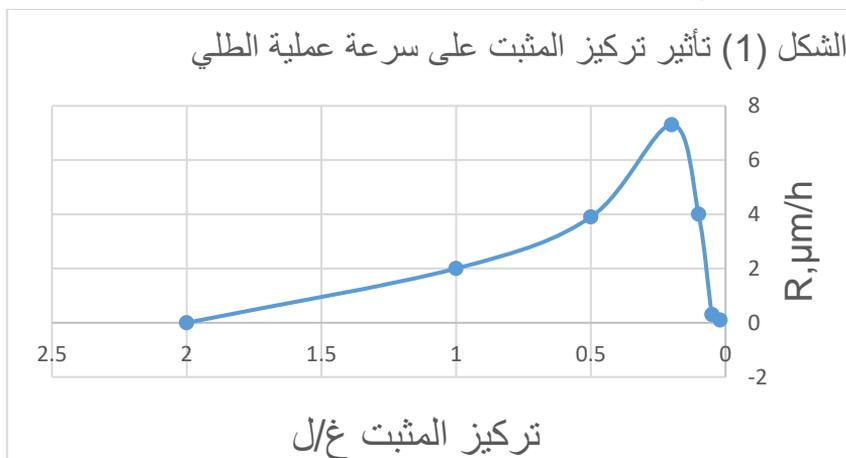
تمت العملية عند درجة حرارة 90°C و لمدة ساعة واحدة و كمية تحميل تساوي متر واحد من الليف الكربوني مع تغيير تركيز المادة المثبتة ضمن المجال (-0.02-2.0) غ/ل.

تم عرض النتائج في الجدول (1)

جدول (1) تأثير تركيز المثبت على سرعة عملية الطلي

SnCl <sub>2</sub> , g/l	0.02	0.05	0.1	0.2	0.5	1	2.0
R, μm/h	0.1	0.3	4	7.3	3.9	2	0

عند ازدياد تركيز كلوريد القصدير تزداد سرعة الترسيب حتى التركيز 0.2 غ/ل بسبب زيادة سرعة تفكك هيدريد الصوديوم و البور بواسطة المثبت ثم تتناقص بسرعة حتى الصفر تقريباً عند الوصول للتركيز 2.0 غ/ل و ذلك بسبب ادمصاص شوارد القصدير على الليف الكربوني و شغل و تعطيل المراكز الفعالة للبالاديوم. الشكل (1)



## 2. تأثير تركيز المادة المرجعة (NaBH<sub>4</sub>):

تم اختيار التركيب الكيميائي التالي لمكونات الحوض من أجل لدراسة تأثير المادة المرجعة على سماكة الطبقة المترسبة و سرعة عملية الطلي:

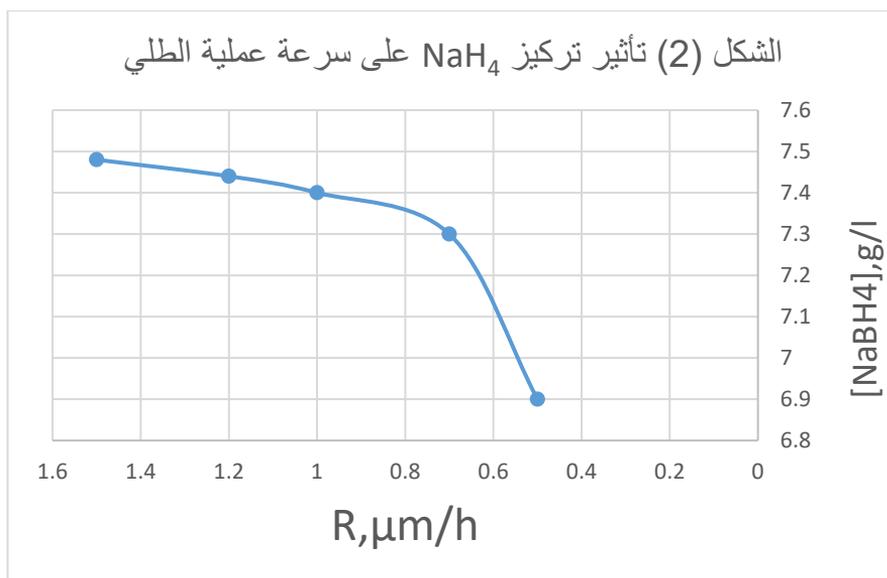
المادة	NiCl <sub>2</sub> .6H <sub>2</sub> O	NaOH	Ethylenediamine	SnCl <sub>2</sub>
التركيز	30 g/l	40 g/l	90 ml/l	0.2 g/l

تمت العملية عند درجة حرارة 90°C و لمدة ساعة واحدة و كمية تحميل تساوي متر واحد من الليف الكربوني مع تغيير تركيز المادة المرجعة ضمن المجال (0.5-1.5 g/l). تم عرض النتائج في الجدول (2)

جدول (2) تأثير تركيز NaH<sub>4</sub> على سرعة عملية الطلي

NaBH <sub>4</sub> , g/l	0.5	0.7	1	1.2	1.5
R, μm/h	6.9	7.3	7.4	7.44	7.48

لا تؤثر كمية  $NaBH_4$  سماكة الطبق المترسبة بشكل ملحوظ ضمن المجال المدروس من التراكيز ( الشكل 2) لذلك تم اختيار التركيز 0.7 غ/ل لدراساتنا اللاحقة على اعتبار أن زيادة تركيز هيدريد الصوديوم و البور ستزيد من نسبة البور في الطبقة المترسبة.



### 3. تأثير تركيز دي إيتيلين أمين:

لدراسة تأثير دي إيتيلين أمين على طبيعة العملية تم تغيير تركيز هذه المادة ضمن المجال ( 60-140) مل/ل واجراء التليبس عند درجة حرارة  $90^{\circ}C$  و لمدة ساعة واحدة و كمية تحميل تساوي متر واحد من الليف الكربوني في الحوض الطلي:

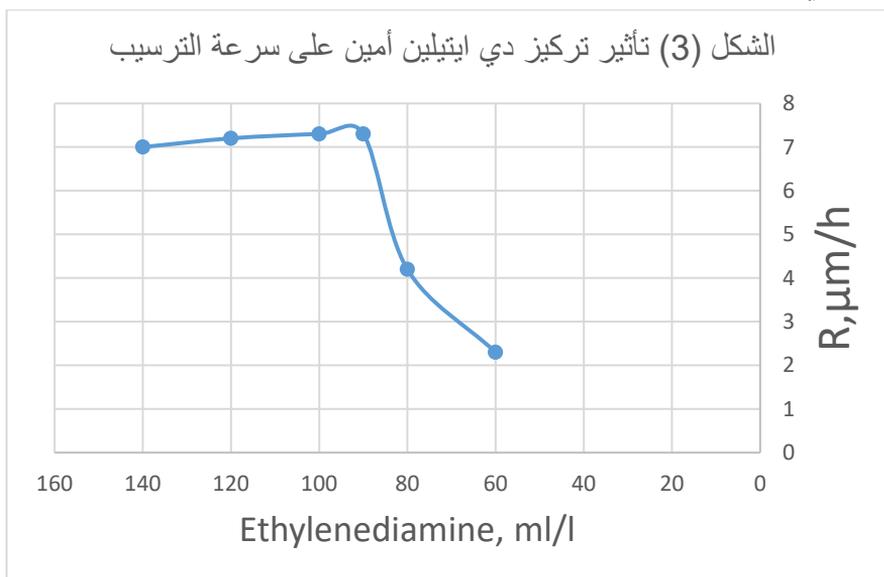
المادة	$NiCl_2 \cdot 6H_2O$	NaOH	$NaBH_4$	$SnCl_2$
التركيز	30 g/l	40 g/l	0.7 g/l	0.2 g/l

تم عرض النتائج في الجدول (3)

الجدول (3) تأثير تركيز دي إيتيلين أمين على سرعة الترسيب

Ethylenediamine, ml/l	60	80	90	100	120	140
R, μm/h	2.3	4.2	7.3	7.3	7.2	7

تزداد سرعة التلبس زيادة ملحوظة عند رفع تركيز إيتيلين دي أمين في التركيب الكيميائي للحوض بسبب الطبيعة المخليبة لهذه المادة التي تمنع تحويل شوارد النيكل إلى هيدروكسيد النيكل في الأوساط شديدة القلوية. تبلغ سرعة التلبس حدها الأعظمي عند التركيز 90 مل/ل ثم تتخفض بعد ذلك بكل ضئيل. شكل 3.



#### 4. تأثير الـ $pH$ المحلول على سرعة الترسيب:

لدراسة تأثير كمية هيدروكسيد الصوديوم ( الـ  $pH$  ) على سرعة الطلي للنيكل تم حساب سرعة طلي النيكل على الألياف الكربونية من حوض التلبس التالي:

	$NiCl_2 \cdot 6H_2O$	$NaBH_4$	Ethylenediamine	$SnCl_2$
التركيز	30 g/l	0.7 g/l	90 ml/l	0.2 g/l

عند درجة حرارة  $90^\circ C$  و لمدة ساعة واحدة و كمية تحميل تساوي متر واحد من الليف الكربوني. عند تغيير تركيز  $NaOH$  ضمن المجال (20-120) غ/ل تزداد سماكة طبقة التلبس بشكل مضطرب كما هو متوقع حيث كان مجال سماكة طبق التلبس عند الشروط السابقة من 4.7 و حتى 15 ميكرومتر.

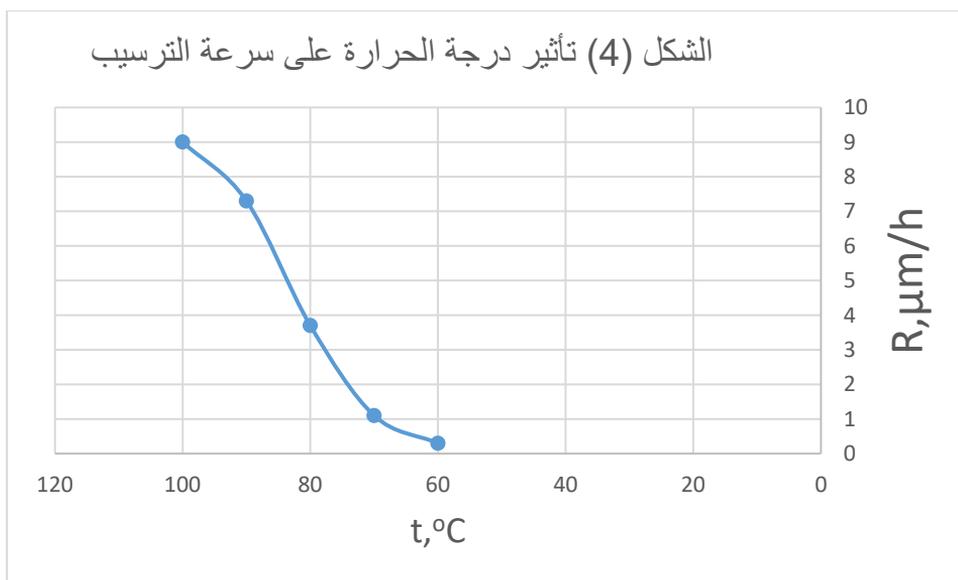
## 5. تأثير درجة الحرارة على سرعة الترسيب

كما هو معلوم تزداد سرعة الترسيب للنيكل من المحاليل القلوية الحاوية على هيدريد البور بزيادة درجة الحرارة لذلك تمت دراسة العملية عند مجال من درجات الحرارة من  $60^{\circ}\text{C}$  و حتى  $100^{\circ}\text{C}$  باستخدام حوض الترسيب المماثل في الشروط و التركيب للحوض الوارد في الفقرة 4. تم تسجيل النتائج في الجدول 4.

الجدول (4) تأثير درجة الحرارة على سرعة الترسيب

T, °C	60	70	80	90	100
R, $\mu\text{m}/\text{h}$	0.3	1.1	3.7	7.3	9.0

كما هو ملاحظ تزداد سرعة التلييس بزيادة درجة الحرارة. (الشكل 4)



## 6. تأثير كمية التحميل على سرعة الترسيب

تمت دراسة تأثير كمية التحميل على سرعة عملية التلييس باستخدام حوض الترسيب المماثل في الشروط و التركيب للحوض الوارد في الفقرة 4 عند درجة الحرارة  $90^{\circ}\text{C}$

وذلك بتغيير طول الليف الكربوني من 1 متر و حتى 5 م. تم تسجيل النتائج في الجدول 5.

طول الليف الكربوني، م	1	2	3	4	5
R, $\mu\text{m}/\text{h}$	7.3	3.7	1.6	1.38	1.2

تنخفض سماكة الطبقة المترسب بزيادة درجة التحميل بشكل واضح.

### الاستنتاجات و التوصيات

- 1) استخدام كلوريد القصدير كمثبت لمحلول التلبيس الكيميائي للنيكل على الألياف الكربونية عوضاً عن أملاح الرصاص السامة و الضارة بالبيئة أعطى نتائج جيدة خلال الفترة المدروسة و كانت أفضل سرعة تلبيس عند التركيز 0.2 غ/ل
- 2) لا تؤثر كمية هيدريد الصوديوم و البور في المجال المدروس من التراكيز على سرعة الترسيب بينما تبلغ هذه السرعة قيمة حدية تساوي 90 مل/ل عند تغيير تركيز دي ايتلين أمين.
- 3) ارتفاع كل من درجة الحرارة و الـ pH يزيد من سماكة طبقة نيكول-بور وتبلغ حتى 15 ميكرومتر خلال ساعة من الزمن.
- 4) الاقتراح بدراسة تفصيلية للخواص الميكانيكية و البنيوية لطبقة التلبيس من المحلول المقترح على الألياف الكربونية.

## References

- 1) L. Bonin, V. Vitry, F. Delaunois, 2020, Inorganic salts stabilizers effect in electroless nickel–boron plating: Stabilization mechanism and microstructure modification, j Surface & Technology 401 126276.
- 2) L. Bonin, V. Vitry, F. Delaunois, 2019, Replacement of Lead stabilizer in electroless Nickel-Boron baths: Synthesis and characterization of coatings from bismuth stabilized bath, Journal Pre-proof.
- 3) MORDECHAY SCHLESINGER and Milan Paunovic, 2010, Modern Electroplating, Fifth Edition Copyright John Wiley & Sons, Inc. p 44-458
- 4) Mustafa Anik, Erhan Körpe, Esin Şen, 2008, Effect of coating bath composition on the properties of electroless nickel–boron films, Surface & Coatings Technology 202 (2008) 1718–1727.
- 5) Muslum Yunacti , AlexandreMégret , Mariana Henriette Staia , Alex Montagne and Véronique Vitry, 2021, Characterization of Electroless Nickel–Boron Deposit from Optimized Stabilizer–Free Bath, Coatings 2021, 11, 576.
- 6) Q. Barati , Seyed Mohammad Mehdi Hadavi, 2020, Electroless Ni–B and Composite Coatings: A Critical Review on Formation mechanism, Properties, Applications and Future trends, Journal Pre-proof.

