

مجلة جامعة حمص

سلسلة العلوم الهندسية الميكانيكية
والكهربائية والمعلوماتية



مجلة علمية محكمة دورية

المجلد 47 . العدد 6

1447 هـ - 2025 م

الأستاذ الدكتور طارق حسام الدين رئيس جامعة حمص

المدير المسؤول عن المجلة

أ. د. وليد حمادة	رئيس تحرير مجلة جامعة حمص للعلوم الإنسانية
د. نعيمة عجيب	رئيس تحرير مجلة جامعة حمص للعلوم الطبية والهندسية والأساسية والتطبيقية

د.محمد فراس رمضان	عضو هيئة التحرير
د. مضر سعود	عضو هيئة التحرير
د. ممدوح عبارة	عضو هيئة التحرير
د. موفق تلاوي	عضو هيئة التحرير
د. طلال رزوق	عضو هيئة التحرير
د. أحمد الجاعور	عضو هيئة التحرير
د. الياس خلف	عضو هيئة التحرير
د. روعة الفقس	عضو هيئة التحرير
د. محمد الجاسم	عضو هيئة التحرير
د. خليل الحسن	عضو هيئة التحرير
د. هيثم حسن	عضو هيئة التحرير
د. أحمد حاج موسى	عضو هيئة التحرير

تهدف المجلة إلى نشر البحوث العلمية الأصيلة، ويمكن للراغبين في طلبها
الاتصال بالعنوان التالي:

رئيس تحرير مجلة جامعة حمص

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

. هاتف / فاكس : 2138071 31 963 ++

. موقع الإنترنت : www.homs-univ.edu.sy

. البريد الإلكتروني : journal.homs-univ.edu.sy

ISSN: 1022-467X

شروط النشر في مجلة جامعة حمص

الأوراق المطلوبة:

- 2 نسخة ورقية من البحث بدون اسم الباحث / الكلية / الجامعة) + word / CD
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 (كتابة 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 حمص

1. دفع رسم نشر (50000) ل.س أربعون ألف ليرة سورية عن كل بحث لكل باحث يريد نشره في مجلة جامعة البعث.
2. دفع رسم نشر (200000) ل.س مئة 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 (15000) ل.س ستة آلاف ليرة سورية رسم موافقة على النشر من كافة الباحثين.

المحتوى

الصفحة	اسم الباحث	اسم البحث
38-11	م. هبة تدمري أ.د. إبراهيم الشامي أ.د. عمار زقزوق	مقارنة تأثير الترميز الموضوعي في أنظمة كشف التسلل المعتمدة على المحولات القابلة للتفسير
60-39	م. هديل عبد الرحمن أ.د. محمد فاضل سكر	منظومة لغوية للتحليل الصرفي والنحوي للجمل العربية باستخدام التعلم العميق AraBERT - Transformer
98-61	م. محمد الهلامي د. أكرم مرعي د. ماهر عباس	تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات إنترنت الأشياء التي تعتمد على بروتوكول التوجيه (RPL)
142-99	م. محمد الهلامي د. أكرم مرعي د. ماهر عباس	تحسين اختبار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL) بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء
158-143	م. محمد رامي الجندي د. عبّاد كاسوحة	تأثير تغيير درجة حرارة الارجاع على بنية وماتنة الكرات الطاحنة المصنوعة من الحديد الصب عالي الكروم

193-159	م.إيلياس شلوف أ.د.ناصر أبو صالح د.أسامة ناصر	دراسة سرعة أداء أنظمة الربط العلانقي المستخدمة في تطوير التطبيقات
---------	---	--

مقارنة تأثير الترميز الموضعي في أنظمة كشف التسلسل المعتمدة على المحولات القابلة للتفسير

م. هبة تدمري* أ.د. إبراهيم الشامي* أ.د. عمار زقزوق*

الملخص

مع تزايد تعقيد التهديدات السيبرانية، برزت نماذج المحولات (Transformers) كحل فعال لكشف الهجمات في الشبكات الحاسوبية، استناداً إلى نجاحها في معالجة اللغة والرؤية الحاسوبية. يُعد الترميز الموضعي مكوناً أساسياً في بنية المحولات لفهم العلاقات بين الميزات وتحسين دقة الكشف. تهدف هذه الدراسة إلى تقييم تأثير أساليب الترميز الموضعي في أداء المحولات في عملية الكشف عن الهجمات من خلال مقارنة ثلاثة نماذج: بدون ترميز (NoPos) وبترميز جيبّي (SinPos) وبترميز قابل للتعلم (LearnPos)، تم اختبار النماذج على مجموعتي بيانات CICIDS2017 و UNSW-NB15 مع معالجة اختلال التوازن باستخدام focal loss والأوزان اللوغاريتمية. أظهر نموذج LearnPos تفوقاً بدقة 98.55% في مجموعة بيانات CICIDS2017 و 97.64% في مجموعة بيانات UNSW-NB15 متفوقاً على SinPos و NoPos. كما بيّن تحليل التفسيرية باستخدام أوزان الانتباه وتقنية LIME (Local Interpretable Model-agnostic Explanations) أن LearnPos يركّز على سمات الهجوم، ما يعزز شفافية قراراته وفعاليته في كشف الفئات النادرة مثل Web Attack، ما يجعله مناسباً لتطبيقات الأمن السيبراني الذكي.

الكلمات المفتاحية: المحولات - التفسير الموضعي - كشف هجمات الشبكة - LIME - أوزان الانتباه - الأمن السيبراني.

م. هبة تدمري طالبة دراسات عليا دكتوراه في كلية الهندسة الميكانيكية والكهربائية -قسم هندسة التحكم الآلي والحواسيب.

أ.د. إبراهيم الشامي أستاذ في كلية الهندسة الميكانيكية والكهربائية -قسم هندسة التحكم الآلي والحواسيب.

أ.د. عمار زقزوق أستاذ في كلية الهندسة الميكانيكية والكهربائية -قسم هندسة التحكم الآلي والحواسيب.

Comparing the Impact of Positional Encoding in Interpretable Transformer-Based on Intrusion Detection Systems

Abstract

With the increasing complexity of cyber threats, transformer models have emerged as an effective solution for network intrusion detection, leveraging their success in natural language processing and computer vision. Positional encoding is a critical component in transformer for understanding feature relationships and enhancing detection accuracy. This study evaluates the impact of positional encoding strategies on transformer performance by comparing three approaches: no positional encoding (NoPos), fixed sinusoidal encoding (SinPos), and learnable positional encoding (LearnPos). The models were tested on the CICIDS2017 and UNSW-NB15 datasets, with class imbalance addressed using focal loss and logarithmic weighting. LearnPos achieved superior performance with 98.55% accuracy on CICIDS2017 and 97.64% on UNSW-NB15, outperforming SinPos and NoPos. Interpretability analysis using attention weights and LIME revealed that LearnPos focuses on attack-related features, enhancing decision transparency and effectiveness in detecting rare attack categories, such as Bot and Web Attack, making it ideal for intelligent cybersecurity applications.

Keywords: Transformers - Positional Encoding - Network Attack Detection - LIME - Attention Weights - Cybersecurity.

1- مقدمة:

أدى التسارع الكبير في التطور التكنولوجي لمؤسسات القطاعين العام والخاص إلى تطور التهديدات السيبرانية وانتشارها، ما يشكل خطراً متزايداً على الاستقرار الاقتصادي والاجتماعي. وفقاً لتقرير المنتدى الاقتصادي العالمي لأمن المعلومات 2025، تتوقع 72٪ من المؤسسات تصاعداً كبيراً في الهجمات، بينما يرى 66٪ أن الذكاء الاصطناعي سيكون عنصراً محورياً في تحسين الدفاع السيبراني [1].

تُعدّ هجمات DDoS والبوت نت من أبرز التهديدات، إذ قد تتسبب بخسائر مالية وتعطيل للخدمات، فقد تعجز أنظمة الكشف التقليدية المعتمدة على قواعد ثابتة عن رصد الهجمات المتقدمة أو غير المعروفة [2]، لذا ظهرت الحاجة إلى حلول ذكية أكثر تكيفاً مع تطور التهديدات.

برزت نماذج المحولات كحل واعد ضمن تقنيات التعلم العميق لكشف التسلسل، وذلك بفضل قدرتها على التعامل مع البيانات المعقدة واستخلاص الأنماط من تسلسل الحزم. لكن تطبيقها في المجال الشبكي يواجه تحديات، خاصة فيما يتعلق بترميز الموضع (Positional Encoding)، إذ أن بيانات الشبكة لا تتبع ترتيباً زمنياً صريحاً كما في النصوص، ما يستدعي تطوير استراتيجيات ترميز فعالة لفهم العلاقات البنوية بين الميزات [3].

بناءً على ذلك، يقترح هذا البحث نموذجاً قائماً على المحولات يعالج كل سجل بيانات بشكل مستقل، مستفيداً من آلية الانتباه الذاتي لاكتشاف التفاعلات بين الميزات داخل كل عينة، كما يُجري مقارنة بين ثلاث استراتيجيات ترميز موضعي: بدون ترميز والترميز الجيبي والترميز القابل للتعلم، على مجموعتي بيانات واقعتين (CICIDS2017 و UNSW-NB15)، ويُدمج ذلك مع تقنيات معالجة عدم التوازن (مثل Focal Loss والأوزان اللوغاريتمية)، بالإضافة إلى أدوات تفسير (LIME وتحليل أوزان الانتباه) لتعزيز شفافية النموذج وتقديم رؤى قابلة للتفسير لمتخصصي الأمن السيبراني.

تتقسم الورقة على النحو الآتي: يقدم القسم الثاني مشكلة البحث، ويعرض القسم الثالث والرابع هدف البحث وأهميته، أما القسم الخامس فيبين النموذج المقترح، بينما يوضح القسم السادس الإعدادات التجريبية ومجموعات البيانات، ويناقش القسم السابع النتائج وتفسيراتها، ويختتم القسم الثامن بالاستنتاجات وآفاق العمل المستقبلي.

2- مشكلة البحث:

رغم تفوق نماذج المحولات في مجالات عدة، إلا أن تطبيقها في أنظمة كشف التسلل الشبكي ما زال يواجه تحديات، أبرزها كيفية تمثيل المعلومات الموضعية داخل سجلات الشبكة. يُعد ترميز الموضع عنصراً جوهرياً في هذه النماذج، لكن أثر أنواعه المختلفة على دقة النتائج وتفسيرها لم يُدرس بعد بشكل كافٍ، ما يستدعي تحليلاً معمقاً لتأثير هذه الاستراتيجيات في سياق الأمن السيبراني.

3- هدف البحث:

يهدف هذا البحث إلى دراسة تأثير ثلاث استراتيجيات ترميز الموضع (بدون والترميز الجيبي والترميز القابل للتعلم) على أداء نماذج المحولات في كشف الهجمات، مع التركيز على الدقة والتعامل مع الفئات النادرة، وتفسير القرارات باستخدام أدوات مثل LIME وخرائط الانتباه.

4- أهمية البحث:

يُظهر هذا العمل أهمية كبيرة لأنه واحد من الدراسات القليلة التي تدرس كيفية تأثير أساليب ترميز الموضع على أداء المحولات في كشف الهجمات عبر الشبكة الحاسوبية، كما أنه يوفر رؤية تفسيرية عميقة عن سلوك النموذج وأهم الميزات التي تؤثر فيه الشيء الذي يدعم تطوير أنظمة أذكى وأكثر موثوقية في بيئات أمنية حساسة.

1-5 التقدم في كشف التسلل باستخدام المحولات:

تُستخدم نماذج المحولات بشكل متزايد في أنظمة كشف التسلل على الشبكات بفضل قدرتها على النقاط العلاقات المعقدة بين الميزات في البيانات عالية الأبعاد، كما أبرزت استطلاعات حديثة في التعلم الآلي [2].

اقترحت الدراسة [4] نموذج RTIDS باستخدام ترميز موضعي ثابت لتحسين أداء كشف التسلل، إذ حقق النموذج درجة F1 بنسبة 99.17% على مجموعة بيانات CICIDS2017 و 98.48% على CIC-DDoS2019، ورغم فعاليته ضد الهجمات الشائعة، فقد واجه صعوبات في اكتشاف الهجمات النادرة، وهي مشكلة شائعة في نماذج المحولات.

قدمت الدراسة [5] إطار عمل (Flow Transformer)، وهو إطار يعتمد على المحولات لكشف التسلل القائم على التدفق. يعتمد النموذج على بيانات التدفق لتحليل حركة الشبكة، مستفيداً من آليات الانتباه لالتقاط العلاقات المعقدة والسلوكيات طويلة المدى. أظهر النموذج أداءً متميزاً على مجموعات بيانات مثل NSL-KDD و CSE-CIC-IDS2018، مع تقليل التعقيد الحسابي وتحسين سرعة التدريب والاستدلال مقارنةً بالنماذج التقليدية.

أما الدراسة [6] فقد اقترحت نموذج IDS-INT الهجين الذي يدمج المحولات مع CNN-LSTMs مستخدماً التعلم المنقول وتقنية SMOTE لمعالجة عدم توازن الفئات. حقق النموذج دقة 99.21% على مجموعتي بيانات UNSW-NB15 و CICIDS2017، مع تحسين الكفاءة الحسابية لتقليل استهلاك الموارد.

قدمت الدراسة [7] نموذج محولات محسّن باستخدام ترميز موضعي معدل لتعزيز تعلم التبعيات الزمنية، وقد حقق النموذج دقة تقارب 98% على CICIDS2017 ضد هجمات مثل DDoS، مع سرعة تدريب محسنة تجعله مناسباً لسيناريوهات الهجمات النادرة رغم أنهم لم يتناولوا الدراسة التفسيرية.

2-5 دور ترميز الموضع في أداء المحولات:

يُعدّ الترميز الموضعي حجر الأساس في قدرة المحولات على فهم الترتيب داخل البيانات، إلا أن بيانات الشبكة تختلف عن النصوص بغياب ترتيب زمني واضح بين الميزات ما يجعل فعالية الترميز تحدياً رئيساً. استخدمت الدراسة [4] الترميز الثابت بنجاح نسبي، في حين اقترحت الدراسة [8] أسلوباً أكثر دقة من خلال ترميز موضعي على مستوى الميزات (FPE)، ما أدى إلى تحسين اكتشاف الهجمات النادرة. من جهة أخرى، طوّر نموذج في [9] متعدد المقاييس مع ترميز موضعي، لكنه واجه صعوبات في التصنيف متعدد الفئات، لا سيما التمييز بين الأنماط النادرة والشائعة.

3-5 أدوات التفسيرية وتحليل الانتباه:

تبرز التفسيرية كعامل حاسم في قبول نماذج كشف التسلل في البيئات الواقعية. اعتمد [10] على LIME و SHAP لشرح قرارات نموذج يعتمد على المحولات، وأظهرت النتائج أن SHAP يوفر تفسيرات عامة، بينما يقدم LIME رؤية محلية سريعة. بالمثل، استخدم [11] هذه الأدوات ضمن بيئات IoMT، ما عزز فهم تأثير الميزات، بينما قدم [12] دراسة تحليلية عامة للتفسيرية، لم تتناول أوزان الانتباه بعمق، ما يبرز الحاجة لاستكشاف تكامل أدوات التفسير مع آلية الانتباه الذاتي. رغم التقدّم الكبير، لا تزال هناك نقطتان بحثيتان أساسيتان هما:

1- الحاجة إلى دراسة مقارنة منهجية بين استراتيجيات ترميز الموضع في بيانات الشبكة [13].

2- قلة استخدام الانتباه الذاتي كأداة تفسيرية إلى جانب LIME.

من هنا، يركّز هذا البحث على تحليل تأثير ترميز الموضع على أداء المحولات ودمج أدوات التفسير (LIME) والانتباه لتعزيز شفافية النماذج، مع معالجة مشكلة اختلال التوازن باستخدام Focal Loss والأوزان اللوغاريتمية.

6- الإطار النظري:

1-6 المحولات:

نستخدم في هذا البحث المحولات القياسية (Encoder)، إذ إنه كافٍ لتصنيف سجلات حركة الشبكة والمكوّنة من:

- آلية الانتباه الذاتي: تُقيّم أهمية كل ميزة نسبةً إلى الأخرى، وذلك باستخدام الاستعلامات (Q) والمفاتيح (K) والقيم (V)، وتُحسب أوزان الانتباه وفق المعادلة (1):

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \quad (1)$$

- الطبقة الأمامية (Feed-Forward): طبقة شبكية مكونة من وحدتين كثيفتين، مزودة بتطبيع الطبقة والاتصالات المتبقية.

رغم تفوق المحولات في معالجة اللغة الطبيعية ورؤية الحاسوب، إلا أنها تواجه تحدياً في التعامل مع بيانات الشبكة غير المتسلسلة لعدم امتلاكها إدراكاً متأصلاً للترتيب على عكس RNNs و LSTM. لمعالجة ذلك، يُستخدم ترميز الموضع (Positional Encoding: PE) لإضافة معلومات الترتيب، وهو أمر جوهري لضمان التعلم الفعّال في بيانات الشبكة التي تنفقر إلى تسلسل طبيعي [14].

2-6 دور ترميز الموضع:

لا تأخذ المحولات ترتيب الرموز (الميزات) بالحسبان، لذا يُستخدم ترميز الموضع لإضافة معلومات الترتيب إلى التضمينات، ما يمكّن النموذج من فهم السياق [15].

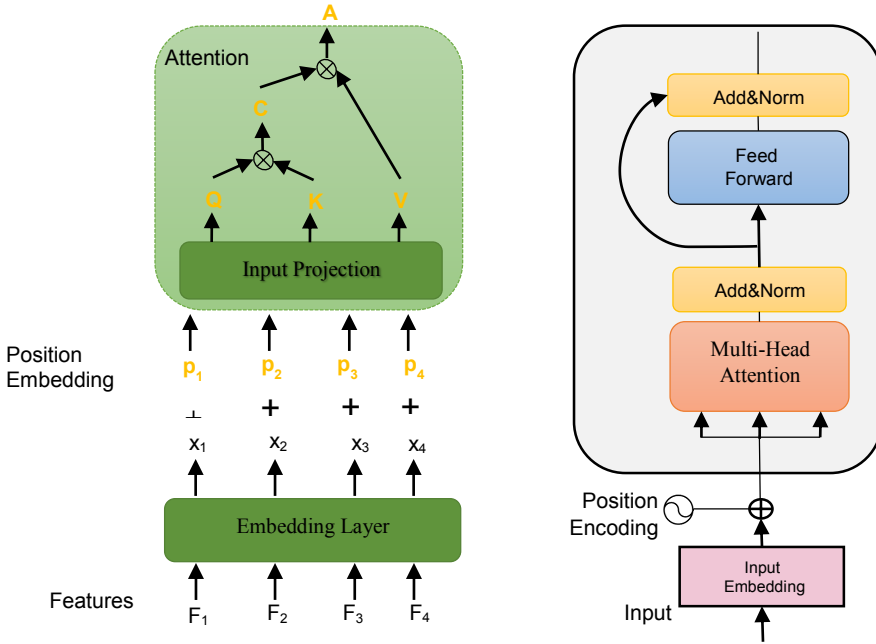
يركز بحثنا على الترميز الموضعي المطلق (Absolute Positional Encoding: APE)، إذ يُضاف متجه لكل رمز يمثل موضعه [16]، وهنا سندرس نوعين من APE هما:

- الترميز الجيبي: يُولد متجهات غير قابلة للتدريب باستخدام دوال جيبيّة وتجيبيّة، كما في المعادلة (2)، ما يساعد على تعميم التسلسلات الطويلة [13].

$$p_{pos,2i} = \sin\left(\frac{pos}{1000^{2i/d}}\right), p_{pos,2i+1} = \cos\left(\frac{pos}{1000^{2i/d}}\right), \quad (2)$$

- الترميز القابل للتعلم: يُخصص متجه موضعي قابل للتحديث لكل موقع، ويتعلم النموذج من خلاله أنماطاً موضوعية خاصة بالمهمة [16].

يوضح الشكل (1) بنية المحول، إذ يتم دمج تضمينات الميزات (F) مع تضمينات الموضع (P) للحصول على الإدخال النهائي (x) المرسل إلى وحدة التشفير [16].



الشكل (1): بنية المحول مع ترميز موضعي مطلق [14].

7- مواد البحث وطرائقه:

في بحثنا هذا، سنعمل على تقويم تأثير ثلاث استراتيجيات لترميز الموضع على أداء المحولات في كشف الهجمات باستخدام مجموعتي بيانات CICIDS2017 و UNSW-NB15.

1-7 مجموعة البيانات:

CICIDS2017: تمثل بيئة واقعية مع هجمات متنوعة مثل DDoS و Web و Bot، وتتميز بتوزيع نسبي بين الهجمات النادرة والشائعة [17].

UNSW-NB15: طُورت بيئة هجومية متنوعة تشمل فئات مثل Analysis و Worms، وتُستخدم لتقويم التعميم [20].

اعتماد هاتين المجموعتين يُمكن من اختبار النموذج على بيانات متنوعة من حيث التوزيع والخصائص البنيوية، ما يساعد في قياس قدرة النموذج على التكيف مع أنماط هجومية مختلفة.

تم دمج الفئات النادرة وفقاً لأساليب سابقة في معالجة بيانات التسلسل، كما يبين الجدولين (1) و (2) [17]. حيث تم تطبيع جميع الميزات لضمان التوافق مع متطلبات النموذج، وتقسيم البيانات إلى 70% للتدريب و 15% للتحقق و 15% للاختبار باستخدام العينة الطبقية للحفاظ على نسب الفئات.

الجدول (1): توزيع الفئات في CICIDS2017. الجدول (2): توزيع الفئات بعد المعالجة المسبقة.

Class	Sample Count	Class	Sample Count
BENIGN	2,096,484	BENIGN	2,096,484
DoS	193,748	DoS	193,748
DDoS	128,016	DDoS	128,016
Port Scan	90,819	Port Scan	90,819
Brute Force	9,152	Brute Force	9,152
Web Attack	2,143	Web Attack	2,143
Bot	1,953	Bot	1,953
Other	47	Infiltration	36
		Heartbleed	11

تمت معالجة مجموعة البيانات (UNSW-NB15) بالطريقة نفسها وتقسيم البيانات بالنسب ذاتها، كما حفظنا هذه التقسيمات للمجموعتين لتدريب النماذج عليها.

يوضح الجدول (3) توزيع الفئات فيها.

الجدول (3): توزيع العينات الأصلية في UNSW-NB15.

Class	Sample Count
Normal	2218760
Generic	215481
Exploits	44525
Fuzzers	24246
Dos	16353
reconnaissance	13987
Analysis	2677
Backdoor	2329
Shellcode	1511
Worms	174

2-7 معالجة عدم توازن الفئات:

للتخفيف من مشكلة عدم توازن الفئات خاصة بالنسبة للهجمات النادرة مثل البوتات وهجمات الويب، تم تطبيق دالة خسارة التركيز (Focal Loss) التي تركز على الأمثلة صعبة التصنيف من خلال تقليل وزن الأمثلة السهلة [18]، حيث $(\gamma=0.2)$: معلم التركيز، $(\alpha=0.25)$: الوزن للفئة المهيمنة)، كما اعتمدنا على تطبيق أوزان الفئات اللوغاريتمية وفق المعادلة (3)، حيث C هو إجمالي عدد العينات، و f_i هو تكرار الفئة i .

$$w_i = \log\left(\frac{C}{f_i}\right) \quad (3)$$

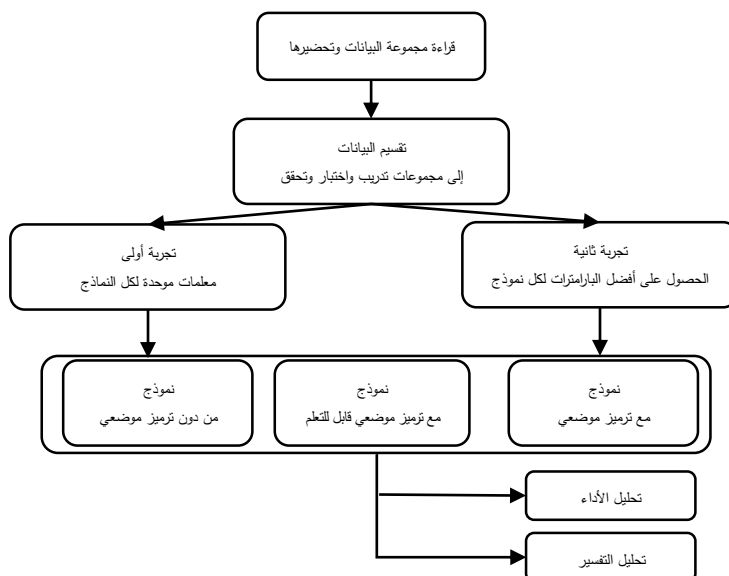
3-7 استخدام خوارزمية LIME في تفسير قرارات نموذج كشف الهجمات:

استخدامنا في بحثنا خوارزمية LIME لتفسير قرارات نموذج المحولات، كونها لا تعتمد على بنية النموذج الأصلي، بل تُنشئ نموذجاً خطياً مبسطاً حول كل حالة اختبار لتحديد الميزات الأكثر تأثيراً في التصنيف تعتمد على اختيار الميزات لتقدير أهميتها، ما يوضح سبب تصنيف تدفق معين كتهديد أو نشاط طبيعي، إذ أثبتت فعاليتها في تفسير نماذج معقدة كـ Random Forest و Decision Trees، حتى في مجالات حساسة مثل كشف خطاب الكراهية [21]، وقياساً على ذلك سنستخدمها هنا لتوضيح أثر الميزات الشبكية (مثل عدد الحزم أو البروتوكول) على قرار النموذج، ما يُعزز موثوقيته، ويُضيف طابعاً تفسيرياً مهماً في السياقات الأمنية.

4-7 الإعداد التجريبي:

كما يوضح الشكل (2)، تم اعتماد مسارين لتقييم نماذج المحولات:

1. إعداد موحد لجميع النماذج (بدون ترميز وبترميز جيبى وبترميز قابل للتعلم) باستخدام المعلومات الفائقة نفسها، والتي تم عرضها في الجدول (4).
2. إعداد محسن باستخدام أداة Keras Tuner لتخصيص المعلومات الفائقة (مثل عدد الطبقات والرؤوس و d_model و dff) لكل نموذج لتحقيق أفضل أداء، موضحة في الجدول (5).



الشكل (2): سير العمل المقترح لتقييم نماذج المحولات باستخدام إعدادات موحدة ومحسنة.

على خلاف دراسات اعتمدت على تقنيات اختيار الميزات، فقد تم الاحتفاظ بجميع الميزات الأصلية دون تعديل، وذلك بهدف تقويم تأثير بنية النموذج والمعلومات فقط بالاعتماد على قدرة المحول على تمثيل التفاعلات المعقدة بين الميزات داخلياً [22].

الجدول (4): المعلومات الفائقة المشتركة عبر النماذج.

Parameter	Value
Number of layers	2
Number of heads	4
Hidden dimension (d_model)	64

Feed-forward size (dff)	128
Dropout rate	0.1
Learning rate (Adam)	1e-4

الجدول (5): أفضل المعلمات الفائقة لكل نموذج.

CICIDS2017					UNSW-NB15			
Model	Layers	Heads	d_model	dff	Layers	Heads	d_model	dff
NoPos	2	8	128	256	3	8	32	128
LearnPos	3	8	32	128	3	4	64	64
SinPos	3	8	32	128	2	8	64	256

8- النتائج والمناقشة:

1-8 مقارنة الأداء العام:

تُظهر نتائج التقويم على مجموعتي البيانات CICIDS2017 و UNSW-NB15 تفوقاً واضحاً لنموذج الترميز الموضعي القابل للتعليم (LearnPos) على النموذجين الآخرين: بدون ترميز موضعي (NoPos) ومع الترميز الجيبي (SinPos)، وذلك سواءً باستخدام الإعدادات الموحدة أو المعلمات المُحسّنة.

CICIDS2017: إذ سجّل LearnPos دقة تصنيف بلغت 98.55% في الإعدادات الموحدة و 98.16% في الإعدادات المُحسّنة، مع تحقيقه لأعلى F1-score في الفئات النادرة مثل:

— Web Attack: ب 0.8000 مقابل 0.2441 في NoPos.

— Bot: ب 0.3480 مقابل 0.2233 في NoPos.

كما بلغ معدل ROC-AUC 0.9996، ما يعكس قدرة ممتازة على التمييز بين الفئات المختلفة.

يعرض الجدول (6) مقارنة تفصيلية لأداء النماذج حسب الفئة باستخدام كل من الإعدادات الموحدة والمحسّنة، إذ يتفوق LearnPos بشكل ملحوظ في الفئات الشائعة (مثل BENIGN و DoS) والنادرة (مثل Bot و Other).

الجدول (6). أداء النماذج حسب الفئة (F1-score) باستخدام المعطيات الموحدة والمحسنة
مجموعة بيانات CICIDS2017.

Class	SinPos (Unified)	SinPos (Optimized)	LearPos (Unified)	LearnPos (Optimized)	NoPos (Unified)	NoPos (Optimized)
BENIGN	0.9839	0.9862	0.9912	0.9889	0.9657	0.9774
Bot	0.2174	0.3125	0.1430	0.3480	0.0942	0.2233
Brute Force	0.9429	0.9553	0.9719	0.9764	0.7169	0.9515
DDoS	0.9940	0.9829	0.9993	0.9932	0.9666	0.9744
DoS	0.9721	0.9846	0.9922	0.9871	0.9124	0.9253
Other	0.0000	0.5714	0.3500	0.6000	0.0000	0.0000
Port Scan	0.8440	0.8537	0.9542	0.8537	0.8023	0.8369
WebAttack	0.2824	0.3105	0.8000	0.5717	0.2441	0.3035

UNSW-NB15: حقق LearnPos أداءً متفوقاً على مستوى: Macro F1: 0.5824 مقارنة بـ SinPos (0.5699) و NoPos (0.4674)، كما أظهر أداءً جيداً في الفئات الصعبة مثل: (Shellcode: F1 = 0.7970) و (Fuzzers: F1 = 0.5815). وقد قدم توازناً عاماً في التصنيف بين الفئات، في حين أن NoPos رغم دقته الإجمالية المرتفعة، إلا أنه تراجع بشكل كبير في الفئات قليلة التمثيل.

يوضح الجدول (7) الأداء التفصيلي لكل نموذج حسب الفئة، ما يؤكد تفوق LearnPos خاصة في الفئات التي تتطلب قدرة أعلى على التمييز، إذ يعكس تفوق LearnPos مرونة معمارية تجعله قادراً على التكيف مع بنى بيانات متنوعة، سواءً أكانت غير متوازنة (CICIDS2017) أو متوازنة نسبياً (UNSW-NB15)، ويبرز هذا أن تصميم النموذج يلعب دوراً أكبر من مجرد تحسين المعلمات الفائقة.

الجدول (7): أداء النماذج حسب الفئة (F1-score) باستخدام المعطيات الموحدة والمحسنة
مجموعة بيانات UNSW-NB15.

Class	SinPos (Unified)	SinPos (Optimized)	LearPos (Unified)	LearnPos (Optimized)	NoPos (Unified)	NoPos (Optimized)
analysis	0.1721	0.1848	0.1582	0.1692	0.1366	0.1438
backdoor	0.1232	0.1448	0.1384	0.1757	0.0110	0.0056
dos	0.4515	0.4535	0.4502	0.4552	0.03687	0.4374

مقارنة تأثير الترميز الموضعي في أنظمة كشف التسلل المعتمدة على المحولات القابلة للتفسير

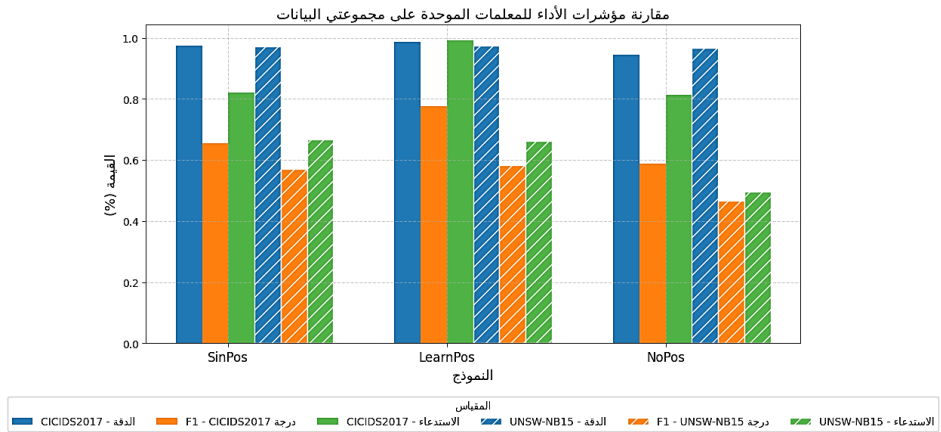
exploits	0.6380	0.6556	0.6537	0.6586	0.5808	0.5874
fuzzers	0.5755	0.5704	0.5766	0.5815	0.4472	0.4999
generic	0.9878	0.9874	0.9877	0.9886	0.9741	0.9791
normal	0.9932	0.9932	0.9932	0.9932	0.9916	0.9908
reconnaissance	0.7502	0.8122	0.803	0.8056	0.5912	0.6122
shellcode	0.7334	0.7621	0.7569	0.7970	0.4552	0.4730
worms	0.2737	0.209	0.3056	0.3226	0.1167	0.2414

رغم ذلك، لوحظ تراجع أداء LearnPos في فئة Web Attack من $F1 = 0.8000$ إلى 0.5717 بعد استخدام المعلمات المحسنة. يُرجّح أن يكون ذلك بسبب تأثير دالة الخسارة Focal Loss، والتي تعيد توزيع الانتباه لصالح الفئات النادرة، ما يقلل من التركيز على الفئات متوسطة التكرار كما وثّقت بعض الدراسات [18-19-20].

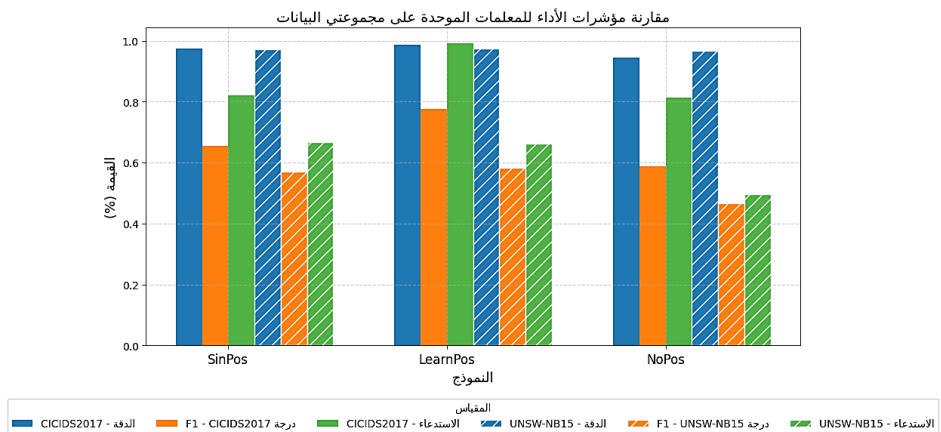
يعرض الشكل (3) مقارنة بين النماذج الثلاثة من حيث الدقة ودرجة $F1$ والاسترجاع باستخدام الإعدادات الموحدة، فيما يوضح الشكل (4) المقاييس نفسها مع أفضل المعلمات. أما الشكلان (5) و (6) فيبينان أخطاء التصنيف (FN و FP) لفئات حرجة مثل Web Attack و Fuzzers، إذ يظهر LearnPos بأقل عدد من الأخطاء.

2-8 تحليل التعقيد الحسابي:

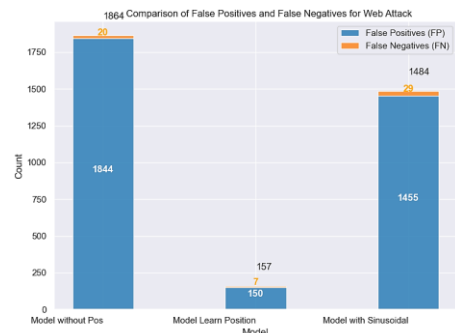
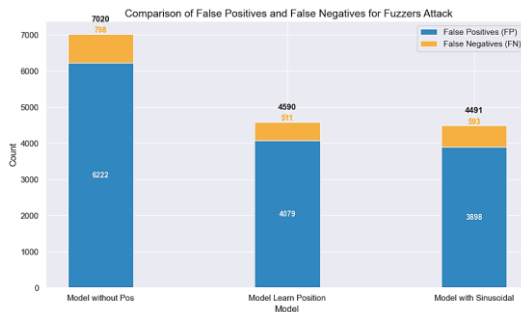
يعرض الجدول (8) مقارنة شاملة للتعقيد الحسابي بين النماذج الثلاثة باستخدام الإعدادات الموحدة للتدريب على مجموعتي البيانات CICIDS2017 و UNSW-NB15، وتُبين النتائج أن إدراج الترميز الموضعي — وخاصة الترميز القابل للتعلم (LearnPos) — يؤدي إلى زيادة ملحوظة في التكلفة الحسابية.



الشكل (3): مقارنة الدقة و F1 والاسترجاع تحت إعدادات موحدة.



الشكل (4): مقارنة الدقة و F1 والاسترجاع تحت إعدادات المحسنة.



الشكل (6): مقارنة أخطاء التصنيف (FP و FN) لفئة هجمات fuzzier عبر النماذج.

الشكل (5): مقارنة أخطاء التصنيف (FP و FN) لفئة هجمات الويب عبر النماذج.

الجدول (8): تحليل التعقيد الحسابي.

UNSBW			CICIDS2017			المقياس
SinPos	LearnPos	NoPos	SinPos	LearnPos	NoPos	
67,592	72,072	67,592	67,592	72,072	67,592	عدد البارامترات
5,655	6,769	4,283	6,615	10,124	4,745	الزمن الكلي للتدريب (بالثواني)
94.25	112.82	71.38	110.25	168.73	79.08	الزمن الكلي للتدريب (بالدقائق)
282.75	338.45	214.15	315.00	326.58	237.25	متوسط زمن العصور (epoch) بالثواني

يرجع ذلك إلى أن LearnPos يتطلب تحديثاً مستمراً لمتجهات الموضع أثناء كل حقبة تدريب بخلاف الترميز الجيبي (SinPos) أو غياب الترميز (NoPos)، اللذين يعتمدان على قيم ثابتة لا تحتاج إلى تعديل خلال عملية التعلم.

رغم هذه الزيادة في الزمن الكلي للتدريب وعدد المعاملات (Parameters)، فإن التحسينات الكبيرة في الأداء، خصوصاً في تصنيف الفئات النادرة، تجعل من LearnPos خياراً مُبرراً — خاصة في سياقات أمنية حرجة تتطلب دقة وموثوقية عاليتين في الكشف عن الهجمات.

3-8 تحليل أوزان الانتباه

لفهم آليات اتخاذ القرار في نماذج المحولات المستخدمة في كشف التسلل الشبكي، فقد تم تحليل أوزان الانتباه عبر خوارزمية مخصصة (خوارزمية 1) تقوم بحساب متوسط الأوزان عبر جميع الرؤوس والطبقات، ثم تُحدد أهم الميزات التي يُركز عليها النموذج خلال تصنيفه لمختلف أنواع الهجمات، وقد تم ذلك ضمن إطار الذكاء الاصطناعي التفسيري (XAI)، باستخدام خرائط حرارية ومخططات شريطية لتصور توزيع الانتباه وتأثير الترميز الموضعي عليه.

تشير الدراسات السابقة [14] إلى أن آليات الانتباه في نماذج المحولات قادرة على تحسين تمثيل البيانات من خلال التركيز على عناصر ذات دلالة مثل أنماط المنافذ أو خصائص تدفق الحزم، وهذا ما يُساهم في اكتشاف الأنماط المعقدة في حركة مرور الشبكة.

في هذا السياق، أُجري مع الحفاظ على المعلمات (الموحدة) نفسها لتدريب النماذج، ما يضمن مقارنة عادلة لتركيز الانتباه ومدى تفسيره لسلوك النموذج.

Input:

- Trained transformer model M
- Input sample X (shape: [batch_size, sequence_length, features])
- Feature names list F = [f₁, f₂, ..., f_n]

Output:

- Attention matrix A (feature-to-feature interaction)
- Feature importance vector I (as percentage)

Steps:

1. Forward X through the model M with attention tracking enabled:
→ Obtain attention_weights from M(X, return_attention=True)
2. Stack all attention heads and layers:
→ W ← stack(attention_weights)
(shape: [num_layers, num_heads, seq_len, seq_len])
3. Average over all layers and heads:
→ W_mean ← mean(W, axis=(0, 1))
(shape: [seq_len, seq_len])
4. Construct attention matrix A using W_mean:
→ A ← DataFrame(W_mean, rows=F, columns=F)
5. Compute total attention received per feature:
→ total ← sum(A, axis=0)
6. Normalize into percentage importance:
→ I ← 100 × total / sum(total)
7. Return:
→ A, I

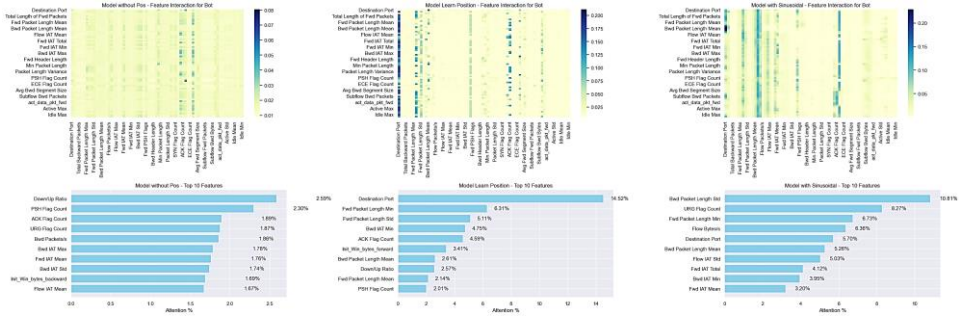
خوارزمية 1: استخراج مصفوفة الانتباه ونسب التأثير لكل ميزة.

1-3-8 التحليل على مجموعة بيانات CICIDS2017:

• هجوم Bot:

يُظهر الشكل (7) أن نموذج LearnPos ركّز بنسبة كبيرة على ميزة منفذ الوجهة (14.52%)، وهي ميزة حاسمة في الكشف عن الأنماط المتكررة للاتصال المرتبط بشبكات البوت. كما أظهر انتباهاً موجهاً نحو الحد الأدنى لطول الحزمة الأمامية (6.31%) وعدد أعلام ACK (4.59%)، ما يعكس فهماً منظماً لسلوك الهجوم.

مقارنة تأثير الترميز الموضعي في أنظمة كشف التسلل المعتمدة على المحولات القابلة للتفسير

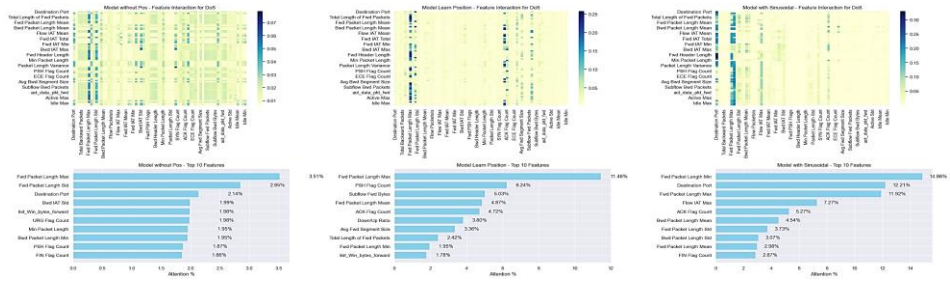


الشكل (7): خرائط حرارية الانتباه لهجوم Bot.

في المقابل، أظهر نموذج NoPos توزيع انتباه مشتت مركّزًا بنسبة 2.59% فقط على نسبة التحميل/الرفع (Down/Up Ratio)، وهي ميزة ضعيفة الدلالة ما أدى إلى انخفاض أداء الكشف. أما نموذج SinPos فركّز على ميزات مثل الانحراف المعياري لطول الحزم العكسية (10.81%) وعدد أعلام URG (8.27%)، لكن افتقاره إلى التركيز على المنافذ تسبب في انخفاض F1-Score بنسبة 8% مقارنة بـ LearnPos.

• هجوم DoS:

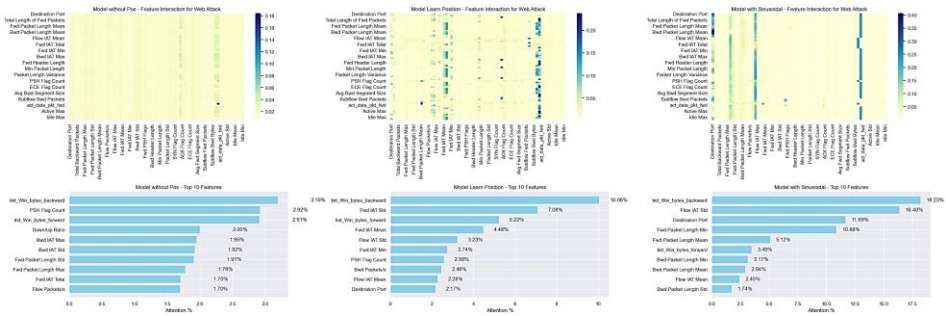
أظهر LearnPos فهماً تفسيرياً واضحاً من خلال تركيزه على ميزات مثل الحد الأقصى لطول الحزمة (11.46%) وعدد أعلام PSH (6.24%)، مما يعكس قدرته على النقاط أنماط هجمات الفيضان وساعده في خفض معدلات السلبات الكاذبة. في المقابل ركّز SinPos على Flow IAT Max (7.27%) ومنفذ الوجهة (12.21%)، لكنه لم يُظهر تركيزاً مستقراً. أما NoPos، فلم تتجاوز أقوى ميزاته نسبة 3.51%، ما يشير إلى تشتت واضح في الانتباه، كما هو موضح في الشكل (8).



الشكل (8): خرائط حرارية الانتباه لهجمات DoS.

• هجوم الويب (Web Attack):

تفوق LearnPos تفسيرياً عبر تركيزه على ميزة التباين في وقت الوصول الأمامي (Fwd IAT Std) بنسبة (7.08%)، والتي تساعد في رصد الاضطرابات الزمنية المرتبطة بسلوك هجمات الويب، رغم أن SinPos أعطى وزناً مرتفعاً لميزة Init_Win_bytes_backward (18.23%)، إلا أن هذه الميزة غير مرتبطة سلوكياً بشكل مباشر بهجمات الويب، ما يُفسر أدائه المتواضع. أما NoPos فافتقر إلى التنظيم وركز على ميزات غير تفسيرية مثل PSH Flag Count و Down/Up Ratio وهذا ما يعرضه الشكل(9).



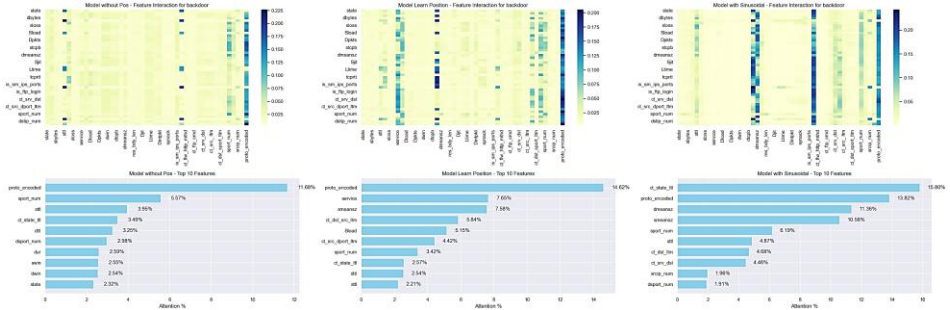
الشكل (9): خرائط حرارية الانتباه لهجمات الويب.

2-3-8 التحليل على مجموعة بيانات UNSW-NB15:

• هجمات (backdoor):

أظهر نموذج LearnPos نقوفاً في كشف هجمات Backdoor، إذ ركّز على ميزات تفسيرية مثل proto_encoded (ترميز البروتوكول مثل TCP/UDP بنسبة 14.62%) و service (نوع الخدمة ك HTTP/SSH)، ما يعكس قدرته على التقاط أنماط الاتصال الخفي. في المقابل، ركّز SinPos على ct_state_ttl (عدد حالات TTL في التدفق الزمني بنسبة تركيز 15.80%)، ما يبرز فعاليته في تمثيل الخصائص الزمنية لكنه افتقر إلى التركيز على ميزات السياق مثل المنافذ، ما أدى إلى انخفاض F1-score (0.1232) مقارنة بـ LearnPos (0.1360). أما NoPos فأظهر توزيع انتباه مشتت (2.59% على نسبة التحميل/الرفع)، ما تسبب في أداء ضعيف (F1-score: 0.0110).

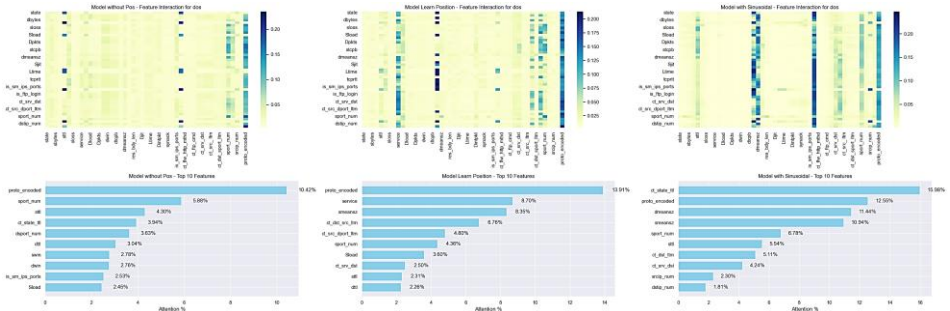
يوضح الشكل (10) توزيع الانتباه المتوازن لـ LearnPos، ما يبرر تفوقه في كشف الأنماط الخبيثة.



الشكل (10): خرائط حرارية الانتباه لهجمات backdoor.

• هجمات DoS:

كرر LearnPos نمط التركيز على proto_encoded و service، بينما برز SinPos مجدداً من خلال ct_state_ttl، بينما فشل النموذج NoPos في تشكيل تركيز تفسيري واضح، ما قلل من فعاليته انظر الشكل (11).



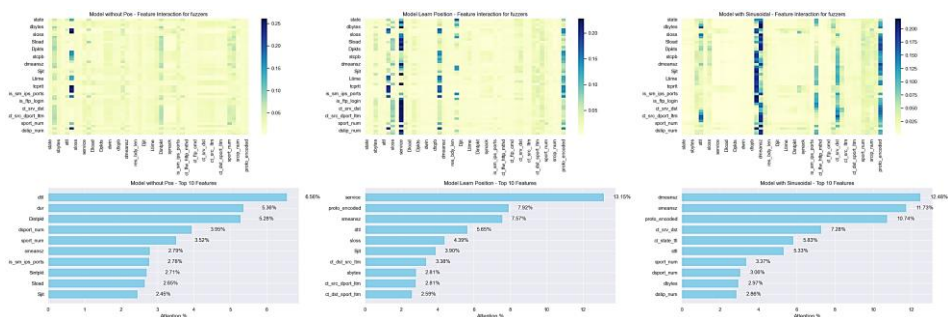
الشكل (11): خرائط حرارية الانتباه لهجمات dos.

• هجوم Fuzzers:

أظهر LearnPos مرونة تفسيرية قوية من خلال تركيزه على service و proto_encoded، ما مكّنه من التعامل مع أنماط البيانات العشوائية ومن تمييز البروتوكولات والخدمات المستهدفة بالبيانات العشوائية. بينما ركّز نموذج SinPos على smeansz (متوسط حجم حزم المصدر) و dmeansz (متوسط حجم حزم الوجهة)، ما يعكس حساسيته لتغيرات حجم

الحزم، لكنه كان أقل اتساقاً بسبب ضعف تركيزه على ميزات السياق. أما NoPos فأظهر توزيع انتباه مشتت ما يفسر أدائه الضعيف.

يبين الشكل (12) تفوق LearnPos في النقاط الأنماط العشوائية لهجمات Fuzzers بفضل فهمه للسياق الوظيفي للبروتوكولات والخدمات.



الشكل (12): خرائط حرارية الانتباه لهجمات Fuzzers.

يُظهر هذا التحليل أن استخدام الترميز الموضعي القابل للتعليم (LearnPos) لا يحسّن مؤشرات الأداء العددي فقط بل يُعزز تفسير قرارات النموذج بشكل واضح. ويوفر LearnPos تنظيمًا بنيويًا أعلى لتركيز الانتباه عبر فئات هجومية متنوعة ومجموعات بيانات مختلفة.

في المقابل، تعاني النماذج الأخرى (NoPos و SinPos) من ضعف في التنظيم والانتباه التفسيري، ما يبرز أهمية LearnPos في تطوير نماذج تفسيرية قابلة للتعميم في سياقات أمن الشبكات.

4-8 تفسير النموذج عبر LIME:

يُعد السؤال "ما الذي جعل النموذج يتخذ هذا القرار؟" حاسماً للتحقق من صحة نماذج التصنيف في كشف التسلسل على الشبكات. في هذا السياق، تمت مقارنة حالتين متباينتين لاختبار قدرات النماذج التفسيرية هما:

الحالة الأولى: تصنيف واضح لعينة Web Attack (من CICIDS2017).

تم اختيار عينة مصنفة كهجوم Web Attack من مجموعة بيانات (CICIDS2017) لتحليل آلية اتخاذ القرار، وقد صنفت النماذج العينة بثقة عالية (تجاوزت 90%)، وأظهرت LIME كما يوضح الشكل (13) أن السمات المؤثرة تضمنت مؤشرات زمنية مثل Flow IAT Mean و Flow

Packets/s، والتي تعكس سلوكاً هجومياً نمطياً. لوحظ تطابقاً واضحاً بين نتائج LIME وتحليل أوزان الانتباه، ما يشير إلى تفسير واضح للقرار واستناد النموذج إلى أنماط قابلة للفهم.



الشكل (13): مقارنة تفسيرات LIME وأوزان الانتباه لهجوم Web Attack عبر النماذج الثلاثة.

هذا التوافق بين نتائج LIME والتحليل الداخلي للنموذج يعزز موثوقية التفسير، ويؤكد أن النموذج لم يتخذ قراره بشكل عشوائي، بل استند إلى إشارات واضحة وقابلة للفهم ضمن البيانات.

الحالة الثانية: تصنيف صحيح منخفض الثقة Backdoor:

كشفت نتائج تفسير النموذج باستخدام LIME عن غياب سمات تفسيرية مباشرة، حيث أظهرت جميع السمات المفعلة تأثيراً ضعيفاً جداً واقتصرت على "منطقة التأثير الصفري"، مما يعكس افتقار النموذج إلى إشارات واضحة تدعم قراره. وعلى الرغم من أن العينة صُنِّفت بشكل صحيح، إلا أن الثقة بقيت منخفضة نسبياً. يُظهر الشكل (14) هذا التوجه حيث يتوزع الانتباه عبر سمات متعددة منخفضة الأهمية دون نمط واضح أو تركيز مهيمن، لا سيما في نمودجي Learned Pos و Sin Pos، مما يُشير إلى اعتماد النماذج على إشارات خفية وغير مستقرة يصعب تفسيرها.

Model: No Pos

Model: Learned Pos

Model: Sin Pos



الشكل (14): مقارنة تفسيرات LIME وأوزان الانتباه لهجوم back door عبر النموذج الثلاثة.

تُظهر هذه النتائج أن قابلية تفسير قرارات النموذج تختلف باختلاف طبيعة الهجوم، إذ تكون الهجمات النمطية أو الزمنية (مثل Web Attack) أسهل في التفسير، بينما تتطلب الهجمات المتخفية (مثل Backdoor) أدوات تفسير أعمق نظراً لتعقيد أنماطها.

9- الاستنتاجات والتوصيات:

أظهرت الدراسة أهمية ترميز الموضع في تعزيز أداء نماذج المحولات ضمن أنظمة كشف التسلل، وخاصة في كشف الهجمات المعقدة أو النادرة. تفوق نموذج LearnPos الذي يستخدم ترميزاً موضعياً قابلاً للتعليم، من حيث الدقة و F1-score على نمودجي NoPos و SinPos في كل من CICIDS2017 و UNSW-NB15. كما ساهم استخدام Focal Loss وأوزان الفئات في تحسين التعامل مع عدم توازن البيانات، ما انعكس في أداء مميز على الفئات النادرة مثل Bot و Web Attack.

عززت أدوات التفسير (LIME وأوزان الانتباه) شفافية النموذج مظهرًا اعتماداً على سمات واضحة (كمنفذ الوجهة و Flow IAT)، ما يدل على أن LearnPos يجمع بين الفعالية التنبؤية وقابلية التفسير.

انطلاقاً من النتائج الجيدة التي تم الحصول عليها في هذا البحث، يمكننا اقتراح مجموعة النقاط التي قد تسهم في تحسين أداء أنظمة كشف التسلل المعتمدة على المحولات:

— تحسين بنية المحول (مثل زيادة العمق أو الرؤوس أو استخدام آليات انتباه مخصصة).

— تطوير ترميزات هجينة أو موجهة بالمجال.

— اعتماد أساليب ديناميكية لمعالجة عدم التوازن (مثل: SMOTE أو GAN).

تؤكد الدراسة أن استخدام ترميز موضعي قابل للتعليم لا يحسن فقط أداء النموذج، بل يعزز أيضاً شفافيته وملاءمته للتطبيق العملي في الأمن السيبراني.

10- المراجع:

1. World Economic Forum. (2025). Global Cybersecurity Outlook 2025. Retrieved from: <https://www.weforum.org/publications/global-cybersecurity-outlook-2025/>
2. Momand, A., Jan, S. U., & Ramzan, N. (2023). A systematic and comprehensive survey of recent advances in intrusion detection systems using machine learning: Deep learning, datasets, and attack taxonomy. Journal of sensors, 2023(1), 6048087.
3. Corea, P. M., Liu, Y., Wang, J., Niu, S., & Song, H. (2024, July). Explainable AI for comparative analysis of intrusion detection models. In 2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom) (pp. 585–590). IEEE.
4. Wu, Z., Zhang, H., Wang, P., & Sun, Z. (2022). RTIDS: A robust transformer-based approach for intrusion detection system. IEEE Access, 10, 64375–64387.
5. Manocchio, L. D., Layeghy, S., Lo, W. W., Kulatilleke, G. K., Sarhan, M., & Portmann, M. (2024). Flow transformer: A transformer framework for flow-based network intrusion detection systems. Expert Systems with Applications, 241, 122564.
6. Ullah, F., Ullah, S., Srivastava, G., & Lin, J. C. W. (2024). IDS–INT: Intrusion detection system using transformer-based transfer learning for imbalanced network traffic. Digital Communications and Networks, 10(1), 190–204.
7. Liu, Y., & Wu, L. (2023). Intrusion detection model based on improved transformer. Applied Sciences, 13(10), 6251.

8. Çavşı Zaim, H., & Yolaçan, E. N. (2025). FPE–Transformer: A Feature Positional Encoding–Based Transformer Model for Attack Detection. *Applied Sciences*, 15(3), 1252.
9. Xi, C., Wang, H., & Wang, X. (2024). A novel multi–scale network intrusion detection model with transformer. *Scientific Reports*, 14(1), 23239.
10. Gaspar, D., Silva, P., & Silva, C. (2024). Explainable ai for intrusion detection systems: Lime and shap applicability on multi–layer perceptron. *IEEE Access*.
11. Kalakoti, R., Sharma, M., & Bhattacharya, S. (2024). Explainable transformer–based intrusion detection for IoMT networks. *Journal of Medical Systems*, 48(5), Article 20. <https://doi.org/10.1007/s10916-024-02054-5>.
12. Patil, S., Varadarajan, V., Mazhar, S. M., Sahibzada, A., Ahmed, N., Sinha, O., ... & Kotecha, K. (2022). Explainable artificial intelligence for intrusion detection system. *Electronics*, 11(19), 3079.
13. Dufter, P., Schmitt, M., & Schütze, H. (2022). Position information in transformers: An overview. *Computational Linguistics*, 48(3), 733–763.
14. Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
15. Shaw, P., Uszkoreit, J., & Vaswani, A. (2018). Self–attention with relative position representations. *arXiv preprint arXiv:1803.02155*.

16. Zhao, Liang, Xiachong Feng, Xiaocheng Feng, Weihong Zhong, Dongliang Xu, Qing Yang, Hongtao Liu, Bing Qin, and Ting Liu. (2023). Length extrapolation of transformers: A survey from the perspective of positional encoding. arXiv preprint arXiv:2312.17044.
17. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSp, 1(2018), 108–116.
18. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980–2988).
19. Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. Journal of big data, 6(1), 1–54.
- Moustafa, N., & Slay, J. (2015, November). UNSW–NB15: a comprehensive data set for network intrusion detection systems (UNSW–NB15 network data set). In 2015 military communications and information systems conference (MilCIS) (pp. 1–6). IEEE.

20. م. ربيع محي الدين الكردي، د. كمال سلوم، د. وسيم رمضان (2024). تفسير نموذج مُدرب لاكتشاف خطاب الكراهية في التغريدات العربية. مجلة جامعة حمص، سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، (1)46.
21. م. علي ياسين، د. كمال سلوم، د. وسيم رمضان (2022). تحديد عتبة التصنيف المثلى ديناميكياً في أنظمة الكشف المبكر عن الشذوذ القائمة على التعلم العميق. سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، (4)44.

منظومة لغوية للتحليل الصرفي والنحوي للجملة العربية

باستخدام التعلم العميق - AraBERT Transformer

م. هديل عبد الرحمن*، أ.د. محمد فاضل سكر**

* طالبة دكتوراة، قسم الذكاء الصناعي واللغات الطبيعية، كلية الهندسة المعلوماتية، جامعة حلب

** أستاذ في قسم الذكاء الصناعي واللغات الطبيعية، كلية الهندسة المعلوماتية، جامعة حلب

الملخص

تُعد معالجة النصوص العربية آلياً من المهام المعقدة بسبب الخصائص الصرفية والنحوية الفريدة للغة، مثل الإلصاق ومرونة ترتيب الكلمات، خاصة في النصوص غير المشكولة. في هذا البحث، نقترح منظومة لغوية تعتمد على نموذج لغوي متقدم قادر على فهم السياق اللغوي بشكل عميق، مما يساهم في استخراج 13 حقلاً نحوياً وصرفياً بدقة عالية. تعتمد المنظومة على هيكلية متعددة الرؤوس تتيح تحليل عدة خصائص لغوية في آن واحد، وقد أظهرت المنظومة أداءً عالياً واستقراراً ملحوظاً عند اختبارها على قاعدة بيانات مخصصة، متفوقة بذلك على عدة نماذج لغوية حديثة. تعكس هذه النتائج فعالية النموذج المقترح في دعم تطبيقات المعالجة الآلية للغة العربية مثل الترجمة وتصحيح الأخطاء.

الكلمات المفتاحية: AraBert v2، المحولات، إعراب الجملة العربية، التحليل الصرفي،

التحليل النحوي

A Linguistic System for Morphological and Syntactic Analysis of Arabic Sentences Using Deep Learning: AraBERT – Transformer

Hadeel Abdulrahman*, Mohammed Fadel Sukkar**

*Postgraduate Student (PhD.), Dept. of Artificial Intelligence and Natural Languages, Faculty of Informatics Engineering, University of Aleppo

**Prof., Dept. of Artificial Intelligence and Natural Languages, Faculty of Informatics Engineering, University of Aleppo

Abstract

Processing Arabic text automatically is a complex task due to the language's unique morphological and syntactic characteristics, such as agglutination and flexible word order, especially in unvowelized texts.

In this research, we propose a linguistic system based on an advanced linguistic model capable of deeply understanding linguistic context, contributing to the accurate extraction of 13 syntactic and morphological fields. The system relies on a multi-headed architecture that enables the simultaneous analysis of multiple linguistic features. The system demonstrated high performance and remarkable stability when tested on a dedicated dataset, outperforming several state-of-the-art linguistic models. These results reflect the effectiveness of the proposed model in supporting Arabic language processing applications such as translation and error correction.

Keywords: AraBERT v2, Transformers, Arabic sentence parsing, Morphological analysis, Syntactic analysis

1. مقدمة

تمتاز اللغة العربية بثرائها الصرفي والاشتقائي، حيث أن كلمة واحدة يمكن أن تأخذ عدداً كبيراً من الأشكال بناءً على الاشتقاق والتصريف، وعلى الرغم من كونها مصدر قوة للغة العربية إلا أنها

تشكل تحدياً كبيراً للأنظمة الآلية. فعلى سبيل المثال، كلمة "كتب" قد تعني "كُتِبَ" (فعل ماضٍ)، أو "كُتِبَ" (فعل مبني للمجهول)، أو "كُتِبَ" (اسم جمع)، وكل ذلك دون تغيير في الحروف الساكنة الأساسية [1]، حيث يكون معدل الغموض في التفسير النحوي أعلى في الكلمات غير المشكولة بمتوسط 8.7، مقارنةً بمتوسط 5.6 للكلمات المشكولة [2]. بالإضافة إلى ذلك، تمتاز اللغة العربية بخصائص فريدة تجعل عملية التحليل صعبة، كما في الإلصاق من خلال إضافة السوابق واللاحق، حيث يمكن أن تلتصق بالكلمات أدوات التعريف وحروف الجر والضمائر، مما يؤدي إلى بنى كلمات وجمل معقدة وغير تقليدية. علاوة على ذلك يمكن ترتيب الكلمات في اللغة العربية بشكل مرن، فعلى الرغم من وجود ترتيب شائع وهو (فعل - فاعل - مفعول به) على سبيل المثال، عادةً ما توضع الكلمة المراد التأكيد عليها في بداية الجملة، وفي بعض الحالات يكون الهدف من التقديم والتأخير لأغراض بلاغية أو نحوية، الأمر الذي يؤدي إلى ارتباك نحوي يصعب حله باستخدام القواعد النحوية [2].

كل العوامل السابقة تجعل مهمة تحديد مواقع الفعل والفاعل والمفعول وإسناد الحالات الإعرابية صعبة. بالتالي فإن عملية الإعراب الكامل لكلمات الجملة يتطلب تكامل الجانب الصرفي وتحديد المواضع النحوية بدقة وهو أمر صعب في النصوص غير المشكولة. لذلك جذب إعراب الجمل العربية اهتماماً متزايداً في بحوث معالجة اللغات الطبيعية، حيث تهدف عملية الإعراب إلى تحديد البنية النحوية للجملة والتي تمثل معلومات ضرورية لمجموعة واسعة من تطبيقات معالجة اللغات الطبيعية مثل الترجمة الآلية، تصحيح الأخطاء النحوية، تلخيص النصوص واستخراج المعلومات [2]

رغم التقدم الكبير الذي أحرزته أبحاث إيجاد الصيغ النحوية للجملة العربية، إلا أنه لا تزال هناك فجوة واضحة ولاسيما في:

- غياب منظومة تقدم اعراباً متكاملًا يدمج بين الصرف والنحو بجميع تفصيلاتها كزمن الفعل، نوع الاسم، التعريف والتذكير، السوابق واللاحق.
- الاعتماد على النصوص المشكولة، وتجاهل النصوص غير المشكولة

وعليه فإن مساهمة هذا البحث تتجلى في النقاط التالية:

- بناء قاعدة بيانات مخصصة لدعم هذه المهمة، تتألف من الجمل العربية المشكولة وغير المشكولة لقوالب محدودة ويترافق مع كل جملة 13 حقل صرفي ونحوي
- تحليل صرفي ونحوي للجملة العربية البسيطة غير المشكولة، من خلال الاستفادة من محول AraBERT v2 وتوظيف رؤوس تصنيف متعددة لإيجاد 13 حقلاً صرفياً ونحويًا للجملة العربية المدخلة.

بناءً على ما سبق، نقدم خلال هذا البحث إجابة لسؤال البحث، وهو "هل يمكن تطوير منظومة لغوية قادرة على فهم السياق للنصوص العربية غير المشكولة، وتقديم تحليل صرفي ونحوي متكامل وبدقة عالية؟"، وذلك بالاستناد إلى أن دمج نموذج لغوي عربي مدرب مسبقاً (مثل AraBERT) مع طبقات محول مخصصة ورؤوس تصنيف متعددة، سيساهم في رفع دقة الإعراب الآلي وتحسين موثوقيته في النصوص ذات البنية المعقدة.

2. هدف البحث

يهدف هذا البحث إلى تصميم وتطوير نظام آلي متكامل لإجراء تحليل صرفي ونحوي شامل لجملة عربية بسيطة غير مشكولة، من خلال بناء نموذج لغوي هجين يعتمد على بنية المحولات حيث يدمج قوة نموذج AraBERT المدرب مسبقاً لاستخلاص السمات السياقية العميقة للكلمات، مع طبقات محول إضافية لمعالجة هذه السمات، ورؤوس تصنيف متعددة لاستخراج 13 حقلاً إعرابياً وصرفياً مختلفاً بشكل متزامن. تسعى المنظومة المقترحة إلى تحقيق دقة عالية في المهام المتعددة التالية:

- تحديد الفعل وجذره وزمنه
 - تحديد الاسم ونوعه النحوي وحالة التعريف والتنكير للاسم بالإضافة إلى إعرابه الكامل
 - تصنيف الكلمات داخل الجملة من حيث الوظيفة (فاعل، مفعول به، صفة، مبتدأ، ...)
- وذلك بالاعتماد على قاعدة بيانات مصممة خصيصاً لهذا الغرض. كما ويستهدف البحث في تطبيقاته المستقبلية دعم عدد من المجالات مثل:

- إنشاء جمل ونصوص عربية سليمة صرفياً ونحويًا
- منصات التعليم الإلكتروني للغة العربية، من خلال تقديم إعراب دقيق ومتكامل.

3. الدراسات المرجعية

تتوعد الأبحاث في مجال إيجاد الصيغ النحوية للجمال العربية بدءاً من المعربات الإحصائية التقليدية وصولاً إلى نماذج التعليم العميق الحديثة والمحولات، حيث نستعرض في هذا القسم أبرز الأبحاث والأدوات المستخدمة مع مناقشة نقاط القوة والقصور في كل مجموعة منها.

النماذج الإحصائية المعتمدة على القواعد: وهي من أقدم أساليب التحليل النحوي، وقد نُشرت العديد من الأبحاث والأدوات في هذا المجال، وكان أولها المحلات الإحصائية التقليدية مثل محلل ستانفورد Stanford Parser [3]، والذي اعتمد على نموذج القواعد النحوية الخالية من السياق Context-Free Grammars (CGF) الاحتمالي والمدرّب على بيانات شجرة (PATB) Penn Arabic Tree Bank. أما معرب CAMELParser [4] فقد عمل على تقديم نظام متكامل لتحليل الاعتماديات النحوية بين الكلمات بما يتوافق مع السمات الصرفية الغامضة سياقياً، وقد استخدم أداة متطورة لإزالة الغموض الصرفي، كما وحسن نتائجه بالاعتماد على كل من MaltParser [5] مع تقنية SVM محققين دقة عالية نسبياً في تحليلات النحو العربية التقليدية. وبذلك فإن نقاط القوة في هذه النماذج كانت بالاستناد الى قواعد لغوية واضحة، في حين أن نقاط القصور تمثلت في عدم القدرة على التعميم للأنماط اللغوية غير التقليدية.

الحزم البرمجية المتكاملة: تقدم مجموعة من المهام الصرفية والنحوية، كما في [6] Farasa، والتي دمجت تقنيات تحليل الصرف والاعتمادية، حيث تقوم بمجموعة من المهام مثل تقطيع الكلمة tokenization وإيجاد أجزاء الكلام POS، وتحليل الاعتمادية باستخدام محلل MaltParser. في حين أن فريق CAMEL طور مجموعة من الأدوات الشاملة تدعى CAMEL tools والتي توفر مجموعة من الأدوات المساعدة في المعالجة المسبقة، النمذجة الصرفية، بالإضافة إلى إيجاد أجزاء الكلام [7]. كما برز نظام MADAMIRA للتحليل الصرفي وإزالة الغموض [8] حيث جمع بين مميزات الأنظمة السابقة، وحقق كل من MADAMIRA ونظام CAMEL دقة مقاربة في مهمة POS للنصوص العربية [9] بنسبة

97.1% MADAMIRA، و 97.2% في CAMEL من أجل العربية الفصحى، و 91.7% MADAMIRA بمقابل 91.8% في CAMEL من أجل اللهجة المصرية. وبذلك فإن تعدد الوظائف وسهولة الاستخدام تعتبر كنقاط قوة في حين أن ضعف الأداء على النصوص غير المشكولة تعتبر من نقاط القصور.

نماذج تعلم الآلة: بدأ التحليل النحوي يشهد تحسناً ملموساً، في [10]، تم استخدام مصنف يعتمد على Conditional Random Field CRF للتقطيع وإيجاد الوظيفة النحوية في آن واحد، وفي [11] تفوقت طريقة SVM-Rank المعتمدة على المميزات التقليدية على نموذج Bi-LSTM، إلا أن الجمع بين تمثيلات الكلمة embeddings وبعض المميزات المستخرجة يدوياً ساعد في تحسين أداء الشبكة العصبونية. تمتاز هذه النماذج بقدرتها على التعامل مع التسلسلات الطويلة، إلا أنها تحتاج إلى حجم بيانات كبير.

نماذج التعلم العميق والمحولات: قدمت الدراسة [12] أول نموذج محول يقوم على آلية الانتباه الذاتي self-attention، وتبعه طرح نموذج [13] BERT وهو نموذج لغوي عميق ثنائي الاتجاه، مدرب على كميات هائلة من النصوص لإنجاز مهمتي نمذجة اللغة المقنعة Masked Language Modeling (MLM)، ومهمة التنبؤ بالجملة التالية Next sentence prediction (NSP). ثم تلاه تطوير نموذج AraBERT [14] المدرب على عدد كبير من النصوص العربية، وقد حقق أداءً متقدماً في فهم اللغة العربية وتحليل المشاعر وتصنيف النصوص، كما في [15] حيث تفوق نموذج AraBERT في الكشف عن الأخبار الزائفة محققاً 78.90% متوسط f1-score. في حين تفوق fine-tuned AraBERT في [16] بمهمة تصنيف المواضيع المكتوبة باللغة العربية واللهجة السورية تحديداً على الخوارزميات الأخرى بمعدلات f1-score, accuracy بمقدار 99%. وجدت الدراسة [17] أن AraBERTv2 كان الأفضل بالاختبار على جميع الأشجار المعيارية treebanks في مهمة الإعراب الاعتمادي، حيث حقق دقة UAS بلغت 84.0% و LAS 80.3% على شجرة PADT، و 86.4/87.5 على شجرة CATiB. في [18] طُوِّر فريق CAMEL نموذج CamelParser2.0 مفتوح المصدر لتحليل الاعتمادية العربية، وقد حقق هذا النموذج دقة عالية في إيجاد الوظيفة النحوية للجمل العربية الغير مشكولة مع درجات LAS تتجاوز 91%.

وبذلك نجد بأن نقاط القوة للمحولات والنماذج المدربة مسبقاً في دعم المهام المتعددة والتفوق على النماذج التقليدية في حين أن القصور يتجلى في أن معظم النماذج تفتقر إلى تحليل إعرابي شامل يجمع بين الجوانب الصرفية والنحوية والدلالية

4. مواد وطرق البحث

4.1. قاعدة البيانات المستخدمة

يعتبر تصميم قاعدة بيانات مناسبة عنصراً مهماً في تدريب وتقييم أداء النماذج اللغوية ولاسيما في اللغة العربية، لغناها بالتركيب الصرفية والنحوية، وقد افتقرت قواعد البيانات التقليدية بشكل عام إلى تمثيل رسمي ومنظم مناسب، ففي أغلب الأحيان يتم تخزين التحليلات النحوية بشكل نثري، الأمر الذي يتطلب مستوى عالٍ من الخبرة لدى المستخدم في قواعد اللغة العربية، مما يجعل عملية تصميم نظام الاعراب الآلي أمراً معقداً.

خلال هذا البحث تم تصميم قاعدة البيانات من خلال توليد الجمل ومكوناتها الصرفية والنحوية يدوياً من قبل الباحث، حيث اجتزأت هذه الكلمات من قاعدة بيانات Flickr8k المعربة [19] والمخصصة لتوصيفات الصور باللغة العربية وشكلنا منها جملاً تتبع للقوالب المقترحة، مما يضمن تنوع الجمل وتوازنها ضمن القوالب النحوية. تتبع هذه المرحلة، عملية تدقيق من قبل خبرة لغوية لتشكيل الجمل واعرابها، وجذور الأفعال، بالإضافة إلى سوابق الأفعال ولواحقها وكذلك الأمر بالنسبة للأسماء.

تتألف قاعدة البيانات من 3960 جملة قصيرة وغير مشكولة، أما أعمدة قاعدة البيانات فكانت عبارة عن المعلومات الصرفية والنحوية لكل جملة وبشكل مفصل، وهي: الجملة مشكولة، الجملة بدون تشكيل، الفعل، جذر الفعل، سوابق الفعل، لواحق الفعل، زمن الفعل، نوع الفعل (مفرد مذكر، مفرد مؤنث، مثنى مذكر، ...)، الاسم، نوع الاسم النحوي (فاعل، مفعول به، ...)، نوع الاسم (مفرد مذكر، مفرد مؤنث، ...)، اعراب الاسم الكامل.

تنتهي جمل قاعدة البيانات المقترحة إلى 66 قالباً نحوياً، بحيث أننا قمنا بتوليد 60 جملة بدون وجود تكرارات بين الجمل من أجل كل قالب نحوي مقترح، كما وتنتهي القوالب النحوية

البالغ عددها 66 إلى 11 قالباً نحوياً أساسياً، مع الحرص على التنوع في الضمائر، أنماط التأنيث والتذكير، الإفراد والجمع، وتراكيب الاسم والفعل، وبذلك فإن القوالب كانت على الشكل:

- فعل ماضٍ/ فعل مضارع + فاعل
- مع الاخذ بعين الاعتبار جميع حالات الفاعل (مفرد مذكر/ مفرد مؤنث/ مثنى مذكر/ مثنى مؤنث/ جمع مذكر/ جمع مؤنث)
- فعل ماضٍ/ فعل مضارع/ فعل أمر + مفعول به
- مع الاخذ بعين الاعتبار جميع حالات الفعل (مفرد مذكر/ مفرد مؤنث/ مثنى مذكر/ مثنى مؤنث/ جمع مذكر/ جمع مؤنث)
- فعل ماضٍ/ فعل مضارع/ فعل أمر + ظرف زمان
- مع الاخذ بعين الاعتبار جميع حالات الفعل (مفرد مذكر/ مفرد مؤنث/ مثنى مذكر/ مثنى مؤنث/ جمع مذكر/ جمع مؤنث)
- فعل ماضٍ/ فعل مضارع/ فعل أمر + حال
- مع الاخذ بعين الاعتبار جميع حالات الفعل (مفرد مذكر/ مفرد مؤنث/ مثنى مذكر/ مثنى مؤنث/ جمع مذكر/ جمع مؤنث)

ويبين الجدول (1) أمثلة عن القوالب النحوية المستخدمة:

الجدول 1- القوالب النحوية المقترحة

ال قالب النحوي 1	فعل ماضٍ + فاعل مفرد مذكر
ال قالب النحوي 2	فعل ماضٍ + فاعل مفرد مؤنث
ال قالب النحوي 3	فعل ماضٍ + فاعل مثنى مذكر
...	...
ال قالب النحوي 11	فعل مضارع + فاعل جمع مذكر سالم
...	...
ال قالب النحوي 21	فعل مضارع مثنى مذكر + مفعول به
...	...
ال قالب النحوي 66	فعل أمر جمع مؤنث + حال

على سبيل المثال قمنا بتمثيل جملة "عملن التصميم" في قاعدة البيانات المقترحة كما في الجدول (2)

الجدول 2-مثال عن تمثيل الجمل في قاعدة البيانات المقترحة

عملن التصميم			الجملة
الفعل	عملن	الاسم	التصميم
جذر الفعل	عمل	-	-
سوابق الفعل	Null	سوابق الاسم	ال
لواحق الفعل	ن	لواحق الاسم	Null
زمن الفعل	ماض	نوع الاسم النحوي	مفعول به
نوع الفعل	جمع مؤنث	نوع الاسم	مفرد مذكر
اعراب الفعل	فعل ماض مبني على السكون لاتصاله بنون النسوة، والنون ضمير متصل مبني على الفتحه في محل رفع فاعل	اعراب الاسم	مفعول به منصوب وعلامة نصبه الفتحة الظاهرة على اخره

الجدير بالذكر أن اعتماد القوالب النحوية ساعد على ضبط البنية اللغوية للجمل والتحكم في مكوناتها، إلا أن هذا النهج قد يقلل من تنوع التراكيب اللغوية، مما قد يؤثر على قدرة النموذج على التعامل مع النصوص الطبيعية بكفاءة. وانطلاقاً من ذلك، يقترح هذا البحث في الأعمال المستقبلية:

- توسيع قاعدة البيانات بإدخال جمل من مصادر طبيعية مثل الكتب التعليمية والمقالات
- دمج الجمل المعقدة والتراكيب البلاغية تدريجياً في مراحل لاحقة من التدريب.

4.2. تصميم المنظومة

4.2.1. الدراسة التحليلية - المحولات [12]

تُعد بنية المحولات من أبرز التطورات في مجال الشبكات العصبونية، حيث تخلت عن آليات التكرار التقليدية المستخدمة في الشبكات العصبونية التكرارية RNN والذاكر الطويلة قصيرة الأمد LSTM، واعتمدت بشكل كلي على آلية الانتباه الذاتي لمعالجة التسلسلات بكفاءة عالية [12]، سواء في اللغة أو الصور.

أولاً: البنية العامة للمحول: يتكون نموذج المحول الأساسي من جزئين رئيسيين وهما المشفر Encoder وفك التشفير Decoder، وكل منهما يتألف من عدة طبقات متماثلة.

في طبقات المشفر، نجد وحدتين فرعيتين أساسيتين:

- آلية الانتباه الذاتي متعدد الرؤوس Multi-Head Self Attention
 - شبكة تغذية أمامية Feed Forward Neural Network
- أما فك التشفير، فيحتوي على هاتين الوحدتين، بالإضافة إلى آلية انتباه Encoder-Decoder Attention

ثانياً: الانتباه الذاتي: تعمل على حساب أوزان الانتباه بين كل كلمة في الجملة وبقية الكلمات، بغض النظر عن المسافة بين الكلمات، مما يساعد النموذج على فهم العلاقات السياقية القصيرة والطويلة داخل التسلسل، وذلك من خلال تحويل كل عنصر في التسلسل الى ثلاث متجهات:

- Q (الاستعلام Query)
- K (المفتاح Key)
- V (القيمة Value)، وفق المعادلة (1) التالية:

$$Attention(Q, K, V) = softmax(QK^T / \sqrt{d_k})V \dots \dots (1)$$

- QK^T درجة التشابه بين كل Key و Query
- $\sqrt{d_k}$ لتقادي القيم الكبيرة في SoftMax

ثالثاً: الانتباه متعدد الرؤوس: بدلاً من استخدام Attention واحدة، يستخدم المحول عدة رؤوس انتباه Heads تعمل بشكل موازي. كل رأس يتعلم علاقات سياقية مختلفة، ثم تُجمع نتائج الرؤوس وتُمرر عبر مصفوفة نهائية، كما في العلاقتين (2) و (3)

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^o \dots \dots (2)$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \dots \dots (3)$$

رابعاً: مزايا المحولات:

- التوازي الكامل أثناء التدريب، لعدم وجود اعتماد زمني كما في RNN
- كفاءة في التقاط العلاقات البعيدة داخل التسلسل اللغوي
- سهولة التوسع إلى نماذج ضخمة كما في BERT و GPT، وقد أدى نجاح المحولات إلى ظهور نماذج لغوية قوية مثل: BERT [13] نموذج ثنائي الاتجاه يعتمد على المحولات لفهم السياق من كلا الجانبين، و GPT-3 [20] نموذج توليدي ضخم يستخدم طبقات المحول لإنتاج نصوص شبيهة بالبشر

4.2.2. بنية النموذج المقترح:

يعتمد النموذج المقترح بشكل أساسي على بنية هجينة تجمع بين قوة نموذج AraBERT v2 المدرب مسبقاً في تمثيل السمات السياقية العميقة للنصوص العربية، مع طبقات محول إضافية لمعالجة هذه السمات، وفي النهاية رؤوس تصنيف متعددة بما أن المسألة تتدرج ضمن مجال Multi Task Classification، يهدف هذا التصميم إلى تحليل الجملة العربية واستخلاص الحقل الإعرابية النحوية والصرفية لكل كلمة من كلمات الجملة.

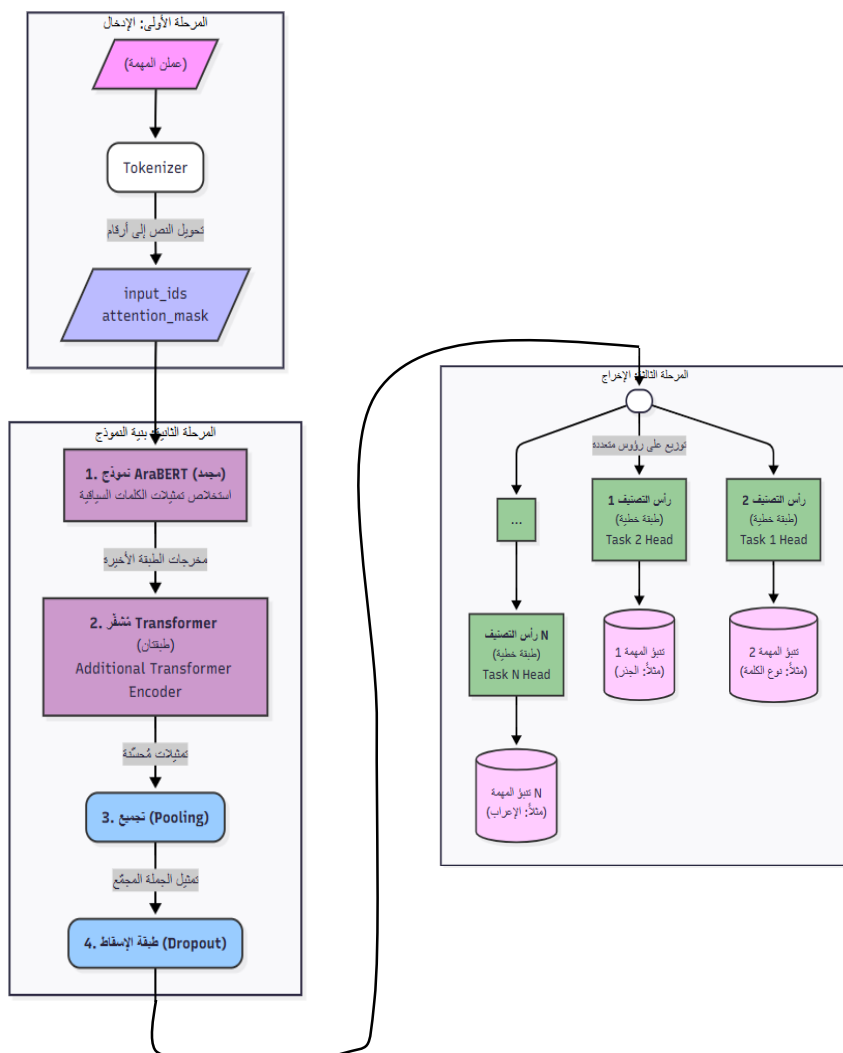
يعود سبب الاعتماد على نموذج AraBERT v2 لكونه أحدث إصدار متوفر من سلسلة نماذج AraBERT حتى تاريخ إعداد البحث، ويتميز بتحسينات ملحوظة في الأداء والاستقرار مقارنة بالإصدارات السابقة، كما أنه مخصص للغة العربية مما يضمن تمثيلاً أكثر دقة للسياق اللغوي والتراكيب النحوية مقارنة بالنماذج متعددة اللغات.

يتكون النموذج المقترح من المكونات التالية:

المرحلة الأولى: التقطيع النصي باستخدام المجزئ Tokenizer الخاص بنموذج AraBERT

المرحلة الثانية، والتي تتألف من:

1. مرحلة تمثيل الكلمات: باستخدام نموذج AraBERT v2 المدرب مسبقاً لاستخراج تمثيل الكلمات، تم تجميد أوزان النموذج الأساسي، واستخراج التمثيلات من خرج آخر طبقة خفية



للمنموذج، وبذلك نحصل على التمثيلات الدلالية بطول 768 توكن.

2. محول سياقي Transformer Encoder، دخل هذا المحول هو شعاع تمثيلات كلمات الجملة الناتجة عن المرحلة السابقة، ويتألف من طبقتين، كل طبقة من نوع Transformer Encoder layer، حيث تتميز كل طبقة بما يلي:

الشكل 1- بنية النموذج المقترح

○ عدد مداخل هذه الطبقة يساوي 768 وهو متوافق مع طول شعاع التضمين الناتج

عن AraBERT

○ آلية انتباه ذاتي متعدد الرؤوس، تتألف من 8 رؤوس انتباه ذاتي لالتقاط العلاقات السياقية المتنوعة.

○ شبكة تغذية أمامية تحتوي على 512 عصبون

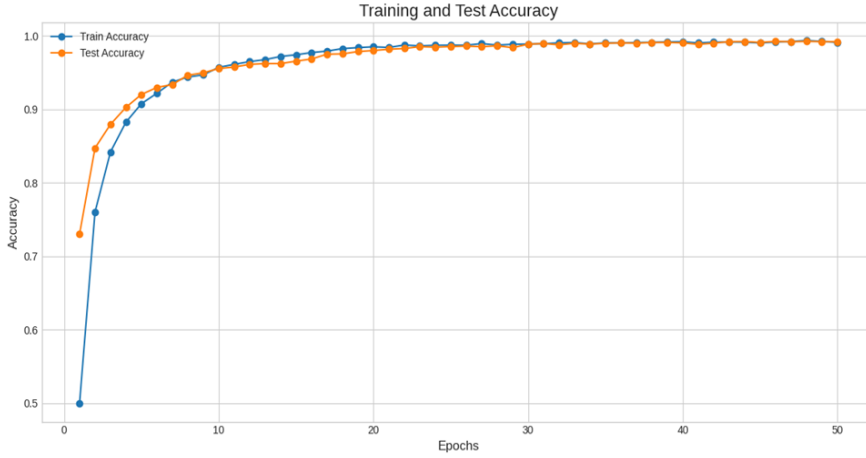
3. طبقة التجميع pooling layer، لدمج المعلومات السياقية الناتجة للجملة في متجه واحد بطول ثابت ويساوي 768 توكن

4. طبقة الاسقاط Dropout للحد من مشكلة overfitting وتحسين التعميم Generalization، وذلك بعد مرحلة التجميع على التمثيل الناتج، بحيث يتم اسقاط 10% من الوحدات الناتجة Units بشكل عشوائي.

المرحلة الثالثة: مرحلة الإخراج وهي عبارة عن مجموعة رؤوس مستقلة بعدد الحقول الإعرابية المستهدفة (الفعل، الجذر، الزمن، ...) والبالغ عددها 13 مخرجاً، حيث استخدمنا طبقة خطية Linear لكل حقل لتوليد التنبؤ المناسب حسب عدد فئاته. ويبين الشكل (1) بنية النظام المقترح

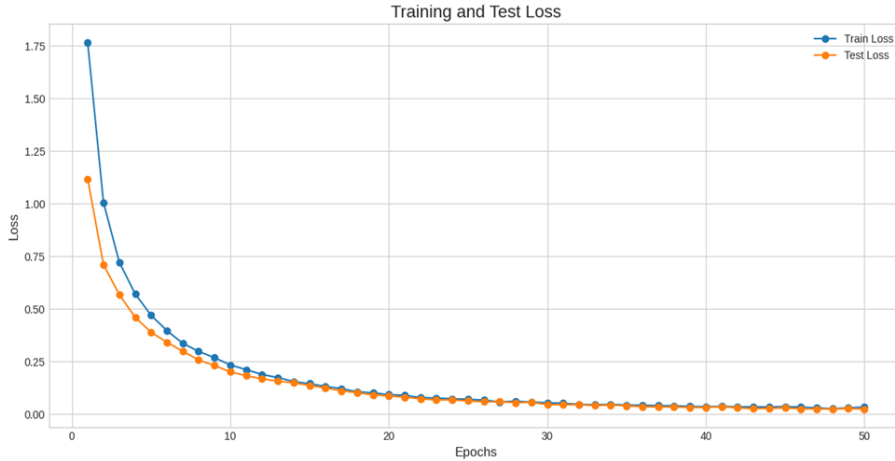
5. النتائج والمناقشة

تم تدريب النموذج باستخدام استراتيجية متعددة المهام حيث يتم توليد مخرجات متعددة متوازية من رؤوس مختلفة، يمثل كل منها تصنيفاً مستقلاً لأحد الحقول النحوية أو الصرفية وهي الفعل، جذر الفعل، سوابق الفعل، لواحق الفعل، اعراب الفعل بالإضافة الى نوع الفعل وزمنه، وكذلك الامر بالنسبة للاسم. يتم حساب خسارة loss كل مخرج بشكل مستقل باستخدام Cross Entropy، ثم يُؤخذ متوسط هذه الخسائر لتحديث النموذج. كما ويتم استخدام الدقة Accuracy للمقارنة بين التوقعات والتصنيفات الحقيقية لجميع المهام. أُنجزت عملية التدريب على دفعات بحجم ثابت batch size = 32، مع تطبيق الانتشار العكسي والتحديث باستخدام خوارزمية التحسين Adam، وكان عدد دورات التدريب المستخدمة 50 دورة، كما وتم تقسيم مجموعة البيانات إلى تدريب واختبار، بنسبة 80% تدريب و20% للاختبار، ويبين الشكل (2) مخطط الدقة لمرحلي التدريب والاختبار،



الشكل 2- قيم Accuracy للتدريب والاختبار

حيث نلاحظ بأن قيمة الدقة التي تم الحصول عليها بنهاية الاختبار هي 99.29%، في حين أن آخر قيمة تم الحصول عليها اثناء التدريب 99.31%



الشكل 3- قيم Loss للتدريب والاختبار

في حين يمثل الشكل (3) قيم الخطأ خلال عملية التدريب والاختبار، ونلاحظ بأن نسبة الخطأ التي تم التوصل إليها بنهاية عملية الاختبار كانت 0.0233، في حين أن نسبة الخطأ التي تم التوصل اليه بمرحلة التدريب 0.0280

من الشكل 2 و 3 يمكن أن نستنتج أن:

- التعلم سريع: بدأ النموذج المقترح بالتعلم بسرعة كبيرة، حيث قفزت الدقة من 50% إلى 84% في 3 دورات فقط. هذا يدل على أن المهمة قابلة للتعلم وأن بنية النموذج مناسبة
- تقارب الخسارة Loss Convergence: تناقصت كل من خسارة التدريب وخسارة الاختبار بشكل مستمر ومتناسق، ولم يُلاحظ وجود تباعد بينهما، وهذا يؤكد عدم وجود Overfitting
- الاستقرار في النهاية: في آخر 15 دورة، بدأ التحسن في الدقة يتباطأ، وهذا دليل أن النموذج وصل إلى نقطة قريبة جداً من الأداء الأمثل على هذه البيانات وهي 99.29% على بيانات الاختبار

قمنا بمقارنة نتائج النموذج المقترح مع نماذج لغوية رائدة وهي Gemini 2.5 pro، GPT-4 turbo، بالإضافة إلى Felo v3.1.2 على مجموعة جمل الاختبار، حيث أعددنا أمراً واحداً prompt لإدخاله للنماذج الثلاثة، وكان على الشكل التالي:

- إعطاء ذات الأمر للنماذج الثلاثة الجاهزة بأن الدخل هو مجموعة جمل الاختبار فقط والبالغ عددها 792 جملة وسيتم إدخالها دفعة واحدة، أما المخارج المطلوبة فهي (الفعل، جذر الفعل، سوابق الفعل، لواحق الفعل، زمن الفعل، نوع الفعل، اعراب الفعل، الاسم، سوابق الاسم، لواحق الاسم، نوع الاسم النحوي، نوع الاسم، اعراب الاسم) مع إعطاء مثال مفصل عن الخرج المطلوب لكل مخرج
- تهيئة النماذج الثلاثة بخمسة جمل من قاعدة البيانات حيث تم ادخال الجملة مع الخرج المطلوب وهو ما يعرف "Few-shot"

وكانت النتائج (الدقة) كما هو موضح في الجدول (3)

الجدول 3- مقارنة نتائج المنظومة المقترحة مع كل من GPT, Felo, Gemini

Felo	Gemini pro	GPT 4 turbo	المنظومة المقترحة	Acc%
%58.48	%67.79	%59.19	%99.29	

ويبين الجدول (4) نظرة أعمق على نتائج كل من GPT، Gemini، Felo، والمنظومة المقترحة لنبيين أداء كل مخرج بشكل مفصل. كما ويبين الشكل (4) دقة كل مخرج للمنظومة المقترحة بشكل منفصل خلال مرحلة الاختبار.

الجدول 4- أداء كل مخرج من مخارج المنظومة (Accuracy)

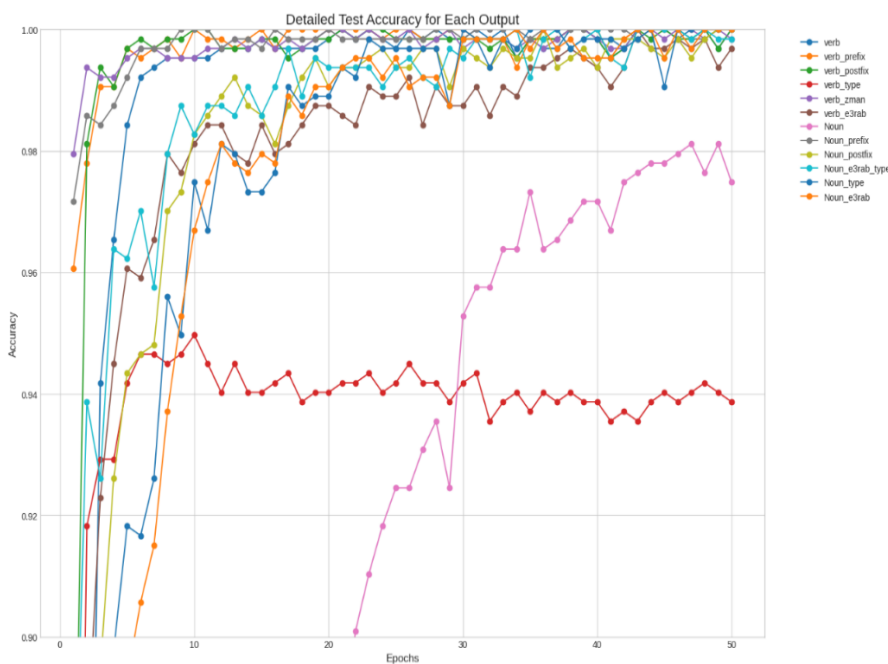
Felo	Gemini pro	GPT 4 turbo	المنظومة المقترحة	
%98.11	%100	%100	%100	الفعل
%99.53	%99.69	%61.48	%100	جذر الفعل
%45.13	%100	%79.09	%99.69	سوابق الفعل
%82.70	%61.95	%73.74	%100	لواحق الفعل
%78.14	%0.16	%27.36	%94.03	نوع الفعل
%54.72	%100	%79.09	%99.84	زمن الفعل
%0	%36.32	%0	%99.21	اعراب الفعل
%99.37	%99.69	%100	%97.80	الاسم
%46.54	%100	%100	%100	سوابق الاسم
%52.99	%71.70	%52.99	%99.84	لواحق الاسم
%40.57	%58.49	%32.55	%99.53	نوع الاسم النحوي
%62.42	%46.23	%30.66	%100	نوع الاسم
%0	%7.08	%32.55	%99.84	اعراب الاسم

ويوضح الجدول (5) مؤشرات الأداء، وهي زمن التدريب، زمن الاختبار بالإضافة إلى حجم النموذج المقترح من أجل 50 دورة تدريب، وتجدر هنا الإشارة إلى أن نماذج GPT، Gemini، Felo هي نماذج مبتكرة من شركات مرموقة ومدرية على ملايين النصوص التي لا يمكن حصرها وهذه النماذج غير قابلة لإعادة التهيئة والتدريب. تم تنفيذ عملية تدريب النموذج باستخدام بيئة

Google Colab، مع الاستفادة من وحدة المعالجة المخصصة للتعلم العميق (TPU V2) لتسريع عملية التدريب وتقليل الزمن الحسابي.

الجدول 5- مؤشرات الأداء

حجم النموذج	زمن الاختبار	زمن التدريب	المنظومة المقترحة
541.07 MB	60.18 ثانية	294.17 ثانية	



الشكل 4- تغيير Accuracy عبر epochs من أجل كل مخرج من مخرج المنظومة المقترحة

قمنا بدراسة المخرجات الصرفية والنحوية لكل من المنظومة المقترحة بالإضافة إلى GPT، Gemini، Felo، وكانت النتائج كما يلي:

أولاً - المنظومة المقترحة: وفقاً لنتائج الجدول (4)، كان حقل "نوع الفعل" هو الأقل دقة في المخرجات حيث بلغت 94.03%، لذلك نورد فيما يلي أمثلة عن جمل أخفق فيها النموذج في تحديد نوع الفعل من حيث التذكير أو التأنيث وكذلك العدد:

- "عملوا متفائلين"، كانت جميع الحقول صحيحة عدا حقل "نوع الفعل"

- التنبؤ الصحيح: جمع مذكر
- تنبؤ النموذج: جمع مؤنث
- "تكتبان الرسالة"، كانت جميع الحقول صحيحة عدا حقل "نوع الفعل"
- التنبؤ الصحيح: مثنى مؤنث
- تنبؤ النموذج: جمع مذكر

في حين أنه اعطى نتائج صحيحة من أجل باقي الجمل وعلى سبيل المثال:

- يكتبان فجرا:
- الفعل: "يكتبان" - الجذر: "كتب" - سوابق الفعل: "ي" - لواحق الفعل: "ان" - نوع الفعل: "مثنى مذكر" - زمن الفعل: "مضارع" - اعراب الفعل: "فعل مضارع مرفوع وعلامة رفعه ثبوت النون، وألف الاثنين ضمير متصل مبني في محل رفع فاعل"
- الاسم: "فجرا" - سوابق الاسم: " " - لواحق الاسم: "ا" - نوع الاسم النحوي: " ظرف زمان" - نوع الاسم: " " - اعراب الاسم: " ظرف زمان منصوب وعلامة نصبه تنوين الفتح الظاهر على آخره"

ثانياً - GPT-4 turbo:

- قام بإعراب جميع الأفعال سواء أكانت فعل ماض، مضارع أو أمر، وسواء أكانت جمع أو مفرد أو مثنى، على أنها "فعل ماض مرفوع وعلامة رفعه الضمة الظاهرة، والفاعل ضمير مستتر تقديره هو".
- في حقل زمن الفعل، لم يتعرف على فعل الأمر، وإنما اعتبره فعل ماض
- أعرب الكلمة الثانية من الجملة "مفعول به منصوب وعلامة نصبه الفتحة الظاهرة على آخره"، علماً أن اعرابات الأسماء تتنوع بين فاعل، مفعول به، ظرف زمان وحال.
- أرجع قيم NULL من أجل لواحق الاسم، علماً أن هناك أسماء لها لواحق مثل "ات"، "ان"، "ون"، ...

ثالثاً - Gemini-2.5 Pro:

- في نوع الاسم النحوي، أعرب الحال على أنه مفعول به
- أعرب ظرف الزمان بشكل صحيح، إلا أنه أخطأ بإعراب المفعول به على أنه فاعل.
- في زمن الفعل بالنسبة للأفعال الماضية فكان التصنيف لها في بعض الأحيان صحيح وفي بعض الأحيان مضارع أو أمر.
- في نوع الفعل، بعضها كان صحيح والبعض خاطئ إلا أنه أرفق جميع الأنواع بعبارة "مفرد مذكر غائب معلوم" وبالتالي لم يتطابق مع المطلوب وهو فقط مفرد مذكر (على سبيل المثال).
- فشل في إعراب تاء التأنيث المتصلة بالفعل وكذلك الأمر بالنسبة لنون النسوة
- كما فشل في إعراب المفعول به جمع المؤنث السالم (المنتهي ات) وقام بإعرابه على أنه فاعل

رابعاً - Felo v3.1.2:

- أخطأ بإعراب جميع الأفعال، فعلى سبيل المثال:
 - درس/ درست/ درسا: فعل ماضٍ مبني على الفتحة الظاهرة على آخره، والفاعل ضمير مستتر تقديره هو وبعض الجمل تقديره هما وهم وهي
 - درست: فعل ماضٍ مبني على الفتحة الظاهرة على آخره، والتاء للتأنيث الساكنة ضمير متصل مبني في محل رفع فاعل
 - يكتب: فعل مضارع مرفوع وعلامة رفعه ثبوت النون، وفاعل مرفوع ضمير مستتر تقديره هو
 - اكتب: فعل أمر مبني على حذف النون، وفاعل مستتر تقديره أنت
 - كتبن: فعل ماضٍ مبني على حذف النون، وفاعل مستتر تقديره هن
 - يكتب مساءً: فعل مضارع مرفوع وعلامة رفعه ثبوت النون
 - أعرب أغلب الأسماء على أنها مبتدأ عدا الحال وظرف الزمان
 - اكتشف الحال وظرف الزمان بشكل صحيح، إلا أنه أخطأ بالعلامة الإعرابية
6. الاستنتاجات والتوصيات:

قدمنا خلال هذا البحث مقترحاً لقاعدة بيانات للجمل العربية وقد شملت قاعدة البيانات المقترحة على أغلب المعلومات الصرفية والنحوية الخاصة بهذه الجمل، وكانت الغاية هي الوصول إلى الإعراب الكامل لمكونات الجملة، كما واقترحنا بنية لمنظومة الإعراب الكامل وإيجاد الخصائص الصرفية والنحوية للجمل العربية باستخدام محول AraBERT v2 المدرب مسبقاً مع إضافة طبقتين من المحولات، وذلك ضمن إطار Multi-Task Classification لمعالجة المهام المتعددة بشكل متوازي، وقد أثبت النتائج تفوق المنظومة المقترحة على كل من ChatGPT، Gemini و Felo، ولاسيما بعد ادخال جمل الاختبار دفعة واحدة الى هذه النماذج، وسنعمل في المستقبل على توسعة قاعدة البيانات لتشمل جمل عربية بعدد مركبات أكبر وقوالب نحوية أشمل.

7. المراجع

- [1]. SHAHROUR, A., *et al.* 2015- Improving Arabic diacritization through syntactic analysis, in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 1309-1315). <https://doi.org/10.18653/v1/D15-1152>
- [2]. KHOUI, N., *et al.* 2024- Comparative Study for Text Chunking Using Deep Learning: Case of Modern Standard Arabic, Computación y Sistemas, 28(2), 517-527. <https://doi.org/10.13053/cys-28-2-4560>
- [3]. DARWISH, K., *et al.* 2020- A panoramic survey of natural language processing in the Arab world, Communications of the ACM, 64(4), 72-81 <https://arxiv.org/pdf/2011.12631>
- [4]. SHAHROUR, A., *et al.* 2016- CamelParser: A system for Arabic syntactic analysis and morphological disambiguation, in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations (pp. 228-232) <https://aclanthology.org/C16-2048/>
- [5]. NIVRE, J., *et al.* 2006- MaltParser: A data-driven parser-generator for dependency parsing, in Proceedings of the Language Resources and Evaluation Conference, pp. 2216–2219. http://www.lrec-conf.org/proceedings/lrec2006/pdf/162_pdf.pdf

- [6]. ABDELALI, A., *et al.* 2016- Farasa: A fast and furious segmenter for Arabic, in Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Demonstrations (pp. 11-16). <https://aclanthology.org/N16-3003/>
- [7]. OBEID, O., *et al.* 2020- CAMEL tools: An open-source python toolkit for Arabic natural language processing, in Proceedings of the 12th Language Resources and Evaluation Conference, pp. 7022–7032. <https://aclanthology.org/2020.lrec-1.868/>
- [8]. PASHA, A. *et al.* 2014- MADAMIRA: A Fast, Comprehensive Tool for Morphological Analysis and Disambiguation of Arabic, in Proceedings of the Ninth International Conference on Language Resources and Evaluation, p 1094–1101. http://www.lrec-conf.org/proceedings/lrec2014/pdf/593_Paper.pdf
- [9]. ALLUHAIBI, R., *et al.* 2021- A Comparative Study of Arabic Part of Speech Taggers Using Literary Text Samples from Saudi Novels, Information, 12(12), 523. <https://doi.org/10.3390/info12120523>
- [10]. GAHBICHE-BRAHAM, *et al.* 2012- Joint segmentation and POS tagging for Arabic using a CRF-based classifier, in LREC (pp. 2107-2113). <https://aclanthology.org/L12-1509/>
- [11]. DARWISH, K. *et al.* 2017- Arabic POS Tagging: Don't Abandon Feature Engineering Just Yet, in Proceedings of the Third Arabic Natural Language Processing Workshop, pages 130–137. <https://aclanthology.org/W17-1316/>
- [12]. VASWANI, A. *et al.* 2017- Attention is all you need, Advances in neural information processing systems, 30. <https://arxiv.org/abs/1706.03762>
- [13]. DEVLIN, J., *et al.* 2019- BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, in Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186). <https://arxiv.org/abs/1810.04805>
- [14]. Antoun, W., *et al.* 2020- ARABERT: Transformer-based model for Arabic Language Understanding, in Proceedings of the 4th Workshop on

Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection, pages 9–15.
<https://aclanthology.org/2020.osact-1.2/>

[15]. MHREZ, A., *et al.* 2024- Evaluating the Performance of Arabic Language Models on the Task of Stance Detection Toward Fake News, *Journal of Homs University Vol 46. 5, P 83,106* [In Arabic].
<https://journal.homs-univ.edu.sy/index.php/Engineering/article/view/5120>

[16]. AHMAD, A., *et al.* 2023- Topic Detection in Arabic Social Media Text Data (Syrian Dialect), *Journal of Homs University Vol 45. 14,*[In Arabic].
<https://journal.homs-univ.edu.sy/index.php/Engineering/article/view/3020>

[17]. AL-GHAMDI, S., *et al.* 2023- Fine-Tuning BERT-Based Pre-Trained Models for Arabic Dependency Parsing, *Applied Sciences*, 13(7), 4225.
<https://doi.org/10.3390/app13074225>

[18]. ELSHABRAWY, A., *et al.* 2023- CamelParser2.0: A State-of-the-Art Dependency Parser for Arabic, in *Proceedings of ArabicNLP 2023* (pp. 170-180). <https://aclanthology.org/2023.arabnlp-1.15.pdf>

[19]. ElJundi, O., *et al.* 2020- Resources and end-to-end neural network models for Arabic image captioning. *VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. SciTePress, pp. 233–241.
https://www.researchgate.net/publication/340044948_Resources_and_End-to-End_Neural_Network_Models_for_Arabic_Image_Captioning

[20]. BROWN, T., *et al.*, 2020- Language Models are Few-Shot Learners, *Advances in neural information processing systems*, 33, 1877-1901.
<https://arxiv.org/abs/2005.14165>

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

م. محمد الهلامي * د. أكرم مرعي ** د. ماهر عباس ***

كلية الهندسة المعلوماتية- جامعة حمص

ملخص

يُتوقع أن يشهد العقد القادم تزايداً في الأجهزة والهواتف المحمولة والمستشعرات والمشغلات التي تتواصل عبر الانترنت لتعزيز نمط الحياة البشرية. يجب أن يضمن تصميم شبكة انترنت الأشياء الكفاءة في استهلاك الطاقة، والسرعة، والموثوقية لضمان استدامة الشبكة على المدى الطويل. يعتبر توجيه وموازنة البيانات بين العقد أحد الجوانب التي تتحكم بشكل كبير في جودة الخدمة في شبكات أنترنت الأشياء. تم تصميم البروتوكول القياسي بروتوكول التوجيه للشبكات منخفضة الطاقة والتي تحتوي فقدان لرزم Routing Protocol For Low-Power And Lossy Networks (RPL) لاختيار أفضل مسار بناءً على مقياس جودة وصلة اتصال واحدة مثل عدد القفزات أو عدد الإرسال المتوقع، يؤدي هذا إلى اختيار مسارات غير موثوقة وبالتالي هدر لموارد الشبكة وزيادة عملية إعادة إرسال البيانات وعدم موازنة الحمل والتي تشكل ما يعرف بالنقاط الساخنة داخل الشبكة. في هذا البحث قمنا بعملية تصميم وتحقيق مقياس جودة وصلة اتصال جديد يسمى (HU_DLQ) Homs University Dynamic Link Quality يقوم باختيار المسار الأفضل من بين مجموعة من المسارات بالاعتماد على مجموعة من المقاييس كما يعمل أيضاً على موازنة الحمل بين عقد الشبكة والتخلص من مشكلة النقاط الساخنة، تمت عملية اختبار المقياس الجديد باستخدام نظام التشغيل ContikiOS والمحاكي Cooja كما تم تقييم أداء المقياس المقترح وذلك من أجل شبكات متباينة الكثافة ومختلفة الطوبولوجيا، أظهرت نتائج المحاكاة تفوق المقياس الجديد (HU_DLQ) على المقياس الأساسي (MORHF_ETX) Minimum

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

Rank Hysteresis Objective Function based on Expected Transmission Count من خلال تحسين استهلاك الطاقة وتقليل من رسائل التحكم وتحسين معدل تسليم الرزم.

الكلمات المفتاحية: الشبكات المنخفضة الطاقة والتي تحتوي ضياع، انترنت الأشياء، التابع الوظيفي (MORHF_ETX)، البرتوكول RPL، عدد الارسل المتوقع (ETX).

*طالب دكتوراه في قسم النظم والشبكات الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص.

**أستاذ مساعد - قسم النظم والشبكات الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص.

***أستاذ مساعد - قسم النظم والشبكات الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص

Design and implementation of a new link quality metric that achieves load balancing within IoT networks based on the RPL routing protocol.

Eng. Mohammad alholami* Dr. Akram almorei** Dr. Maher abbas***
Abstract

The next decade is expected to witness an increase in devices, mobile phones, sensors and actuators that communicate over the Internet to enhance human lifestyle. The design of the IoT network must ensure energy efficiency, speed and reliability to ensure the long-term sustainability of the IoT. Routing and balancing of data between nodes is one of the aspects that greatly controls the quality of service in IoT networks. The standard RPL protocol is designed to choose the best path based on a single link quality metric such as hop count or expected transmission number. Choosing unreliable paths leads to wasted network resources, increased data retransmission and unbalanced load, which forms what are known as hot spots within the network. In this research, we aim to design and implement a new link quality measure called (HU_DLQ) that selects the best path from a set of paths based on a set of measures. It also works to balance the load between network nodes and eliminate the problem of hot spots. The new measure was tested using the operating system (ContikiOS) and the simulator (Cooja). The performance of the proposed algorithm was also evaluated for networks of varying density and different topologies. The simulation results showed that the new measure (HU_DLQ) outperforms the basic algorithm within the (RPL) protocol, which is called (MORHF_ETX), by improving energy consumption, reducing control messages, and improving the packet delivery rate.

Keywords: Low Power and Lossy Networks (LLNs), Internet of Things (IOT), MORHF_ETX, RPL protocol, ETX link metric.

*PhD Student in the Department of Systems & Computer Networks of Faculty of Informatics –Homs University.

**Associated Professor in the Department of Systems & Computer Networks of Faculty of Informatics - Homs University.

1. مقدمة:

يتم ربط مليارات الأجهزة مثل عدادات الطاقة والمياه، والأجهزة المنزلية، والأجهزة القابلة للارتداء، والمستشعرات، والمشغلات بالانترنت لتكوين شبكة اتصالات ضخمة تُعرف بالانترنت الأشياء. من المتوقع أنه بحلول عام 2025، سيكون هناك أكثر من 27 مليار اتصال بالانترنت الأشياء. ستحدث انترنت الأشياء تحولاً في مجالات مراقبة الطقس، والمدن الذكية، والرعاية الصحية، والأتمتة الصناعية، والأتمتة المنزلية[1]. تتكون شبكات انترنت الأشياء وهو الجيل الجديد من شبكات الحساسات اللاسلكية من مستشعرات تعمل بالبطاريات تستخدم متحكمات دقيقة مثل (MSP430)[2] وتتمتع بسعة تخزين محدودة. بخلاف الشبكات التقليدية، يمكن أن تكون حركة المرور في هذه الشبكات واحد إلى واحد أو واحد إلى متعدد أو متعدد إلى واحد. كما أن شبكة الاتصالات هذه تعاني من عدة عيوب مثلاً عدم الاستقرار ومعدلات البيانات المنخفضة وفقدان حزم البيانات العالي ونطاق الإرسال المحدود وتغير طوبولوجيا الشبكة الديناميكي وبسبب هذه القيود العديدة على الأجهزة وتكنولوجيا الاتصالات الأساسية، يتم تصنيف شبكات انترنت الأشياء (Internet Of Things) IOT كشبكات منخفضة الطاقة وفاقة البيانات (Low Power Lossy Network) LNN ونظراً لأن بروتوكول التوجيه هو أحد دعائم هيكلية الشبكة وهو المسؤول عن قرارات التوجيه حيث أن القرارات الخاطئة سوف تسبب هدراً لموارد الشبكة صمم البروتوكول (RPL) من أجل الشبكات ذات الضياع والطاقة المنخفضة والشبكات التي تحتوي عدداً كبيراً جداً من الأجهزة الذكية، والقادرة على الاتصال بالمستخدمين عبر شبكة الانترنت حيث اهتم الباحثون بدراسة وتطوير البروتوكول (RPL) الذي يعتمد على الوظائف الغرضية (Objective Function) OF من أجل إنشاء المسارات داخل الشبكة وتشكيل طوبولوجيا الشبكة. تُستخدم مقاييس جودة وصلات الاتصال من أجل اختيار المسار الأقصر بين عقدتين وتقييم كلفته وتستخدم في داخل الوظائف الغرضية التي تعتمد على مقياس جودة وصلة اتصال وحيد فقط بحيث لا يغطي هذا المقياس جميع خصائص المسار بين العقد، كما تتطلب بعض تطبيقات (RPL) استخداماً لمقياس جودة وصلة اتصال واحد بينما تحتاج تطبيقات أخرى إلى تحديد مجموعة من مقاييس جودة وصلات الاتصال وذلك من أجل اختيار المسار الأمثل على سبيل المثال يمكن الاستفادة من مقياس جودة وصلة الاتصال (ETX)

لتجنب الروابط غير الموثوقة، ومن خلال هذا مقياس يستطيع البروتوكول اختيار المسار بناء على الروابط التي توفر مستوى محدد من الموثوقية واستخدام مقياس الطاقة المتبقية للعقدة كمقياس ثاني من أجل اختيار المسار بين العقد التي تحوي طاقة بنسبة معينة وتجنب العقد ذات الطاقة المنخفضة[14].

في هذا البحث، قمنا بإنشاء مقياس جودة وصلة اتصال جديد وقمنا بتسميته (HU_DLQ)، بحيث يقوم هذا المقياس باختيار أفضل مسار بين العقد بناءً على معايير جودة وصلات متعددة، تقوم هذه المعايير بأخذ نظرة شاملة عن طبيعة الشبكة وخصائصها من حيث (الطاقة المتبقية في العقدة، مقياس جديد مقترح يسمى معدل تمرير الرسائل، مقياس عدد الإرسال المتوقع (ETX) وعدد القفزات ومتوسط طول المسار). حيث ينجح هذا النهج في أخذ نظرة شاملة عن طبيعة الشبكة وتحسين موازنة الحمل داخل عقد الشبكة واختيار المسارات بين العقد التي تملك كميات طاقة أكبر والتقليل من عمليات إعادة الإرسال وفقد البيانات وتخفيف استهلاك الطاقة وإطالة عمر الشبكة ككل[5].

2. مشكلة البحث:

صمّم البروتوكول (RPL) من أجل الشبكات ذات الضياع والطاقة المنخفضة والشبكات التي تحتوي عدداً كبيراً جداً من الأجهزة الذكية، والقادرة على الاتصال بالمستخدمين عبر شبكة الانترنت، اهتم الباحثون بدراسة وتطوير بروتوكول (RPL) الذي يعتمد على الوظائف الغرضية (OF) من أجل انشاء المسارات داخل الشبكة وتشكيل طوبولوجيا الشبكة، على الرغم من أن بروتوكول (RPL) كان قادراً على حل المشاكل الرئيسية، إلا أنه لا يزال هناك بعض القيود التي تحتاج إلى معالجة على سبيل المثال يتخذ بروتوكول (RPL) القياسي قرار التوجيه معتمداً على مقياس جودة وصلة اتصال واحدة داخل الوظائف الغرضية مما يؤدي إلى اختيار مسارات غير مثالية وتحتوي على ضياعات وتشكيل ما يسمى بالنقاط الساخنة والتي يمكن تعريفها بـ(عقدة داخل الشبكة تعاني من الاختناق الشديد في مسألة تمرير الطرود وتبقى طوال الوقت مشغولة) مما يسرع في عملية هدر طاقة العقدة داخل الشبكة وبالتالي تقليل عمر الشبكة ككل.

3. الهدف من البحث:

نهدف في هذا البحث إلى تحسين أداء التوجيه واستهلاك الطاقة في شبكات الانترنت الأشياء وذلك من خلال تصميم وتحقيق مقياس جودة وصلة اتصال جديد يعتمد على الوظيفة الغرضية الجديدة (HU_DLQ_OF) والتي تعتمد على معيار جودة وصلة اتصال مركب والتي تأخذ بعين

الاعتبار أهم السمات التي تمثلها الشبكة ولتحقيق هذا المقياس قمنا باقتراح خوارزمية جديدة تقوم باختيار المسار الأفضل من بين مجموعة مسارات وذلك بالاعتماد على أربع معايير من مقدرات جودة وصلة الاتصال بحيث يتم اختيار المسار بناءً على القيمة الأعلى من حيث كمية الطاقة المتبقية والقيمة الأقل بالنسبة للمقياس معدل تمرير الرزم وللمقياس ETX داخل العقدة ، وعند تساوي كلفة المسارات نقوم بإخذ متوسط طول المسار ونختار القيمة الأقل لكلفة المسار بحيث نضمن أن هذا المسار الذي تم اختياره لا يحوي على مسارات جزئية تشكل عنق زجاجة (bottleneck) داخل الشبكة.

4. أدوات وطرائق البحث:

في هذا البحث تم استخدام حاسوب شخصي يحتوي على معالج من نوع (corei7-4th_generation) ومقدار ذاكرة عشوائية تقدر ب(8GB) ونظام تشغيل من نوع (Ubuntu) وتم تثبيت نظام التشغيل الخاص بأنترنت الأشياء (Contiki_Os) بالإضافة إلى المحاكي الخاص بنظام التشغيل الخاص بأنترنت الأشياء (Cooja Simulator).

5. الشبكات منخفضة الطاقة والتي يكون فيها ضياع (LLN):

مكنت انترنت الأشياء مليارات الكائنات المادية من جمع البيانات عن طريق المراقبة والتحسس والتحكم في البيئات وللقيام بذلك تُوزع تلك الكائنات (الكائنات الذكية أو الأجهزة) وتُدمج في شبكات فرعية مثل شبكات (WSN) Wireless Sensor Network وتُوصَل بالشبكة العالمية الأنترنت باعتباره العمود الفقري لـ (IOT) يتمثل الدور الرئيسي لشبكات (IOT) في تحسس البيانات وجمعها عبر الحساسات الموزعة (العقد) ومعالجتها ونقلها إلى عقدة مستقبل معلومات الشبكة، يطلق على هذا النوع من شبكات (IOT) والتي تتكون من عدد كبير من هذه الأجهزة المدمجة بالشبكات منخفضة الطاقة والتي يكون فيها ضياع (والخسارة) (LLN) [3] من المتوقع أن يكون هناك عدد هائل من التطبيقات المنزلية والتجارية والتي ستستخدم (LLN) كبنية تحتية لها لتحسين نوعية حياتنا. تتخصص شبكات (LLN) بتطبيقات حركة مرور البيانات المنخفضة، حيث تعد حركة المرور البيانات المنخفضة أمراً شائعاً جداً بالنسبة للتطبيقات المنزلية الذكية، والحوسبة المنتشرة في كل مكان حيث يلزم وجود دورة حياة طويلة جداً للحساسات. تدعم (LLN) ثلاثة أنواع من حركة المرور: نقطة إلى نقطة (بين الأجهزة داخل شبكة LLN) من نقطة إلى متعددة النقاط (بين الجهاز

المركزي والأجهزة الأخرى داخل شبكة LLN) أو عدة نقاط إلى نقطة تجميع (بين أجهزة داخل LLN إلى جهاز مركزي) يلزم التوجيه في الأجهزة الموجودة في (LLN) للوصول إلى بعضها البعض بسبب محدودية نطاق الإرسال، حيث يكون بروتوكول التوجيه هو المسؤول عن إدارة المسارات بين عقد الحساسات وإعادة توجيه الطرود على الطرق المكتشفة الأكثر كفاءة، تستخدم العقد جزء جهاز الإرسال والاستقبال لإرسال هذه الطرود أو استقبالها وهو راديو قصير المدى [14]. يعد الوسط الراديوي المستخدم من قبل أجهزة (LLN) قصير المدى عرضة لأخطاء البتات. حيث يكون الضياع في شبكة (LLN) له تأثير قوي على تصميم بروتوكول التوجيه، نظراً لهذه الأسباب تتمثل إحدى المشكلات الصعبة في شبكات (LLN) العثور على أفضل الطرق لتسليم البيانات، مما يعني وجود آلية توجيه فعالة للغاية للعثور على المسارات في الشبكة والاحتفاظ بها. تواجه آلية التوجيه صعوبات تتمثل بطبيعة قيود الموارد لعقد جهاز الحساس والضياع في الوسط الراديوي. يمثل استهلاك الطاقة أيضاً مشكلة كبيرة بالنسبة لأجهزة (LLN)، فمن المهم تشغيلها بمعدل بيانات منخفض جداً وطاقة إرسال تهدف إلى إطالة العمر الافتراضي للحساسات بشكل كبير [3].

6. بروتوكول التوجيه للشبكات منخفضة الطاقة والتي تحتوي فقدان لرزم (RPL):

يعد بروتوكول التوجيه هو المسؤول عن اتخاذ قرار التوجيه الذكي عن طريق اختيار المسار الأفضل بين عدة مسارات [4]. يعد التوجيه في شبكات (LLN) أحد التحديات الرئيسية ليس فقط بسبب الضياع الذي يكون في الوسط الراديوي والموارد المحدودة لعقد الحساسات ولكن أيضاً بسبب متطلبات التوجيه المختلفة [16] ، لذلك ركزت مجموعة عمل التوجيه عبر الشبكات منخفضة الطاقة والتي يكون فيها ضياع (ROLL) Routing Over Low Power and Lossy Networks على تصميم بروتوكول توجيه لينتاسب مع القيود المفروضة في هذه الشبكات حيث قامت مجموعة العمل (ROLL) [4] بتوجيه من قبل مجموعة مهندسي مهام الانترنت (IETF) Internet Engineering Task Force باعتماد وإنشاء البروتوكول (RPL) كبروتوكول توجيه معياري لشبكات (LLN) وتم توصيفه بالملف (RFC 6550) [3] كبروتوكول معياري يستند إلى الإصدار الجديد من بروتوكول عنونة الشبكات (IPv6) للعقد المقيّدة بالموارد وأصبح بروتوكول التوجيه الفعلي لشبكات (IOT) كما يدعم متطلبات التوجيه الدنيا من خلال بناء طوبولوجيا قوية عبر الروابط التي

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

يكون فيها ضياعات، كما تم تطويره وأصبح معياراً أساسياً لبروتوكولات توجيه البيانات وتم استخدامه في العديد من أنظمة التشغيل مثل Contiki OS، TinyOS، كبروتوكول توجيه قياسي [5].

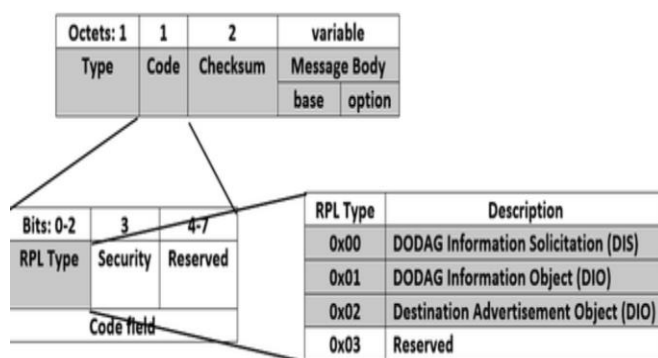
❖ رسائل التحكم في ال (RPL):

تم اقتراح أنواع جديدة من رسائل التحكم داخل البروتوكول (RPL) [3] لبناء طوبولوجيا شبكة (DODAG Destination Oriented Directed Acyclic Graph) وتتكون رسالة التحكم في (RPL) من [14]:

✓ ترويسة (ICMPV6 Internet Control Message Protocol) والتي تتألف من ثلاثة حقول (type، code، CRC checksum).

✓ متن الرسالة التي تحتوي على نص الرسالة وبعض الخيارات.

يتم تمييز نوع طرود التحكم بواسطة حقل الشيفرة (code) كما في الشكل رقم (1)، تتألف هذه الطرود من أربعة أنواع وهي:



الشكل رقم (1) بنية رسائل التحكم في بروتوكول (RPL)

• رسالة المعلومات (DIO) DODAG Information Object:

بداية يقوم جذر (DODAG) ببث رسالة (DIO) باتجاه هابط إلى عقد الجوار من أجل تمكين حركة المرور من نقطة إلى عدة نقاط في اتجاه تنازلي. تحتوي رسائل (DIO) على معرف جذر شجرة (DODAG_ID)، معرف مثيل (RPL_Instance_ID) ومجموعة من المقاييس منها

(مقاييس التوجيه، رتبة العقدة، الوظيفة الغرضية (OF)) تُرسل هذه الرسائل بشكل دوري مع رقم تسلسلي تراكمي من أجل بدء عملية اختيار الأب الأفضل للعقد [14].

• رسالة التماس المعلومات (DIS) DODAG Information Solicitation:

يتم استخدام رسالة التحكم (DIS) بواسطة أي عقدة لجلب رسائل تحكم من نوع (DIO) بشكل صريح من العقد المجاورة، وتُستخدم أيضا من أجل انضمام العقد إلى شجرة (DODAG) وفي حال لم تتمكن من تلقي رسائل من نوع (DIO) بعد فاصل زمني محدد مسبقا يتم إعادة بث هذه الرسالة من جديد [14].

• رسالة إعلان الوجهة (DAO) Destination Advertisement Object:

تخدم هذه الرسائل أساساً غرض نشر معلومات الوجهة في الاتجاه الصاعد نحو جذر الشبكة، تُستخدم هذه المعلومات لبناء المسارات الصاعدة بين العقد المرتبطة وجذر (DODAG) [14].

• الإقرار باستلام رسالة إعلان الوجهة (DAO-ACK):

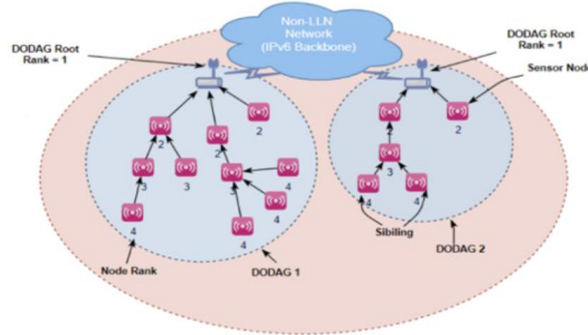
عند استلام رسائل تحكم من نوع (DAO)، يمكن للعقدة إرسال رسالة (DAO-ACK) إلى مرسل رسالة (DAO) للاعتراف باستلام الرسالة بشكل صحيح ولإقرار باستعداد العقدة للعمل كقفزة تالية باتجاه جذر [14].

❖ طريقة تشكيل الطوبولوجيا داخل البروتوكول (RPL):

يقوم RPL بإنشاء طوبولوجيا الشبكة من خلال مخطط يسمى رسم بياني موجه (Directed Acyclic Graph DAG) والذي يتكون من واحد أو أكثر من الرسم البياني والذي يدعى رسم بياني لا حلقي موجه إلى الوجهة [6] (DODAG)، يمثل كل (DODAG) شجرة توجيه تم إنشاؤها بواسطة عقدة الجذر (Root Node) والمعروفة أيضا باسم عقدة مستقبل معلومات الشبكة. عند البدء بتهيئة المسارات وتشكيل البنية، تبدأ عقدة الجذر بتكوين الشبكة وبناء (DODAG) عن طريق بث رسالة تحكم من نوع (DIO) كما يوضح الشكل رقم (2)، تحتوي هذه الرسالة على العديد من المعلومات مثل معلومات رتبة العقدة (للسماح للعقد بأخذ مواقعها في DODAG ولمنع حلقات التوجيه) [3]، نمط التشغيل والوظائف الغرضية (OF) والمقاييس التي تحوي معلومات حول البنية المطلوب تشكيلها، تقوم العقد المجاورة لهذه العقدة بالاستماع إلى هذه الرسالة وإعادة إرسالها إلى العقد المجاورة لها وتحديد فيما إذا كانت تريد الانضمام إلى هذه البنية

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

أم لا، وتقرر ذلك بناءً على قيود محددة، في الوقت الذي تصبح به هذه العقدة جزءاً من البنية، فسيكون لها مسار للأعلى نحو الجذر وعندها تحسب العقدة رتبته وتقوم بتحديث جدول الجيران الخاص بها، وتختار الأب الأفضل الذي سيتم استخدامه لإعادة توجيه الرسائل إلى جذر (DODAG) وعادة يكون الأب الأفضل هو الجار الذي يملك الرتبة الأدنى المحسوبة وفقاً للوظيفة الغرضية الصفرية (OF_0)، وبعدها تقوم هذه العقدة بإعادة إرسال رسائل التحكم إلى العقد المجاورة وبهذه الطريقة تستمر العملية حتى بناء كامل الشبكة [14]. يدعم بروتوكول (RPL) وجود عدة بنى ضمن الشبكة الواحدة والتي من الممكن أن تكون تابعة لعقدة جذر أو أكثر وتدعى هذه البنية (DAG Instance) ويمكن لعقدة الواحدة أن تنتمي إلى بنية واحدة أو أكثر بنفس الوقت حيث يكون لكل عقدة (ID Instance) واحد لكل بنية تنتمي إليها، تكمن فائدة هذه الخاصية في بناء مسارات حسب الهدف المطلوب من التطبيق مثلاً: مسارات للبيانات العادية يتجنب العقد التي تعمل على البطارية، ومسار لبيانات ذات الأهمية القصوى والذي يكون فيه التأخير بحدوده الدنيا وتسلق مسارات ذات ووثوقية عالية وطاقة عقدة كبيرة [21].



الشكل رقم (2) شجرة DODAG.

7. الوظائف الغرضية في بروتوكول الانترنت (RPL):

تعتمد عملية إنشاء شجرة (DODAG) على الوظائف الغرضية (OF) بشكل أساسي والتي تأخذها الشبكة بعين الاعتبار [6] ، يعتمد استخدام الوظيفة الغرضية بواسطة بروتوكول التوجيه (RPL) على ثلاث مكونات:

- المقياس المحدد، والذي يمثل المقياس الذي تستخدمه العقدة من أجل إنشاء المسار بين العقد.

- تكلفة المسار والذي يمثل جودة مسار الاتصال من بين مجموعة من المسارات بين العقد ويعتمد على المقياس المحدد.
- أفضل أب والذي يمثل أفضل مسار بين عقدة ابن وعقدة أب من حيث جودة وصلة الاتصال أو الطاقة المتبقية بالعقدة الاب أو العقدة الأقل من حيث كمية استقبال وأرسال البيانات. وضع الباحثون [7] [3] الوظائف الغرضية المعيارية في (RPL) كما يلي:

✓ الوظيفة الغرضية الصفرية (OF_0) :

تعتمد على عدد القفزات كمقياس لتكلفة المسار. يتم اختيار الأب المرشح الذي لديه أقل عدد من القفزات، مما يعني أن العقدة الأقرب إلى الجذر ستختار كأب مفضل. تلائم هذه الوظيفة الغرضية (OF0) الشبكات منخفضة الطاقة والتي يكون فيها ضياع لكن مقياس العقدة هذا قد يوفّر جودة روابط سيئة لأن العقد تختار المسار مع الحد الأدنى لعدد القفزات حتى لو كان غير موثوق [14]، وبالتالي تزداد عمليات إعادة الإرسال وضياع المزيد من الطرود [21].

✓ الوظيفة الغرضية (MORHF Based On ETX) :

تستخدم الوظيفة الغرضية (ETX_OF) مقياس جودة وصلة الاتصال والذي يسمى عدد الإرسال المتوقع (ETX) كمقياس توجيه لتحديد أفضل مسار من بين مجموعة مسارات، ويعرف بالحد الأقصى لعدد عمليات إعادة الإرسال اللازم حتى يتم إقرار لاستلام الطرود بنجاح [21]. يمثل هذا المقياس بمعنى آخر كمية الضياعات في المسار بحيث كلما كانت كمية الضياعات كبيرة ازدادت قيمة المقياس ويكون المسار سيئ وكلما انخفضت قيمة (ETX) يكون هناك ضياعات قليلة ويكون المسار أفضل، تحسب قيمة (ETX) بالعلاقة الأساسية التالية [14] [16]:

$$ETX = \frac{1}{dF * dR} \quad (1)$$

حيث:

df: تمثل نسبة التسليم وتعرف احتمال الطرود الواردة الناجحة من المصدر إلى الوجهة.

dr: تمثل نسبة التسليم العكسي وهو احتمال الإقرار الوارد الناجح من الوجهة إلى المصدر.

تُقاس جودة الرابط من حيث الموثوقية من خلال قيمة (ETX)، تشير كلفة (ETX) المنخفضة إلى أن الروابط في الشبكة أكثر موثوقية.

كل من الوظيفتين الغرضيتين (OF0,MORHF) لهما القصور نفسه وذلك بسبب اعتمادهم على مقياس واحد فقط من أجل اختيار المسار [5]. تقتصر هذه الوظائف إلى النهج الشمولي الضروري لاختيار المسار الأفضل من حيث التأخير الزمني أو تجاوز العقد ذات الازدحام العالي. للتغلب على ذلك وتوفير حلول تؤمن جودة الخدمة (QoS)، قدمت مواصفات (RPL) المرونة للمصمم لإدراج المزيد من المقاييس في عملية اتخاذ القرار وسنرى في الأجزاء اللاحقة كيف يمكن تحقيق ذلك.

8. الدراسة المرجعية:

نوضح في هذا الجزء من البحث دراسة استقصائية عن أبحاث تحسين أداء جودة وصلة الاتصال والتوابع الوظيفية المتوفرة في (RPL)، تهدف معظم الحلول إلى تحسين أداء التوجيه وفعالية الطاقة وإدارة موارد العقدة.

اقترح الباحثون في الورقة البحثية [7] وظيفة غرضية (COM_OF) جديدة تأخذ بعين الاعتبار مزيجاً من مقاييس (العمر المتوقع للعقدة، عدد الأبناء للعقدة، الطاقة المتبقية ومقياس جودة المسار EXT وذلك من أجل توزيع الحمل بين العقد المختلفة وتم مقارنة الحل المقترح مع كل من وظائف الهدف القياسية (OF0 و MRHOF) وباستخدام المحاكى (Cooja) أعطت النتائج انخفاض في استهلاك الطاقة وتحسين من معدل تسليم الرزم، تم استخدام مقياس عدد الأبناء حيث هذا المقياس لا يعكس كمية انشغال العقدة بجميع الحالات ولا يؤدي إلى موازنة دقيقة للحمل كما تم استخدام مقاييس أداء قليلة من أجل تقييم كامل الحل المقترح.

اقترح الباحثون [8] مقياس توجيه مركب للبروتوكول (RPL) يعتمد على مؤشر تفريغ البطارية وعدد الأبناء للعقدة من أجل اختيارها كأفضل أب للتوجيه البيانات وتم تجربة الحل باستخدام المحاكى (COOJA) وتم تقييم الحل عن طريق مقاييس فقط (عمر الشبكة، معدل تسليم الرزم) وعدد قليل من العقد ويقتصر الحل على عدم استخدام أي مقاييس جودة وصلات المسار بين العقد واستخدام مقياس عدد الأبناء الذي لا يعكس الصورة الحقيقية لكمية الحمل داخل العقدة.

في حين الورقة البحثية [9] أشارت إلى اقتراح وظيفية غرضية جديدة (RPL-FZ) تعتمد على المقاييس (الطاقة المتبقية بالعقدة، مقدار تأخير استلام الرزم (Delay)، عدد الارسل

المتوقع (ETX))، تم الجمع بين المقاييس باستخدام تقنية المنطق الضبابي للحصول مقدار درجة جودة لكل عقدة أب وبعدها يتم اختيار أفضل أب يحقق أعلى درجة جودة، تم محاكاة الشبكة باستخدام محاكي الشبكات (COOJA) ونظام التشغيل (Contiki) وأظهرت النتائج تحسين الأداء من ناحية تسليم الرزم وتوفير قليل بالطاقة بالمقارنة مع الوظيفة الغرضية (OF0 و MRHOF)، جميع الحلول التي تعتمد على المنطق الضبابي تحتاج إلى كمية استهلاك طاقة عالية مما يؤدي إلى زيادة استهلاك الطاقة.

في حين اقترح المؤلفون [11] وظيفة غرضية جديد للبروتوكول (RPL) تعتمد على اختيار العقد الأقل في استهلاك للطاقة كأفضل عقدة أب، وتم تقييم الأداء عن طريق عمر الشبكة ومعدل تسليم الرزم PDR (Packet Delivery Ratio) فقط.

الباحثون في الورقة البحثية [12] قدموا وظيفة غرضية تعتمد على (عدد القفزات والطاقة المتبقية داخل العقدة) وتم تقييم الأداء باستخدام المحاكى (COOJA) بالنسبة لاستهلاك الطاقة ومعدل تسليم الرزم PDR فقط.

اقترح الباحثون [13] وظيفة غرضية تسمى (WSM - OF) تعتمد على أربعة معايير في اختيار الأب الأفضل من بين مجموعة الآباء المرشحين بالاعتماد على (عدد الإرسال المتوقع ETX، الطاقة المتبقية، عدد الأولاد، مقياس جودة وصلة الاتصال LQL)، باستخدام المحاكى (Cooja) أظهرت النتائج تحسن عمر الشبكة مع تقليل من استهلاك الطاقة بشكل طفيف.

المؤلفون في [14] أشاروا إلى إنشاء وظيفة غرضية تستخدم عملية كفاءة الطاقة وتعتمد على مقياس مركب (الطاقة المتبقية بالعقدة ومؤشر تفريغ البطارية وعدد الإرسال المتوقع ETX) أدت هذه الأعمال إلى تحسين عمر الشبكة وتقليل من استهلاك الطاقة.

في حين تناولت الورقة البحثية [16] عملية تصميم تابع وظيفي مركب يعتمد على ثلاثة مقاييس في عملية اختيار العقد الأب الأفضل وذلك عن طريق اختيار مجموعة من العقد المرشحة لتصبح عقدة أب بالاعتماد على مقياس قوة الإشارة المستقبلية (RSSI) والتي تكون ضمن عتبة معينة ثم يتم اختيار العقدة الأفضل بالاعتماد على مقياس عدد الإرسال المتوقع (ETX) وكمية الطاقة المتبقية أظهرت النتائج تحسن في عملية تسليم الرزم (PDR) والتقليل من زمن تأخر زمن الوصول البيانات (Latency) .

9. المقياس المقترح الجديد (HU_DLQ) لتحسين التوجيه وموازنة الحمل:

من أجل تحسين أداء التوجيه وحل مشكلة توزيع الحمل داخل عقد الشبكة تم اقتراح مقياس جديد يأخذ بعين الاعتبار كل من المقاييس التالية (حالة الطاقة المتبقية للعقد، مقياس جديد يسمى معدل تمرير الرزم، المقياس (ETX) ، وعدد القفزات ومتوسط طول المسار ويمكن حساب المقاييس المستخدمة في المقياس المقترح بالشكل التالي:

✓ **الطاقة المتبقية (Residual Energy):** يتم حساب الطاقة المتبقية للعقدة المرشحة حتى تكون عقدة أب بالعلاقة رقم (2) [7] [1]:

$$E_{res}(n) = E_{ini}(n) - E_{cons}(n) \quad (2)$$

حيث:

E_{res} : يشير إلى الطاقة المتبقية للعقدة (n).

E_{ini} : يشير إلى الطاقة الابتدائية للعقدة (n).

E_{cons} : يشير إلى كمية الطاقة المستهلكة للعقدة (n).

✓ **معدل تمرير الرزم (packet forwarding ratio):**

لضمان توازن الحمل بين عقد الشبكة وعدم اختيار عقد ذات حمل عالي تشكل نقاط ساخنة داخل الشبكة وتؤدي إلى تدمير البيانات وزيادة كمية إعادة الإرسال واستنزاف طاقة العقدة بشكل كبير، تم اقتراح مقياس جديد يسمى معدل تمرير الرزم داخل العقدة ويقصد به كمية الرسائل التي تقوم باستقبالها عقدة ما وتقوم بإعادة تمرير هذه الرزم إلى عقدة أخرى ذات مستوى أعلى داخل الشبكة حيث تكون هذه العقدة مشغولة دائماً بعملية استقبال وتمرير رزم البيانات طوال الوقت ويمكن حساب معدل تمرير الرزم للعقدة المرشحة كعقدة أب خلال فترة زمنية معينة بالعلاقة المقترحة رقم (3).

$$packet\ forwarding = (\sum packet_{recv} - \sum packet_{send}) / 100 \quad (3)$$

$packet_{send}$: يشير كمية البيانات المرسله من قبل العقدة نفسها والتي تم تحسبها.

$packet_{recv}$: يشير إلى كمية البيانات المستقبلية من قبل العقدة التي تشكل أبناء لهذه العقدة.

✓ **عدد الإرسال المتوقع (ETX):** من أجل جودة عالية للروابط وتجنب فقدان الطرود، يتم حساب عدد الإرسال المتوقع بالعلاقة [16] (4).

$$ETX = \frac{1}{dF * dR} \quad (4)$$

حيث:

df: تمثل نسبة التسليم وتعرف احتمال الطرود الواردة الناجحة من المصدر إلى الوجهة.

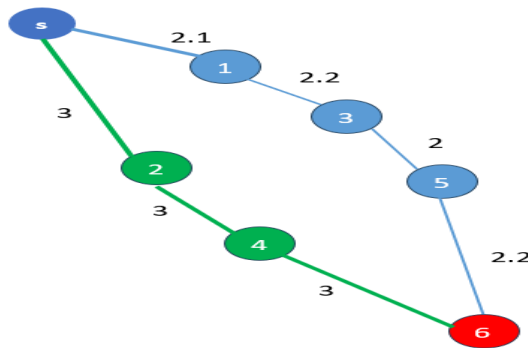
dr: تمثل نسبة التسليم العكسي وهو احتمال الإقرار الوارد الناجح من الوجهة إلى المصدر.

✓ **عدد القفزات (Hop):** يتم حساب عدد القفزات التي تسلكها رسالة ال (DIO) حتى الوصول للعقدة التي يتم حساب تكلفة مسارها، حيث يتم زيادة عدد القفزات بمقدار (1) عند كل عقدة تسلكها رسالة ال (DIO).

✓ **متوسط تكلفة المسار (Average Path Cost):** نقوم بحساب متوسط تكلفة المسار وذلك عند تساوي مسارين لهما نفس التكلفة ولكن أحد المسارات يحوي على مسارات جزئية تحوي على ضياعات عالية ويمكن حساب متوسط طول المسار وذلك بتقسيم تكلفة المسار على عدد القفزات التي تشكل هذا المسار كما بالعلاقة رقم (5)[18].

$$Average_{path_{cost}} = \frac{path_{cost}}{hop_{count}} \quad (5)$$

بفرض لدينا المساريين في الطوبولوجيا كما يوضح الشكل (3) وتحتاج العقدة (6) حساب تكلفة مسارها فيكون تكلفة المسارات:



الشكل رقم (3) طوبولوجيا شبكة ذات مسارات متساوية.

- $P1_{pathCost} = 2.1 + 2.2 + 2 + 2.2 = 9$
- $P2_{pathCost} = 3 + 3 + 3 = 9$

نلاحظ تساوي قيم المسارات للعقدة رقم (6) مع العلم أن المسار الأول (p1) أفضل من المسار الثاني (p2) حيث المسار الثاني يحوي على وصلات جزئية تحوي ضياعات أكثر فعند هذه الحالة نقوم بحساب متوسط طول المسار فيكون:

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

- $P1_cost=9/4=2.3$
- $P2_cost=9/3=3$

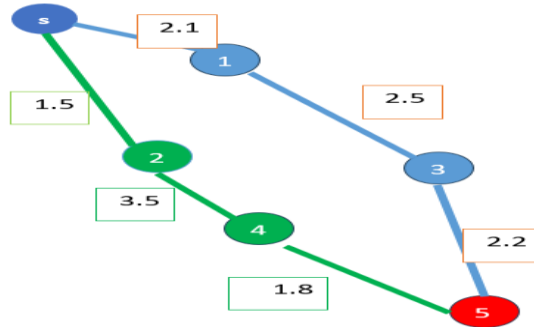
فنقوم باختيار المسار ذو المتوسط الأقل وهو (p1).

✓ **معامل تخفيف للمسارات (Setamol):** يقوم هذا المعامل بعملية مراقبة وكشف التغيرات العالية التي تطرأ على المسارات الجزئية والتي تشكل عنق زجاجة داخل المسار، حيث نقوم بالرجوع إلى هذا المعامل في حالة عند تساوي متوسط تكلفة المسارات حيث نختار المسار الذي يحوي قيمة أعلى من المسار الآخر وعند تساوي قيم هذا المعامل مع المسار الآخر نختار أحد المسارات، ونقوم بحساب قيم المتحول بالعلاقة المقترحة رقم (6).

معامل التخفيف (Setamol) = (تكلفة المسار الجزئي للعقدة) - (تكلفة المسار الجزئي القادم إلى هذه

العقدة) (6)

نفرض لدينا الطوبولوجيا كما يوضح الشكل (4) وتحتاج العقدة (5) حساب تكلفة مسارها



الشكل (4) يبين طوبولوجيا شبكة ذات معدل مسارات متساوي

- ✓ $P1_cost=2.1+2.5+2.2=6.8$
- ✓ $P2_cost=1.5+3.5+1.8=6.8$

نلاحظ تساوي تكلفة المسارات نقوم عندها بحساب متوسط طول المسار فيكون:

- ✓ $P1_cost=6.8/3=2.2$
- ✓ $P2_cost=6.8/3=2.2$

عند هذه الحالة نلاحظ تساوي متوسط تكلفة المسارات فنقوم بالرجوع إلى معامل تخفيف المسار من أجل اختيار المسار الأفضل حيث نلاحظ أن المسار الثاني يحوي على مسار جزئي بين العقدة (3 و 4) يشكل عنق زجاجة، فنقوم بحساب معامل تخفيف المسار:

- ✓ $P1_setamol=(2.2-(2.5-(2.1-0)))=1.9$
 - ✓ $P2_setamol=(1.8-(3.5-(1.5-0)))=-1.2$
- Best_path=max(p1_setamol,p2_setamol)=p1

نلاحظ أن قيمة المسار (p1) أعلى من قيمة المسار (p2) عندها نختار المسار الأفضل وهو P1

10. توصيف خطوات المقياس (HU_DLQ):

يمكن توصيف خطوات الأساسية للمقياس الجديد كما يلي:

1. يبدأ جذر الشبكة (node sink) ببث رسائل (DIO) بشكل دوري إلى جميع العقد المجاورة وذلك من أجل تشكيل المسارات الهابطة وتحتوي هذه الرسالة على المقاييس التي تم اقتراحها من حيث طاقة العقدة ومعدل تمرير الرزم والمقياس ETX وعدد الففقات.
2. تتعرف كل عقدة على العقد المجاورة لها وعلى جذر الشبكة من خلال رسائل DIO التي تم استقبالها دورياً ويتم قراءة وتحديث معلومات هذه الرسالة مثل (مقدار الطاقة المتبقية للعقدة، المقياس ETX، معدل تمرير الرزم للعقدة وعدد الففقات) وبعدها يتم حساب رتبة العقدة، وإعادة بث هذه الرسالة من جديد للعقد المجاورة.
3. تقوم العقدة بحساب رتيبها بالاعتماد على العلاقة (7،8،9) وذلك من أجل أخذ موقعها وعدم تشكيل حلقات داخل طوبولوجيا الشبكة وبعدها تقوم بإنشاء مجموعة من الآباء المرشحين من العقد المجاورة وذلك من أجل تشكيل مسارات صاعدة إلى عقدة الجذر [8].

$$RANK(N) = RANK(PN) + RankIncrease \quad (7)$$

$$RankIncrease = MinHopRankIncrease * step \quad (8)$$

$$Step = \alpha_1 * ETX(N, PN) + \alpha_2 * fw_ratio \quad (9)$$

$$\alpha_i \in R^+, i \in \{1,2\} \text{ and } \sum_{i=1}^3 \alpha_i = 1$$

حيث:

$RANK(N)$: يشير إلى رتبة عقدة المرسل.

$RANK(PN)$: يشير إلى رتبة العقدة الأب.

$Min_Hop_Rank_Increase$: يشير إلى مقدار الحد الأدنى للزيادة في الرتبة والقيمة الافتراضية هي (256).

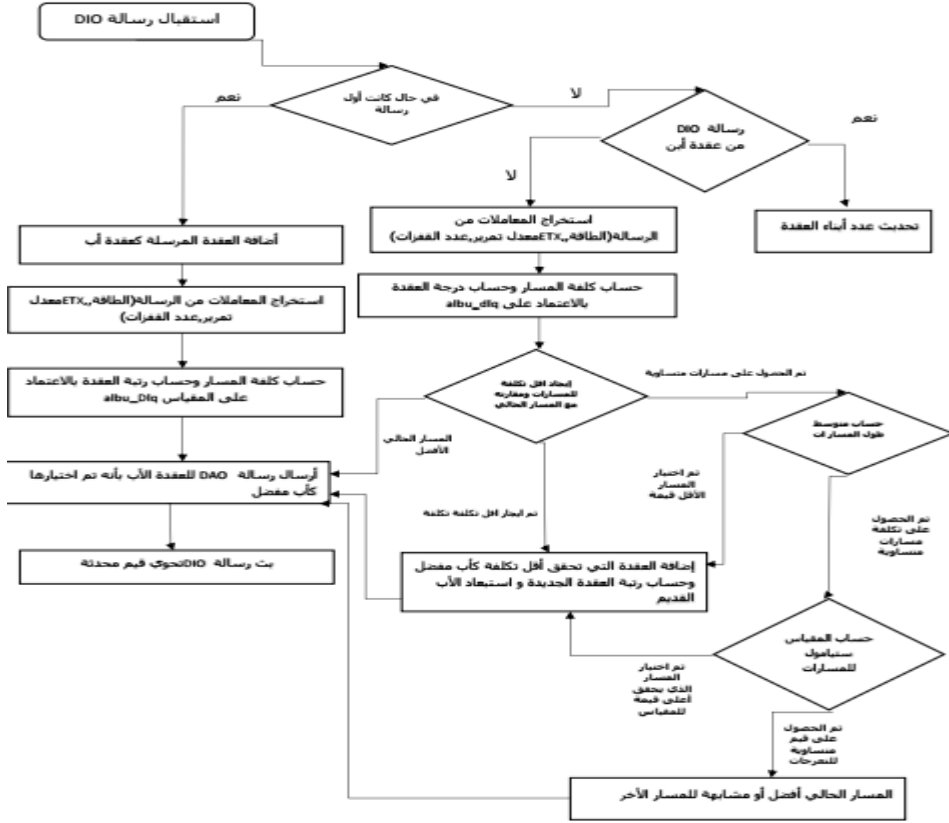
$Step$: يشير إلى مقدار الخطوة في زيادة الرتبة.

α_i : يشير إلى مقدار التتبع في الزيادة بقيم مجموعها يساوي (1).

4. تقوم كل عقدة بعملية حساب تكلفة المسارات الصاعدة وذلك من أجل اختيار عقدة أب من بين مجموعة من الآباء المرشحين واختيار المسار الذي يحقق أقل تكلفة اعتماداً على المقاييس التي تم ذكرها سابقاً وذلك بالاعتماد على أقل قيمة للمقياس ETX وأقل قيمة لمعدل تمرير الرزم وكمية طاقة متبقية ضمن عتبة معينة.

5. في حال تساوي تكلفة المسارات لعقدتين أب يتم أخذ متوسط تكلفة المسار الأدنى وذلك عن طريق تقسيم تكلفة المسار على العدد القفزات القادم منها هذا المسار وذلك من أجل اختيار المسار الذي لا يحوي أي وصلات جزئية تحتوي على ضياعات عالية.
6. في حال تساوي متوسط طول المسار لعقدتين مرشحتين يتم اختيار المسار الذي يحوي القيمة الأعلى لمعامل تخفيف التعرجات للمسار (Setamol).
7. بعد عملية اختيار عقدة الأب، ستتم إعادة توجيه البيانات إلى العقدة الأب الأفضل ويتم إرسال رسالة من نوع (DAO) للعقدة التي تم اختيارها وانتظار رسالة (ACK).
8. تلقي طرد إقرار بالاستلام (Ack) من طبقة MAC، تعني أن الطرد تم استلامه بنجاح بواسطة العقدة الأب.
9. إذا لم يتم تلق أي إقرار بالاستلام من طبقة MAC، هذا يعني أن الطرد قد ضاعت ويجب إعادة تكرار العملية بالانتقال مرة أخرى إلى الخطوة 2.

❖ المخطط الصندوقي لطريقة عمل المقياس المقترح:



11.مقاييس الأداء:

في هذا البحث استخدمنا خمسة مقاييس أداء قياسية وذلك من أجل تقييم أداء جودة اختيار المسار داخل البروتوكول (RPL) في الخوارزمية المقترحة، وهذه المقاييس هي:

✓ وقت تقارب الشبكة (Network Convergence Time):

تحتاج العقد في شبكة انترنت الاشياء إلى تشكيل طوبولوجيا من أجل التواصل بين العقد. لذلك فإن وقت إعداد الشبكة هو مقياس حاسم يجب تقييمه لأي بروتوكول توجيه، يتم تعريف وقت التقارب في (RPL DAG) على أنه مقدار الوقت الذي تحتاجه جميع العقد التي يمكن الوصول إليها (من حيث الراديو) في الشبكة للانضمام إلى طوبولوجيا (DAG). يجب اعتبار هذا التقارب بمثابة وقت التقارب الأولي في شبكة RPL ذات العقد الثابتة من أجل الحصول على وقت التقارب في شبكة RPL، نحدد وقت إرسال DIO الأول من عقد العميل وانضمام DIO الأخير إلى DAG.

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

يتم الحصول على وقت التقارب عن طريق طرح وقت انضمام DIO الأخير من وقت إرسال DIO الأول إلى DAG كما بالعلاقة رقم (10) [10].

$$\text{Convergence}_{\text{Time}} = (\text{Last_DIO_Join_DAG}) - (\text{First}_{\text{DIO}}_{\text{Send}}) \quad (10)$$

✓ حمل رسائل التحكم (Control Traffic Overhead) للشبكة:

يشمل هذه المقياس على رسائل DIO و DIS و DAO التي يتم إنشاؤها بواسطة كل عقدة، ومن الضروري حصر مقياس الحمل مع مراعاة الموارد النادرة في شبكات (LLN)، يتحكم بروتوكول (RPL) في رسائل التحكم الزائدة عن الحاجة من خلال استخدام مؤقتات التدفق. الهدف من هذا المقياس هو تحليل حمل الخوارزمية المقترحة المذكورة في استخدام موارد العقدة ويمكن حساب حمل رسائل التحكم عن طريق العلاقة رقم (11) [10]:

$$\text{control_traffic_overhead} = \sum_{i=1}^n \text{DIO}(I) + \sum_{i=1}^m \text{DIS}(k) + \sum_{i=1}^z \text{DAO}(i) \quad (11)$$

✓ استهلاك الطاقة (Energy Consumption):

من أجل حساب استهلاك الطاقة سوف نستخدم آلية نظام Power trace المتاحة في نظام التشغيل (ContikiOS) وهو نظام لتحديد ملامح الطاقة على مستوى الشبكة للشبكات اللاسلكية منخفضة الطاقة التي تقدر استهلاك الطاقة لوحدة المعالجة المركزية ونقل الحزم والاستماع. تحتفظ هذه الآلية بجدول للمدة الزمنية التي كان فيها مكون مثل وحدة المعالجة المركزية وجهاز الإرسال اللاسلكي قيد التشغيل. بناءً على هذا الحساب، يتم حساب الطاقة المستهلكة بالميلي واط عن طريق العلاقة رقم (12,13) حيث فولت التغذية (3 فولت) كما يلي [1][13]:

$$\text{power(mW)} = \text{Current} * \text{voltage} \quad (12)$$

ومن أجل حساب التيار المستهلك نستخدم العلاقة رقم (13) [7][13].

$$\text{Current} = \frac{tx * \text{current}_{tx} + Rx * \text{current}_{rx} + \text{Cpu} * \text{Current}_{cpu} + \text{Cpu}_{idel} * \text{Current}_{cpuidel}}{\text{Rtimer_ARCH_Second}} \quad (13)$$

حيث قيم:

TX: زمن تشغيل المرسل الراديوي في العقدة بوحدة أجزاء الساعة (نبضة).

RX: زمن تشغيل المستقبل الراديوي في العقدة بوحدة أجزاء الساعة (نبضة).

CPU: زمن تشغيل وحدة المعالجة المركزية بوحدة أجزاء الساعة (نبضة).

cpu_{idel}: زمن تشغيل وحدة المعالجة المركزية في وضع الطاقة المنخفض بوحدة أجزاء الساعة (نبضة).

Rtimer_ARCH_Second: وهو عدد نبضات المعالج في الثانية الواحدة (32768 نبضة).

نُبين في الجدول رقم (1) قيم التيارات المأخوذة من النشرة التعريفية للعقدة (Sky Mote) [15] وهي كما يلي:

الجدول (1) قيم التيارات المأخوذة من النشرة التعريفية للعقدة.

نوع الوحدة	كمية الاستهلاك
$current_{cpu}$	1.8 Ma
$current_{rx}$	21.8 Ma
$current_{tx}$	19.5 mA
$current_{idel}$	54.5 μ A

✓ زمن تأخر وصول الحزمة (Packet Latency):

يتم تعريف زمن تأخر وصول الحزمة بأنه مقدار الوقت الذي تستغرقه حزمة البيانات المرسلّة من عقدة ما للوصول إلى العقدة الهدف أو عقدة مصب الشبكة (Sink Node) حيث نقوم بحساب متوسط زمن الوصول لجميع حزم البيانات في الشبكة من جميع العقد إلى عقدة الجذر [13] ويتم حساب قيمة المقياس عن طريق العلاقة رقم (14) وحساب معدل زمن وصول الحزمة بالعلاقة [15] [9] [10].

$$Total_Latency = \sum_{i=1}^n (recv_time(i) - send_time(i)) \quad (14)$$

$$Averagy_Latency = Total_Latency / Total_{packets_received} \quad (15)$$

حيث (n) هو العدد الإجمالي للحزم المستلمة بنجاح. يتم توفير معلومات التوقيت حساب متوسط زمن حيث الوصول نقسم إجمالي زمن الوصول من المعادلة (14) على إجمالي عدد الحزم المستلمة في عقدة الجذر كما تبين العلاقة رقم (15).

✓ متوسط معدل تسليم الحزمة (PDR Average):

ويتم تعريفه على أنه عدد الحزم المستلمة في عقدة مصب الشبكة (sink node) إلى عدد الحزم المرسلّة إلى مصب الشبكة من بقية العقد [9] حيث نقوم بأخذ متوسط قيمة (PDR) لجميع الحزم التي تم استلامها بنجاح في عقدة المصب (Sink) ولحساب متوسط نسبة تسليم الحزمة (PDR)، نقيس عدد الحزم المرسلّة من جميع العقد إلى عقدة المصب (Sink) ونقسمها على عدد الحزم المستلمة بنجاح في عقدة المصب (Sink) كما تبين بالعلاقة رقم (16) [7] [10].

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

$$\text{Average PDR} = \left(\frac{\text{Total}_{\text{packet}_{\text{received}}}}{\text{Total}_{\text{packet}_{\text{send}}}} \right) * 100 \quad (16)$$

12. القسم العملي:

تركز معظم الأبحاث على المحاكاة بدلاً من التجارب وذلك بسبب توفر العتاد البرمجي وسهولة الوصول إليه، على الرغم من أن التقييم المعتمد على عمليات المحاكاة لا يعكس الواقع لأنه لا يأخذ بعين الاعتبار مجموعة من القيود مثل تغير المناخ والعوائق الطبيعية والتفاعلات البشرية ولكنه الخيار الوحيد المتوفر في كثير من الأحيان للطلبة والباحثين.

✓ نظام التشغيل (ContikiOS):

يعتبر نظام التشغيل (ContikiOS) والذي تم تطويره من قبل المعهد السويدي لعلوم الكمبيوتر البرنامج المستخدم على نطاق واسع في شبكات IoT وWSN وذلك لأنه مفتوح المصدر وهو نظام التشغيل الأكثر استخداماً في مجال الأبحاث الخاصة بشبكات الحساسات اللاسلكية وشبكات انترنت الأشياء كما يستخدم (Contiki) في العديد من الأنظمة، مثل المراقبة الصناعية وأنظمة الإنذار ومراقبة المنزل عن بعد ومراقبة الإشعاعات. يؤمن نظام التشغيل (ContikiOS) والمكتوب بلغة البرمجة (C) التي تتميز بسرعة تنفيذها للأوامر حلاً للمتحكمات صغيرة الحجم والمنخفضة الكلفة والطاقة إمكانية الاتصال بالإنترنت، كما يُستخدم نظام التشغيل (Contiki) عدة أنواع من المعالجات مثل سلسلة (Atmel ، Atmega ، Texas Instruments) كما أنه يدعم كل من بروتوكول عنونة الشبكة IPv4 و IPv6 ، كما يدعم عدة مقاييس حديثة لاسلكية موفرة للطاقة مثل بروتوكول (RPL) وشبكات (LowPan6) ويدعم أيضاً تقنية (Rime Stack) وهي تقنية لتخفيف حمل طبقات الاتصالات على شبكات الحساسات، حيث تحوي طبقات أقل من طبقات الاتصالات التقليدية الموصّفة ببروتوكول TCP/IP وتكون الطبقات فيها أبسط وذات ترويسة أصغر ببضعة بايتات [19].

✓ محاكي الشبكات (COOJA):

سوف نستخدم محاكي الشبكات (COOJA) وذلك بسبب كونه المحاكي الأكثر استخداماً في الأبحاث وأجراء التجارب في شبكات منخفضة الطاقة بنسبة تصل إلى (61%) [20] وهو محاكي قائم على لغة (java) ويعتمد على نظام التشغيل (ContikiOS) بشكل أساسي ويمكن اعتباره أداة هجينة تُستخدم كأداة محاكاة للعقد الافتراضية (Simulator) وأداة محاكاة للأجهزة الحقيقية

(Emulator) بمعنى آخر أن أدائه أقرب للواقع لأنه يقوم بتشغيل حقيقي للأجهزة الموجودة في الشبكة مما يجعل النتائج التي نحصل عليها دقيقة أكثر وتحاكي الواقع [21].
✓ تهيئة الشبكة:

في هذه التجارب سوف نستخدم عقد من نوع (Sky mote) والتي تتميز بمعدل نقل عالي للبيانات (250kb/s) وذاكرة وصول عشوائي تقدر (10kb)، قمنا بأثناء مجموعة من الملفات كما قمنا بالتعديل على بعض هذه الملفات بالبروتوكول (RPL) وإضافة توابع جديدة على الملفات ومن هذه الملفات نذكر:

(rpl.h, rpl-icmp6.c, rpl-conf.h, HU_DLQ.C, energest.c, rpldag.c, uip6.c)

ولإجراء التجارب قمنا باستخدام طوبولوجيا عشوائية ومنتظمة مؤلفة من نموذج تطبيق UDP Contiki يتألف من ملفين الملف (udp-client.c) وهو الملف الذي تم تنزيله على جميع عقد الشبكة (client node) بحيث تقوم العقدة (client node) بإرسال رسالة نصية ضمن فاصل زمني محدد إلى موجه الشبكة أو جذر الشبكة (DODAG ROOT) والذي يحوي على الملف (udpsrvr.c) عند وصول المعلومات المرسل من عقدة الزبائن (client node) يقوم موجه الشبكة بإعلام العقدة التي تم استلام المعلومات منها بأن المعلومات تم استلامها بشكل صحيح ومن أجل اختبار الخوارزمية المقترحة تم إجراء التجارب على شبكات مختلفة الكثافة :

✓ منخفضة الكثافة (20 عقد).

✓ متوسطة الكثافة (40 عقدة).

✓ عالية الكثافة (60 عقدة).

ولإدخال الضياع في الوسط اللاسلكي استخدمنا النموذج (UDGM) والذي يعرف الضياع بناء على المسافات النسبية بين العقد في الوسط الراديوي [10] ، يمكننا من خلاله ضبط نسبة الإرسال والاستقبال، حيث قمنا بالمحافظة على نسبة الإرسال (TX-100%) (لأننا نهتم بالضياع عند طرف المستقبل، بينما قمنا بتغيير نسبة الاستقبال (RX) والتي تُعرف بأنها نسبة نجاح استقبال الطرود عند المستقبل على نسبة (RX 80%)، وتشير إلى الضياع في الوسط اللاسلكي، كما تم تحديد نطاق الإرسال على 30 متر ونطاق التداخل على 50 متر، تم استخدام كل من التطبيق الملحق مع المحاكي (Cooja) المسمى (Contiki Simulation Script Editor) لقياس وقت المحاكاة

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

وإيقاف المحاكاة بعد الوقت المحدد، ينشئ هذا البرنامج المساعد ملف سجل (COOJA_log file) لجميع مخرجات المحاكاة والتي سنقوم بتحليلها في نهاية المحاكاة باستخدام برنامج نصي مكتوب بلغة (Perl) وباستخدام برنامج (MS Excel) نقوم برسم المخططات البيانية. قمنا بتشغيل المحاكاة لمدة (40) دقيقة لكل سيناريو وتم إعادته لمدة أربع مرات وأخذ متوسط النتائج، يوضح الجدول رقم (2) المعاملات المستخدمة في التجارب المختلفة [10].

جدول رقم (2) ضبط بارامترات التجارب

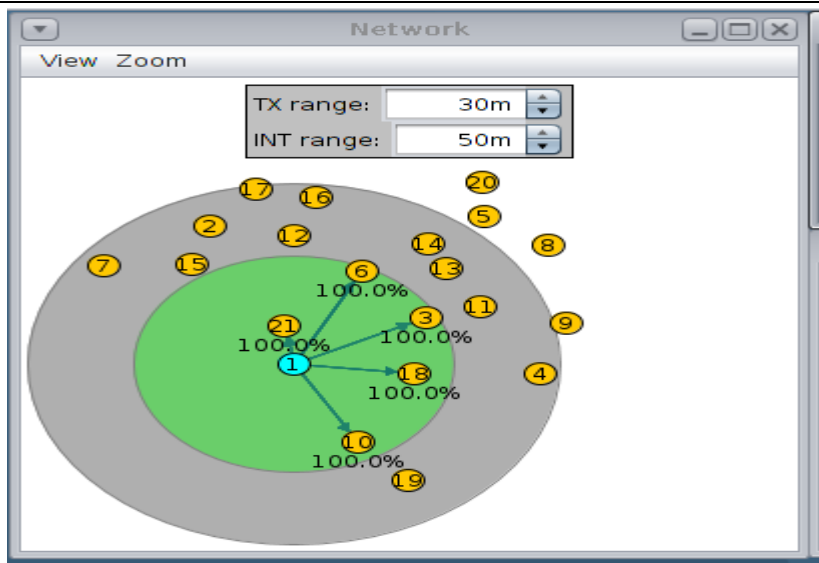
Parameters	Description
No of Nodes	20,40,60
Simulation Model	UDGM (Unit Disk Graph Model)
Simulation time	40 Mins.
Data Rate	250kps
Node Arrangement Random	Random, Grid
Operating System	Contiki2.7
Simulator	Cooja
OF	OF_0MORHF_ETX, HU_DLQ_OF
DIO Min	12
DIO Doublings	8
Radio Duty-Cycling Interval	16HZ
RX Ratio	80%
TX Ratio	100%
TX Range	30m
Interference Range	50m
Packet Analyzer	Perl code

13. النتائج ومناقشتها:

قمنا في هذه السيناريو بتقييم أداء المقياس المقترح الجديد مع الوظائف الغرضية الأساسية (OF_0,MORHF_ETX) وذلك من أجل عدد عقد مختلف (20,40,60) مع طوبولوجيا عشوائية النتائج:

✓ السيناريو الأول:

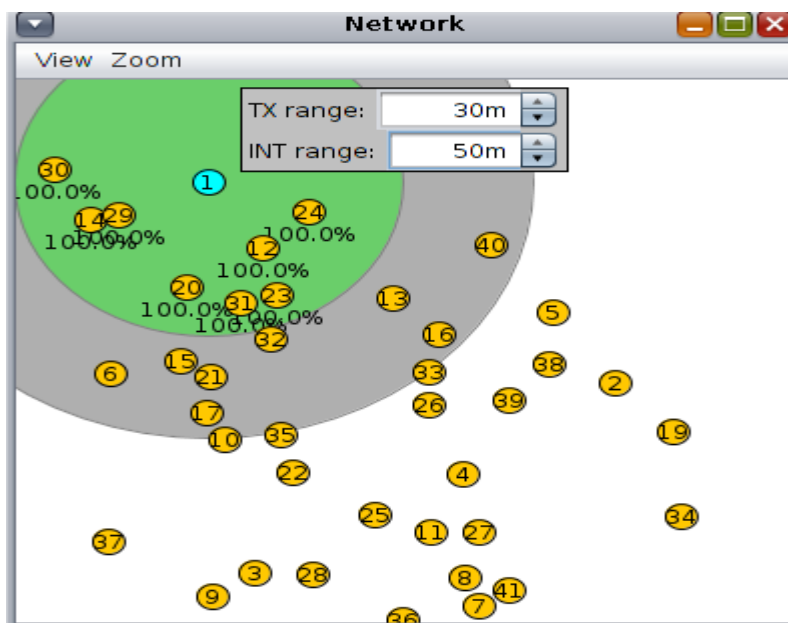
في هذا السيناريو قمنا بمقارنة الأداء من أجل عدد عقد منخفض (20عقد) مع توزيع عشوائي كما يوضح الشكل (5) بعدها قمنا بتسجيل النتائج من أجل تحليلها باستخدام مقياس الأداء.



الشكل رقم (5) يبين توزع العقد العشوائي (20 عقد)

✓ السيناريو الثاني:

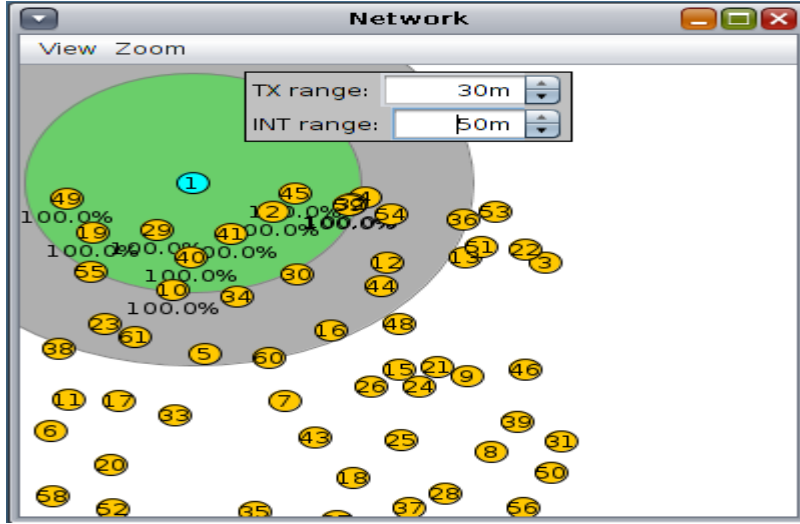
قمنا بهذه السيناريو بزيادة عدد العقد من أجل كثافة متوسطة (40 عقدة) وتوزع عشوائي كما يبين الشكل (6) وبعد الانتهاء قمنا بتسجيل النتائج من أجل تحليلها باستخدام مقاييس الأداء.



الشكل رقم (6) توزع العقد العشوائي (40 عقدة).

✓ السيناريو الثالث:

قمنا بإعادة السيناريو مع زيادة عدد العقد من أجل كثافة مرتفعة (60 عقدة) وتوزع عشوائي كما يبين الشكل (4) وذلك من أجل زيادة كثافة الرسائل المتناقلة بين العقد وبعد الانتهاء، قمنا بتسجيل النتائج من أجل تحليلها باستخدام مقاييس الأداء.



الشكل رقم(6) توزع العقد المنتظم والعشوائي (60عقدة).

14. مناقشة النتائج:

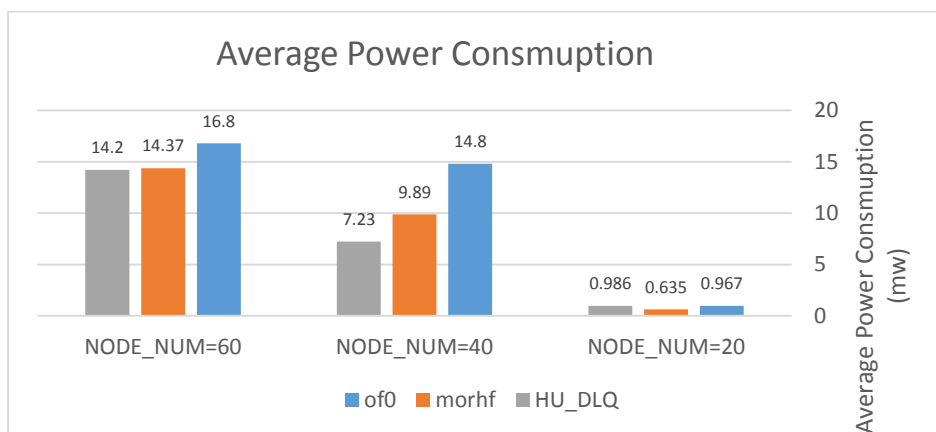
✓ معدل استهلاك الطاقة:

تم تحليل النتائج وتسجيل قيم معدل استهلاك الطاقة كما يبين الجدول (2) والشكل رقم (7). بالنسبة لمقياس معدل استهلاك الطاقة كانت كمية استهلاك الطاقة كما يوضح الشكل رقم(7) نلاحظ بالنسبة للمقياس الجديد كانت كمية استهلاك أكبر من المقياس (MORHF_ETX,OF_0) من أجل عدد عقد (20) عقدة وسبب هذه الزيادة هي اعتماد الخوارزمية المقترحة على أكثر من عملية مقارنة واستدعاء توابع وتعريف متحولات جديد مما يزيد في عملية استهلاك طاقة معالج العقدة بكمية ضئيلة، إما عند زيادة عدد العقد (40عقدة وأكثر) نلاحظ تفوق المقياس المقترح في عملية التقليل من استهلاك الطاقة وذلك بسبب اختيار مسارات ذات وثوقية عالية وتحوي معدل تمرير رزم اقل كما أن زيادة كثافة الشبكة يؤدي إلى عمل مقياس متوسط طول المسار بشكل أكبر من الشبكات متوسطة

الكثافة مما يخفض من عملية اختيار وصلات ذات معدل ضياع عالي ويخفف هذا النهج عملية إعادة ارسال البيانات وتأخيرها وتقليل من عملية استهلاك الطاقة .

الجدول رقم(2) معدل استهلاك الطاقة

HU_DLQ	morhf	of0	Average Power Consumption (mw)
0.986	0.635	0.967	NODE_NUM=20
7.23	9.89	14.8	NODE_NUM=40
12.2	14.37	16.8	NODE_NUM=60



الشكل رقم (7) يوضح معدل استهلاك الطاقة لجميع العقد.

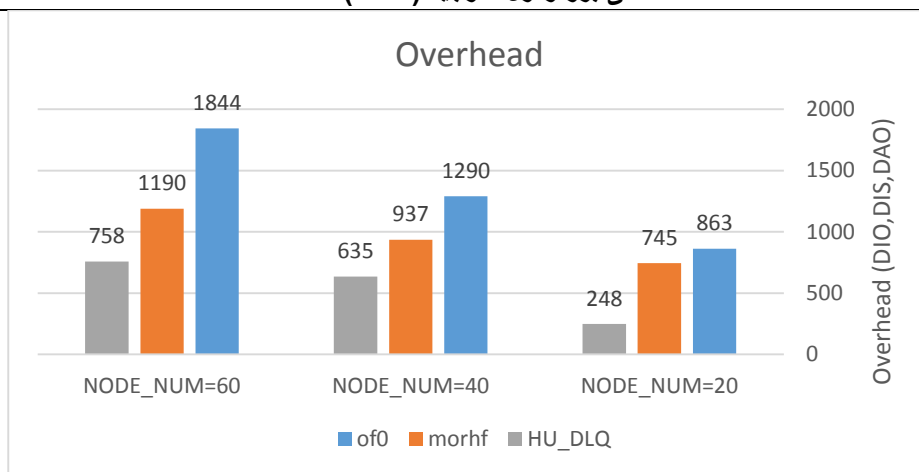
✓ مقياس حمل رسائل التحكم:

بالنسبة لمقياس حمل رسائل التحكم، كانت النتائج كما يبين الشكل (8) والجدول رقم(3) نلاحظ تفوق المقياس الجديد في تقليل من عدد رسائل التحكم (DIS,DIO,DAO) وذلك من أجل عدد عقد مختلف وتوزع عشوائي وذلك يعود بسبب اختيار المقياس مسارات أكثر وثوقيه واختيارها عقدة ذات طاقة كبيرة وتجاوز العقد ذات حركة البيانات الكثيفة والطاقة المنخفضة والتي تؤدي إلى خروج بعض العقد عن الخدمة بسبب نفاذ طاقتها وزيادة عدد رسائل التحكم بسبب تغير الطوبولوجيا الدائم.

الجدول رقم(3)

HU_DLQ	MORHF	OF0	OVERHEAD (DIO,DIS,DAO)
248	745	863	NODE_NUM=20
635	937	1290	NODE_NUM=40
758	1190	1844	NODE_NUM=60

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)



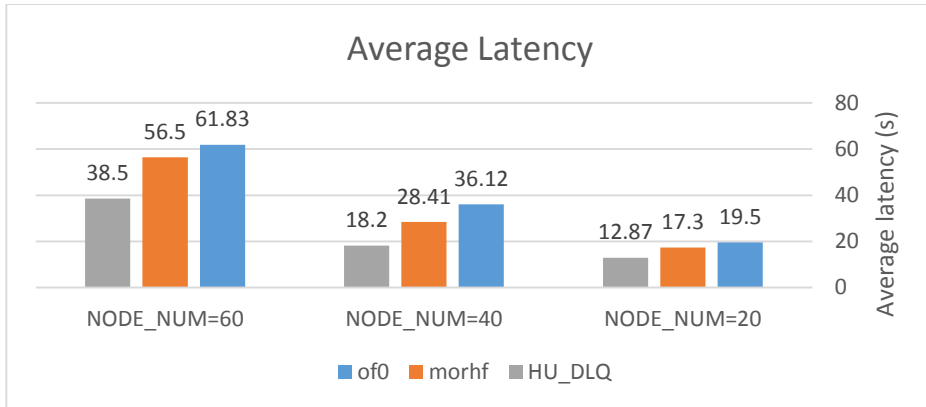
الشكل رقم (8) يبين حمل رسائل التحكم.

✓ متوسط تأخر الرزم:

بالنسبة لمقياس متوسط معدل تأخر الرزم كانت النتائج كما يبين الشكل (9) والجدول رقم (4)، نلاحظ تفوق المقياس المقترح في التقليل من معدل زمن التأخير في عملية تسليم البيانات وذلك من أجل عدد عقد كبير أكبر من (20 عقدة) ، وذلك يعود إلى تجاوز العقد ذات معدل تمرير الرسائل العالي والذي يزيد من زمن انتظار الرسائل في ذاكرة العقدة حتى وصولها إلى العقدة الجذر، كما أن عملية اختيار مسارات لا تحوي على ضياعات عالية وذلك عن طريق أخذ متوسط طول المسار تؤدي إلى تقليل من زمن الوصول الرسائل إلى العقد الجذر.

الجدول رقم (4) متوسط تأخر الرزم

<i>HU_DLQ</i>	<i>morhf</i>	<i>of0</i>	<i>Average latency (s)</i>
12.87	17.3	19.5	NODE_NUM=20
14.2	28.41	36.12	NODE_NUM=40
26.5	56.5	61.83	NODE_NUM=60



الشكل رقم (9) يبين متوسط معدل تأخر الرزم.

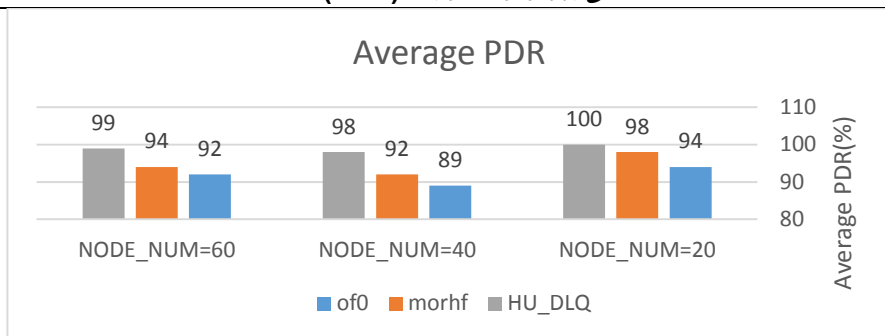
✓ مقياس متوسط تسليم الرزم:

بالنسبة لمقياس متوسط تسليم الرزم كانت النتائج كما يبين الشكل (10) والجدول رقم (5)، نلاحظ تفوق المقياس الجديد في عملية تسليم الرزم بالمقارنة مع كل الوظائف الغرضية (OF_0, MORHF_ETX) وذلك من أجل عدد عقد مختلفة وذلك بسبب وثوقيه اختيار المسارات وتجاوز المسارات التي تحوي نقاط ساخنة عن طريق مقياس معدل تمرير الرزم وتجاوز العقد التي تشكل عنق زجاجة عن طريق مقياس تخفيف المسارات (Setamol) مما يخفف من عملية ضياع الرزم

الجدول رقم (5) متوسط تسليم الرزم

HU_DLQ	MORHF	OF0	AVERAGE PDR(%100)
100	98	94	NODE_NUM=20
98	92	89	NODE_NUM=40
99	94	92	NODE_NUM=60

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)



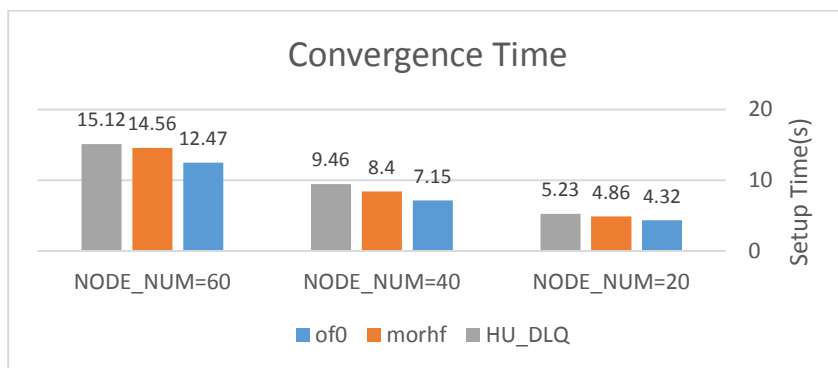
الشكل رقم (10) يوضح متوسط معدل تسليم الرزم.

✓ مقياس وقت تقارب الشبكة:

بالنسبة لمقياس وقت تقارب الشبكة كانت النتائج كما يبين الشكل (11)، نلاحظ أن المقياس الجديد يستهلك وقت أكبر بمقدار (2-3 ثانية) بالمقارنة مع الوظائف الغرضية الأساسية للبروتوكول (RPL) وذلك بالنسبة لعدد عقد مختلفة، ويرجع ذلك إلى اعتماد المقياس الجديد على عدة مقاييس داخل رسائل الانضمام إلى الطوبولوجيا الشبكة (DIO) مما يساهم ذلك في عملية زيادة في وقت التقارب داخل الشبكة بشكل كامل.

الجدول رقم (6) وقت تقارب الشبكة

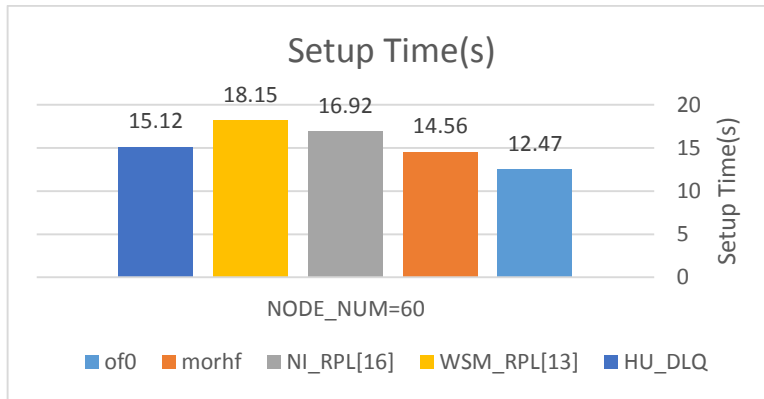
HU_DLQ	morhf	of0	Setup Time(s)
5.23	4.86	4.32	NODE_NUM=20
9.46	8.4	7.15	NODE_NUM=40
15.12	14.56	12.47	NODE_NUM=60



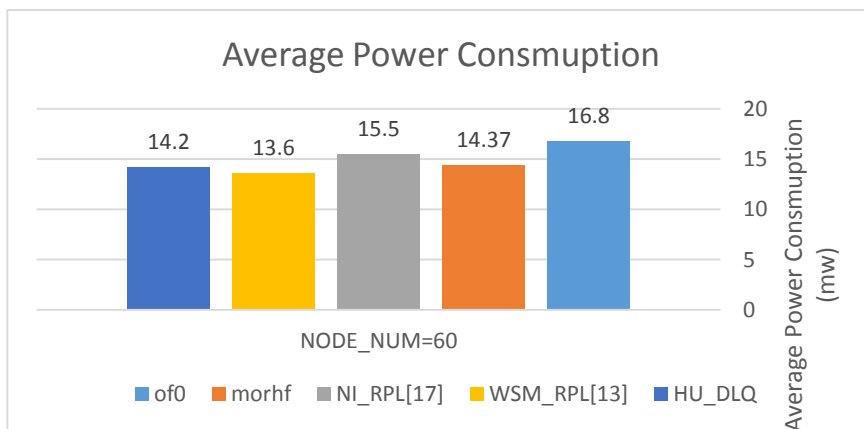
الشكل رقم (11) يبين وقت تقارب الشبكة

15. مقارنة المقياس الجديد (HU_DLQ) مع بعض الدراسات السابقة:

في هذه الفقرة قمنا بعمل مقارنة مع بعض الدراسات السابقة التي اقترحت تقنيات انشاء اتواع وظيفية ومقاييس جودة وصلة اتصال مركبة وتم قياس إدائها بنفس مقاييس الأداء المقترحة وذلك من أجل عدد عقد ثابت (60عقدة) وطوبولوجيا عشوائية، ومن هذه الاعمال الدراسة المرجعية رقم [13] التي اقترحت أنشاء تابع وظيفي مركب (WSM-OF) يعمل على موازنة الحمل داخل عقد الشبكة كما اقترحت الدراسة المرجعية رقم [16] انشاء تابع وظيفي يسمى (NI-RPL) مركب يعمل على تحسين جودة اختيار وصلات الاتصال داخل عقد الشبكة فكانت النتائج كما توضح الاشكال رقم (12,13,14,15,16)

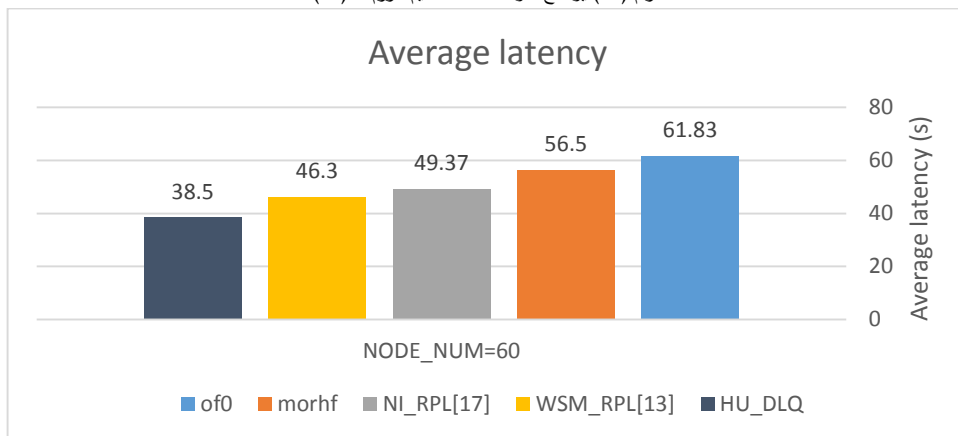
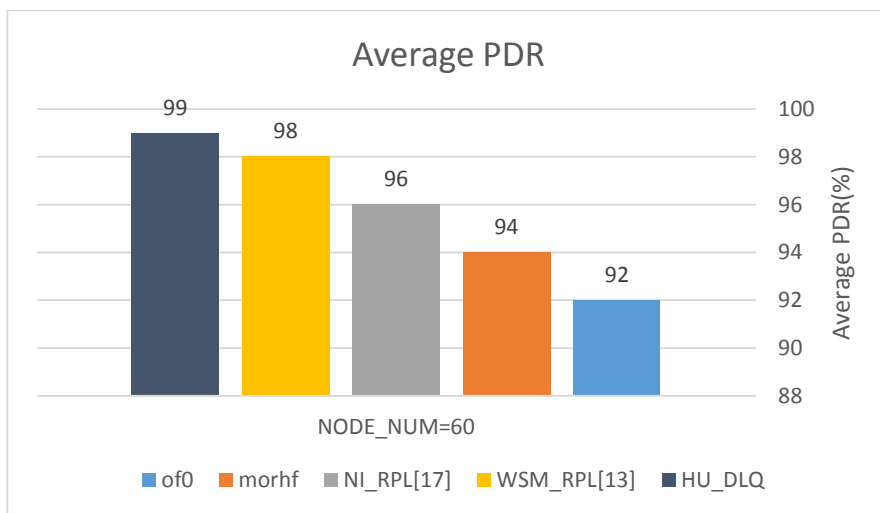
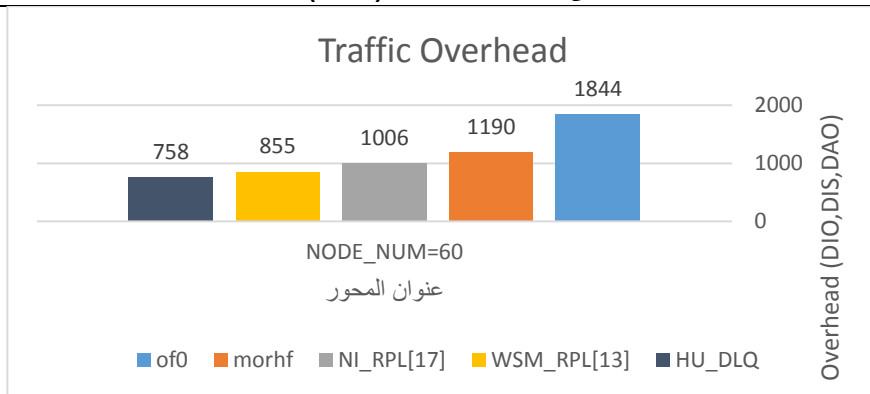


الشكل (12) يبين وقت تقارب الشبكة



الشكل رقم (13) يوضح معدل استهلاك الطاقة (60) عقدة.

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)



- ✓ بالنسبة لمقياس وقت تقارب الشبكة منا يبين الشكل (12) نلاحظ ان كل من التوابع الوظيفية (MORHF,OF0) اعطى قيم أفضل من التوابع الوظيفية المركبة بزمين يقدر ب (2) ثانية بشكل وسطي وذلك بسبب طبيعة تصميم هذه التوابع التي تختار مسارات الاتصال دون عملية فحص بعض المقاييس مثل (كمية الطاقة المتبقية، معدل تمرير الرزم، عدد القفزات، عدد أبناء هذه العقدة) وبالنسبة للتوابع الوظيفية المركبة (NI_RPL,WSM_RPL,HU_DLQ) فهناك تفاوت زمني بسيط وهذا بسبب اعتماد هذه التوابع الوظيفية على فحص واختيار المسارات والذي بدوره يأخذ زمن يقدر ب (ميلي ثانية)، كما نلاحظ أن المقياس المركب الجديد اعطى زمن تقارب أفضل بالمقارنة مع التوابع الوظيفية المركبة (NI-RPL,WSM_RPL) ويعود هذا الشيء الى دقة اختيار مقاييس وصلة الاتصال التي توفر في استخدام اكثر من مقياس والتي تأخذ نظرة شمولية لمسارات توجيه البيانات داخل الشبكة .
- ✓ بالنسبة لمعدل استهلاك الطاقة كما يبين الشكل (13) اعطى المقياس نتائج أفضل من حيث عملية توفير الطاقة بمقارنة مع كل من التوابع الوظيفية (MORHF,OF0,NI_RPL) ويعود ذلك التوفير في عملية اختيار مسارات ذات موثوقية عالية ولا تعاني من ازدحام مما يخفف من عملية إعادة ارسال البيانات الذي بدوره يخفف من كمية الطاقة داخل الشبكة ككل.
- ✓ بالنسبة لمقياس حمل رسائل التحكم كما يوضح الشكل (14) اعطى المقياس المقترح (HU_DLQ) أفضل النتائج في عملية التقليل من رسائل التحكم والتي تسبب زيادة استهلاك الطاقة وزيادة حركة البيانات داخل الشبكة مما يؤدي الى زيادة الضياعات وزيادة تأخر وصول رزم البيانات داخل الشبكة حيث ساعد المقياس المقترح في عملية اختيار مسارات ذات ثباتية عالية من حيث كمية الطاقة ومعدل إعادة تمرير الرزم مما يساعد في عملية التخفيف في عملية تبديلات المسارات والذي بدوره يقلل من رسائل التحكم.
- ✓ من الشكل (15) نلاحظ أن المقياس الجديد (HU_DLQ) أعطى أفضل النتائج من حيث عملية معدل تسليم الرزم داخل الشبكة وذلك بمقارنة مع بقية التوابع الوظيفية ويعود ذلك إلى سلوك المقياس في عملية تحسين تسليم البيانات والتقليل من ضياعها عن طريق اختيار المسارات الاقصر التي تحوي على أكبر كمية طاقة وأقل عدد من القفزات كما أن المقياس (متوسط تكلفة المسار) يساعد في عملية التقليل من اختيار المسارات الأطول

تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات انترنت الاشياء التي تعتمد على بروتوكول التوجيه (RPL)

✓ نلاحظ من الشكل (16) أن المقياس الجديد (HU_DLQ) أعطى نتائج أفضل من حيث تأخر وصول الرزم وذلك بالمقارنة مع كل من التتابع الوظيفية (MORHF,OF0,NI_RPL,WSM_RPL) ويعود ذلك إلى تجنب المسار التي تعاني من حركة بيانات عالية مما يخفف ضياع وتأخر البيانات داخل العقد.

16. الاستنتاجات والتوصيات:

يعد التوجيه عاملاً مهماً في شبكة انترنت الأشياء لأنه ينفذ قرارات توجيه الطرود حيث أن اختيار المسارات السيئة في التوجيه يسبب المزيد من عمليات إعادة الإرسال وإهدار الموارد، بناءً على نتائج المحاكاة وتحليل الأداء نجد أن المقياس الجديد (HU_DLQ) يقدم أداءً متفوقاً بشكل عام بمقارنة مع الوظائف الغرضية الأساسية في عملية اختيار جودة وصلة الاتصال (MORHF_ETX,OF_0)، يأتي هذا التفوق من عملية اختيار المسار الأفضل بالنسبة لكمية الطاقة و معدل تمرير الرزم الأقل للعقدة الأب وبالتالي اختيار أفضل المسارات لتوجيه الطرود وتوزيع الحمل والطاقة بشكل أفضل داخل عقد الشبكة.

17. التوصيات المستقبلية:

على الرغم من التحسينات التي أجريت في البحث، إلا أننا يمكننا طرح المزيد من التوقعات لتحسين هذا البحث في المستقبل منها:

1. طرح مقاييس جديدة تدخل في تصميم الخوارزمية من أجل تحسين الأداء
2. دراسة أداء الخوارزمية من أجل مقاييس مختلفة وفي سيناريوهات مختلفة مع زيادة عدد العقد.
3. دراسة وتقييم أداء الخوارزمية من أجل عقد متنقلة في الشبكة.
4. تغيير طريقة اختيار الأب المفضل عن طريق طرح خوارزمية جديدة تتبع طريقة التعلم الآلي (Machine Learning).

18. جدول الاختصارات:

Ipv6 Routing Protocol For Low-Power And Lossy Networks.	RPL	بروتوكول توجيه الشبكات المنخفضة الطاقة والتي تحتوي ضياعات
ALBAATH UNIVERSITY DYNAMIC LINK QUALITY	HU_DLQ	مقياس جودة الوصلة الديناميكي الجديد باسم جامعة حمص
Minimum Rank Hysteresis Objective Function	MORHF	الوظيفية الغرضية لاستعادة استقرار الرتبة
Expected Transmission Count	ETX	عدد الإرسال المتوقع
Internet Of Things	IOT	الإنترنت الأشياء
Low Power And Lossy Network	LNN	الشبكات منخفضة الطاقة والتي يكون فيها ضياع
Objective Function	OF	تابع وظيفي
ZERO Objective Function	Of0	الوظيفة الغرضية الصفرية
Wireless Sensor Network	WSN	شبكات الحساسات اللاسلكية
Routing Over Low Power And Lossy Networks	ROLL	مجموعة عمل التوجيه عبر الشبكات منخفضة الطاقة والتي يكون فيها ضياع
Internet Engineering Task Force	IETF	فريق مهام هندسة الإنترنت
Destination Oriented Directed Acyclic Graph	DODAG	رسم بياني لا حلقي موجه إلى الوجهة
Internet Control Message Protocol	ICMPV6	بروتوكول رسائل التحكم الإصدار السادس
DODAG Information Object	DIO	رسالة المعلومات
DODAG Information Solicitation	DIS	رسالة التماس المعلومات
Destination Advertisement Object	DAO	رسالة إعلان الوجهة
Directed Acyclic Graph	DAG	رسم بياني موجه
Packet Delivery Ratio	PDR	معدل تسليم الرزم
Unit Disk Graph Model	UDGM	نموذج رسم بياني لاقتطاع القرص الوحدوي
Ipv6 Routing Protocol For Low-Power And Lossy Networks.	RPL	بروتوكول توجيه الشبكات المنخفضة الطاقة والتي تحتوي ضياعات

19.المراجع باللغة الإنكليزية:

- [1] N. Gupta, A. Pughat, and V. Sharma, 2021-A critical analysis of RPL objective functions in internet of things paradigm,” *Peer Peer Netw Appl*, vol. 14, no. 4, pp. 2187–2208, Jul. doi: 10.1007/s12083-021-01180-9.
- [2] MSP430,<https://www.ti.com/design-development/embedded-development/msp430-mcus.html>.
- [3] “rfc6550”, <http://www.rfc-editor.org/info/rfc6550>.
- [4] S.-Iong. Ao, Oscar.Castillo, Craig. Douglas, D. Dagan. Feng, and Alexander. Korsunsky,2016-Routing Issues in Internet of Things,*International MultiConference of Engineers and Computer Scientists : IMECS 2016*, March.
- [5] I. Kechiche, I. Bousnina, and A. Samet, 2020-A review of RPL objective function based enhancement approaches,” *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 5, pp. 201–211, doi: 10.25046/AJ050525.
- [6] M. Qasem, A. Al-Dubai, I. Romdhani, B. Ghaleb, and W. Gharibi, 2016- A new efficient objective function for routing in Internet of Things paradigm,” in *2016 IEEE Conference on Standards for Communications and Networking, CSCN 2016*, Institute of Electrical and Electronics Engineers Inc., Dec,doi: 10.1109/CSCN.2016.7785168.
- [7] K. Venugopal and T. G. Basavaraju,2022-ACombined Metric Objective Function for RPL Load Balancing in Internet of Things,” *International Journal of Internet of Things*, vol., no. 1, pp. 22–31, doi: 10.5923/j.ijit.20221001.02.
- [8] S. Sennan and S. Palanisamy,2017-Composite metric based energy efficient routing protocol for internet of things,” *International Journal of Intelligent Engineering and Systems*, vol. 10, no. 5, pp. 278–286, doi: 10.22266/ijies2017.1031.30.
- [9] S. Kuwelkar and H. G. Virani,2018-Design of an Efficient RPL Objective Function for Internet of Things Applications,[Online]. Available: www.ijacsa.thesai.org
- [10] Ali H. A performance evaluation of rpl in contiki: A cooja simulation based study. School of Computing, Blekinge Institute of Technology. 2012 Oct.
- [11] J. P. Yunis and D. Dujovne,2014-Energy efficient routing performance evaluation for LLNs using combined metrics,” in *2014 IEEE Biennial Congress of Argentina (ARGENCON)*, IEEE, Jun. pp. 741–746. doi: 10.1109/ARGENCON.2014.6868581.
- [12] A. Hassan, S. Alshomrani, A. Altalhi, and S. Ahsan, 2016-Improved routing metrics for energy constrained interconnected devices in low-power and lossy networks, *Journal of Communications and Networks*, vol. 18, no. 3, pp. 327–332, Jun., doi: 10.1109/JCN.2016.000048.
- [13] P. S. Subramani and S. Bojan,2023-Weighted Sum Metrics – Based Load Balancing RPL Objective Function for IoT, *Annals of Emerging Technologies in Computing*, vol. 7, no. 2, pp. 35–55, Doi: 10.33166/AETiC.2023.02.004.
- [14] Mardini, K., & Abdulsamad,2023-Designing a Novel RPL Objective Function Testing RPL Objective Functions Performance (Dissertation). Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-121754>
- [15] sky mote datasheet, <http://www.crew-project.eu>.
- [16] Dharmalinga,A, Latha P. Additive metric composition-based load aware reliable routing protocol for improving the quality of service in industrial internet of things. *Int. Arab J. Inf. Technol.*. 2023 Nov 1;20(6):954-64.
- [18] Sanmartin P, Rojas A, Fernandez L, Avila K, Jabba D, Valle S. Sigma routing metric for RPL protocol. *Sensors*. 2018 Apr 21;18(4):1277.

المراجع باللغة العربية:

- [19] العمر منال، 2022- تحسين أداء خوارزمية التقطير في إنترنت الأشياء، مجلة جامعة البعث سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، المجلد 44، العدد 8، صفحة 65-92.
- [20] الكلزلي طريف، 2023- تحسين زمن تبديل العقد في بروتوكول RPL ضمن نظام التشغيل Contiki، مجلة جامعة البعث سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، المجلد 54، العدد 32، صفحة 73-92.
- [21] الشعار رهنف، 2020- تطوير خوارزمية لتحسين أداء التوجيه في إنترنت الأشياء، رسالة ماجستير، المعهد العالي للعلوم التطبيقية والتكنولوجيا، الجمهورية العربية السورية.

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)

بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء

م. محمد الهلامي* د. أكرم مرعي** د. ماهر عباس***

كلية الهندسة المعلوماتية- جامعة حمص

ملخص

يجب أن يضمن تصميم شبكة إنترنت الأشياء كفاءة الطاقة والسرعة والموثوقية لضمان الاستدامة طويلة الأجل لإنترنت الأشياء. اقترحت مجموعة عمل (IETF Roll) بروتوكول توجيه (RPL) للشبكات منخفضة الطاقة والتي تعاني من الخسارة كمعيار توجيه فعلي مع مراعاة استدامة شبكات إنترنت الأشياء المقيدة. تم تصميم بروتوكول RPL القياسي لاختيار أفضل مسار بناءً على مقياس جودة رابط واحد مثل عدد القفزات (Hop) أو عدد الإرسال المتوقع (ETX)، وبالتالي فإن اختيار المسارات غير الموثوقة يؤدي إلى إهدار موارد الشبكة وزيادة حالات إعادة إرسال البيانات والحمل غير المتوازن بين العقد.

نهدف في هذا البحث إلى استخدام إمكانيات تقنية التعلم الآلي الخاضع للإشراف لتحسين عملية التوجيه داخل البروتوكول (RPL)، قمنا بدمج التعلم الآلي الخاضع للإشراف داخل بروتوكول التوجيه المخصص للشبكات منخفضة الطاقة والتي تعاني من الخسارة (RPL) أدى هذا العمل إلى نسخة مقترحة من البروتوكول يسمى HUML_RPL يعتمد على تقنية التعلم الآلي الخاضع للإشراف من أجل اختيار المسارات وتحسين عملية التوجيه. النسخة المقترحة يستخدم نموذج شجرة القرار (Decision Tree) للتدريب على مجموعة بيانات التدريبية التي تم توليدها ويتنبأ باحتمالية الوصول بنجاح إلى أفضل عقدة أب، تم دمج هذا النموذج داخل كل العقد في الشبكة وذلك من أجل اختيار المسار الأفضل لتوصيل الحزم بفعالية. تم اختبار أداء النسخة المقترحة

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء

من البروتوكول (HUML_RPL) في محاكي الشبكات Cooja لنظام التشغيل Contiki Os وتقييم أدائه لشبكات ذات كثافات متفاوتة. أظهرت نتائج المحاكاة أن النسخة المقترحة (HUML_RPL) يتفوق على الخوارزمية الأساسية داخل بروتوكول (RPL)، والتي تسمى (MORHF_ETX,OF0)، من خلال تحسين استهلاك الطاقة وتقليل رسائل التحكم وتحسين معدل تسليم الحزم وخفض زمن الوصول في الشبكة.

الكلمات المفتاحية: البروتوكول RPL، إنترنت الأشياء، التابع الوظيفي (MORHF)، التعلم الآلي الخاضع للإشراف

*طالب دكتوراه في قسم النظم والشبكات الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص.
**استاذ مساعد - قسم النظم والشبكات الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص.
***استاذ مساعد - قسم النظم والشبكات الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص

Improving Link Quality Selection in Routing Protocol (RPL)

Using Supervised Machine Learning Technologies in IOT Networks

Eng. Mohammad alholami* Dr. Akram almorei** Dr. Maher abbas***

Abstract

The design of the iot network must ensure energy efficiency, speed and reliability to ensure the long-term sustainability of the IoT. The IETF ROLL working group proposed the internet protocol version_6 (ipv6) Routing protocol for Low power Lossy Networks (RPL) as de-facto routing standard keeping in mind the sustainability of constrained IoT networks.

The standard RPL protocol is designed to choose the best path based on a single link quality metric such as hop count or expected transmission number so Choosing unreliable paths leads to wasted network resources, increased data retransmission and unbalanced load.

In This research explores the potential of Supervised Machine Learning technology to enhance IoT routing protocol, we integrated Machine Learning into the well-established Routing Protocol for Low-Power and Lossy Networks (RPL), resulting in an advanced (new) version called HUML_RPL based on Supervised Machine Learning technology.

The new version uses decision tree model for trained on a dataset of routing metrics, predicts the probability of successfully reaching a destination node and Each node in the network uses this model to choose the route with the highest probability of effectively delivering packets.

The new version of protocol (HUML_RPL) is tested in the COOJA simulator of Contiki OS. The performance evaluated for networks of varying density. The simulation results showed that the new version of (HUML_RPL) outperforms the basic algorithm within the (RPL) protocol, which is called (MORHF_ETX), by improving energy consumption, reducing control messages, and improving the packet delivery rate and lowers latency in a network.

Keywords: RPL protocol, Internet of Things (IOT), MORHF, supervised machine learning.

**PhD Student in the Department of Systems & Computer Networks of Faculty of Informatics –Homs University.*

***Associated Professor in the Department of Systems & Computer Networks of Faculty of Informatics - Homs University.*

****Associated Professor in the Department of Systems & Computer Networks of Faculty of Informatics - Homs University.*

(1) مقدمة:

يتم ربط مليارات الأجهزة مثل عدادات الطاقة والمياه، والأجهزة المنزلية، والأجهزة القابلة للارتداء، والمستشعرات، والمشغلات بالإنترنت لتكوين شبكة اتصالات ضخمة تُعرف بالإنترنت الأشياء. من المتوقع أنه بحلول عام 2030، سيكون هناك أكثر من 27 مليار اتصال بالإنترنت الأشياء. ستحدث إنترنت الأشياء تحولاً في مجالات مراقبة الطقس، والمدن الذكية، والرعاية الصحية، والأتمتة الصناعية، والأتمتة المنزلية[1]. تتكون شبكات إنترنت الأشياء وهو الجيل الجديد من شبكات الحساسات اللاسلكية من مستشعرات تعمل بالبطاريات تستخدم متحكمات دقيقة مثل (MSP430)[2] وتتمتع بسعة تخزين محدودة. بخلاف الشبكات التقليدية، يمكن أن تكون حركة المرور في هذه الشبكات واحد إلى واحد أو واحد إلى متعدد أو متعدد إلى متعدد. كما أن شبكة الاتصالات هذه تعاني من عدة عيوب مثلاً عدم الاستقرار ومعدلات البيانات المنخفضة وفقدان حزم البيانات العالي ونطاق الإرسال المحدود وتغير طوبولوجيا الشبكة الديناميكي وبسبب هذه القيود العديدة على الأجهزة وتكنولوجيا الاتصالات الأساسية، يتم تصنيف شبكات إنترنت الأشياء

IOT (Internet Of Things) كشبكات منخفضة الطاقة وفائدة البيانات (Low Power Lossy) [27] LNN (Network) ونظراً لأن بروتوكول التوجيه هو أحد دعائم هيكلية الشبكة وهو المسؤول عن قرارات التوجيه حيث أن القرارات الخاطئة سوف تسبب هدراً لموارد الشبكة ، صمم البروتوكول (RPL) من أجل الشبكات ذات الضياع والطاقة المنخفضة والشبكات التي تحتوي عدداً كبيراً جداً من الأجهزة الذكية والقادرة على الاتصال بالمستخدمين عبر شبكة الانترنت حيث أهتم الباحثون بدراسة وتطوير البروتوكول (RPL) الذي يعتمد على الوظائف الغرضية (Objective Function) OF من أجل إنشاء المسارات داخل الشبكة وتشكيل طوبولوجيا الشبكة. على الرغم من أن بروتوكول RPL كان قادراً على حل المشاكل الرئيسية، إلا أنه لا يزال هناك بعض القيود التي تحتاج إلى معالجة على سبيل المثال، يتخذ بروتوكول (RPL) القياسي قرار التوجيه معتمداً على مقياس جودة وصلة اتصال وحيد، مما يؤدي إلى اختيار مسارات غير مثالية وبالتالي تقليل عمر الشبكة [3].

في هذا البحث، نقترح نسخة جديدة محسنة من بروتوكول التوجيه (RPL) يسمى HUML_RPL (HOMS_Univercity_MachineLearning_RPL)، يقوم البروتوكول باختيار أفضل مسار لعقدة الاب بناءً على تقنية التعلم الآلي الخاضع للإشراف وبالاعتماد على معايير جودة وصلة اتصال متعددة. ينجح هذا النهج في التغلب على قيود مقياس جودة وصلة الاتصال الواحد في بروتوكول (RPL) القياسي. تم اختيار كل من عدد القفزات HC (Hop Count) ، العدد المتوقع لإعادة الإرسال ETX (Expected Transmission Count) ، الطاقة المتبقية (Residual Energy) ، معدل تمرير الرزم (packet forwarding ratio) كمعايير لتدريب نموذج التعلم الآلي الخاضع لإشراف وذلك من أجل اختيار الأب الأمثل وباستخدام برنامج WEKA تم تدريب واختبار أربعة نماذج (Model) من خوارزميات التعلم الآلي الخاضع للإشراف وهي (شجرة القرار Decision Tree) ، الانحدار الخطي متعدد الحدود (multinomial logistic regression) ، الخلايا العصبية متعددة الطبقات (MultiLayer_Artificial_Neural_Network) ، آلة المتجهات الداعمة (Support vector machine) على بيانات تدريبية تقدر بـ (43600 عينة) تم توليدها عن طريق محاكي الشبكات COOJA ، حيث يقوم النموذج المدرب بحساب درجة جودة لكل عقدة أب مرشحة واختيار أفضل أب يحصل على أعلى درجة جودة كأب مفضل نحو عقدة الجذر وبعد

عملية التدريب وأجراء المقارنة باستخدام مقاييس التقييم لخوارزميات التعلم الآلي تم دمج أفضل خوارزمية وهي شجرة القرار (Decision Tree) داخل بنية بروتوكول التوجيه (RPL) وبعدها تم مقارنة النتائج مع كل من التوابع الوظيفية الأساسية (MORHF,OF0)، أظهرت النتائج تحسن في عملية استهلاك الطاقة وتحسين معدل تسليم الرزم والتقليل من التأخير الزمني بالإضافة إلى التقليل من رسائل التحكم.

(2) مشكلة البحث:

أهتم الباحثون بدراسة وتطوير البروتوكول (RPL) الذي يعتمد على الوظائف الغرضية (OF) من أجل إنشاء المسارات داخل الشبكة وتشكيل طوبولوجيا الشبكة على الرغم من أن بروتوكول (RPL) كان قادراً على حل المشاكل الرئيسية، إلا أنه لا يزال هناك بعض القيود التي تحتاج إلى معالجة على سبيل المثال يتخذ بروتوكول (RPL) القياسي قرار التوجيه معتمداً على مقياس جودة وصلة اتصال واحدة داخل الوظائف الغرضية مما يؤدي إلى اختيار مسارات غير مثالية وتحتوي على ضياعات وتشكيل ما يسمى بالنقاط الساخنة والتي يمكن تعريفها بـ(عقدة داخل الشبكة تعاني من الاختناق الشديد في مسألة تمرير الطرود وتبقى طوال الوقت مشغولة) وتسبب هذه النقاط الساخنة عملية تقليل عمر الشبكة ككل.

(3) الهدف من البحث:

نهدف في هذا البحث إلى تحسين أداء التوجيه واستهلاك الطاقة في شبكات أنترنت الأشياء وذلك من خلال تطوير بروتوكول التوجيه (RPL) الخاص بشبكات أنترنت الأشياء وذلك بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف.

حيث تم تحقيق مقياس جودة وصلة اتصال جديد يعتمد على الوظيفية الغرضية الجديدة (HUML_OF) والتي تعتمد على نموذج تعلم آلي يعتمد على تقنية شجرة القرار (Decision Tree) وذلك من أجل اختيار عقدة الـأب الأفضل من بين مجموعة العقد المرشحة لتكون عقدة أب مفضلة للعقدة الحالية.

(4) أدوات وطرائق البحث:

في هذا البحث تم استخدام حاسوب شخصي يحتوي على معالج من نوع (corei7-4th_generation) ومقدار ذاكرة عشوائية تقدر بـ(8GB) ونظام تشغيل من نوع (Ubuntu)

يحتوي نظام التشغيل الخاص بأنترنت الأشياء (Contiki_Os) بالإضافة إلى المحاكى الخاص بنظام التشغيل الخاص بأنترنت الأشياء (Cooja Simulator) مع تنزيل كل من برنامج التعلم الآلي (Weka) وبرنامج (EmbML) من أجل توليد نموذج تعلم آلي مدرب متوافق مع محدودية موارد العقد داخل الشبكة.

5) بروتوكول التوجيه للشبكات منخفضة الطاقة والتي تحتوي فقدان لرزم (RPL):

يعد بروتوكول التوجيه هو المسؤول عن اتخاذ قرار التوجيه الذكي عن طريق اختيار المسار الأفضل بين عدة مسارات[4]. يعد التوجيه في شبكات (LLN) أحد التحديات الرئيسية ليس فقط بسبب الضياع الذي يكون في الوسط الراديوي والموارد المحدودة لعقد الحساسات ولكن أيضا بسبب متطلبات التوجيه المختلفة [5] ، لذلك ركزت مجموعة العمل الخاصة بشبكات منخفضة الطاقة والتي يكون فيها ضياع Routing Over Low Power and Lossy Networks (ROLL) على عملية تصميم بروتوكول توجيه جديد يتناسب مع القيود المفروضة على هذه الشبكات حيث قامت مجموعة العمل (ROLL) [4] بتوجيه من قبل مجموعة مهندسي مهام الانترنت (IETF) Engineering Task Force بإعتماد وإنشاء بروتوكول التوجيه (RPL) كبروتوكول توجيه معياري لشبكات (LLN) وتم توصيفه بالملف (RFC 6550)[6] كبروتوكول معياري يستند إلى النسخة المقترحة من بروتوكول عنوان الشبكات (IPv6) للعقد المقيّدة بالموارد وأصبح بروتوكول التوجيه الفعلي لشبكات (IOT) كما يدعم متطلبات التوجيه الدنيا من خلال بناء طبولوجيا قوية عبر الروابط التي يكون فيها ضياعات، كما تم تطويره وأصبح معياراً أساسياً لبروتوكولات توجيه البيانات وتم استخدامه في العديد من أنظمة التشغيل مثل Contiki OS ، TinyOS ، كبروتوكول توجيه قياسي[5] .

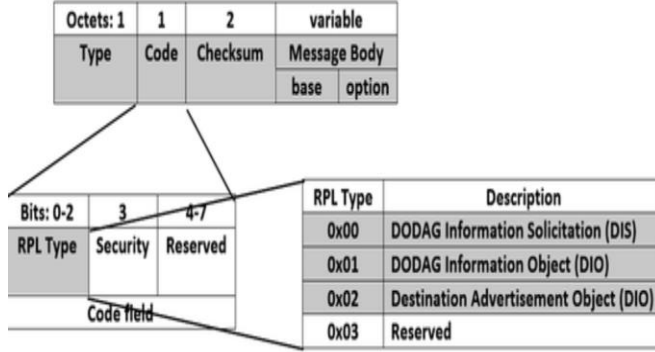
❖ رسائل التحكم في الـ (RPL):

تم اقتراح أنواع جديدة من رسائل التحكم داخل البروتوكول (RPL) [3] لبناء طبولوجيا شبكة (DODAG Destination Oriented Directed Acyclic Graph) وتتكون رسالة التحكم في (RPL) من [7] [26]:

✓ ترويسة (ICMPV6 Internet Control Message Protocol) والتي تتألف من ثلاثة حقول (type، code، CRC checksum) .

✓ متن الرسالة التي تحتوي على نص الرسالة وبعض الخيارات.

يتم تمييز نوع طرود التحكم بواسطة حقل الشيفرة (code) كما في الشكل رقم (1)، تتألف هذه الطرود من أربعة أنواع وهي:



الشكل رقم (1) بنية رسائل التحكم في بروتوكول (RPL)

• رسالة المعلومات DODAG Information Object (DIO):

بداية يقوم جذر (DODAG) ببث رسالة (DIO) باتجاه هابط إلى عقد الجوار من أجل تمكين حركة المرور من نقطة إلى عدة نقاط في اتجاه تنازلي. تحتوي رسائل (DIO) على معرف جذر شجرة (DODAG_ID)، معرف مثيل (RPL_Instance_ID) ومجموعة من المقاييس منها (مقاييس التوجيه، رتبة العقدة، الوظيفة الغرضية (OF)) تُرسل هذه الرسائل بشكل دوري مع رقم تسلسلي تراكمي من أجل بدء عملية اختيار الأب الأفضل للعقد [7].

• رسالة التماس المعلومات DODAG Information Solicitation (DIS):

يتم استخدام رسالة التحكم (DIS) بواسطة أي عقدة لجلب رسائل تحكم من نوع (DIO) بشكل صريح من العقد المجاورة، وتُستخدم أيضاً من أجل انضمام العقد إلى شجرة (DODAG) وفي حال لم تتمكن من تلقي رسائل من نوع (DIO) بعد فاصل زمني محدد مسبقاً يتم إعادة بث هذه الرسالة من جديد [7].

• رسالة إعلان الوجهة Destination Advertisement Object (DAO):

تخدم هذه الرسائل أساساً غرض نشر معلومات الوجهة في الاتجاه الصاعد نحو جذر الشبكة، تُستخدم هذه المعلومات لبناء المسارات الصاعدة بين العقد المرتبطة وجذر (DODAG) [7].

• الإقرار باستلام رسالة إعلان الوجهة (DAO-ACK):

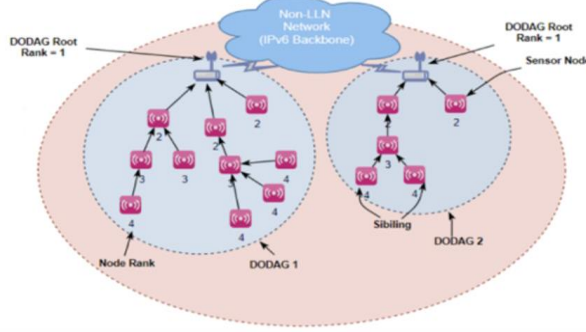
عند استلام رسائل تحكم من نوع (DAO)، يمكن للعقدة إرسال رسالة (DAO-ACK) إلى مرسل رسالة (DAO) للاعتراف باستلام الرسالة بشكل صحيح وللإقرار باستعداد العقدة للعمل كقفزة تالية باتجاه جذر [7].

❖ طريقة تشكيل الطوبولوجيا داخل البروتوكول (RPL):

يقوم RPL بإنشاء طوبولوجيا الشبكة من خلال مخطط يسمى رسم بياني موجه (Directed Acyclic Graph DAG) والذي يتكون من واحد أو أكثر من الرسم البياني والذي يدعى رسم بياني لا حلقي موجه إلى الوجهة [6] (DODAG)، يمثل كل (DODAG) شجرة توجيه تم إنشاؤها بواسطة عقدة الجذر (Root Node) والمعروفة أيضا باسم عقدة مستقبل معلومات الشبكة. عند البدء بتهيئة المسارات وتشكيل البنية، تبدأ عقدة الجذر بتكوين الشبكة وبناء (DODAG) عن طريق بث رسالة تحكم من نوع (DIO) كما يوضح الشكل رقم (2)، تحتوي هذه الرسالة على العديد من المعلومات مثل معلومات رتبة العقدة (للسماح للعقد بأخذ مواقعها في DODAG ولمنع حلقات التوجيه) [3]، نمط التشغيل والوظائف الغرضية (OF)، تقوم العقد المجاورة لهذه العقدة بالاستماع إلى هذه الرسالة وإعادة إرسالها إلى العقد المجاورة لها وتحديد فيما إذا كانت تريد الانضمام إلى هذه البنية أم لا، وتقرر ذلك بناءً على قيود محددة، في الوقت الذي تصبح به هذه العقدة جزءاً من البنية، فسيكون لها مسار للأعلى نحو الجذر وعندها تحسب العقدة رتبته وتقوم بتحديث جدول الجيران الخاص بها، وتختار الأب الأفضل الذي سيتم استخدامه لإعادة توجيه الرسائل إلى جذر (DODAG) وعادة يكون الأب الأفضل هو الجار الذي يملك الرتبة الأدنى المحسوبة وفقاً للوظيفة الغرضية الصفرية (OF_0)، وبعدها تقوم هذه العقدة بإعادة إرسال رسائل التحكم إلى العقد المجاورة وبهذه الطريقة تستمر العملية حتى بناء كامل الشبكة [7]. يدعم بروتوكول (RPL) وجود عدة بنى ضمن الشبكة الواحدة والتي من الممكن أن تكون تابعة لعقدة جذر أو أكثر وتدعى هذه البنية (DAG Instance) ويمكن لعقدة واحدة أن تنتمي إلى بنية واحدة أو أكثر بنفس الوقت حيث يكون لكل عقدة (ID Instance) واحد لكل بنية تنتمي إليها، تكمن فائدة هذه الخاصية في بناء مسارات حسب الهدف المطلوب من التطبيق مثلاً: مسارات للبيانات العادية يتجنب العقد

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL) بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء

التي تعمل على البطارية، ومسار لبيانات ذات الأهمية القصوى والذي يكون فيه التأخير بحدوده الدنيا وتسلك مسارات ذات وثوقية عالية وطاقة عقدة كبيرة[3].



الشكل رقم (2) شجرة DODAG

❖ الوظائف الغرضية في بروتوكول الانترنت (RPL):

تعتمد عملية إنشاء شجرة (DODAG) على الوظائف الغرضية (OF) بشكل أساسي والتي تأخذها الشبكة بعين الاعتبار [6] ، يعتمد استخدام الوظيفة الغرضية بواسطة بروتوكول التوجيه (RPL) على ثلاث مكونات:

- المقياس المحدد، والذي يمثل المقياس الذي تستخدمه العقدة من أجل إنشاء المسار بين العقد.
- تكلفة المسار والذي يمثل جودة مسار الاتصال من بين مجموعة من المسارات بين العقد ويعتمد على المقياس المحدد.
- أفضل أب والذي يمثل أفضل مسار بين عقدة أبن وعقدة أب من حيث جودة وصلة الاتصال أو الطاقة المتبقية بالعقدة الاب أو العقدة الأقل من حيث كمية استقبال وأرسال البيانات. وضع الباحثون [7] [3] الوظائف الغرضية المعيارية في (RPL) كما يلي:

✓ الوظيفة الغرضية الصفرية (OF_0) :

تعتمد على عدد القفزات كمقياس لتكلفة المسار. يتم اختيار الأب المرشح الذي لديه أقل عدد من القفزات، مما يعني أن العقدة الأقرب إلى الجذر ستختار كأب مفضل. تلائم هذه الوظيفة الغرضية (OF0) الشبكات منخفضة الطاقة والتي يكون فيها ضياع. لكن مقياس العقدة هذا قد يؤثر جودة روابط سيئة لأن العقد تختار المسار مع الحد الأدنى لعدد القفزات حتى لو كان غير موثوق [7] ، وبالتالي تزداد عمليات إعادة الإرسال وضياع المزيد من الطرود.

✓ الوظيفة الغرضية (MORHF Based On ETX):

تستخدم الوظيفية الغرضية (ETX_OF) مقياس جودة وصلة الاتصال والذي يسمى عدد الإرسال المتوقع (ETX) كمقياس توجيه لتحديد أفضل مسار من بين مجموعة مسارات، ويعرف بالحد الأقصى لعدد عمليات إعادة الإرسال اللازم حتى يتم إقرار لاستلام الطرود بنجاح [3]. يمثل هذا المقياس بمعنى آخر كمية الضياعات في المسار بحيث كلما كانت كمية الضياعات كبيرة ازدادت قيمة المقياس ويكون المسار سيئ وكلما انخفضت قيمة (ETX) يكون هناك ضياعات قليلة ويكون المسار أفضل، تحسب قيمة (ETX) بالعلاقة رقم (1) [5]:

$$ETX = \frac{1}{dF * dR} \quad (1)$$

حيث:

dF: تمثل نسبة التسليم وتعرف احتمال الطرود الواردة الناجحة من المصدر إلى الوجهة.
dR: تمثل نسبة التسليم العكسي وهو احتمال الإقرار الوارد الناجح من الوجهة إلى المصدر.
تُقاس جودة الرابط من حيث الموثوقية من خلال قيمة (ETX)، تشير كلفة (ETX) المنخفضة إلى أن الروابط في الشبكة أكثر موثوقية.
كل من الوظيفتين الغرضيتين (OF0, MORHF) لهما القصور نفسه وذلك بسبب اعتمادهم على مقياس واحد فقط من أجل اختيار المسار [27]. تفنقر هذه الوظائف إلى النهج الشمولي الضروري لاختيار المسار الأفضل من حيث التأخير الزمني أو تجاوز العقد ذات الازدحام العالي. للتغلب على ذلك وتوفير حلول تؤمن جودة الخدمة (QoS)، قدمت مواصفات (RPL) المرنة للمصمم لإدراج المزيد من المقاييس في عملية اتخاذ القرار وسنرى في الأجزاء اللاحقة كيف يمكن تحقيق ذلك.

6) التعلم الآلي (Machine Learning (ML:

يعتبر مجال التعلم الآلي واحد من أهم مجالات الذكاء الاصطناعي الأربعة وهي: تعلم الآلة، معالجة اللغات الطبيعية، الرؤية الحاسوبية، ومجال الروبوتات، ان الهدف الاساسي من مجال التعلم الآلي هو تصميم نموذج آلي تستطيع الآلة من خلاله التصرف والتفكير والرؤية والتكلم مثل الإنسان. وذلك بناء على البيانات التي يتم تدريب هذه الآلة عليها.

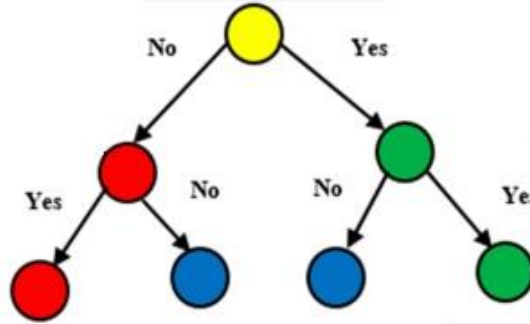
يمكن تعريف مجال التعلم الآلي على أنه مجال في علم الحاسوب يستند على الإحصاء والتقنيات البرمجية حيث يكمن الغرض الرئيسي من هذا المجال في تطوير أنظمة ذكية تستطيع تحسين نفسها تدريجياً وبشكل آلي عندما تتعرف على البيانات والتجارب الجديدة بدون الحاجة للإعداد اليدوي [8].

6-1 التعلم الآلي الخاضع للإشراف Supervised ML :

يعد التعلم الآلي الخاضع للإشراف من الأقسام المهمة في التعلم الآلي إن لم يكن أهمها، حيث يتم تدريب خوارزميات هذا النوع على البيانات المصنفة، مما يعني أن البيانات تتضمن متغيرات الإدخال (features) ومتغير الإخراج المقابل لها (السمة)، الهدف من التعلم الخاضع للإشراف هو بناء نموذج يمكنه التنبؤ بدقة بمتغير المخرجات في ضوء بيانات الإدخال الجديدة كما يمكن تقسيم مهام هذا التعلم عند التقسيم في البيانات إلى نوعين رئيسيين هما التصنيف والتنبؤ. التصنيف هو نوع من خوارزميات التعلم الآلي التي يمكن استخدامها للتنبؤ بمتغير فئوي بناءً على واحد أو أكثر من متغيرات التنبؤ، حيث يتم تدريب النموذج على مجموعة من البيانات المصنفة ثم يتم استخدامه للتنبؤ بالبيانات الجديدة غير المرئية و بشكل عام، تعد خوارزميات التصنيف أداة قوية للتنبؤات ويمكن استخدامها في مجموعة واسعة من الصناعات والتطبيقات، بينما يستخدم التنبؤ لفهم العلاقة بين المتغيرات التابعة والمستقلة وبشكل عام يتم استخدامه من أجل التنبؤ بالتوقعات، مثل إيرادات المبيعات لنشاط تجاري معين، ومن أهم خوارزميات التعلم الآلي المستخدمة حالياً [9][8]:

6-1-1 شجرة القرار (Decision Tree) : تُستخدم خوارزمية أشجار القرار بهدف إنشاء نموذج يتنبأ بقيمة المتغير

المستهدف، إذ يتم تعلم قواعد القرار البسيطة التي تم استنتاجها من ميزات البيانات و يبدأ عمل هذه الخوارزمية من عقدة الجذر كما يبين الشكل (3) التي تطرح سؤالاً محدداً حول البيانات، ووفقاً للإجابة تُوجه البيانات إلى مختلف الفروع ومنها إلى العقد الداخلية والتي تطرح المزيد من الأسئلة وتوجه البيانات إلى الفروع اللاحقة، وتبقى هذه العملية مستمرة حتى تصل إلى عقدة النهاية التي لا تتفرع مرة أخرى [8].

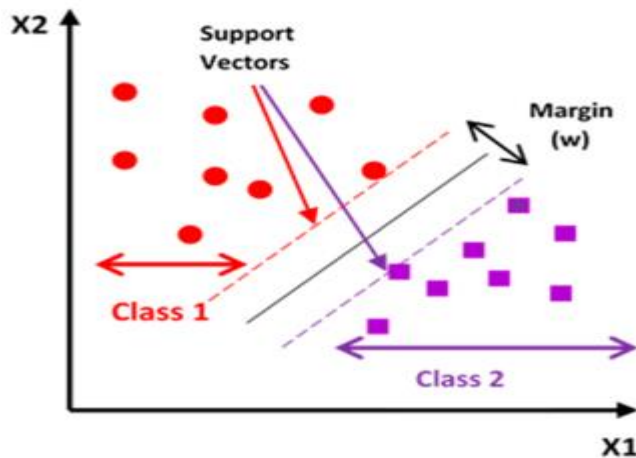


الشكل (3) يبين بنية أشجار القرار [8]

2-1-6 خوارزمية آلة المتجهات الداعمة (Support Vector Machine): عبارة عن أسلوب

يجمع بين النظرية

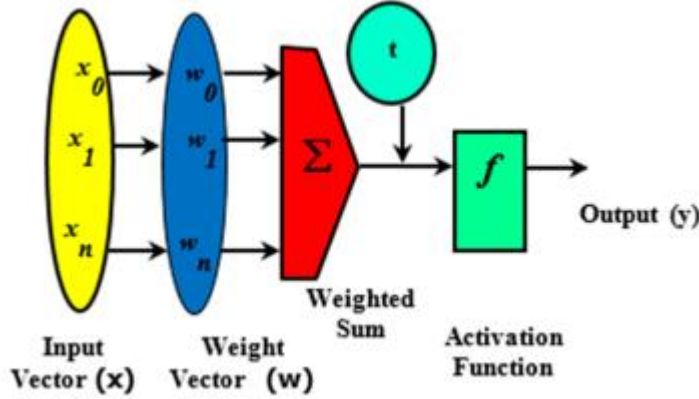
الإحصائية والتعلم الموجه، وتمتلك هذه الخوارزمية طريقتها الفريدة في التنفيذ مقارنة بخوارزميات التعلم الآلي الأخرى وفي الآونة الأخيرة، أصبحت شائعة للغاية بسبب قدرتها على التعامل مع العديد من المتغيرات المستمرة والقوية ويمكن استخدامها في مسائل التصنيف (Classification) أو الانحدار (Regression) [8]. الفكرة الرئيسية لعمل هذه الخوارزمية كما يبين الشكل (4) هي إيجاد أفضل مستوى فاصل (separating Hyperplane) الذي يفصل بين المجموعات (classes) وذلك عن طريق المحاولة لجعل الهامش (margin) أكبر ما يمكن والهامش (margin) هي المسافة بين بيانات المختلفة المراد تصنيفها والتي تعتبر محور عملية التدريب في آلة المتجهات الداعمة [8].



الشكل (4) خوارزمية آلة المتجهات الداعمة [8]

3-1-6 الشبكات العصبية الاصطناعية متعددة الطبقات (Multilayer Artificial Neural Network): تتكون

الشبكات العصبية الاصطناعية متعددة الطبقات (Multilayer ANN) من طبقات من العقد، حيث تحتوي على طبقة للمدخلات، وطبقة مخفية واحدة أو أكثر، وطبقة للمخرجات. كل عقدة، أو خلية عصبية اصطناعية، ترتبط بأخرى ويكون لها حمل وحد مرتبطين بها. إذا كانت مخرجات أي عقدة فردية أعلى من قيمة الحد المحددة، فإنه يتم تفعيل هذه العقدة، وإرسال البيانات إلى الطبقة التالية من الشبكة، خلاف ذلك لن يتم تمرير أي بيانات إلى الطبقة التالية من الشبكة. يمكن اعتبار كل عقدة على أنها نموذج انحدار خطي خاص بها، يتكون من بيانات المدخلات والأوزان والميل (أو الحد) والمخرجات، وبشكل عام تتكون الشبكة العصبية من ثلاث طبقات هي طبقة الإدخال، الطبقة المخفية وطبقة الإخراج [8]، في هذا البحث سوف يتم استخدام الشبكات العصبية القائمة على التغذية الأمامية، أو المستقبلات ذات الطبقات المتعددة (MultiLayer Precptron) MLP ويوضح الشكل (5) بنية عمل الشبكة العصبية.

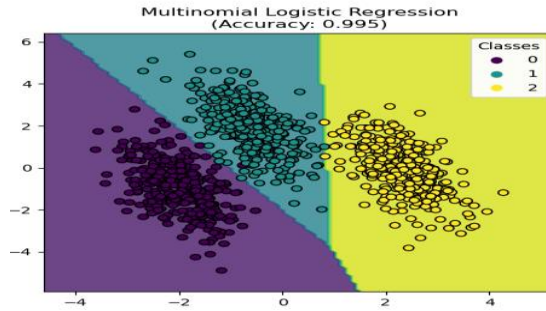


الشكل (5) يبين طبيعة عمل الشبكة العصبية [8]

4-1-6 الانحدار اللوجستي متعدد الحدود (Multinomial Logistic Regression):

الانحدار اللوجستي المتعدد الحدود هو أسلوب تصنيف يعمم الانحدار اللوجستي (Logistic Regression) على المشاكل متعددة الفئات، أي مع أكثر من نتيجتين منفصلتين محتملتين، أي أنه نموذج يستخدم للتنبؤ باحتمالات النتائج المختلفة لمتغير تابع موزع تصنيفياً، مع الأخذ في

الاعتبار مجموعة من المتغيرات المستقلة (والتي قد تكون ذات قيمة حقيقية، أو ثنائية، أو تصنيفية) [24] [9]، ونلاحظ من الشكل (6) طريق تصنيف البيانات عن طريق هذا المقياس.



الشكل (6) الانحدار متعدد الحدود [24]

2-6 مقاييس تقييم أداء نموذج التعلم الآلي:

في هذا الجزء من البحث، سنناقش بعض مقاييس التقييم الأكثر شيوعاً المستخدمة في تقييم نماذج التعلم الآلي وذلك لتأكد من صحة نتائج الخوارزمية ومنها:

2-6-1 مصفوفة الارتباك (Confusion Matrix): تستخدم مصفوفة الارتباك لتقييم أداء نموذج

التصنيف [16] حيث

توضح عدد التنبؤات الصحيحة والغير الصحيحة التي قدمها النموذج مقارنة بالنتائج الفعلية. عادة ما يتم تمثيل المصفوفة على شكل جدول كما يبين الشكل (7) حيث يتم حساب هذه المكونات بناء على توقعات النموذج والنتائج الفعلية [16]:

✓ **إيجابية صحيحة (TP):** هذه القيمة تعبر عن عدد الحالات الإيجابية الفعلية التي توقعها النموذج بشكل صحيح على أنها إيجابية.

✓ **سلبية صحيحة (TN):** تعبر عن عدد الحالات السلبية الفعلية التي توقعها النموذج بشكل صحيح على أنها سلبية.

✓ **إيجابية خاطئة (FP):** تعبر عن عدد الحالات السلبية الفعلية التي توقعها النموذج بشكل غير صحيح على أنها إيجابية.

✓ **سلبية خاطئة (FN):** تعبر عن عدد الحالات الإيجابية الفعلية التي توقعها النموذج بشكل غير صحيح على أنها سلبية.

		PREDICTED classification				
		Classes	a	b	c	d
ACTUAL classification	a	TN	FP	TN	TN	
	b	FN	TP	FN	FN	
	c	TN	FP	TN	TN	
	d	TN	FP	TN	TN	

يبين الشكل (7) مصفوفة الارتباك [16]

2-2-6 الدقة (Accuracy): الدقة هي مقياس التقييم الأكثر استخداماً في مسائل التعلم الآلي

الخاضع للإشراف

حيث تقيس نسبة نقاط البيانات المصنفة بشكل صحيح من إجمالي عدد نقاط البيانات ويتم

حسابها بالعلاقة [16] رقم (2):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

3-2-6 الانضباط (Precision): يعرف بأنه نسبة التنبؤات الإيجابية الحقيقية من جميع

التوقعات الإيجابية حيث يقيس قدرة النموذج على التنبؤ بالعينات الإيجابية بشكل صحيح ويتم

حسابه بالعلاقة [16] رقم (3):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

4-2-6 الاستدعاء (Recall): ويعرف أيضاً باسم الحساسية، هو نسبة التنبؤات الإيجابية

الصحيحة من جميع العينات الإيجابية الفعلية، كذلك يقيس قدرة النموذج على تحديد العينات

الإيجابية بشكل صحيح ويتم حسابه بالعلاقة [16] رقم (4):

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

6-2-5 درجة (F1): ويعبر عنه بالمتوسط التوافقي للدقة والاستدعاء كذلك يعتبر هذا المقياس شائع الاستخدام عندما

يكون كل من الدقة والاستدعاء مهمين يتم حساب درجة (F1) بالعلاقة رقم [16] (5):

$$F1_{Score} = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

حيث:

✓ TP هو عدد الإجابات الحقيقية.

✓ TN هو عدد السلبات الحقيقية.

✓ FP هو عدد الإجابات الخاطئة.

✓ FN هو عدد السلبات الخاطئة.

6-2-6 متوسط الخطأ المطلق (Mean absolute error): يستخدم لقياس دقة

النموذج وبالتالي حساب معدل

الخسارة (loss) وذلك عن طريق إيجاد متوسط الفروقات بين القيم الحقيقية والمتوقعة باستخدام النموذج. كلما قلت القيمة الناتجة من الدالة كلما زادت دقة النموذج، ويكون استخدامه مفيداً في حالة قياس دقة النماذج المطبقة على البيانات التي تحتوي على قيم شاذة أو متطرفة وبحسب بالعلاقة [17] رقم (6):

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (6)$$

6-2-7 جذر متوسط الخطأ التربيعي (Root Mean Squared Error): يعد خطأ الجذر

التربيعي المتوسط في

التعلم الآلي مقياساً شائعاً يستخدم لقياس دقة النموذج التنبئي فهو يقيس الاختلافات بين القيم المتوقعة والقيم الفعلية، ويربع الأخطاء، ويأخذ المتوسط، ثم يجد الجذر التربيعي. يوفر خطأ الجذر التربيعي المتوسط فهماً واضحاً لأداء النموذج، حيث تشير القيم المنخفضة إلى دقة تنبؤية أفضل وبحسب بالعلاقة [17] رقم (7):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad (7)$$

حيث:

(n) حجم العينة.

(y) القيمة الحقيقية.

(x) القيمة المتوقعة.

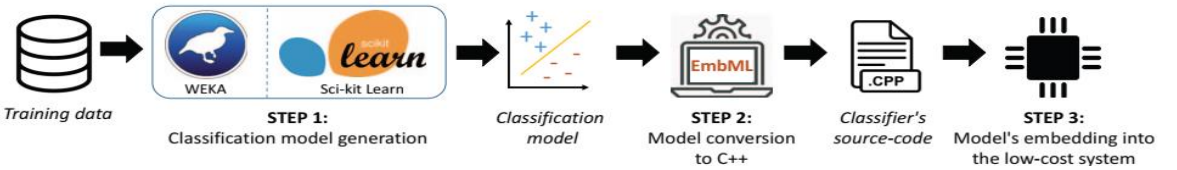
3-6 برنامج تدريب نماذج التعلم الآلي (Weka):

هي أداة برمجية مفتوحة المصدر وشائعة الاستخدام في استخراج البيانات وتطبيق خوارزميات التعلم الآلي، والتي تم تطويرها في جامعة (Waikato University) في نيوزيلندا، تم تصميم برنامج (Weka) لتوفير مجموعة شاملة من الأدوات من أجل تحليل البيانات والنمذجة التنبؤية وهي حالياً تحظى بشعبية كبيرة في البيئات الأكاديمية والبحثية بسبب مرونتها وسهولة استخدامها. تم تصميمها لمساعدة المستخدمين على تحليل مجموعات البيانات الكبيرة وتطبيق خوارزميات التعلم الآلي المختلفة مثل التجميع (Clustering) والتصنيف (Classification) والانحدار (Regression) واستخراج البيانات [20].

4-6 برنامج (EmbML):

هي أداة لتحويل النماذج المدربة إلى ملفات كود مصدر بلغة (C, C++) بحيث يمكن دمج وتشغيل النموذج المحول داخل الأنظمة المضمنة، كما تقبل هذه الأداة مجموعة من النماذج المدربة باستخدام برنامج WEKA أو برنامج Scikit-Learn ، وذلك لأنها شائعة ومفتوحة المصدر وتوفر مجموعة واسعة من خوارزميات التعلم الآلي.

يوضح الشكل (8) تسلسل خطوات عملية تحويل نموذج التعلم الآلي المولد باستخدام برنامج (Weka) إلى ملفات كود مصدرية [21].



الشكل (8) خطوات استخدام برنامج [21] EmbML

5-6 كيفية عمل التنبؤ في نموذج التعلم الآلي الخاضع للإشراف:

تحتاج نماذج التعلم الآلي إلى تدريب نفسها قبل أن تتمكن من التنبؤ بدقة جيدة تنقسم مجموعة البيانات المستخدمة للتنبؤات إلى مجموعتين: مجموعة اختبار ومجموعة تدريب، تتكون مجموعة التدريب من بيانات مُسمّاة تُستخدم لتدريب نموذج التعلم الآلي وزيادة دقة التنبؤ بينما تحتوي مجموعة الاختبار على بيانات ربما لم يراها نموذج التعلم الآلي المُدرَّب من قبل، والتي تُستخدم لقياس مدى قوة تنبؤات النموذج في تجاربنا تكون نسبة الانقسام 80% لمجموعة التدريب و20% لمجموعة الاختبار.

7 الدراسة المرجعية:

نوضح في هذا الجزء من البحث دراسة استقصائية عن أبحاث تحسين أداء جودة وصلة الاتصال والتوابع الوظيفية المتوفرة في (RPL) وذلك باستخدام تقنيات التعلم الآلي الخاضع للإشراف، تهدف معظم الحلول إلى تحسين أداء التوجيه وفعالية الطاقة وإدارة موارد العقدة وتحسين معدل تسليم الرزم وتقليل من التأخير الزمني.

المؤلفون في [10] استخدموا تقنية التعلم الآلي الخاضع للإشراف وتم استخدام خوارزمية (Neighborhood-Based Nonnegative Matrix) وذلك من أجل تقييم جودة وصلة الاتصال، تم استخدام برنامج (MATLAB) من أجل تدريب النموذج المقترح كما تم تقييم أداء الخوارزمية عن طريقة مجموعة من المقاييس (recall-accuracy-F1_score) أظهرت النتائج زيادة في عدد معدل تسليم الرزم، الدراسة لم تذكر نوع البروتوكول الذي تم استخدامه كما تم تدريب النموذج فقط عن طريق برنامج (MATLAB) مع توفر الكثير من البرنامج التي تستخدم في تدريب واختبار النموذج بشكل أكثر سهولة.

اقترح الباحثون في [11] تحسين اختيار جودة وصلة الاتصال عن طريق استخدام تقنية التعلم الآلي حيث اقترحوا استخدام ثلاثة من خوارزميات التعلم وهي (الغابات العشوائية) Random (Forst)، الانحدار المنطقي (Logistic Regression)، آلة المتجهات الداعمة (SVM)، آلة المتجهات الداعمة الخطية (Linear SVM) وذلك من أجل تدريب نموذج تعلم يحوي على مقياسين من مقياس جودة الاتصال وهي (قوة الإشارة المستقبلية) (RSSI)، ومعدل تسليم الرزم (PDR Packet Delivery Ratio)، تم تقييم أداء الخوارزميات المقترحة عن طريقة كل من الدقة ومصفوفة التشويش، اقترحت

الدراسة استخدام خوارزمية الغابات العشوائية (Random Forest) في عملية التنبؤ بجودة وصلات الاتصال، الدراسة لم توضح قياس الأداء للبروتوكول المستخدم ومقارنة النتائج عند استخدام تقنية التعلم الآلي أو طريقة دمج خوارزمية الغابات العشوائية داخل بنية البروتوكول.

المؤلفون في الدراسة المرجعية [12] استخدموا تقنية التعلم الآلي الخاضع للإشراف من أجل تدريب نموذج تعلم آلي يحوي على أربعة ميزات (Feature) وهي (الطاقة المتبقية، والتأخير، عدد القفزات، مقياس عدد الارسل المتوقع) وذلك من أجل ثلاثة من خوارزميات التعلم وهي (الغابات العشوائية (Random Forst)، الانحدار الخطي (Linear Regression)، خوارزمية أقرب جار (K-Nearest Neighbor) تم استخدام برنامج (Scikit Learn) من أجل تدريب واختبار النموذج، تم تقييم إداء الخوارزميات عن طريق مقياس الدقة (Accuracy) وزمن التدريب (Training Time) وزمن الحساب (Computation time)، أعطت النتائج الدقة الأفضل خوارزمية اقرب الجيران تم دمج وتحقيق الخوارزمية اقرب الجيران (k-nearest neighbor) داخل بروتوكول التوجيه (RPL) وأظهرت النتائج تحسن كل من مقياس معدل تسليم الرزم وكمية الطاقة المستهلكة وذلك بالمقارنة مع كل من التابع الوظيفي (MORHF,OF0) وذلك من اجل عدد عقد مختلفة.

الورقة البحثية [13] ناقشت عملية تحسين اختيار العقدة الاب الأفضل داخل بروتوكول التوجيه (RPL) وذلك عن طريق استخدام نموذج التعلم الآلي الخاضع للإشراف من أجل تدريب نموذج يحتوي على ثمانية ميزات (Feature) وهي (عدد الارسل المتوقع، عدد القفزات، قوة الإشارة المستقبلية، جودة القناة، كثافة الشبكة، معدل ضياع البيانات على مستوى طبقة النقل، مقدار الحمل على رتل العقدة، عدد الرزم المسلمة) تم استخدام خوارزمية تعزيز التدريب (GRADIENT BOOSTING) من أجل تدريب واختبار النموذج، تم دمج النموذج المدرب داخل البروتوكول (RPL) تم استخدام محاكي الشبكات (OMNET++) من أجل دمج النموذج المدرب و تقييم النتائج باستخدام مقاييس التقييم (معدل تسليم الرزم، معدل تأخر وصول البيانات، حمل رسائل التحكم) ، نتائج المقارنة أعطت تحسن معدل تسليم الرزم وتقليل من معدل تأخر وصول البيانات وتقليل من رسائل التحكم، الورقة البحثية لم تأخذ مقياس كمية الطاقة عند استخدام النموذج المقترح

كما تم استخدام خوارزمية تعزيز التدريب (GRADIENT BOOSTING) بشكل مباشر دون مقارنتها مع بعض الخوارزميات باستخدام مقاييس تقييم إداء خوارزميات التعلم الآلي المعروفة. المؤلفون في الورقة البحثية [14] اقترحوا تحسين اختيار جودة وصلة الاتصال داخل البروتوكول (RPL) باستخدام الخوارزميات الجينية الفوضوية وتم إنشاء نسخة محسنة تسمى (RPL_CGA) (Chaotic Genetic Algorithm) وذلك بالاعتماد على كل من (حجم رتل الانتظار داخل العقدة (Queue Length)، عدد القفزات (HOPS)، عدد الارسل المتوقع (ETX)، التأخير الزمني من النهاية للنهاية (End-To-End Delay)، الطاقة المتبقية) حيث تم إعطاء لكل مقياس وزن معين وباستخدام الخوارزمية الجينية الفوضوية (CGA) تم اختيار أفضل عقدة أب من أجل القفزة التالية (Next Hop) والتي تملك افضل جودة لوصلة الاتصال وباستخدام محاكي الشبكات (OPNET) تم محاكاة العمل وأعطت النتائج تحسن كل من معدل تسليم الرزم (PDR) والتقليل من متوسط التأخير الزمني (Average End-to-End Delay) بالمقارنة مع البروتوكول التابع الوظيفي الأساسي (MORHF).

الدراسة المرجعية [15] اقترحت استخدام الشبكات العصبونية الاصطناعية من أجل تحسين عملية التوجيه. اقترحت الدراسة إنشاء تابع وظيفي (ANN_OF_RPL) جديد يعتمد على الشبكات العصبونية متعددة الطبقات من أجل اختيار جودة وصلة الاتصال الأفضل من بين مجموعة من الوصلات المقترحة تم استخدام كل من مقاييس (عدد القفزات، الطاقة المتبقية، عدد الارسل المتوقع، التأخير الزمني من النهاية للنهاية) من أجل تدريب الخوارزمية العصبونية والتي تقوم بحساب درجة جودة لكل عقدة أب مفضلة وذلك بالاعتماد على قيم المقاييس المقترحة وباستخدام محاكي الشبكات (COOJA) تم تقييم النتائج مع التوابع الوظيفية الاساسية (MORHF,OF_0) أعطت النتائج تحسين في عملية تسليم الرزم وتحسين كمية استهلاك الطاقة والتخفيف من التأخير الزمني.

8) مقاييس جودة وصلة الاتصال المستخدمة في تدريب نماذج التعلم الآلي:

من أجل تحسين أداء التوجيه وحل مشكلة توزع الحمل داخل عقد الشبكة تم استخدام كل من مقاييس جودة وصلة الاتصال وذلك من أجل توليد بيانات تدريبية لخوارزميات التعلم الآلي ومن هذه المقاييس:

1-8 الطاقة المتبقية (Residual Energy): يتم حساب الطاقة المتبقية للعقدة المرشحة حتى تكون عقدة أب بالعلاقة رقم (8) [18] [1]:

$$E_{res}(n) = E_{ini}(n) - E_{cons}(n) \quad (8)$$

حيث:

E_{res} : يشير إلى الطاقة المتبقية للعقدة (n).

E_{ini} : يشير إلى الطاقة الابتدائية للعقدة (n).

E_{cons} : يشير إلى كمية الطاقة المستهلكة للعقدة (n).

2-8 معدل إعادة توجيه الرزم (packet forwarding ratio):

لضمان توازن الحمل بين عقد الشبكة وعدم اختيار عقد ذات حمل عالي تشكل نقاط ساخنة داخل الشبكة وتؤدي إلى تدمير البيانات وزيادة كمية إعادة الإرسال واستنزاف طاقة العقدة بشكل كبير، تم اقتراح مقياس جديد يسمى معدل تمرير الرزم داخل العقدة ويقصد به كمية الرسائل التي تقوم باستقبالها عقدة ما وتقوم بإعادة تمرير هذه الرزم إلى عقدة أخرى ذات مستوى أعلى داخل الشبكة حيث تكون هذه العقدة مشغولة دائماً بعملية استقبال وتمرير رزم البيانات طوال الوقت ويمكن حساب معدل تمرير الرزم للعقدة المرشحة كعقدة أب خلال فترة زمنية معينة بالعلاقة المقترحة رقم (9) .

$$packet\ forwarding = (\sum packet_{receive} - \sum packet_{send}) / 100 \quad (9)$$

$packet_{send}$: يشير كمية البيانات المرسلة من قبل العقدة نفسها والتي تم تحسبها.

$packet_{receive}$: يشير إلى كمية البيانات المستقبلية من قبل العقدة التي تشكل أبناء لهذه العقدة.

3-8 عدد الإرسال المتوقع (ETX): من أجل جودة عالية للروابط وتجنب فقدان الطرود، يتم

حساب عدد الإرسال المتوقع بالعلاقة [4] (1).

4-8 عدد القفزات (Hop): يتم حساب عدد القفزات التي تسلكها رسالة ال (DIO) حتى الوصل

للعقدة التي يتم

حساب تكلفة مسارها، حيث يتم زيادة عدد القفزات بمقدار (1) عند كل عقدة تسلكها رسالة ال

(DIO) [19] .

9 توليد البيانات التدريبية (Trace Set):

تم توليد بيانات تدريبية تقدر ب (43180) عينة لنموذج شبكة مؤلف من (40) عقدة اتصال موزعة بشكل عشوائي لمدة (45) دقيقة مع برنامج مؤلف من نموذج تطبيق UDP Contiki يتألف من ملفين الملف (udp-client.c) وهو الملف الذي تم تنزيله على جميع عقد الشبكة (client node) بحيث تقوم العقدة (client node) بإرسال رسالة نصية ضمن فاصل زمني محدد الى موجه الشبكة أو جذر الشبكة (DODAG ROOT) والذي يحوي على الملف (udpserver.c) عند وصول المعلومات المرسله من عقدة الزبائن (client node) يقوم موجه الشبكة بأعلام العقدة التي تم استلام المعلومات منها بأن المعلومات تم استلامها بشكل صحيح، كما تم تقسيم جميع قيم مقاييس جودة وصلة الاتصال إلى درجات بحيث يعطي درجة جودة لكل مقياس تبعاً للقيم المولدة خلال فترة جميع بيانات تدريب النموذج ودرجة جودة نهائية لجميع مقاييس جودة وصلة الاتصال، كما يبين الجدول رقم (1).

✓ تم تقسيم قيم المقياس عدد الارسل المتوقع (ETX) الى أربع مستويات تمثل جودة المقياس (250,200,150,100).

✓ تم تقسيم قيم مقياس الطاقة المتبقية إلى ثلاثة مستويات تمثل الطاقة المتبقية بالعقدة وهي (250,170,100).

✓ تم تقسيم مقياس معدل توجيه البيانات إلى ثلاثة مستويات تمثل كمية البيانات التي تم إعادة توجيهها وهي (250,175,100).

✓ تم تقسيم مقياس عدد القفزات إلى ثلاثة مستويات تمثل مقدار القفزات للوصول على عقدة الجذر وهي (250,170,90).

كما تم تقسيم درجة الجودة النهائية لوصلة الاتصال أو ما يسمى بال (Class) إلى خمسة مستويات تمثل قيم الجودة النهائية وهي (الأسوأ، سيئة، جيدة، جيدة جداً، الأفضل) والتي تمثل القيم التالية على الترتيب (1000,800,700,500,300).

**تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء**

الجدول رقم (1) يبين قيم مقاييس جودة وصلة الاتصال

ETX	Forwarding ratio	Hop count	Residual Energy	درجة الجودة النهائية (quality)
250	250	250	250	1000
250	250	250	90	800
200	175	250	250	700
100	250	250	170	500
150	100	250	250	300

كما يوضح الشكل (9) عينة من ملف البيانات التدريبية المولدة (Trace set)

```

trace_num_quality.arff |
1 @relation 'etx_trace'
2 @attribute etx numeric
3 @attribute forwarding numeric
4 @attribute hop numeric
5 @attribute renergy numeric
6 @attribute class {'1000','800','700','500','300'}
7 @data
8 200,250,250,250,1000
9 100,250,250,90,500
10 200,250,250,250,1000
11 100,250,250,90,500
12 100,250,250,90,500
13 200,250,250,250,1000
14 200,250,250,250,1000
15 100,250,250,90,500
16 100,250,250,90,500
17 100,250,250,90,500
18 100,250,250,90,500
19 100,250,250,90,500
20 100,250,250,90,500
21 200,250,250,250,1000
22 200,250,250,250,1000
23 200,250,250,250,1000
24 100,250,250,90,500
25 200,250,250,250,1000
26 200,250,250,250,1000
27 100,250,250,90,500
28 200,250,250,250,1000

```

الشكل (9) عينة من ملف البيانات التدريب

بعد الحصول على ملف البيانات التدريبية (Trace set) تم استخدام برنامج (Weka) من أجل تدريب خوارزميات التعلم الآلي الخاضع للإشراف باستخدام تقنية التصنيف فكانت النتائج كما يبين الشكل (10).

Accuracy=1

classified	best	very_good	good	bad	worst
best	444	0	0	0	0
very_good	0	67	0	0	0
good	0	0	427	0	0
bad	0	0	0	7143	0
worst	0	0	0	0	554

مصفوفة التشويش لخوارزمية شجرة القرار

Accuracy=0.9386

classified	best	very_good	good	bad	worst
best	444	0	0	0	0
very_good	0	67	0	0	0
good	0	0	427	0	0
bad	0	0	0	7079	94
worst	0	0	0	217	337

مصفوفة التشويش لخوارزمية المتجهات الداعمة

Accuracy=0.870

classified	best	very_good	good	bad	worst
best	444	0	0	0	0
very_good	0	67	0	0	0
good	0	0	427	0	0
bad	0	0	0	6888	255
worst	0	0	0	217	337

مصفوفة التشويش لخوارزمية الانحدار اللوجستي متعدد الحدود

Accuracy=0.968

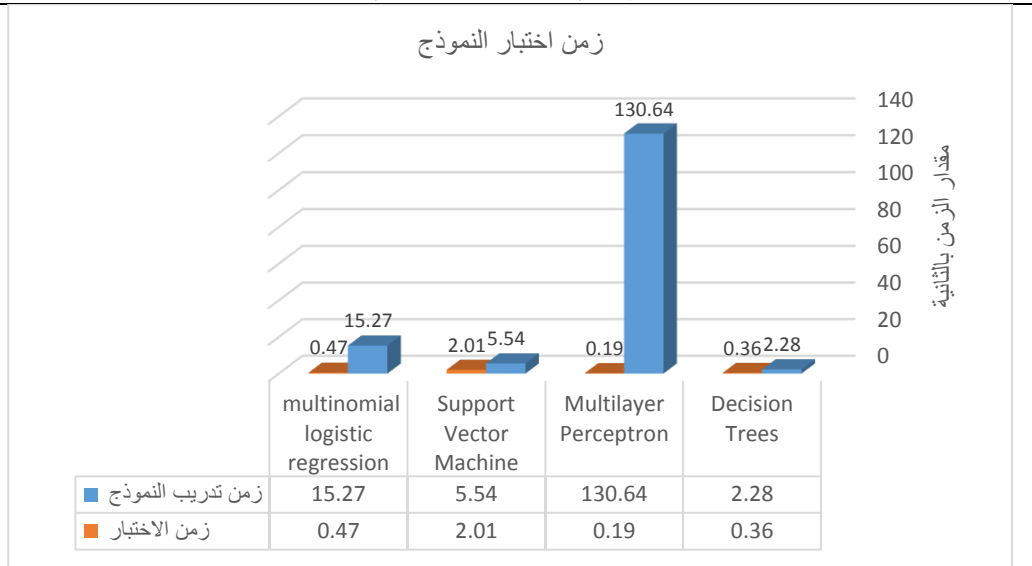
classified	best	very_good	good	bad	worst
best	444	0	0	0	0
very_good	0	67	0	0	0
good	0	0	427	0	0
bad	0	0	0	7049	94
worst	0	0	0	0	554

مصفوفة التشويش لخوارزمية الشبكات العصبونية ذات طبقتين

الشكل (10) مصفوفة التشويش لخوارزميات التعلم الآلي

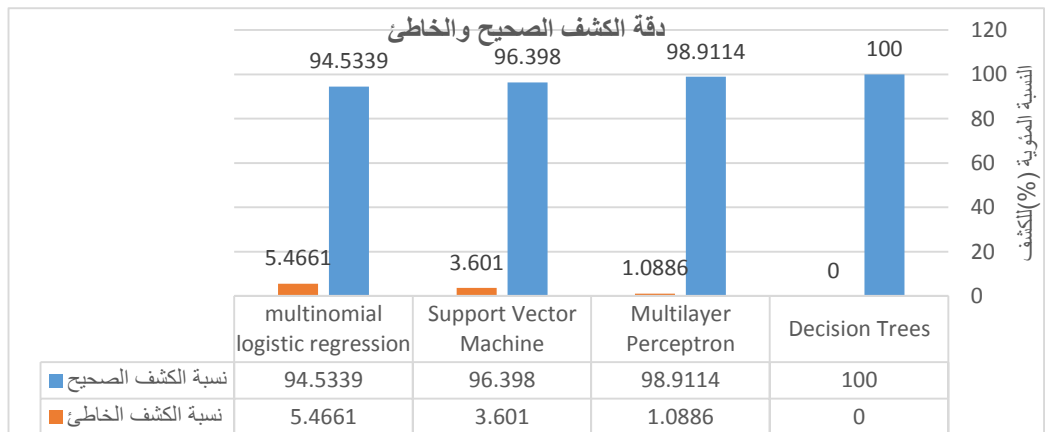
نلاحظ من خلال مصفوفة التشويش أن خوارزمية أشجار القرار تفوقت على بقية الخوارزميات من حيث دقة الكشف، وبالنسبة لزمّن تدريب واختبار النماذج على البيانات التدريبية كانت النتائج كما يبين الشكل (11)

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات أنترنت الأشياء



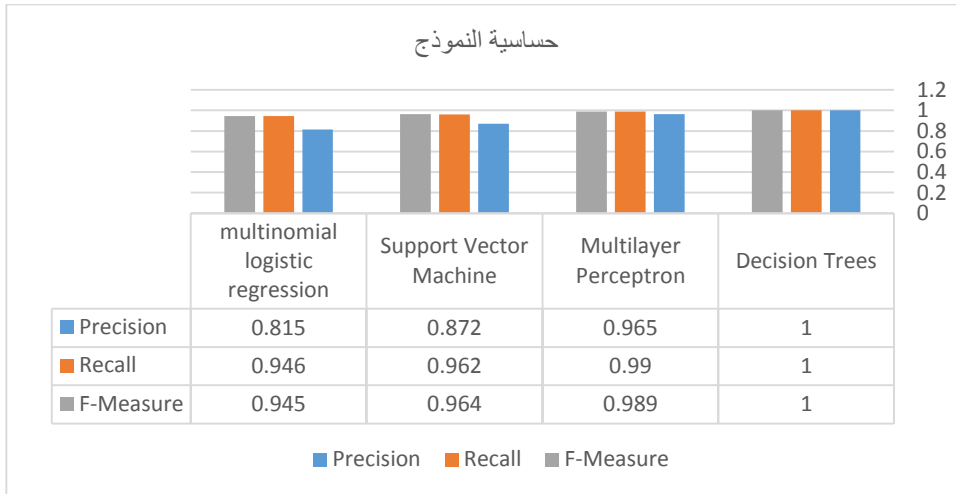
الشكل (11) يبين زمن اختبار النموذج

نلاحظ من الشكل (11) أن زمن التدريب كان أقل ما يمكن بالنسبة لخوارزمية شجرة القرار (Decision Tree) وذلك بالمقارنة مع بقية الخوارزميات إما بالنسبة لزمن اختبار النموذج على بيانات الاختبار حققت خوارزمية الشبكات العصبونية (ML_ANN) أقل زمن بمقارنة مع خوارزمية (Multinomial Logistic Regression, Decision Tree) ولكن سيئة هذه الخوارزمية احتياجها لزمن تدريب كبير جداً وبالنسبة لكل من مقياس دقة كشف على بيانات الاختبار كانت النتائج كما يبين الشكل (12):



الشكل (12) يبين نسبة دقة الكشف لنماذج التعلم الآلي

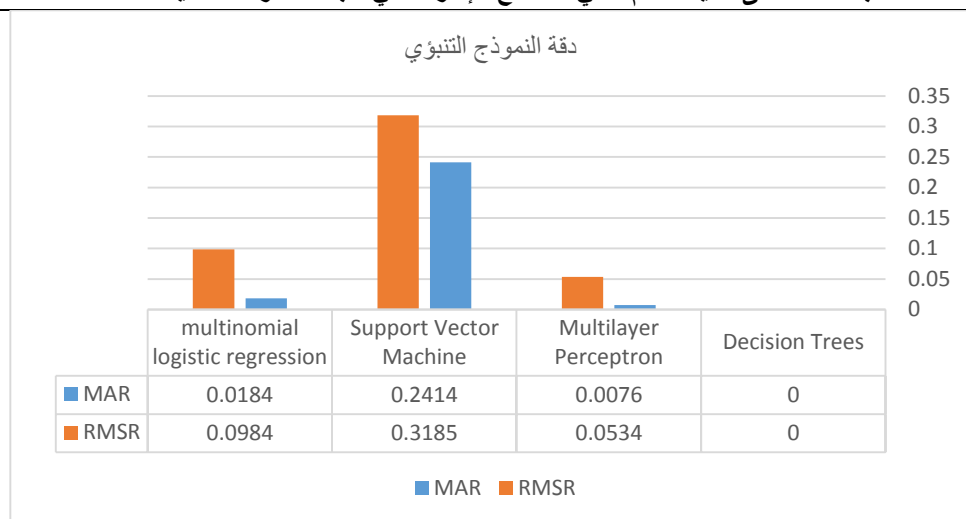
نلاحظ من الشكل (12) تفوق خوارزمية شجرة القرار في عملية الكشف عن البيانات والتي تدخل في عملية تقييم وكشف صلات الاتصال متغيرة السمات، حيث نلاحظ من الشكل أن خوارزمية شجرة القرار تفوقت وأعطت نسبة كشف صحيحة (100%).
إما بالنسبة لمقاييس حساسية النموذج وهي (الانضباط (Precision)، الاستدعاء (Recall)، ودرجة (F1_Measure)) كانت النتائج كما يبين الشكل (13).



الشكل (13) مقاييس حساسية خوارزميات التعلم الآلي

نلاحظ من الشكل (13) تفوق خوارزمية شجرة القرار بالنسبة للمقاييس الحساسية حيث كل ما كانت القيمة أقرب إلى الـ (1) كانت حساسية النموذج في عملية التميز بين البيانات التدريبية أكثر دقة وبالنسبة لدقة النموذج التنبؤي من أجل قوة الكشف كانت النتائج كما يبين الشكل رقم (14) وذلك لكل من مقياسي التنبؤ (MAR, RMSR).

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL) بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات أنترنت الأشياء



الشكل (14) يبين الدقة التنبؤية لخوارزميات التعلم الآلي

نلاحظ من الشكل (14) أن خوارزمية شجرة القرار أعطت أفضل النتائج وذلك بالنسبة لكل من القياس (MAR RMSR) حيث أعطت الخوارزمية قيم صفرية وهذا يدل على أن دقة التنبؤ والكشف على بيانات الاختبار كانت بنتيجة عالية.

❖ الاستنتاجات:

- ✓ بالنسبة لزمن التدريب والاختبار ودقة الكشف (Accuracy) تفوقت خوارزمية شجرة القرار على بقية الخوارزميات المدروسة بشكل كبير.
- ✓ بالنسبة لطريقة عمل الخوارزمية في كشف جودة وصلات الاتصال في شبكات أنترنت الأشياء حققت خوارزمية شجرة القرار أعلى دقة بالمقارنة مع بقية الخوارزميات المدروسة وهذا ما بينه الشكل (12).
- ✓ بالمقارنة بين الخوارزميات المدروسة عن طريق مقاييس الأداء نلاحظ أن تفوق خوارزمية شجرة القرار بشكل كبير وهذا مناسب لدراستنا عن طريق عمل تصنيف (Classification) لوصلات الاتصال وكشف الوصلات ذات الموثوقية العالية عن الوصلات السيئة.

(10) تطوير بروتوكول التوجيه (RPL) باستخدام خوارزمية أشجار القرار (Decision Tree):

في هذه الجزء من البحث تم استخدام برنامج (EmBML) وذلك لتحويل النموذج المدرب باستخدام خوارزمية أشجار القرار المولد عن طريق برنامج (WEKA) إلى برنامج متوافق مع موارد المتحكمات الصغيرة بلغة (C) وتم دمج هذا البرنامج مع بنية البروتوكول (RPL) داخل نظام التشغيل (Contiki OS)، كما تم التعديل على الملف (make File) الخاص بروتوكول (RPL) وذلك مع أجل إضافة الملف الخاص بالنموذج المولد حيث يصبح الملف بالشكل التالية:

```
CONTIKI_SOURCEFILES += rpl.c rpl-dag.c rpl-icmp6.c rpl-timers.c  
huml_rpl.c rpl-ext-header.c micro_tree.c
```

بعد عملية دمج الملف داخل بروتوكول التوجيه (RPL) وتشغيل المحاكى يمكننا تلخيص الخطوات الرئيسية المستخدمة في عملية اختيار المسار الأمثل باستخدام تقنية شجرة القرار (Decision Tree) بما يلي:

1. عند استلام رسالة (DIO) للاحد العقد داخل الشبكة، يتم استخراج قيم مقاييس جودة وصلة الاتصال (ETX, Remain Energy, Hop Forwarding ratio)
2. يتم إضافة عقدة مرسل رسالة (DIO) إلى قائمة الآباء المرشحين.
3. يتم استدعاء التابع (classify) الخاص بخوارزمية أشجار القرار لتحديد الأب الأنسب من بين قائمة الآباء المرشحين حيث يقوم التابع بحساب درجة الجودة لكل عقدة أب.
4. يتم اختيار عقدة الأب التي تملك أعلى درجة جودة ومقارنته بدرجة جودة عقدة الأب الحالية.
5. نتحقق إذا كانت درجة الجودة للعقدة المرشحة كأب أفضل تتجاوز درجة جودة عقدة الأب الحالية بقيمة عتبة معينة تم تحديدها بقيمة أكبر من (150) وهي القيمة التي تعبر عن قيمة الزيادة في درجة جودة الاب وتم اختيارها بالاعتماد على متوسط قيم التصنيف.
6. إذا كانت درجة الجودة للعقدة المرشحة زائد قيمة العتبة أكبر من درجة جودة العقدة الحالية، يتم اختيار هذه العقدة كأفضل عقدة أب أو تبقى عقدة الأب الحالي كأفضل عقدة أب.
7. نقوم بحساب رتبة عقدة الأب الجديدة باستخدام معادلة الرتبة بالعلاقة (11,12,13). من أجل المحافظة على ترتيب العقد داخل طوبولوجيا الشبكة [19].

$$RANK(N) = RANK(PN) + RankIncrease \quad (11)$$

$$RankIncrease = MinHopRankIncrease * step \quad (12)$$

$$Step = \alpha_1 * ETX(N, PN) + \alpha_2 * fw_ratio \quad (13)$$

$$\alpha_i \in R^+, i \in \{1,2\} \text{ and } \sum_{i=1}^3 \alpha_i = 1$$

8. أرسل رسائل DAO للعقدة المجاورة من أجل أعلامها بالتغيرات.

9 مقاييس تقييم الأداء في بروتوكول التوجيه (RPL):

في هذا البحث استخدمنا خمسة مقاييس أداء قياسية [19] وذلك من أجل تقييم أداء جودة اختيار المسار داخل البروتوكول (RPL) وذلك من أجل تقييم خوارزمية التعلم الآلي المقترحة بالمقارنة مع الوظائف الغرضية الأساسية (MORHF, OF0)، وهذه المقاييس هي:

✓ وقت تقارب الشبكة (Network Convergence Time):

تحتاج العقد في شبكة إنترنت الأشياء إلى تشكيل طوبولوجيا من أجل التواصل بين العقد. لذلك فإن وقت إعداد الشبكة هو مقياس حاسم يجب تقييمه لأي بروتوكول توجيه، يتم تعريف وقت التقارب في لـ (RPL DAG) على أنه مقدار الوقت الذي تحتاجه جميع العقد التي يمكن الوصول إليها (من حيث الراديو) في الشبكة للانضمام إلى طوبولوجيا (DAG). يجب اعتبار هذا التقارب بمثابة وقت التقارب الأولي في شبكة RPL ذات العقد الثابتة من أجل الحصول على وقت التقارب في شبكة RPL، نحدد وقت إرسال DIO الأول من عقد العميل وانضمام DIO الأخير إلى DAG. يتم الحصول على وقت التقارب عن طريق طرح وقت انضمام DIO الأخير من وقت إرسال DIO الأول إلى DAG كما بالعلاقة رقم (14)[19].

$$\text{Convergence}_{\text{Time}} = (\text{Last_DIO_Join_DAG}) - (\text{First}_{\text{DIO}}_{\text{Send}}) \quad (14)$$

✓ حمل رسائل التحكم (Control Traffic Overhead) للشبكة:

يشمل هذه المقياس على رسائل DIO و DIS و DAO التي يتم إنشاؤها بواسطة كل عقدة، ومن الضروري حصر مقياس الحمل مع مراعاة الموارد النادرة في شبكات (LLN)، يتحكم بروتوكول (RPL) في رسائل التحكم الزائدة عن الحاجة من خلال استخدام مؤقتات التدفق. الهدف من هذا المقياس هو تحليل حمل الخوارزمية المقترحة المذكورة في استخدام موارد العقدة ويمكن حساب حمل رسائل التحكم عن طريق العلاقة رقم (15)[19]:

$$\text{control_traffic_overhead} = \sum_{i=1}^n \text{DIO}(I) + \sum_{i=1}^m \text{DIS}(k) + \sum_{i=1}^z \text{DAO}(i) \quad (15)$$

✓ استهلاك الطاقة (Energy Consumption):

من أجل حساب استهلاك الطاقة سوف نستخدم آلية نظام Power trace المتاحة في نظام التشغيل (ContikiOS) وهو نظام لتحديد ملامح الطاقة على مستوى الشبكة للشبكات اللاسلكية منخفضة الطاقة التي تقدر استهلاك الطاقة لوحدة المعالجة المركزية ونقل الحزم والاستماع. تحتفظ هذه الآلية بجدول للمدة الزمنية التي كان فيها مكون مثل وحدة المعالجة المركزية وجهاز الإرسال اللاسلكي قيد التشغيل. بناءً على هذا الحساب، يتم حساب الطاقة المستهلكة بالميلي وات عن طريق العلاقة رقم (16,17) حيث فولت التغذية (3 فولت) كما يلي [1][22]:

$$power(mW) = Current * voltage \quad (16)$$

ومن أجل حساب التيار المستهلك نستخدم العلاقة رقم (13)

$$Current = \frac{tx * current_{tx} + Rx * current_{rx} + Cpu * Current_{cpu} + Cpu_{idel} * Current_{cpuidel}}{Rtimer_ARCH_Second} \quad (17)$$

حيث قيم:

TX: زمن تشغيل المرسل الراديوي في العقدة بوحدة أجزاء الساعة (نبضة).

RX: زمن تشغيل المستقبل الراديوي في العقدة بوحدة أجزاء الساعة (نبضة).

CPU: زمن تشغيل وحدة المعالجة المركزية بوحدة أجزاء الساعة (نبضة).

cpu_{idel}: زمن تشغيل وحدة المعالجة المركزية في وضع الطاقة المنخفض بوحدة أجزاء الساعة (نبضة).

Rtimer_ARCH_Second: وهو عدد نبضات المعالج في الثانية الواحدة (32768 نبضة).

نبين في الجدول رقم (3) قيم التيارات المأخوذة من النشرة التعريفية للعقدة (Sky Mote) [23] وهي كما يلي:

الجدول (3) قيم التيارات المأخوذة من النشرة التعريفية للعقدة.

نوع الوحدة	كمية الاستهلاك
current _{cpu}	1.8 mA
current _{rx}	21.8 mA
current _{tx}	19.5 mA
current _{idel}	54.5 μ A

✓ زمن تأخر وصول الحزمة (Packet Latency):

يتم تعريف زمن الوصول على أنه مقدار الوقت الذي تستغرقه حزمة البيانات المرسل من عقدة ما للوصول إلى العقدة الهدف أو عقدة مصب الشبكة (Sink Node) حيث نقوم بحساب متوسط زمن الوصول لجميع حزم البيانات في الشبكة من جميع العقد إلى عقدة الجذر ويتم حساب قيمة المقياس عن طريق العلاقة رقم (18) وحساب معدل زمن وصول الحزمة بالعلاقة (19) [19].

$$\text{Total_Latency} = \sum_{i=1}^n (\text{recv_time}(i) - \text{send_time}(i)) \quad (18)$$

$$\text{Averagy_Latency} = \text{Total_Latency} / \text{Total_packets_received} \quad (19)$$

حيث (n) هو العدد الإجمالي للحزم المستلمة بنجاح.

✓ متوسط معدل تسليم الحزمة (PDR Average):

ويتم تعريفه على أنه عدد الحزم المستلمة في عقدة مصب الشبكة (sink node) إلى عدد الحزم المرسل إلى مصب الشبكة من بقية العقد حيث نقوم بأخذ متوسط قيمة (PDR) لجميع الحزم التي تم استلامها بنجاح في عقدة المصب (Sink) ولحساب متوسط نسبة تسليم الحزمة (PDR)، نقيس عدد الحزم المرسل من جميع العقد إلى عقدة المصب (Sink) ونقسمها على عدد الحزم المستلمة بنجاح في عقدة المصب (Sink) كما تبين بالعلاقة رقم (20) [19].

$$\text{Average PDR} = \left(\frac{\text{Total_packet_received}}{\text{Total_packet_send}} \right) * 100 \quad (20)$$

10 القسم العملي:

تركز معظم الأبحاث على المحاكاة بدلاً من التجارب وذلك بسبب توفر العتاد البرمجي وسهولة الوصول إليه، على الرغم من أن التقييم المعتمد على عمليات المحاكاة لا يعكس الواقع لأنه لا يأخذ بعين الاعتبار مجموعة من القيود مثل تغير المناخ والعوائق الطبيعية والتفاعلات البشرية ولكنه الخيار الوحيد المتوفر في كثير من الأحيان للطلبة والباحثين [25].

1-10 نظام التشغيل (Contiki_OS):

يعتبر نظام التشغيل (ContikiOS) والذي تم تطويره من قبل المعهد السويدي لعلوم الكمبيوتر البرنامج المستخدم على نطاق واسع في شبكات IoT وWSN وذلك لأنه مفتوح المصدر، وهو نظام التشغيل الأكثر استخداماً في مجال الأبحاث الخاصة بشبكات الحساسات اللاسلكية وشبكات انترنت الأشياء، كما يستخدم (Contiki) في العديد من الأنظمة، مثل المراقبة الصناعية وأنظمة الإنذار ومراقبة المنزل عن بعد ومراقبة الإشعاعات [26] [19].

10-2 محاكي الشبكات (COOJA):

سوف نستخدم محاكي الشبكات (COOJA) وذلك بسبب كونه المحاكي الأكثر استخداما في الابحاث وأجراء التجارب في شبكات منخفضة الطاقة بنسبة تصل إلى (61%) وهو محاكي قائم على لغة (java) ويعتمد على نظام التشغيل (ContikiOS) بشكل أساسي ويمكن اعتباره أداة هجينة تُستخدم كأداة محاكاة للعقد الافتراضية (Simulator) وأداة محاكاة للأجهزة الحقيقية (Emulator) بمعنى آخر أن أدائه أقرب للواقع لأنه يقوم بتشغيل حقيقي للأجهزة الموجودة في الشبكة مما يجعل النتائج التي نحصل عليها دقيقة أكثر وتحاكي الواقع [7].

10-3 تهيئة الشبكة والتجارب:

في هذه التجارب سوف نستخدم عقد من نوع (sky mote) [23] والتي تتميز بمعدل نقل عالي للبيانات (250kb/s) وذاكرة دائمة ب(128KB) وذاكرة وصول عشوائي تقدر (35kb)، قمنا بالتعديل على بعض هذه الملفات داخل البروتوكول (RPL) وإضافة توابع جديدة على الملفات ومن هذه الملفات نذكر:

RPL.h, RPL-icmp6.c, RPL-)

(conf.h, HUML_OF.C, energest.c, RPLdag.c, uip6.c, microtree.c

ولأجراء التجارب قمنا باستخدام طوبولوجيا عشوائية مؤلفة من نموذج تطبيق UDP Contiki يتألف من ملفين الملف (udp-client.c) وهو الملف الذي تم تنزيله على جميع عقد الشبكة (client node) بحيث تقوم العقدة (client node) بإرسال رسالة نصية ضمن فاصل زمني محدد الى موجه الشبكة أو جذر الشبكة (DODAG ROOT) والذي يحوي على الملف (udpsrvr.c) عند وصول المعلومات المرسله من عقدة الزبائن (client node) يقوم موجه الشبكة بأعلام العقدة التي تم استلام المعلومات منها بأن المعلومات تم استلامها بشكل صحيح ومن أجل اختبار الخوارزمية المقترحة تم إجراء التجارب على شبكات مختلفة الكثافة :

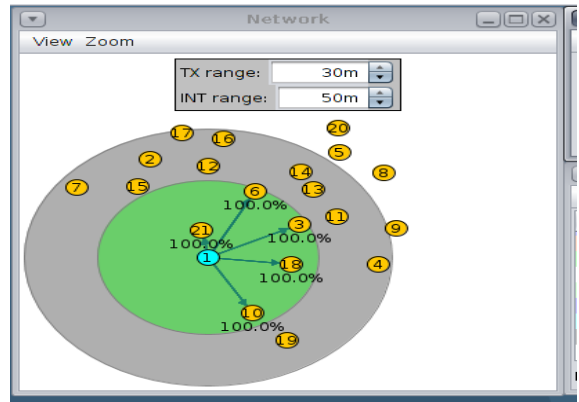
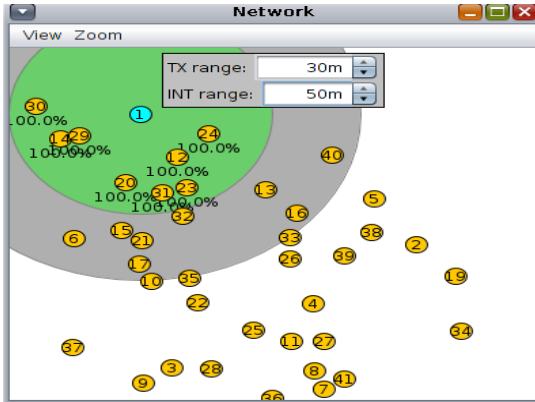
- ✓ منخفضة الكثافة (20 عقد) كما يبين الشكل (15).
- ✓ متوسطة الكثافة (40 عقدة) كما يبين الشكل (16).
- ✓ عالية الكثافة (60 عقدة) كما يبين الشكل (17).

**تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات أنترنت الأشياء**

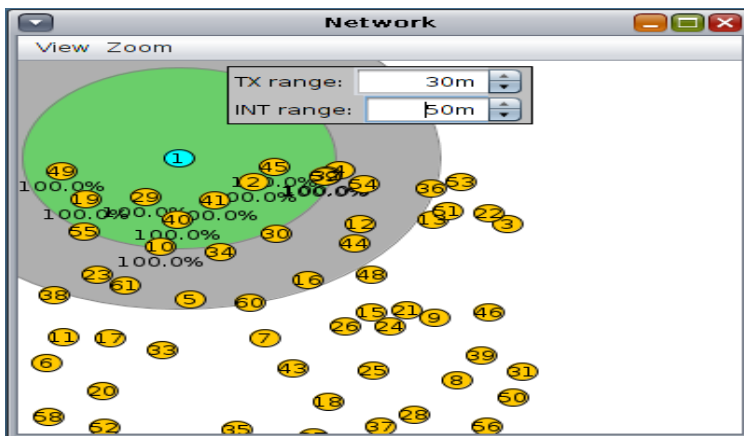
وللإدخال الضياع في الوسط اللاسلكي استخدمنا النموذج (UDGM) والذي يعرف الضياع بناء على المسافات النسبية بين العقد في الوسط الراديوي [27]، يمكننا من خلاله ضبط نسبة الإرسال والاستقبال، حيث قمنا بالمحافظة على نسبة الإرسال (TX-100%) (لأننا نهتم بالضياع عند طرف المستقبل، بينما قمنا بتغيير نسبة الاستقبال (RX) والتي تُعرف بأنها نسبة نجاح استقبال الطرود عند المستقبل على نسبة (RX 80%)، ونشير إلى الضياع في الوسط اللاسلكي، كما تم تحديد نطاق الإرسال على 30 متر ونطاق التداخل على 50 متر، تم استخدام كل من التطبيق الملحق مع المحاكى (Cooja) المسمى (Contiki Simulation Script Editor) لقياس وقت المحاكاة وإيقاف المحاكاة بعد الوقت المحدد، ينشئ هذا البرنامج المساعد ملف سجل (COOJA_log file) لجميع مخرجات المحاكاة والتي سنقوم بتحليلها في نهاية المحاكاة باستخدام برنامج نصي مكتوب بلغة (Perl) وباستخدام برنامج (MS Excel) نقوم برسم المخططات البيانية. قمنا بتشغيل المحاكاة لمدة (45) دقيقة لكل سيناريو وتم إعادته لمدة أربع مرات وأخذ متوسط النتائج، يوضح الجدول رقم (3) المعاملات المستخدمة في التجارب المختلفة.

جدول رقم (3) ضبط بارامترات التجارب

Parameters	Description
No of Nodes	20,40,60
Simulation Model	UDGM (Unit Disk Graph Model)
Simulation time	45 Mins.
Data Rate	250kps
Node Arrangement Random	Random
Operating System	Contiki2.7
Simulator	Cooja
OF	MORHF_ETX, OF0, HUML_OF
DIO Min	12
DIO Doublings	8
Radio Duty-Cycling Interval	16HZ
RX Ratio	80%
TX Ratio	100%
TX Range	30m
Interference Range	50m
Packet Analyzer	Perl code



الشكل رقم (15) توزيع العقد العشوائي (20 عقدة).
 الشكل رقم (16) توزيع العقد العشوائي (40 عقدة).



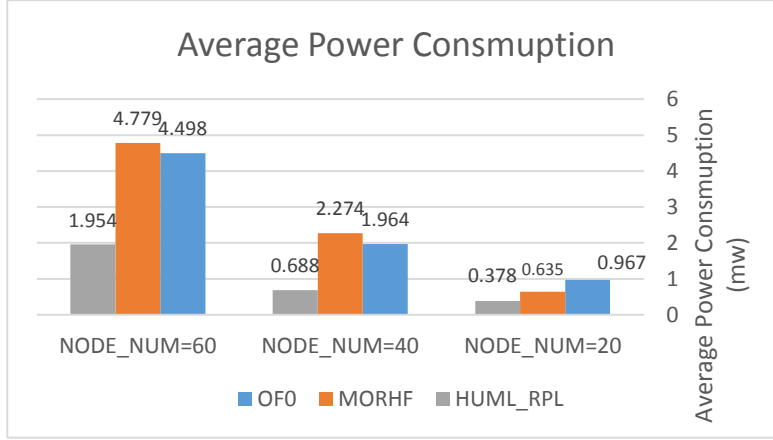
الشكل رقم (17) توزيع العقد العشوائي (60 عقدة).

4-10 مناقشة النتائج:

✓ معدل استهلاك الطاقة:

تم تحليل النتائج وتسجيل قيم معدل استهلاك الطاقة كما يبين الشكل (18). بالنسبة لمقياس معدل استهلاك الطاقة نلاحظ بالنسبة للمقياس الجديد (HUML_RPL) كانت كمية استهلاك أقل من المقياس (MORHF_ETX, OF_0) من أجل جميع السيناريوهات نلاحظ تفوق المقياس الجديد المقترح في عملية التقليل من استهلاك الطاقة وذلك بسبب اختيار مسارات ذات وثوقية عالية وتحوي معدل تمرير رزم أقل كما أن عملية اختيار المسارات تتم بصورة أسرع من الوظيفة الغرضية

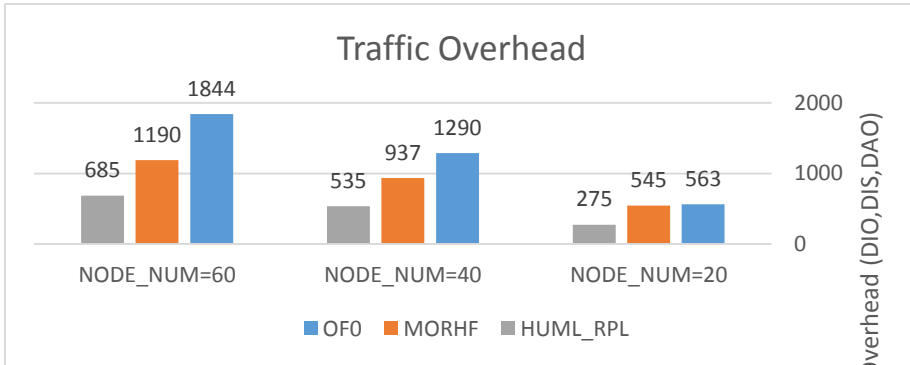
(MORHF) وذلك بسبب حذف بعض التوابع التي تقوم بمقارنة الأبناء المفضلين بشكل كامل مما يخفض من عملية استهلاك الطاقة بالشبكة ككل.



الشكل رقم (18) يوضح معدل استهلاك الطاقة لجميع العقد.

✓ مقياس حمل رسائل التحكم:

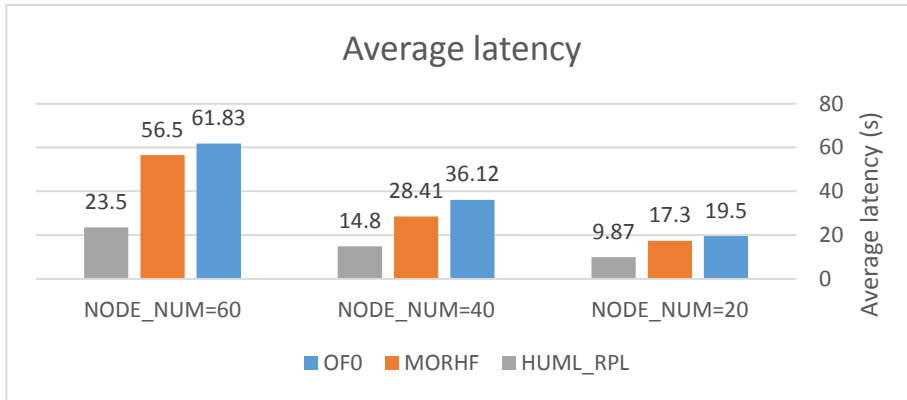
بالنسبة لمقياس حمل رسائل التحكم، كانت النتائج كما يبين الشكل (19) نلاحظ تفوق المقياس الجديد في تقليل من عدد رسائل التحكم (DIS,DIO,DAO) وذلك من أجل عدد عقد مختلف وتوزع عشوائي وذلك يعود بسبب اختيار المقياس مسارات أكثر وثوقيه واختيارها عقدة ذات طاقة كبيرة وتجاوز العقد ذات حركة البيانات الكثيفة والطاقة المنخفضة والتي تؤدي إلى خروج بعض العقد عن الخدمة بسبب نفاذ طاقتها وزيادة عدد رسائل التحكم بسبب تغير الأبناء المفضلين بشكل دائم.



الشكل رقم (19) يبين حمل رسائل التحكم.

✓ زمن تأخر وصول الرزم:

بالنسبة لمقياس متوسط معدل تأخر الرزم كانت النتائج كما يبين الشكل (20)، نلاحظ تفوق المقياس المقترح في التقليل من معدل زمن التأخير في عملية تسليم البيانات وذلك من أجل عدد عقد مختلفة، وذلك يعود إلى تجاوز العقد ذات معدل تمرير الرسائل العالي والذي يزيد من زمن انتظار الرسائل في ذاكرة العقدة حتى وصولها إلى العقدة الجذر، كما أن عملية اختيار مسارات لا تحتوي على عقد تشكل عنق زجاجة أو تحتوي على معدل تمرير بيانات عالي الذي بدوره يقلل من زمن الوصول الرسائل إلى العقد الجذر.

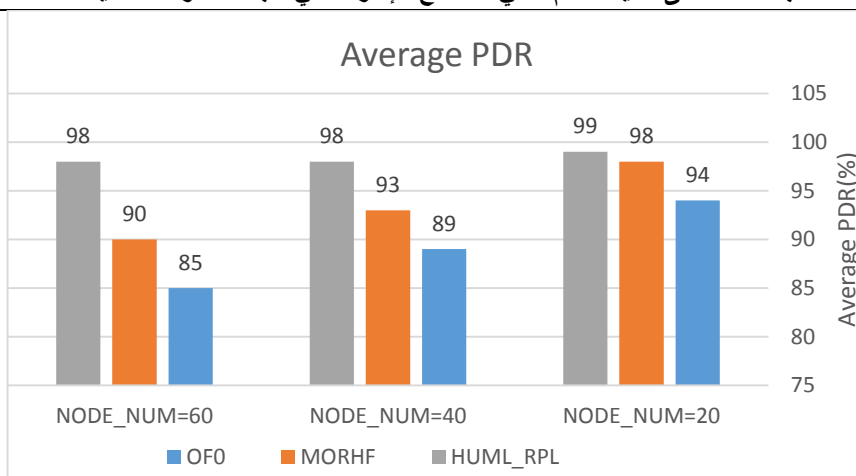


الشكل رقم (21) يبين متوسط معدل تأخر الرزم.

✓ مقياس متوسط معدل تسليم الرزم:

بالنسبة لمقياس متوسط معدل تسليم الرزم كانت النتائج كما يبين الشكل (22)، نلاحظ تفوق المقياس الجديد في عملية تسليم الرزم بالمقارنة مع كل الوظائف الغرضية (OF_0, MORHF_ETX) وذلك من أجل عدد عقد مختلفة وذلك بسبب وثوقه اختيار المسارات وتجاوز العقد المشغولة والتي تزيد من ضياع حزم البيانات وتزيد التأخير الزمني .

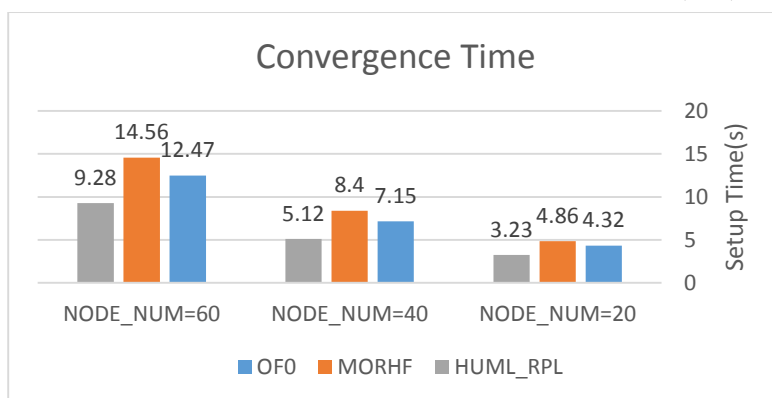
تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء



الشكل رقم (22) يوضح متوسط معدل تسليم الرزم.

✓ مقياس وقت تقارب الشبكة:

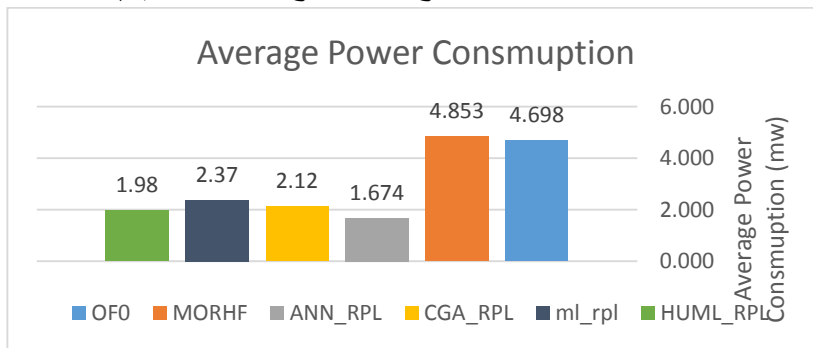
بالنسبة لمقياس وقت تقارب الشبكة كانت النتائج كما يبين الشكل (23)، نلاحظ أن المقياس الجديد يقلل من زمن تقارب الشبكة بمقدار (2-4 ثانية) بالمقارنة مع الوظائف الغرضية الأساسية للبروتوكول (RPL) وذلك بالنسبة لعدد عقد مختلفة، ويرجع ذلك إلى اعتماد المقياس الجديد على خوارزمية تعلم آلي تعمل بشكل أسرع من الخوارزمية التقليدية والتي تحتاج إلى استدعاء أكثر من تابع وظيفي من أجل الانضمام للطوبولوجيا داخل الشبكة مما يساهم ذلك في عملية التقليل في وقت التقارب داخل الشبكة بشكل كامل.



الشكل رقم (23) يبين وقت تقارب الشبكة

5-10 مقارنة المقياس الجديد (HUML_RPL) مع بعض الدراسات السابقة:

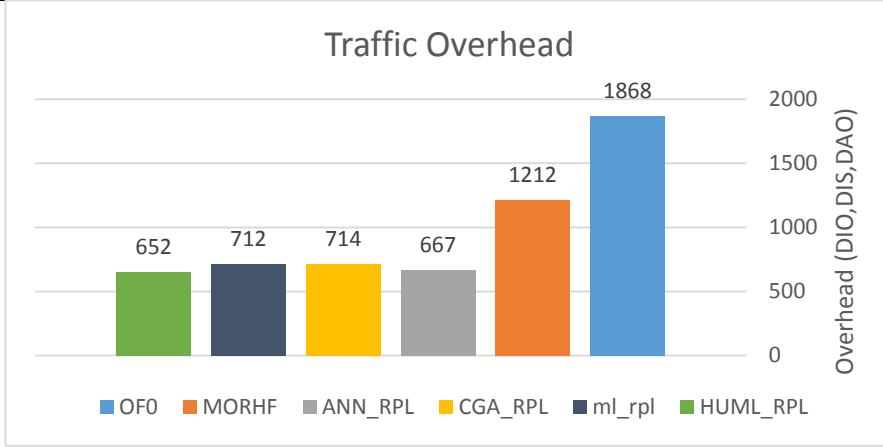
في هذه الفقرة قمنا بعمل مقارنة مع بعض الدراسات السابقة التي اقترحت تقنيات إنشاء توابع وظيفية ومقاييس جودة وصلة اتصال مركبة تعتمد على خوارزميات التعلم الآلي الخاضع للأشراف كما تم قياس أداء الأعمال المقترحة بنفس مقاييس الأداء وذلك من أجل عدد عقد ثابت (70 عقدة) وطوبولوجيا عشوائية، ومن هذه الأعمال الدراسة المرجعية رقم [12] التي اقترحت إنشاء بروتوكول توجيه محسن (ML_RPL) يعتمد على خوارزمية أقرب الجيران (K-Nearest Neighbor)، كما اقترحت الدراسة المرجعية رقم [14] تحسين عملية اختيار وصلات الاتصال داخل عقد الشبكة وتم طرح نسخة محسنة تسمى (CGA-RPL) تعتمد على الخوارزمية الجينية إما الدراسة المرجعية [15] اقترحت استخدام الشبكات العصبونية الصناعية (ANN_RPL) من أجل تحسين عملية اختيار العقدة الأب الأفضل داخل الشبكة فكانت النتائج كما توضح الأشكال رقم (24,25,26,27)



الشكل رقم (24) يوضح معدل استهلاك الطاقة (70 عقدة).

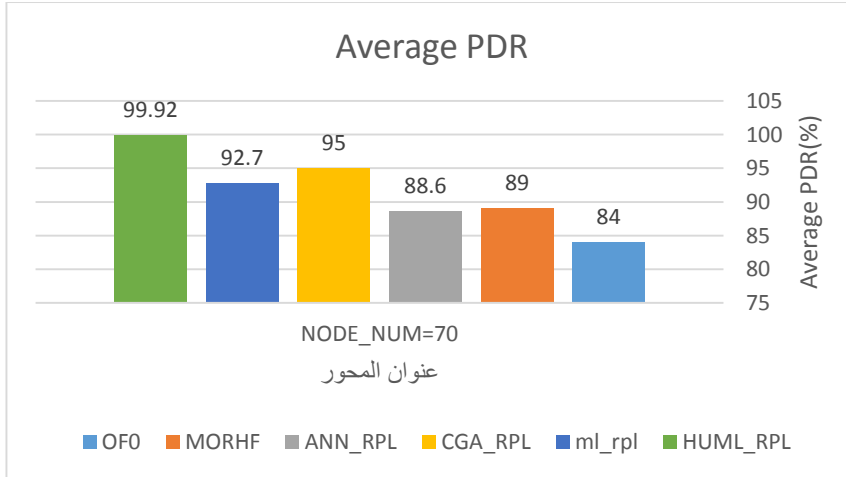
✓ بالنسبة لمعدل استهلاك الطاقة كما يبين الشكل (24) كانت النتائج متقاربة بمقارنة مع كل من المقاييس (HUML_RPL, ANN_RPL, CGA_RPL) إما بالنسبة للمقياس (ML_RPL) الذي يعتمد على خوارزمية أقرب الجيران كانت عملية استهلاك الطاقة بشكل أعلى.

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء



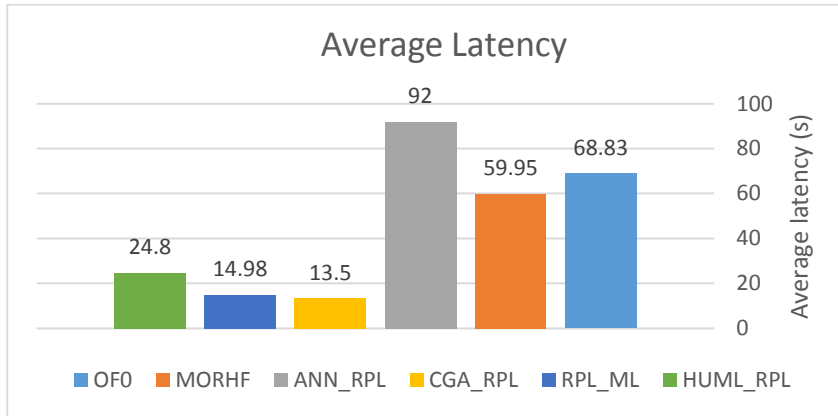
الشكل رقم (25) يبين حمل رسائل التحكم ل(70) عقدة.

✓ بالنسبة لمقياس حمل رسائل التحكم كما يوضح الشكل (25) اعطى المقياس المقترح (HUML_RPL) أفضل النتائج في عملية التقليل من رسائل التحكم والتي تسبب زيادة استهلاك الطاقة وزيادة حركة البيانات داخل الشبكة مما يؤدي الى زيادة الضياعات وزيادة تأخر وصول رزم البيانات داخل الشبكة حيث ساعد المقياس المقترح في عملية اختيار مسارات ذات ثباتيه عالية من حيث كمية الطاقة ومعدل إعادة تمرير الرزم مما ساعد في عملية التخفيف في عملية تبديلات المسارات والذي بدوره يقلل من رسائل التحكم.



الشكل رقم (26) يوضح متوسط معدل تسليم الرزم ل (70) عقدة.

✓ من الشكل (26) نلاحظ أن المقياس الجديد (HUML_RPL) أعطى أفضل النتائج من حيث عملية معدل تسليم الرزم داخل الشبكة وذلك بمقارنة مع بقية التتابع الوظيفية ويعود ذلك في بسبب اعتماد الحل المقترح على خوارزمية أشجار القرار والتي تعتبر من الخوارزميات سريعة التنفيذ كما أن اختيار مقياس معدل توجيه الرزم يساعد في عملية تجاوز العقد المشغولة والتي تحتوي على كثافة تمرير بيانات عالية مما يؤدي إلى ضياع وتأخير بعض الرزم.



الشكل رقم (27) يبين متوسط معدل تأخر الرزم ل(70) عقدة.

✓ نلاحظ من الشكل (27) أن المقياس الجديد (HUML_RPL) أعطى نتائج أفضل من حيث عملية التقليل من زمن تأخر وصول الرزم وذلك بالمقارنة مع كل (MORHF, OF0, ANN_RPL, ML_RPL, CGA_RPL) ويعود ذلك إلى تجنب المسارات التي تعاني من حركة بيانات عالية مما يخفف ضياع وتأخر البيانات داخل العقد بالإضافة إلى اعتماد الحل المقترح على خوارزمية شجرة القرار والتي تعتبر من خوارزميات التعلم الآلي ذات سرعة تنفيذ عالية.

11 الاستنتاجات والتوصيات المستقبلية:

يعد التوجيه عاملاً مهماً في شبكة انترنت الأشياء لأنه ينفذ قرارات توجيه الطرود حيث أن اختيار المسارات السيئة في التوجيه يسبب المزيد من عمليات إعادة الإرسال وإهدار الموارد، بناءً على نتائج المحاكاة وتحليل الأداء نجد أن المقياس الجديد (HUML_OF) والذي يعتمد على تقنية التعلم الآلي الخاضع للأشراف يقدم أداءً متفوقاً بشكل عام بمقارنة مع الوظائف الغرضية الأساسية (MORHF, OF0) في عملية اختيار جودة وصلة الاتصال التي تصلنا إلى العقد الأب

**تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL)
بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات أنترنت الأشياء**

الأفضل، يأتي هذا التفوق من عملية اختيار المسار الأفضل بالنسبة لكمية الطاقة ومعدل توجيه الرزم الأقل للعقدة الأب والقيمة الأفضل للمقياس (ETX) وعدد القفزات الأقل وبالتالي اختيار أفضل المسارات لتوجيه الطرود وتوزيع الحمل والطاقة بشكل أفضل داخل عقد الشبكة. على الرغم من التحسينات التي أجريت في البحث، إلا أننا يمكننا طرح المزيد من التوقعات لتحسين هذا البحث في المستقبل منها:

1. طرح مقاييس جودة وصلة جديدة تدخل في تدريب نموذج التعلم الآلي.
2. تغيير طريقة اختيار الأب المفضل عن طريق طرح تقنية التعلم الآلي التزايدية (Incremental Learning) الذي يقوم بتدريب نفسه خلال فترة عمل الخوارزمية.
3. يمكن تطوير خوارزمية التعلم باستخدام تقنية التعلم العميق (Deep Learning).
4. دراسة أداء الخوارزمية من أجل مقاييس مختلفة وفي سيناريوهات مختلفة مع زيادة عدد العقد.

12 جدول الاختصارات:

Internet Of Things	IOT	انترنت الأشياء
Ipv6 Routing Protocol For Low-Power And Lossy Networks.	RPL	بروتوكول توجيه الشبكات المنخفضة الطاقة والتي تحتوي ضياعات
HOMS UNIVERSITY RPL	HUML_RPL	بروتوكول التوجيه المحسن خاص بجامعة حمص
Expected Transmission Count	ETX	عدد الإرسال المتوقع
Routing Over Low Power And Lossy Networks	ROLL	مجموعة عمل التوجيه عبر الشبكات منخفضة الطاقة والتي يكون فيها ضياع
Internet Engineering Task Force	IETF	فريق مهام هندسة الانترنت
Low Power And Lossy Network	LNN	الشبكات منخفضة الطاقة والتي يكون فيها ضياع
Objective Function	OF	تابع وظيفي
ZERO Objective Function	Of0	الوظيفة الغرضية الصفرية
Minimum Rank Hysteresis Objective Function	MORHF	الوظيفية الغرضية لاستعادة استقرار الرتبة
Destination Oriented Directed Acyclic Graph	DODAG	رسم بياني لا حلقي موجه إلى الوجهة
Internet Control Message Protocol	ICMPV6	بروتوكول رسائل التحكم الإصدار السادس
DODAG Information Object	DIO	رسالة المعلومات
DODAG Information Solicitation	DIS	رسالة التماس المعلومات
Destination Advertisement Object	DAO	رسالة إعلان الوجهة
Directed Acyclic Graph	DAG	رسم بياني موجه
Support Vector Machine	SVM	آلة المتجهات الداعمة

Artificial Neural Network	ANN	الشبكات العصبية الاصطناعية
MultiLayer Precptron	MPL	خلايا متعدد الطبقات
Multinomial Logistic Regression	MLR	الانحدار المنطقي متعدد الحدود
CHAOTIC GENETIC ALGORITHM	CGA	الخوارزميات الجينية الفوضوية
Packet Delivery Ratio	PDR	معدل تسليم الرزم
Unit Disk Graph Model	UDGM	نموذج رسم بياني لاقطاع القرص الحدودي

13 المراجع باللغة الإنكليزية:

- [1] Gupta, Neeti, Anuradha Pughat, and Vidushi Sharma, 2021- "A critical analysis of RPL objective functions in internet of things paradigm." *Peer-to-Peer Networking and Applications* 14, no: 2187-2208.
- [2]MSP430, <https://www.ti.com/design-development/embedded-development/msp430-mcus.html>
- [3] Kechiche, Ines, Ines Bousnina, and Abdelaziz Same,2020-"A Review of RPL Objective Function based Enhancement Approaches, *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 5, pp. 201–211, doi: 10.25046/AJ050525
- [4] Dhumane, Amol, Rajesh Prasad, and Jayashree Prasad,2016-"Routing issues in internet of things: a survey." In *Proceedings of the international multicference of engineers and computer scientists*, vol. 1, pp. 16-18.
- [5] Dharmalingaswamy, Anitha, and Pitchai Latha,2023- "Additive metric composition-based load aware reliable routing protocol for improving the quality of service in industrial internet of things." *Int. Arab J. Inf. Technol.* 20, no. 6 :954-964.
- [6] "rfc6550", <http://www.rfc-editor.org/info/rfc6550>
- [7] Mardini, Khalil, and Emad Abdulsamad,2023-"Designing a Novel RPL Objective Function & Testing RPL Objective Functions Performance", *Bachelor Degree Project*.
- [8] Sharma, Himanshu, Ahteshamul Haque, and Frede Blaabjerg,2021- "Machine learning in wireless sensor networks for smart cities: a survey." *Electronics* 10, no. 9 1012.
- [9] Sarumathi, S., M. Vaishnavi, S. Geetha, and P. Ranjetha,2021- "Comparative Analysis of Machine Learning Tools: A Review." *International Journal of Computer and Information Engineering* 15, no. 6: 354-363.
- [10] Zhao, Yuxin, Shenghong Li, and Jia Hou,2015- "Link quality prediction via a neighborhood-based nonnegative matrix factorization model for wireless sensor networks." *International Journal of Distributed Sensor Networks* 11, no. 10: 828493.
- [11] Sindjoung, Miguel Landry Foko, and Pascale Minet,2019- "Wireless link quality prediction in iot networks." In *2019 8th International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN)*, pp. 1-6. IEEE.
- [12] Kuwelkar, Sonia, and Hassanali G. Virani,2023- "RPL protocol enhancement using K-Nearest Neighbor machine learning approach for IoT application", DOI: <https://doi.org/10.21203/rs.3.rs-2392682/v1>.
- [13] Santos, Carlos Lester Duenas, Ahmad Mohamad Mezher, Juan Pablo Astudillo León, Julian Cardenas Barrera, Eduardo Castillo Guerra, and Julian Meng, 2023- "ML-RPL: Machine

تحسين اختيار جودة وصلة الاتصال داخل بروتوكول التوجيه (RPL) بالاعتماد على تقنية التعلم الآلي الخاضع للإشراف في شبكات إنترنت الأشياء

- learning-based routing protocol for wireless smart grid networks." *IEEE Access* 11: 57401-57414.
- [14] Cao, Yanan, and Muqing Wu, 2018- "A novel RPL algorithm based on chaotic genetic algorithm." *Sensors* 18, no. 11 :3647
- [15] Kuwelkar, Sonia, and Hassanali G. Virani, 2023- "Enhancing the RPL Protocol Using an Artificial Neural Network for Sustainable IoT Infrastructure." *Sustainable Industry, Innovation and Infrastructure*: 95
- [16] Krstinic, Damir, Maja Braovic, Ljiljana seric, and Dunja Bozic-stulic, 2020- "multi-label classifier performance evaluation with confusion matrix." *Computer Science & Information Technology* 1:1-14.
- [17] Hodson, Timothy O, 2022- "Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not." *Geoscientific Model Development Discussions*: 1-10.
- [18] K. Venugopal and T. G. Basavaraj, 2022- "A Combined Metric Objective Function for RPL Load Balancing in Internet of Things." *International Journal of Internet of Things*, Vol. no. 1, doi: 10.5923/j.ijit.20221001.02
- [19] Ali, H, 2012- "A performance evaluation of RPL in Contiki: A cooja simulation-based study." *School of Computing, Blekinge Institute of Technology*.
- [20] Bouckaert, Remco R. Eibe Frank, Hall, Richard Kirkby, Peter Reutemann, Alex Seewald, and David Scuse, 2016- "WEKA manual for version 3-9-1." *University of Waikato: Hamilton, New Zealand*: 1-341.
- [21] Silva, Lucas Tsutsui, Vinicius MA Souza, and Gustavo EAPA Batista, 2019- "EmbML Tool: Supporting the use of supervised learning algorithms in low-cost embedded systems." In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pp. 1633-1637. IEEE.
- [22] Subramani, Poorana Senthilkumar, and Subramani Bojan, 2023- "Weighted Sum Metrics–Based Load Balancing RPL Objective Function for IoT." *Annals of Emerging Technologies in Computing (AETiC)* 7, no. 2: 35-55.
- [23] sky mote datasheet, <http://www.crew-project.eu>.
- [24] <https://scikit-learn.org>

14 المراجع باللغة العربية:

- [25] العمر منال، 2022- تحسين أداء خوارزمية التقطير في إنترنت الأشياء، مجلة جامعة البعث سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، المجلد 44، العدد 8، صفحة 65-92.
- [26] الكلزلي طريف، 2023- تحسين زمن تبديل العقد في بروتوكول RPL ضمن نظام التشغيل Contiki، مجلة جامعة البعث سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، المجلد 54، العدد 32، صفحة 73-92.
- [27] الشعار رهن، 2020- تطوير خوارزمية لتحسين أداء التوجيه في إنترنت الأشياء، رسالة ماجستير، المعهد العالي للعلوم التطبيقية والتكنولوجيا، الجمهورية العربية السورية.

The Effect of Varying Tempering Temperatures on the High Chromium Cast Iron Grinding Ball's Microstructure and Durability

Engineer: Mohamad Rami Al Jundi Dr. Abbad Kassouha*

Abstract

A laboratory work has carried out to observe the effect of the tempering temperature varying on the impact toughness of the White High Chromium Cast Iron, a drop ball test was implemented to determine the ball impact resistance by dropping the ball freely on a steel anvil of 63 HRC. a life of 4210 impacts was achieved after a 1000 C° Austeniteization followed by 500 C° tempering on a local made High Chromium grinding ball with 15.23 % Cr, this tempering process was tested at 550 C° and 600 C° where not reached more than 3200 impacts, the results of the grinding ball were proved that the applied heat treatment affected the microstructure and the hardness of the balls, the carbide volume has shown a critical rule on the impact toughness.

A carbide volume of 18.8 % was decreased till 9.223 % by the 1000 C° Austeniteization and 500 C° tempering, where its hardness became 50 HRC, this is resulted in a martensitic and carbides microstructure with a little amount of retained austinite, this enhanced martensitic microstructure

was shown after reducing the excessive carbide formation that gave the optimum toughness and hardness.

Key words:

Grinding balls , Drop ball , Toughness , Fracture , Hardness .

***PhD Student-Design and Production Engineering- Design and Production Engineering Department-Homs University-Homs-Syria.**

****Assistant professor- Design and Production Engineering Department-Homs University-Homs-Syria**

تأثير تغيير درجة حرارة الارجاع على بنية ومتانة الكرات الطاحنة المصنوعة من الحديد الصب عالي الكروم

م. محمد رامي الجندي د. عبّاد كاسوحة*

ملخص البحث

أجريت دراسة مخبرية للبحث في تأثير تغيير درجة حرارة عملية المعالجة الحرارية بالإرجاع على متانة التحطم للكرات الطاحنة المصنوعة من حديد الصب الأبيض عالي الكروم، كما أجري اختبار السقوط الحر للكرات من خلالها تعريضها للسقوط من ارتفاع 6m على بلاطة فولاذية مقساة حتى 63HRC لتحديد مقاومة الكرة للصدمات المتكررة.

تم تحسين عمر الكرة الطاحنة محلية الصنع ذات محتوى كروم % 15.23 لتصل الى 4210 صدمة قبل التحطم من خلال المعالجة الحرارية برفع درجة الحرارة الى C° 1000 والوصول الى الطور الأوستينيتي ثم السقاية بالزيت متبوعا بعملية ارجاع عند الدرجة C° 500 ، كما أجريت

نفس التجربة بعد تغيير درجة حرارة الارجاع تجريبياً الى 550°C و 600°C لكن عمر الكرة الذي تم الحصول عليه بعد هاتين الحرارتين كان أقل ، حيث لم يتم تجاوز 3200 صدمة قبل التحطم

بينت نتائج الدراسة التجريبية على الكرات الطاحنة تأثير تغيير درجة حرارة الارجاع بعد الطور الأوستينيتي على البنية المجهرية والقساوة للكرات الطاحنة المختبرة كما أظهرت نسبة الكرييدات وتوزعها في البنية دور حاسم ومهم في عمر الكرة الطاحنة على التحطم.

خفضت نسبة الكرييدات من % 18.8 الى % 9.223 من خلال عملية السقاية من الطور الأوستينيتي عند الدرجة 1000°C متبوعاً بالارجاع عند الدرجة 500°C والتي توافقت مع قساوة ناتجة 50 HRC في بنية مارتنسييتية مع وجود الكرييدات و القليل من الأوستينيت المتبقي ، حيث أبدت الكرة أفضل أداء على التحطم من خلال الوصول الى أفضل علاقة بين متانة التحطم و القساوة.

كلمات مفتاحية

الكرات الطاحنة، السقوط الحر، المتانة، التحطم، القساوة.

*طالب دكتوراه - هندسة التصميم والإنتاج - قسم هندسة التصميم والإنتاج - كلية الهندسة الميكانيكية والكهربائية - جامعة حمص - حمص - سوريا rami.al.jundi92@gmail.com

** أستاذ مساعد - قسم هندسة التصميم والإنتاج - كلية الهندسة الميكانيكية والكهربائية - جامعة حمص - حمص - سوريا.

1- Introduction

The grinding process is one of the most critical processes of the cement industry, especially for breaking down the materials till reaching the

proper size, the grinding balls are working under a heavy mechanical stress like the impact and abrasive wear, these stresses lead the grinding balls to breakage and causing the grinding process failure, which affects the process operational time and the total cost.

The balls are widely used in the cement milling and many alloys of steel and cast iron are available in the industrial market for this application ,the High Chromium cast iron has shown one of the best results in the cement milling field but still need to find the suitable range of hardness and fracture toughness as the high chromium cast irons performs as a high hardness and wear resistance which comes from the very hard carbides in the microstructure but this also permits the brittleness to be minimize the durability of the balls in the mill.

Improving the fracture toughness of the grinding balls needs to apply heat treatment like Austenitization for reaching the destabilization of the austenite by precepting the high chromium secondary carbides, at a thermal range of 920-1060 C° with holding for 1-6 hours, this depends on grinding ball chemical composition followed by fast quenching to the room temperature, after this heat treatment with a proper tempering in a range of 450 – 650 C° , due to the destabilization process the austenite that become less from the carbon and chromium and can produce a harder microstructure by the martensitic transformation . the secondary carbides will distribute in the martensitic microstructure with the presence of

retained austenite that is reduced in the tempering which also helping in the residual stress's reduction.¹

Many experimental tests are used for testing the quality of the grinding balls like the Ball-on-Ball test and Ball on Block, both of them aim to study the grinding balls performance under impact conditions and following a continuous observation for the ball behavior starting from plastic deformation if happened till the fracture into two parts or more, these tests help in the ball's durability determination and the needed amount of energy to cause the fracture in the grinding ball. ²

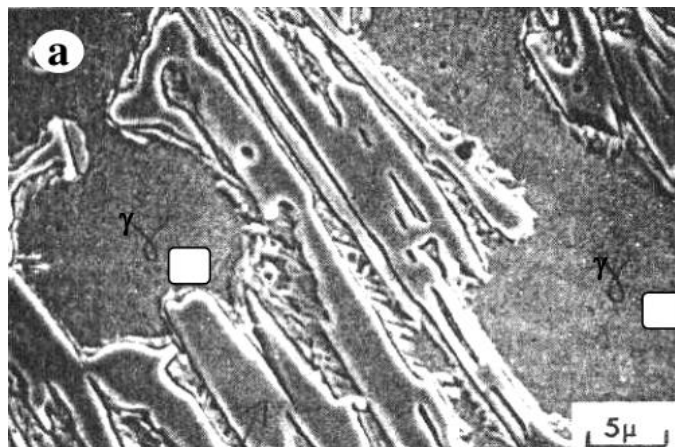


Figure 1- the as-cast microstructure, austenitic matrix with caribes

1. ¹ **Younes, R., Sadeddine, A., Bradai, M. A., Aissat, S., & Benabbas, A.** 2021. Investigation on the Influence of Tempering on Microstructure and Wear Properties of High Alloy Chromium Cast Iron, *Advances in Materials Science*, Vol. 21, 66-73.
2. ² **Wiengmoon, A.** 2010. Carbides in High Chromium Cast Irons, Department of Physics, Faculty of Science, Naresuan University.

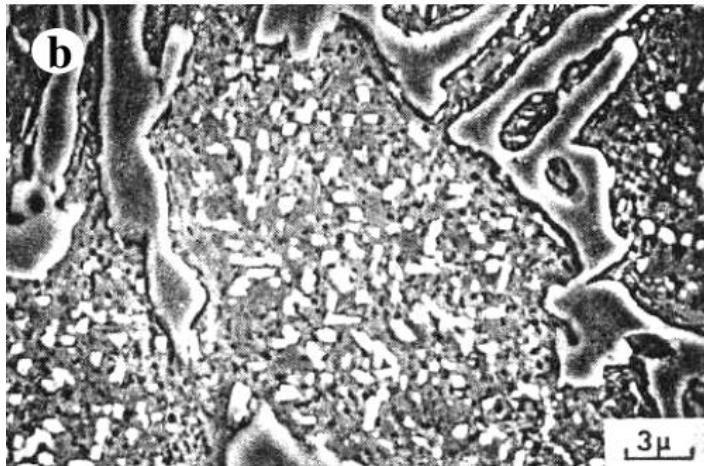


Figure 2- the martensitic microstructure with the secondary carbides in the destabilized matrix

2- Objective of the research

The research aims to improve the working durability of local grinding balls by Optimizing the heat treatment process of the High Chromium Cast Iron grinding balls, especially the tempering temperature for enhancing the balanced relation between the fracture toughness and the hardness.

3- Material and Methodology

3-1- Material and Apparatus:

- 1- Local made grinding balls.
- 2- Grinding papers
- 3- Grinding machine
- 4- automatic Cutting machine
- 5- Nital solution 5%.
- 6- Spectrometer

- 7- Hardness testing machine
- 8- Drop ball testing machine
- 9- (Image J) software, image processor.

3-2- Method

The local high chromium cast iron grinding balls were experimented in a laboratory in aim of studying the tempering effect on the microstructure and its durability also, a series of preparations and experimental processes were used as follows:

3-2-1 Material description:

A high chromium cast iron grinding balls were used with a 15.23 % Cr content, D = 70 mm, local made, quantity of 12 balls.

3-2-1- Material preparation:

The balls were cut for chemical composition analysis by the automatic cutting machine.

- The samples were prepared with a grinding paper starting from size 400 till 600 – 800 – 1000 -1200 -1600 -2000 as a coarse and fine grinding.
- Samples were polished with Alumina powder
- Etching of the samples done with 2% Nital solution (aqueous nitric acid and ethanol) 1.

3-2-2-Microstructure Analysis

A light microscope used to show the microstructure of the samples before and after the heat treatments of the grinding balls.

3-2-2-1- Carbide Fraction % determination

the carbides volume fraction % was determined and measured before and after the heat treatments by using the Image J software by processing the image captured from the light microscope as (B & W) black and white to clarify the selected phase percentage after removal of the other phases.

3-3- Heat Treatment process

The grinding balls were subjected to a heat treatment cycle consists of Austenitization and Tempering processes.

3-3-1- Austenitization

Raising the ball temperature for a 1000 C° - 2 H holding – Oil quenching.
(سامي، 2019)

3-3-2- Tempering

After quenching, three tempering temperatures were selected by heating the grinding balls for 500 C°, 550 C° and 600 C° and keep for room temperature

(علي، 2023) .

3-4- Drop Ball Test

The ball as in commercial state and after heat treatment were tested on the drop ball testing machine to determine the impacts number that the balls can resist before fracture and to observe the ball behavior during the test like plastic deformation or spalling.

The drop ball test machine applies a free fall from 6 m height on a hardened steel anvil 63 HRC and repeating this step till ball fracture.

3-5 Hardness Test

The hardness of the balls in the commercial and heat-treated states were tested for observing the effect of the heat treatment on the ball microstructure especially the carbide's volume fraction.

3-6 Data Analysis

The results of the tests were listed in the tables and graphs plotted to show the relationship between the resulted technical properties.

4- Results and Discussion

The carbide volume fraction was the most important factor has been studied after tests and its relation to the durability of the balls and the hardness.

Three tests was done for the durability life of the grinding balls and their mean was calculated and listed in the results table NO.1 , a life of 1945 impacts was counted for the commercial local ball and this number was improved till

4210 impacts by the heat treatment of (1000 C° - 2 H – oil / 500 C° -1.5 H) where the carbide volume fraction was dropped from 18.8 % till 9.23 % , this was a notable point during the experiment.

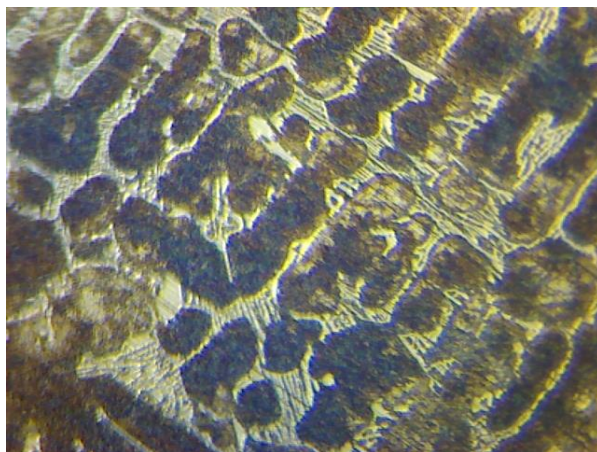


Figure 3 – Microstructure of the ball after heat treatment (X800)
(1000 C° - 2 H – oil / 500 C° -1.5 H)

a 3142 impacts was counted resulted after (1000 C° - 2 H – oil / 550 C° -1.5 H) and the ball that was treated at (1000 C° - 2 H – oil / 600 C° -1.5 H) was only resist a 1845 impacts before fracture , the carbides volume fraction was found for both 550 and 600 C° very close in a range of 12.201 – 12.68 %

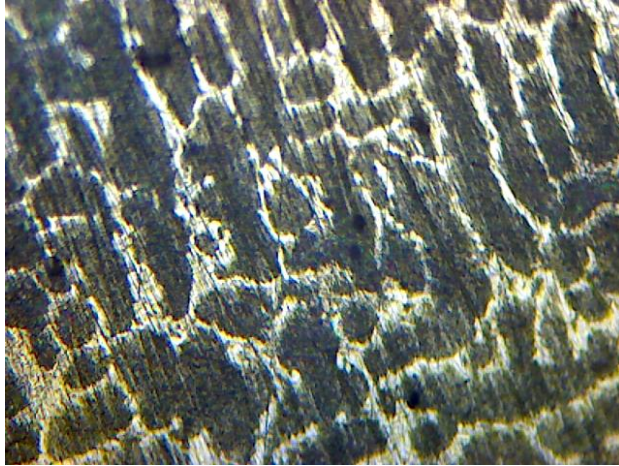


Figure 4 – Microstructure of the ball after heat treatment (X400)
(1000 C° - 2 H – oil / 550 C° -1.5 H)

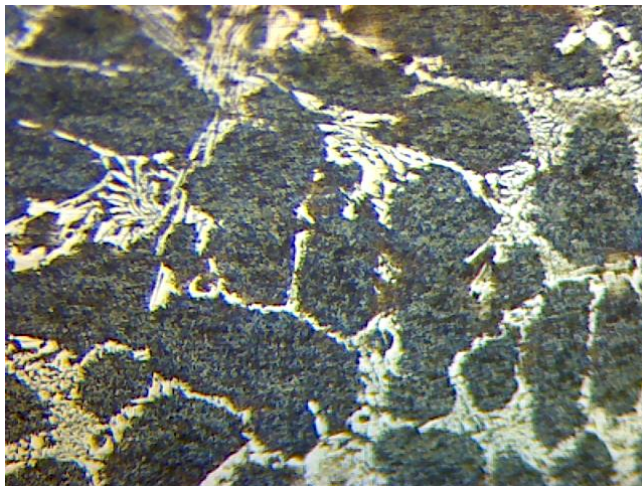


Figure 4 – Microstructure of the ball after heat treatment (X800)
(1000 C° - 2 H – oil / 600 C° -1.5 H)

The hardness of the balls was decreased gradually with heat treatment, started at 54.4 HRC for the commercial ball and found 50 , 49 , and 44.3

HRC respectively with heat treatment with changing the tempering temperature as 500,550, and 600 C°.

Test	Commercial	1000 /500	1000/550	1000/600
1	1855	4050	3144	1800
2	2100	4180	3380	1690
3	1940	4400	2900	2045
Mean	1965	4210	3142	1845
Carbide %	18.8	9.223	12.68	12.201
Hardness (HRC)	54.4	50	49	44.3
Failure type	Fracture	Fracture	Plastic deformation & fracture	Fracture

Table 1 – shows the tests results of the research tests

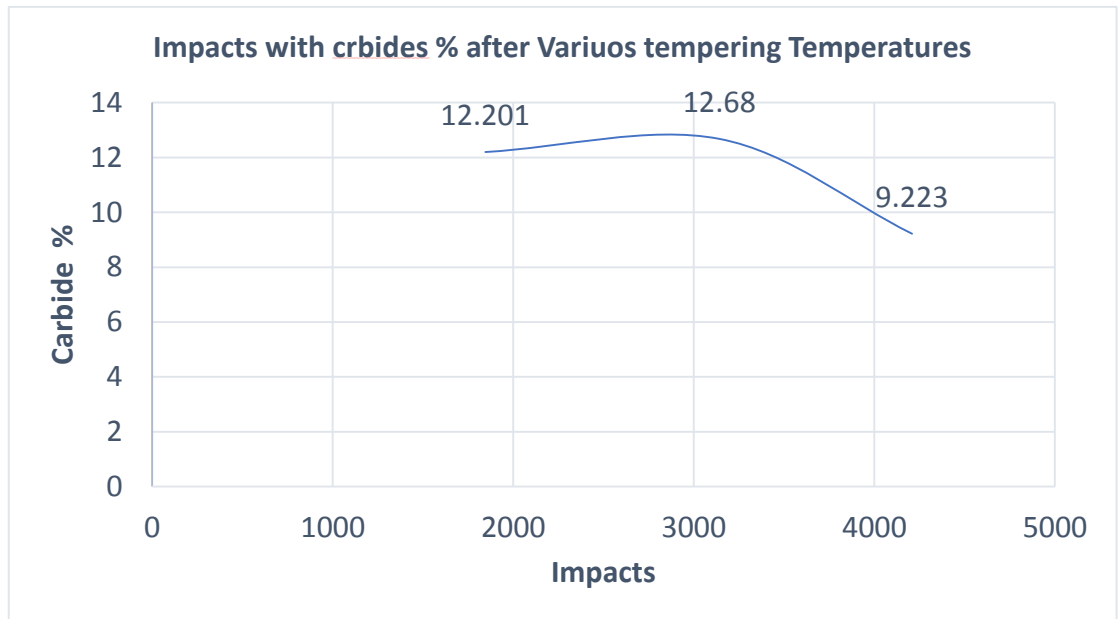
5- Conclusion and Recommendations

The results have shown that the applied heat treatments have a significant effect on the high chromium cast iron grinding ball's microstructure and their mechanical properties.

The most important point was to keep the optimal fracture toughness with an acceptable hardness level to maintain the wear resistance of the balls ,

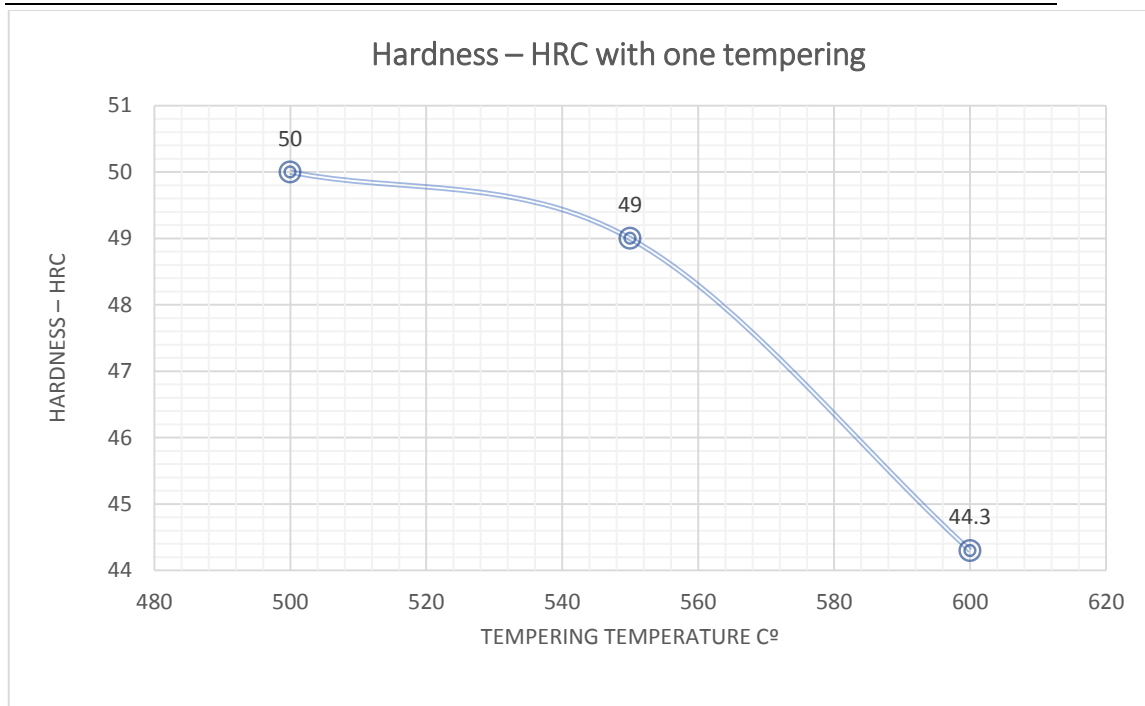
the first heat treatment of the Austenitization by heating the ball till 1000 C° has caused the carbides dissolution and produced an austenitic microstructure so the carbon and the chromium will be dissolved in the austenite, and this microstructure has transformed to martensitic after oil quenching , in the tempering process especially at 500 C° , the microstructure was found as a matrix of martensite and secondary carbides and retained austenite which was controlled during the tempering by transforming most of it to martensite to eliminate its rule of reducing the hardness , this will enhance the fracture toughness , the analyses of the microscopic image of the ball was shown that the carbides has formed finely which make the matrix more able to resist the wear and repetitive impacts and can prevent the brittleness, the carbide fraction volume reduction from 18.8% to 9.23 % indicated a good distribution and dispersion of the carbides in the martensitic matrix , this was with a

helpful hardness value of 50 HRC for optimizing the properties .



At tempering temperature of 550 C° and 600 C°, the increase of the carbides fraction volume was indicated to more carbide's precipitation in the martensitic microstructure as it was ranged between 12.201 and 12.68 % respectively, this was affected the fracture toughness and decreased the number of the impacts till 3142 impacts corresponding to the 550 C° tempering, and a 1845 impacts with the tempering of 600 C°.

The reduced hardness value also indicates that the martensitic matrix was softened, both of excessive carbides perception and hardness decrement was a reason for earlier crack initiation ball fracture and failure.



As a conclusion, the Austenitization at a 1000 C° followed by a tempering at 500 C° is the optimal treatment for maintain the optimized properties of the local grinding balls tested, a better fracture toughness and durability and a good hardness level was resulted in double life of the grinding ball.

- **Recommendation**

- The optimal heat treatment of 1000 C° - 2 H – oil / 500 C° -1.5 H should be adopted for improving the local grinding balls of this composition.
- Testing the effect of repeating the tempering process for two or three times.
- Testing more holding time for the tempering process that may be affect the carbides perception.

- References

1. **Younes, R., Sadeddine, A., Bradai, M. A., Aissat, S., & Benabbas, A.** 2021. Investigation on the Influence of Tempering on Microstructure and Wear Properties of High Alloy Chromium Cast Iron, *Advances in Materials Science*, Vol. 21, 66-73.
2. **Moema, J. S., Papo, M. J., Slabbert, G. A., & Zimba, J.** 2009. Grinding Media Quality Assurance for the Comminution of Gold Ores, *World Gold Conference*, 27-31.
3. **Wiengmoon, A.** 2010. Carbides in High Chromium Cast Irons, *Department of Physics, Faculty of Science, Naresuan University*.
4. **Kootsookos, A., Gates, J. D., & Eaton, R. A.** 1995. "Development of a White Cast Iron of Fracture Toughness $40\text{MPa}\cdot\text{m}^{0.5}$." *International Journal of Cast Metals Research*, Vol. 7, 239-246.
5. **Blickensderfer, R., & Tylczak, J. H.** 1989. "Evaluation of Commercial US Grinding Balls by Laboratory Impact and Abrasion Tests." *Minerals & Metallurgical Processing*, Vol. 6, 60-65.
6. **Filipovic, M., Kamberovic, Z., Korac, M., & Gavrilovski, M.** 2013. Correlation of Microstructure with the Wear Resistance and Fracture Toughness of White Cast Iron Alloys. *Metals and Materials International*, Vol. 19, 473-481.
7. **Gates, J. D., Bennet, P. J., McInnes, L. J., & Tunstall, B. R.** 2015. The Challenge of Accurate Prediction of Industrial Wear Performance from Laboratory Tests. *UQ Materials Performance*, *The University of Queensland*
8. **كاسوحة عباد ، انطكلي سامي ،** 2019، التحقق من تأثير عمليات المعالجة الحرارية على سلوك الاهتراء لفولاذ بوهلر k72 ، *Journal of Engineering Sciences & Information Technology*, v. 3, n. 1, p. 29, doi. 10.26389/AJSRP.S191218
9. **سعيد علي،** 2023 تحسين خواص الفولاذ SUP3 من خلال التغيير في درجة حرارة الإرجاع للمعالجة الحرارية، *مجلة جامعة البعث للعلوم الهندسية* ، مجلد 45 عدد 13 (2023): أبحاث العدد 13.

دراسة سرعة أداء أنظمة الربط العلائقي المستخدمة في تطوير

التطبيقات

اعداد

م. ايلياس شلوف

اشراف

أ.د. ناصر أبو صالح

د. أسامة ناصر

المخلص

تشهد هندسة البرمجيات توسعاً مستمراً في استخدام نماذج الربط العلائقية ORMs عبر أطر عمل متعددة ولغات برمجة متنوعة، حيث تلعب هذه النماذج دوراً محورياً في تسهيل عمليات التفاعل مع قواعد البيانات العلائقية. تهدف هذه الدراسة إلى إجراء مقارنة شاملة بين أحدث وأبرز نماذج ORM المستخدمة حتى عام 2025، من خلال تحليل نظري وتقني لآليات البناء، أساليب التعامل مع الكيانات، وسهولة الاستخدام، إلى جانب تقييم عملي لأداء هذه النماذج عبر مجموعة من استعلامات القواعد البيانات الشائعة والمعقدة. بناءً على تجارب مقارنة مع عدد من أنظمة إدارة قواعد البيانات العلائقية الرائدة، أبرزت النتائج فروقات ملحوظة في سرعة التنفيذ وكفاءة التفاعل مع البيانات بين النماذج المختلفة. في الختام، تقدم هذه الدراسة توصيات مبنية على نتائجها لمساعدة المطورين في اختيار نموذج ORM الأمثل الذي يتناسب مع متطلباتهم التقنية ويساهم في تحسين أداء التطبيقات البرمجية.

الكلمات المفتاحية: قواعد بيانات، لغة الاستعلامات الهيكلية SQL، لغات برمجة، أطر عمل، أنظمة إدارة قواعد البيانات، استعلامات، هندسة برمجيات، ORM (نظام/نموذج الربط العلائقي).

Studying ORM models performance used in developing applications

Abstract

Software engineering has witnessed continuous growth in the use of Object Relational Mapping (ORM) models across various frameworks and programming languages, where these models play a pivotal role in facilitating interactions with relational databases. This study aims to conduct a comprehensive comparison of the latest and most prominent ORM models used up to 2025, through a theoretical and technical analysis of their construction mechanisms, entity handling approaches, and ease of use, alongside a practical evaluation of their performance across a range of common and complex database queries. Based on comparative experiments with several leading relational database management systems, the results revealed significant differences in execution speed and data interaction efficiency among the different models. In conclusion, this study provides recommendations grounded on its findings to assist developers in selecting the most suitable ORM model that aligns with their technical requirements and enhances software application performance.

Keywords: Databases, SQL, Programming languages, Frameworks, Database management systems, Queries, Software engineering, ORM (Object–Relational Mapping).

1. مقدمة

شهدت السنوات الأخيرة تحولاً ملحوظاً في مجال إدارة قواعد البيانات، حيث اتجهت المؤسسات نحو توزيع قواعد البيانات على خوادم متعددة وتحسين الوصول إليها حتى من مواقع جغرافية متباعدة، وذلك لتحسين سرعة الاستجابة وتوافر البيانات. في هذا السياق، برزت الحاجة إلى تسهيل عمليات التعامل مع قواعد البيانات العائقية للمطورين، خاصةً من غير المتخصصين في كتابة استعلامات SQL معقدة، مما أدى إلى ظهور مفهوم نماذج الربط العائقي ORMs.

تتمثل المشكلة الرئيسية التي يعالجها هذا البحث في التحديات المستمرة والمتجددة المتعلقة باختيار نموذج ORM الأمثل، خاصةً مع التطورات السريعة في لغات البرمجة وأطر العمل وأنظمة إدارة قواعد البيانات حتى عام 2025. فبينما توفر نماذج ORM مستوىً عاليًا من التجريد والتسهيل، إلا أن اختلافاتها في الأداء وسهولة الاستخدام وتوافقها مع متطلبات البيانات الحديثة تثير التساؤلات حول مدى ملاءمتها في سياق التطورات التقنية الراهنة.

لذلك، يسلط هذا البحث الضوء على فجوة معرفية قائمة في التقييم الحديث لنماذج ORM، من خلال دراسة مقارنة تجمع بين التحليل النظري واختبارات الأداء العملية، بهدف توفير توصيات مبنية على دلائل علمية تساعد المطورين على اتخاذ قرارات مدروسة في اختيار نماذج ORM التي تلبي متطلبات تطبيقاتهم في ظل التحولات المستمرة التي يشهدها المجال. يُرجى العلم أن التفاصيل التعريفية والفنية المتعلقة بمكونات وبنية نماذج ORM ستُعرض بشكل مفصل في قسم "الإطار النظري" لاحقًا.

2. أهمية البحث وأهدافه

تلعب نماذج الربط العلائقي (Object-Relational Mapping – ORM) دورًا حيويًا في تطوير التطبيقات البرمجية الحديثة، حيث تُسهّل التفاعل بين البرمجة كائنية التوجه وقواعد البيانات العلائقية من خلال تجريد وتعقيد استعلامات SQL، مما يعزز إنتاجية المطورين ويسهّل صيانة الأكواد البرمجية. ومع تنامي تنوع نماذج وأطر ORM، يواجه المطورون تحدي اختيار النموذج الأنسب لمتطلبات مشاريعهم، والتي تتفاوت بناءً على عوامل مثل لغة البرمجة المستخدمة، البيانات التطويرية، متطلبات الأداء، والميزات التقنية الخاصة بكل نموذج.

يركز هذا البحث على إعادة تقييم وتحديث نتائج الدراسات السابقة المتعلقة بنماذج ال ORM، انطلاقًا من تسارع التطور في لغات البرمجة، أطر العمل، محركات قواعد البيانات، ونماذج ORM نفسها حتى عام 2025، وذلك لضمان ملاءمتها للاستخدام الحالي في بيئات التطوير الحديثة.

يهدف البحث لتحقيق ما يلي:

1. **تحليل نظري:** دراسة معمقة لتطور نماذج ORM، خصائصها التشغيلية، وآليات التعامل مع الكيانات عبر أشهر الأطر البرمجية حتى عام 2025.
2. **تقييم عملي:** مقارنة الأداء التشغيلي لهذه النماذج من حيث سرعة تنفيذ الاستعلامات وكفاءتها عبر استعلامات متنوعة، باستخدام أنظمة إدارة قواعد البيانات العلائقية الأكثر انتشارًا.
3. **تقديم توصيات:** استخراج استنتاجات واضحة وأدلة تساعد المطورين في اختيار نموذج ORM الأمثل وفقًا لمتطلبات ومحددات مشاريعهم البرمجية.

3. مواد وطرق البحث

3.1. المفاهيم الأساسية

يستند هذا البحث إلى مجموعة من المفاهيم النظرية الأساسية التي تشكل جوهر دراسة نماذج الربط العلائقي ORM. من بين هذه المفاهيم نمطا Data Mapper و Active Record، حيث يمثل الأول أسلوباً يُفصل بين طبقة الكود البرمجية وقاعدة البيانات بالكامل، بينما يمكّن الثاني كل كائن برمجي من التحكم المباشر في عملياته على الجداول المرتبطة به. فهم الفوارق بين هذه الأساليب يعد خطوة أولى محورية في تحليل مناهج العمل ونماذج الأداء في تقنيات ORM المتنوعة.

3.2. نماذج ORM المشهورة وأطر العمل المستخدمة

تغطي الدراسة مجموعة مختارة من أشهر نماذج الـ ORM المعتمدة في التطوير البرمجي حتى عام 2025، مثل Hibernate و Entity Framework Core و Dapper و Eloquent و Doctrine، بالإضافة إلى نماذج بيئة الاندرويد مثل Room, greenDAO. ويشمل التحليل مقارنة منهجية بين هذه النماذج من حيث التصميم، سهولة الربط مع قواعد البيانات، آليات التفاعل مع الكيانات، وتبني أنماط مختلفة لمعالجة استعلامات البيانات والتحكم في الأداء.

3.3. الأدوات والبيئة البرمجية

اعتمدت الدراسة على تحليل وتلخيص سبع أوراق بحثية مرجعية شكلت عينة ممثلة للدراسات المقارنة في المجال. تمت الاستعانة بأطر عمل برمجية متنوعة (Java, .NET, PHP, Android) وباستخدام أحدث إصدارات بيئات التطوير وأنظمة إدارة قواعد البيانات العلائقية الرائدة حتى عام 2025. تضمن العمل رصد الأداء (زمن الاستجابة واستهلاك الموارد)، ودراسة سهولة الاستخدام، وموثوقية التطبيق العملي للنماذج في سيناريوهات متعددة. بالإضافة إلى ذلك، أُجري تحليل نقدي لمواطن القوة والضعف في الدراسات السابقة، وتمت مراجعة مدى مواكبتها للتقنيات الحديثة وتوجهات التطوير الجارية.

4. الإطار النظري

تُعد نماذج الربط العلائقي ORM من التقنيات الأساسية التي تربط بين البرمجة كائنية التوجه وقواعد البيانات العلائقية. فهذه النماذج توفر طبقة تجريدية تُسهّل عملية التفاعل مع قواعد البيانات عن طريق تمثيل الجداول العلائقية ككائنات داخل لغات البرمجة، مما يقلل الحاجة إلى كتابة استعلامات SQL معقدة.

يتركز الإطار النظري لهذا البحث على دراسة المفاهيم الأساسية التي تقوم عليها نماذج ORM، بما في ذلك آليات تحويل البيانات بين النماذج الكائنية والعلائقية، وتصميم الأطر التي تدعم هذه النماذج، وكذلك تحليل مكونات نموذج ORM الرئيسي، حيث ان:

- الغرض الأساسي من هذا المفهوم كان تمثيل جداول البيانات ك صفوف غرضية في لغات البرمجة دون الحاجة الى معالجة كل بيانات الجداول في كل خطوة.
 - ORM هي اختصار لـ Object-Relational-Mapping (ORM) وهي تقنية برمجية لتحويل البيانات بين قواعد البيانات العلائقية ولغات البرمجة الغرضية التوجه مثل Java من خلال جعل التعامل الأساسي مع الأغراض بدلاً من الجداول والاستعلامات.
 - يستند هذا النموذج على نموذج الاتصال التقليدي مع قواعد البيانات
 - يتم فيه إدارة المناقلات وانشاء المفاتيح بشكل تلقائي.
 - يخفي تفاصيل استعلامات SQL من خلال جعل التعامل مع منطق البرمجة الغرضية التوجه.
- يتكون نموذج ال ORM من الكيانات الأربعة التالية:

1. واجهة برمجة تطبيقات API لإجراء عمليات CRUD الأساسية على الكائنات/الأغراض.
2. لغة أو واجهة برمجة تطبيقات لتحديد الاستعلامات التي تشير إلى صفوف وخصائص الصفوف.
3. وسيلة قابلة للتكوين لتحديد الربط مع البيانات الوصفية.
4. أسلوب للتفاعل مع كائنات المعاملات لإجراء عمليات مثل: فحص التغييرات على البيانات، وجلب النتائج المتأخر، ووظائف تحسين أخرى.

5. الدراسات المرجعية

تتناول الدراسات السابقة نماذج الربط العلائقي (ORM) من زوايا متعددة تجمع بين التحليل النظري والتقييم العملي للأداء، مع اختلاف في نطاق الدراسة والبيئات المستخدمة، وهو ما يسمح برسم صورة شاملة عن تطور هذه النماذج حتى عام 2025.

في البداية، تقدم الدراسة الأولى [1] مراجعة تاريخية نظرية تغطي عشرين عاماً من تطور نماذج ORM في عدة لغات برمجة، مع التركيز على حلول مشكلة "عدم التطابق" بين النظم الغرضية وقواعد البيانات العلائقية. تبرز الورقة أهمية نماذج التصميم الأساسية التي تقوم عليها هذه النماذج وكيف تخدم آليات جلب البيانات المختلفة، مع توفير جدول ملخص

للمقارنة بين هذه النماذج. بالرغم من شمولية الدراسة، تظل محدودة في تقديم تقييم عملي أو نتائج أداء، مما يبرز الحاجة إلى دراسات تجريبية محدثة تأخذ في الاعتبار التقنيات والإصدارات الحديثة.

تكمّن قوة الدراسة الثانية [2] في تقييم الأداء العملي لمجموعة من نماذج ORM داخل بيئة .NET، خصوصاً Entity Framework Core، nHibernate، وDapper، عبر عمليات CRUD على مجموعات بيانات بأحجام متفاوتة. توضح نتائج الدراسة الاستخدام الجزئي الأفضل لكل نموذج حسب نوع العملية، ولا تبرز وجود نموذج متفوق بشكل مطلق؛ مما يؤكد أن الاختيار يعتمد على نوعية العمليات ومتطلبات المشروع. إلا أن الدراسة تفتقر إلى تغطية أطر عمل أخرى أو لغات برمجة غير .NET، كما أنها لا تتعامل مع تعقيدات الاستعلامات المعقدة أو التفاعلات متعددة الأنظمة، وهو ما تعد هذه الورقة البحثية الحالية محاولة لسده.

ركزت الدراسة الثالثة [3] على نموذج Eloquent ORM في بيئة PHP/ Laravel، موضحة خصائصه في دعم أنواع متعددة من العلاقات بين الكيانات، وهو مفيد لفهم الدعم الوظيفي للنماذج في تطبيقات ويب محددة. لكنها تبقى دراسة تطبيقية ضيقة النطاق تفتقر إلى مقارنة أداء مع نماذج أخرى أو تقييم تجريبي متعدد البيئات. لذا تبرز الحاجة إلى استعراض شامل وأعمق يدمج هذه النماذج ضمن مقارنة أوسع وأدوات قياس أداء موحدة.

تناولت الدراسة الرابعة [4] نموذج Hibernate في بيئة Java، مشيرة إلى متطلبات التهيئة، آلية عمل EntityManager، ولغة الاستعلام الخاصة HQL، مع تسليط الضوء على ميزة التخزين المؤقت التي تحسن الأداء بشكل ملحوظ عند تفعيلها. تقدم هذه الدراسة خلفية تقنية ثرية، لكن تحليلها يقتصر على نموذج واحد فقط، ولا تقدم مقارنات مباشرة مع نماذج أخرى. كما تفتقر إلى اختبار عملي متكامل يشمل الاستعلامات المختلفة وحدود الأداء.

في مجال تطبيقات الأندرويد، قدمت الدراسة الخامسة [5] مقارنة عملية بين مكتبات Room و greenDAO، فضلاً عن SQLiteOpenHelper التقليدي. أظهرت النتائج تفوق greenDAO في معظم عمليات CRUD من حيث الأداء، فيما تدعم Room التتابع في الحذف والتعديل ولكن بتكاليف أكبر من حيث استخدام الموارد والوقت. هذه الدراسة تطرح تساؤلات هامة حول تبعية اختيار النموذج لسياق التطبيق ومستوى تعقيد العلاقات بين الكيانات، لكنها تنحصر في بيئة تطوير نموذجية، مما يحد من تعميم النتائج على أطر عمل ولغات برمجة أخرى.

أما الدراسة السادسة [6]، فهي تقدم تحليلاً لميزات ومزايا نموذج Doctrine المستخدم في بيئة PHP/Symfony، مع استعراض لآليات توليد المخططات التلقائية، لغة الاستعلام الخاصة DQL، والتخزين المؤقت. كما تشير إلى التحديات المرتبطة بزيادة التعقيد الطبقي نتيجة التجريد، واحتمالية ضعف الأداء مع الاستعلامات غير المحسنة. مع ذلك، الدراسة تعتمد أساساً على تقييم نظري وتجريبي محدود النطاق، ما يترك مجالاً لبحث أعمق مقارنة بالأطر المختلفة والتقنيات الحديثة.

أخيراً، تقدم الدراسة السابعة [7] مقارنة أداء مختصة بين نماذج ORM الثلاثة الأكثر شيوعاً في Java: Hibernate ، Ebean، و TopLink، مستعينة بالمعيار Benchmark 007 لإجراء اختبارات زمن التنفيذ على مجموعة استعلامات CRUD متنوعة. توضح النتائج تفوق Ebean عموماً بسبب واجهته الخالية من الجلسات، مع تميز Hibernate في إدراج البيانات، وتوصي الدراسة باستخدام Ebean للأداء الأفضل. رغم قدم هذه الدراسة (2012) وتخصيصها لبيئة Java فقط، فإنها تعطي لمحة مهمة عن تطور الأداء عبر الأجيال وتبرز الحاجة لإعادة تقييم مماثل يشمل التحديثات والتقنيات الحديثة.

5.1. ملخص الدراسات السابقة

الرقم المرجعي	عنوان الدراسة	تاريخ النشر	بيئة العمل	المنهجية	أبرز النتائج	نقاط الضعف
1	Twenty years of object-relational mapping	2017	متعددة (Java, .NET, PHP...)	مراجعة نظرية وتحليلية	ملخص شامل لنماذج ORM وآليات التعامل مع عدم التطابق	يفتقر لتقييم عملي وحديث للأداء
2	Performance Comparison of CRUD Methods using NET Object Relational Mappers	2020	NET (EF, nHibernate, Dapper)	تجريبية (وقت التنفيذ، الذاكرة)	لا يوجد متفوق مطلق، الأداء يعتمد على نوع العملية وعدد الكيانات	محدود ببيئة .NET فقط، لا يشمل تعقيدات الاستعلام
3	Eloquent Object Relational Mapping Models for Biodiversity Information System	2017	PHP (Laravel)	دراسة تطبيقية	توضيح دعم العلاقات المعقدة بين الكيانات	عدم وجود مقارنة أداء أو بيانات أخرى

4	Lessons in Persisting Object Data using Object-Relational Mapping	2019	Java (Hibernate)	مراجعة تقنية مع أمثلة تطبيقية	شرح ميزات النموذج من التخزين المؤقت، HQL، ميزات الحماية	تركز على نموذج واحد فقط، لا توجد مقارنة تجريبية
5	Room vs. greenDAO for Android	2022	أندرويد	تجريبية (RAM, CPU, التنفيذ)	GreenDAO أسرع، Room يدعم التتالي لكنه أبطأ ويستهلك موارد أكثر	حصر على بيئة أندرويد فقط، عدم تعميم النتائج
6	Improving data processing performance for web applications using object-relational mapping	2015	PHP (Symfony)	مراجعة نظرية وتجريبية محدودة	شرح الميزات من تبسيط إدارة البيانات، تحسين الأداء مع بعض العيوب	تقييم محدود الأبعاد والمقارنات
7	Object-relational mapping tools and Hibernate	2017	Java	تجريبية (Benchmark 007، زمن التنفيذ)	Ebean الأسرع، Hibernate الأفضل للإدراج، توصية باستخدام Ebean	قديمة لعام 2012، لا تشمل التطورات الأحدث

جدول 1 ملخص مقارنة الدراسات المرجعية

5.2. نقاط الضعف المشتركة والمبررات لإجراء البحث الحالي

على الرغم من القيمة العلمية والعملية لهذه الدراسات، إلا أن لديها عدة نقاط ضعف تبرر القيام بالبحث الحالي:

- **نقص التغطية الشاملة:** تميل معظم الدراسات إلى التركيز على لغات أو بيئات محددة (مثل .NET ، Java ، PHP ، Android)، مع ندرة مقارنة شاملة عبر بيئات تطوير وأطر متعددة في آن واحد.
 - **تفاوت المنهجيات:** تختلف الدراسات بين النظرية الصرفة والتقييم العملي، مع غياب موحد لمنهجيات قياس الأداء والمعايير المستخدمة، مما يصعب استخلاص مقارنات موضوعية دقيقة.
 - **تجاهل التطور التقني الحديث:** بعض الدراسات، مثل [7] سنة 2012 و [1] التي استعرضت التطور حتى 2016، لا تغطي أحدث نماذج ORM أو التطورات الجديدة في لغات البرمجة وأطر العمل حتى عام 2025، وهو أمر حيوي لفهم الفروقات الحالية.
 - **غياب تقييم الأداء في حالات معقدة:** قلة الدراسات التي تدرس تأثير نماذج ORM على استعلامات معقدة، إدارة العلاقات المعقدة، والتعامل مع أحجام بيانات كبيرة بطرق موحدة.
 - **عدم التوازن في التركيز:** بعض الدراسات تعطي تفصيلاً ناعماً (كما في دراسة [2]) فيما تختصر أخرى مواضيع هامة، مثل جوانب الحماية، الأمان، والتوافق مع أنظمة قواعد البيانات الحديثة.
 - **الإغفال عن البيئة العملية الواقعية:** كالتأثير على حجم التطبيقات واستهلاك الموارد، خصوصاً في البيئات منخفضة الموارد كالأجهزة المحمولة، الذي أجازته دراسة [5] لكنه غير مغطى في بقية الدراسات.
- بناءً على ذلك، يهدف البحث الحالي إلى سد هذه الفجوات من خلال مقارنة موسعة ومنهجية تجمع بين التحليل النظري ونماذج الأداء العملي عبر بيئات وأطر متعددة وحديثة، مع تدقيق في استعلامات الأداء المعقدة، واستهلاك الموارد، والتوافق مع متطلبات مطوري البرمجيات في عام 2025. تقدم الورقة توصيات دقيقة وموضوعية تساعد في اختيار نموذج ORM الأمثل وفقاً لاحتياجات الأنظمة المعاصرة، مع الاعتماد على بيئة تجريبية موحدة وأدوات قياس دقيقة تعزز من موثوقية النتائج.

6. النماذج المعمارية لنماذج ORM والمقارنة التطبيقية

في البداية يجب ان نذكر ان نماذج ال ORM الحالية أصبحت تعتمد على أكثر من معمارية في بنائها وهي التي تحدد سهولة الكتابة بالإضافة لسرعة التنفيذ والربط والمقابلة مع قواعد البيانات، وأشهر هذه البنى المعمارية هي:

1. **Active Record**: في هذه البنية يقابل الصف بشكل مباشر جدول قاعدة البيانات ويستطيع أي غرض منه التفاعل معها والقيام بالعمليات بشكل مباشر مثل الإدراج والحذف والتعديل وغيرها، وتتميز هذه البنية بسهولة وسرعة الكتابة للعمليات البسيطة لكنها غير مفضلة في العمليات المعقدة كونها تزيد من درجة الاعتمادية بين الصفوف، وامثلة عن هذه البنية هو ال Django ORM.

2. **Entity Mapper**: في هذه البنية يكون هذا الصف هو الوسيط بين صفوف النماذج التي تمثل الجداول و قاعدة البيانات حيث تكون الأغراض المنشأة من الصفوف هي فقط تحوي على بيانات وليس لها أي وصول لقاعدة البيانات، وتكون مسؤولية ال EM هنا هي مقابلة قيم الحقول في هذه الصفوف وجداول قاعدة البيانات، ومن النماذج التي تستخدم هذه البنية هو نموذج Ktorm.

3. **Entity Manager**: وهي بنية خاصة حصرية فقط ببعض النماذج مثل نموذج ال Hibernate تتحكم بشكل كامل بكل عمليات قاعدة البيانات من انشاء وحذف الجداول والحقول للقيام بالاستعلامات وغيرها واهم ما يميز هذه البنية هو الأداء السريع وتقنيات التخزين المؤقت والجلب الكسول للبيانات وغيرها.

والان بعد ذكر هذه المقارنة النظرية لتوضيح أهمية البنية المعمارية على سهولة كتابة وسرعة أداء النماذج المختلفة، سننتقل للمقارنة بين أشهر النماذج حالياً وهي:

1. **Java Spring Boot**: Hibernate, Eclipse Link, Native Connection

2. **Kotlin Ktor**: Ktorm

3. **PHP**: Eloquent + Laravel, Doctrine + Symfony

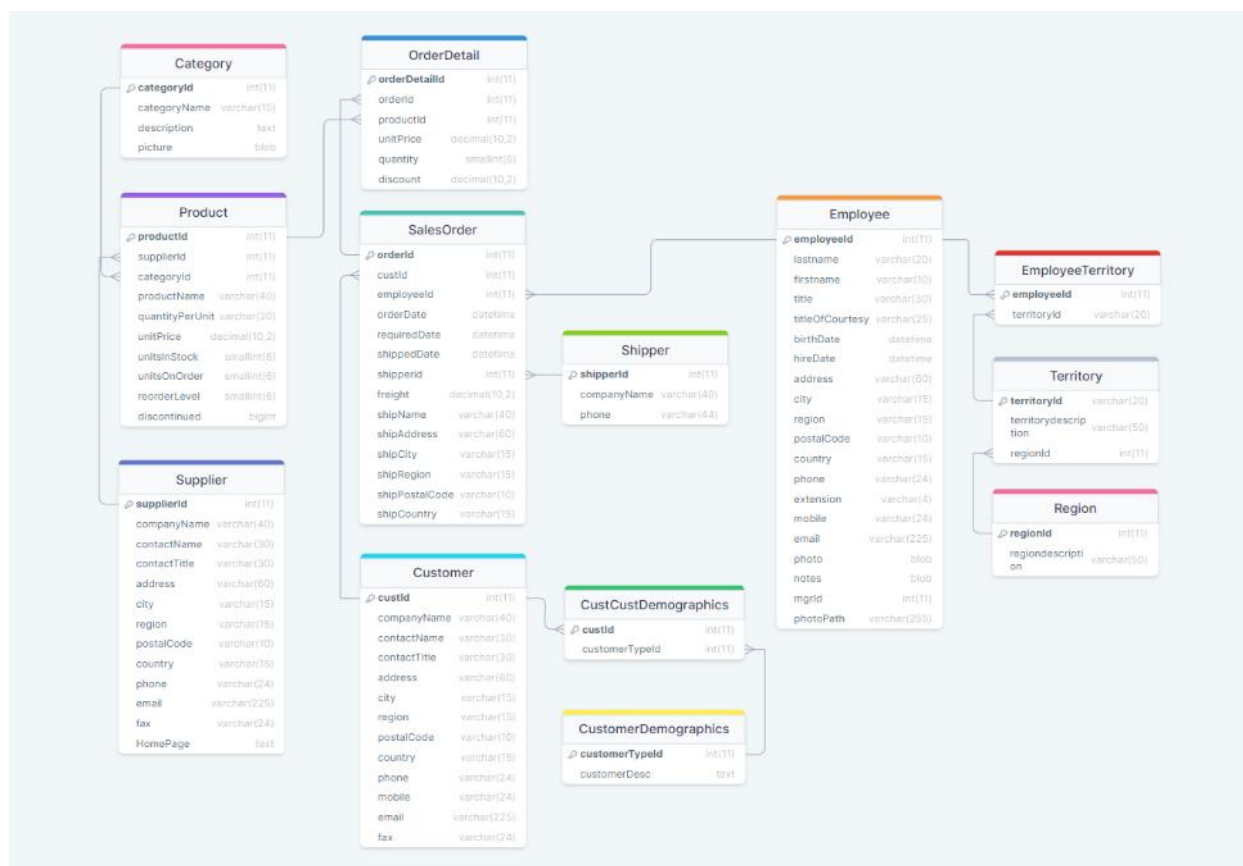
4. **Python**: Django ORM, SQL Alchemy + flask

5. **Node.js**: Sequelize, Prisma, TypeORM (typescript)

6. **.NET Core C#**: Entity Framework, Dapper (Micro ORM)

وتمت المقارنة من خلال تنفيذ 10 استعلامات وقياس زمن التنفيذ لأول استدعاء والمتوسط لتنفيذ كل استعلام 100 مرة متتالية بعد المرة الأولى لرؤية فيما إذا كان هنالك تأثير للتخزين المؤقت على سرعة الاستعلام وبالإضافة لمقارنة النتائج السابقة مع نتائج الاتصال التقليدي في جافا (كونه الأسرع بين باقي اللغات).

كما انه تم استخدام قاعدة بيانات شهيرة **northwind** على نظامي إدارة قواعد بيانات، الأول موجه لقاعدة بيانات **MySQL** والثاني ل **PostgreSQL**، والمخطط التالي يظهر بنية هذه القاعدة والعلاقات بين الجداول فيها.



الشكل 1 المخطط
الهيكلي لقاعدة
البيانات
Northwind

7. بيئة التنفيذ والتجربة العملية

تمت المقارنة على حاسب شخصي وبشكل محلي لكي يكون القياس لزمان التنفيذ فقط دون وجود أي اتصال بعيد مع قواعد البيانات، ومواصفات هذا الحاسب موضحة بالجدول التالي:

Intel Core i7	Ram 16 GB	SSD 1 TB	Windows 11 Pro
12700H	DDR5 5200 MHZ	NVME PCIE 4.0	23H2

وتم تسجيل زمن التنفيذ تلقائياً من خلال صف تم بناؤه بمختلف اللغات البرمجية الموافقة لأطر العمل لقياس زمن تنفيذ كل الاستعلامات على المسلك الرئيسي دون وجود مسالك فرعية والوحدة المستخدمة للقياس هي الميكرو ثانية.

7.1. الاستعلامات المستخدمة للتجربة

1. جلب كل الموظفين (جدول صغير)

```
select * from employee
```

2. جلب كل تفاصيل الطلبات (جدول كبير)

```
select * from orderDetail
```

3. جلب موظف باستخدام رقمه التسلسلي (PK – index)

```
select * from employee where employeeId = ?
```

4. جلب موظف باستخدام اسمه الأول والأخير

```
select * from employee where concat(firstname, ' ', lastname) = ?
```

5. حذف عميل محدد باستخدام رقمه التسلسلي

```
delete from customer where custId = ?
```

6. جلب كل الأماكن ثم تفريغ الجدول ثم إعادة ادخال البيانات المحذوفة

```
select * from territory; // نخزنها في متحول list
delete from territory;
foreach record in list then:
    insert into territory (...) values (...record)
```

7. جلب كل عميل مع اجمالي الطلبات التي قام بدفعها

```
SELECT c.contactName as name, sum(od.unitPrice * od.quantity * (1 - od.discount))
as total
FROM customer c
join salesorder so on cast(so.custId as integer) = c.custId
join orderdetail od on od.orderId = so.orderId
GROUP by c.custId
ORDER by total DESC;
```

8. جلب كل الأماكن السكنية التي يقطنها العملاء (بلد – مدينة) مع اجمالي عدد العملاء القاطنين فيها

```
SELECT c.country, c.city, count(c.custId) as count
FROM customer c
GROUP by c.country, c.city
ORDER by count desc;
```

9. جلب اجمالي مبيعات كل الأصناف

```
SELECT p.productName as name, sum(od.unitPrice * od.quantity * (1 - od.discount))
as total
FROM product p
join orderdetail od on od.productId = p.productId
GROUP by p.productId
ORDER by total DESC;
```

10. جلب الطلبات التي قيمة كل منها بين 5000 و12000

```
SELECT a.* from (
    select so.*, sum(od.unitPrice * od.quantity * (1 - od.discount)) total
    FROM salesorder so
    join orderdetail od on od.orderId = so.orderId
    GROUP by so.orderId
    ORDER by total desc
) a where a.total BETWEEN 5000 and 12000
```

8. التجربة العملية

تم تنفيذ الاستعلامات السابقة على الجهاز الذي تم ذكر مواصفاته بشكل سابق باستخدام الصف البرمجي التالي والذي تم بناء نسخ منه بلغات البرمجة المختلفة لكي تتناسب كل النماذج مع مختلف اطر العمل المدروسة، حيث تم تسجيل اول مدة تنفيذ والمتوسط الحسابي لتنفيذ نفس الاستعلام 100 مرة متتالية، وتم تسجيل النتائج مع قاعدتي البيانات وطرح النتائج في المخططات التالية، حيث يمثل المحور الشاقولي مدة التنفيذ بالميكرو ثانية والمحور الافقي يمثل النموذج الذي تم قياس مدة تنفيذ الاستعلام فيه.

```
public record Task (
    1 usage
    String name, Runnable action
) { }

3 usages
private TaskLogger taskLogger;
3 usages
private TimeUnit timeUnit;

1 usage  Elias Shallouf
public TaskRunner(TaskLogger taskLogger, TimeUnit timeUnit) {...}

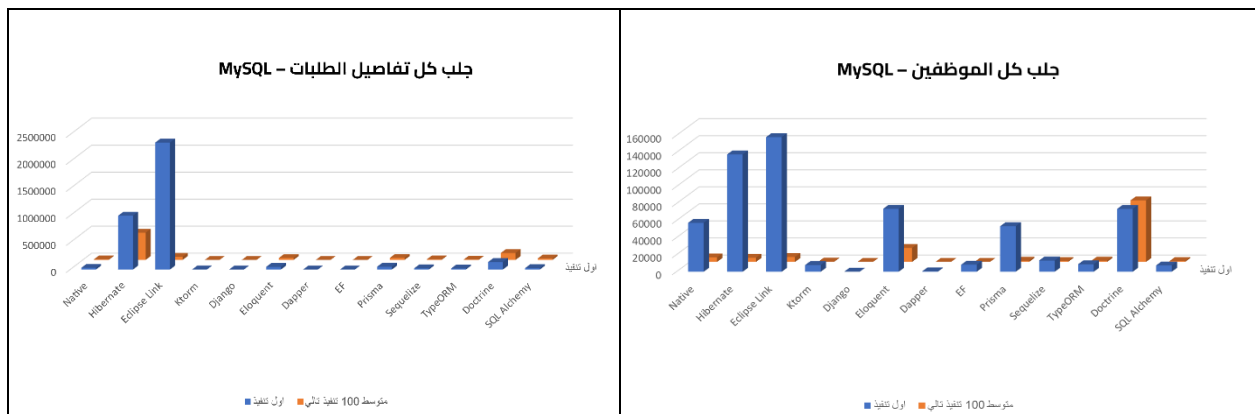
1 usage  Elias Shallouf
public void execute(Task task) {
    long duration = System.nanoTime();

    task.action().run();

    duration = System.nanoTime() - duration;
    taskLogger.log(task, comment: "exec 1", timeUnit.convert(duration));
}
```

الشكل 2 صف TaskRunner المستخدم لقياس ازمدة التنفيذ

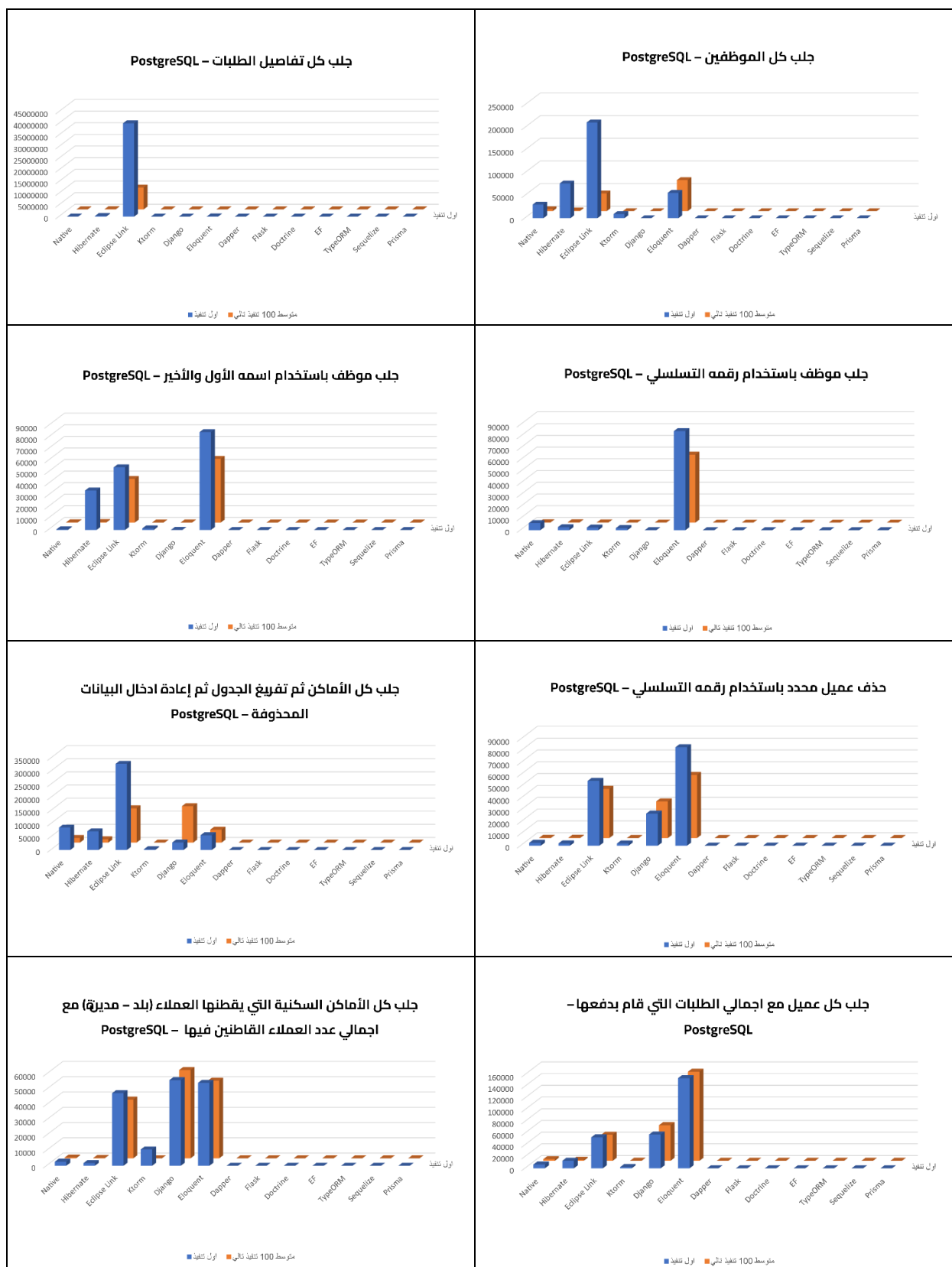
8.1. ازمدة التنفيذ في MySQL

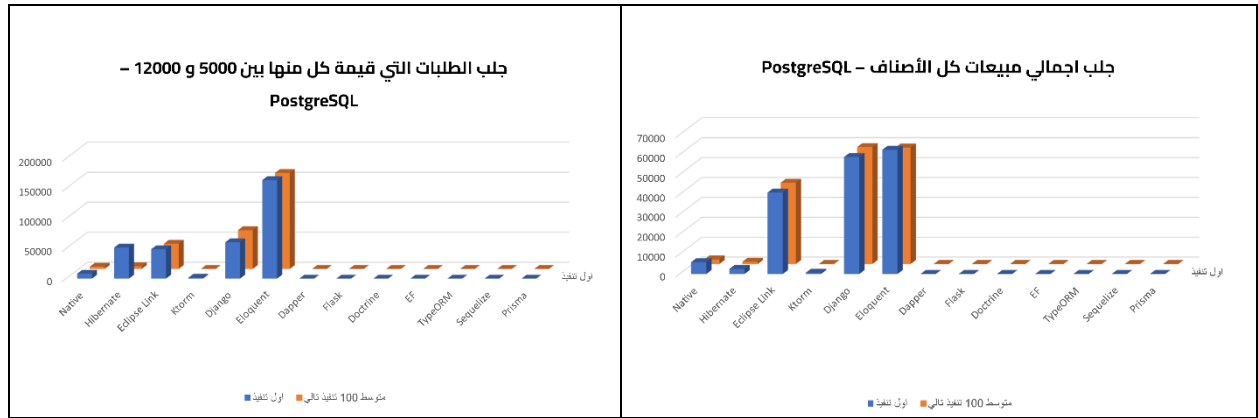




جدول 2 ازمة التنفيذ لكل الاستعلامات مع قاعدة بيانات MySQL

8.2. ازمة التنفيذ في PostgreSQL





جدول 3 ازمة التنفيذ لكل الاستعلامات مع قاعدة بيانات PostgreSQL

8.3. الملاحظات والنتائج المستخلصة بعد القياسات

بعد المقارنات السابقة نأتي للملاحظات عن الأداء مع ذكر اهم مميزات كل نموذج وسيئاته:

النموذج	المميزات	السلبيات	ملاحظات نظرية عن الاداء
Hibernate	معيار JPA، ميزات غنية (التخزين المؤقت، التحميل الكسول/المبكر)، مجتمع كبير	معقد، حمل زائد محتمل، إعدادات مطولة	حمل زائد محتمل، لكن أداء عالي مع الضبط (التخزين المؤقت، تحسين الاستعلامات)
Eclipse Link	معيار JPA، متوافق بدقة، غني بالمميزات	معقد، حمل زائد محتمل	نفس الملاحظات عن الـ Hibernate؛ لكن النتائج الفعلية قد تختلف قليلاً
Ktorm	Kotlin DSL، امن كتابيا مرونة تشبه SQL في Kotlin	مجتمع أصغر من JPA، ميزات أقل شمولاً من ORMs الكاملة	أداء جيد؛ حمل أقل من ORMs الكاملة، أكثر من JDBC
Eloquent	سهل الاستخدام للغاية، تجربة مطور رائعة (Laravel) البساطة	قد يؤدي إلى مشكلة N+1 إذا لم يتم الحذر، مرتبط بشدة غالباً بـ Laravel	سريع جداً لعمليات CRUD البسيطة؛ الأداء يعتمد على التحميل المبكر

Doctrine	غني بالميزات (التخزين المؤقت، التحميل الكسول/المبكر) المستودعات	قد يكون معقدًا، حمل زائد محتمل	نفس ملاحظات Hibernate/EF Core؛ يعتمد على تحسين DQL والتخزين المؤقت للأداء
Django ORM	تكاملي ممتاز مع Django، هجرات مدمجة، إنتاجية عالية، واجهة برمجية موجهة للبايثون بامتياز	مرتبط بشدة بـ Django، احتمال مشكلة N+1 أقل مرونة من SQLAlchemy أحيانًا	جيد عموماً؛ الية جلب ال QuerySets الكسولة تساعد
SQLAlchemy	قوي ومرن تحكم دقيق (Core API)	قد يكون معقدًا بسبب المرونة	أداء عالي، خاصة Core API أداء ORM جيد وقابل للضبط
Sequelize	الأكثر استقرار ل node.js غير متزامن قائم على الوعود، دعم الهجرات	واجهة برمجة التطبيقات تعتبر أحيانًا مطولة/معقدة، البدائل الجديدة نوعاً ما لاقت رواجاً أكثر منه	مقبول عموماً؛ بعض المقارنات تظهره أبطأ من خيارات جديدة مثل Prisma
Prisma	تجربة مطور ممتازة وأمن كتابياً، نظام تهجير رائع، عميل مُولد	نموذج مختلف يتطلب تكيفاً بالأخص لبناء ملف التخطيط	يهدف لأداء عالي؛ غالباً ممتاز في المقارنات
TypeORM	يدعم أنماطاً متعددة، قائم على المحددات مما يجعله الأكثر ملائمة كتابياً ل TypeScript	المحددات مثيرة للجدل المرونة تضيق تعقيداً يوجد تعقيد في واجهته البرمجية	جيد عموماً؛ الأداء يعتمد على نمط الاستخدام Active Record/Data Mapper والإعدادات

أداء جيد جداً ومنافس في الإصدارات الحديثة؛ تحسن كبير عن EF6	يعمل أساساً في بيئة NET. قد يبدو معقداً	تكامل ممتاز مع LINQ مما يجعله امن كتابيا بالإضافة لأدوات مدمجة متعددة دعم من مايكروسوفت تهجير	EF Core
سرعة قريبة من ADO.NET الأصلي؛ من أسرع خيارات الوصول للبيانات في NET.	يتطلب SQL خام (احتمالية الأخطاء)، لا يوجد تتبع تغييرات/وحدة عمل/تخزين مؤقت/تهجير	سريع للغاية، واجهة برمجية خفيفة/بسيطة تحكم كامل في SQL	Dapper

جدول 4 مميزات وسليبيات نماذج الربط العائقي المدروسة مع ملاحظات نظرية عن الاداء

وانطلاقاً من التجربة العملية للتنفيذ قد تم استنتاج بعض الملاحظات كالتالي:

1. نموذج **Eclipse Link** لا يفضل استخدامه ابدا نظرا لوجود مشاكل برمجية مزعجة اثناء عملية التطوير باستخدامه.
2. نموذج **Django** أسرع نموذج في جلب معلومات الجداول بشكل مباشر دون أي عملية ربط مع جداول أخرى لكن لا يقوم بالتخزين المؤقت لذلك سرعة التنفيذ لا تتأثر بعدد المرات التي يتم فيها تنفيذ الاستعلام.
3. نموذج **Ktorm** هو الاسهل في الكتابة من خلال واجهة الجداول المقابلة في Kotlin وافضل النماذج في عملية التخزين المؤقت.
4. **Hibernate** هو النموذج الأكثر استقرارا في جافا ويقدم سرعة تنفيذ قريبة واحيانا اسرع من الاتصال التقليدي.
5. **Eloquent** نموذج جيد لكن بيئة عمل Laravel أصبحت قديمة الطراز نسبيا ونتائجه ليست الأفضل.
6. **PostgreSQL** لديها مشكلة في الربط مع كل الأطر البرمجية وتحتاج لتسمية كل الصفوف والحقول بمحارف صغيرة ولحل هذه المشكلة بسهولة يجب استخدام ميزة إعادة التسمية في حال كان النموذج المستخدم يوفرها (**Lower Case Naming Convention | Strategy**).
7. كما انها لا تدعم العمليات البسيطة مثل الجمع بين الاعمدة ويتم التعويض عنها بالتتابع الموجودة مثل **.concat**

8. ال **Django ORM** في المدارس السابقة قامت بتنفيذ تعليمات الجلب بزمان معدوم عند قياسه وذلك كونها تقوم بهذه العملية على مسلك فرعي وليس الرئيسي.

9. **Entity Framework** اسهل نموذج كتابي ويريح المطور من عبء التعامل مع ال **SQL** بشكل كامل.

10. **Dapper** هو عبارة عن **Micro ORM** أي فقط يستخدم للمقابلة عند عمليات الجلب (صفوف مع قيم بسيطة وليس علاقات بين الجداول على شكل اغراض) ولا يحتوي على أي ميزة لكن يفضل استخدامه لسرعته العالية التي تكافئ الاتصال التقليدي.

11. **Prisma** لا تدعم المقابلة المباشر من خلال الصفوف أي انها تحتاج الى ملف توصيف خاص بها يتم بعد كتابته تحويله باستخدام مفسر لها الى أغراض مكتوبة بلغة ال **JavaScript**.

12. كما انها لا تدعم التوابع الخاصة ب **SQL** وكذلك الامر العمليات الحسابية بين الاعمدة لكنها تدعم فقط العمليات التجميعية على عمود واحد مباشرة (مثل **_sum(price)**) لذلك عند الحاجة لأي استعلام شبه معقد يجب استخدام استعلامات مباشرة باستخدام **queryRaw**.

13. **Sequelize** هو نموذج سهل الكتابة ويدعم توابع ال **SQL** كما يدعم الكتابة المباشر ك **SQL** بحيث يمكن دمجها مع الاستعلامات المكتوبة فيها أيضا.

14. **TypeORM** ابسط من **Sequelize** ويدعم ال **SubQuery** مع حرية كتابة كبيرة دون الحاجة الى خبرة كبيرة مع ال **SQL**.

8.4. مقارنة آليات عمل نماذج ORM المدروسة

في ضوء النتائج العملية التي تم التوصل إليها، يصبح من الضروري التعمق في تحليل الجوانب التقنية والنظرية التي توضح أسباب تفاوت أداء نماذج ORM المختلفة. يعنى هذا القسم بدراسة معمقة للبنية المعمارية لكافة النماذج الخاضعة للتحليل، بداية من آليات بناء الاستعلامات، مروراً بإدارة الكيانات وتحويل النتائج، وانتهاءً بمتطلبات سهولة الاستخدام والتكامل مع لغات البرمجة المختلفة. كما سيتم تسليط الضوء على التحديات التقنية التي تؤثر على الأداء والكفاءة التشغيلية لتلك النماذج، وذلك بتوظيف إطار مقارن يجمع بين النظرية والتطبيق، بما يدعم توجهات اختيار النموذج الأمثل في بيئات التطوير الحديثة.

8.4.1. Hibernate

• نوع النموذج: Data Mapper

- **طريقة العمل:** يعتمد على معيار JPA (Java Persistence API) يستخدم EntityManager لإدارة الكيانات (Entities) ضمن سياق الثبات (Persistence Context) الذي يعمل كوحدة عمل ويتتبع التغييرات. يقوم بربط الجداول بـ POJOs (Plain Old Java Objects) باستخدام التعليقات التوضيحية (Annotations) أو XML. كما يدعم التحميل الكسول (Lazy Loading) والنشط (Eager Loading) والتخزين المؤقت (Caching) للمستوى الأول والثاني.
- **طريقة كتابة الصفوف والتعامل معها:** تُعرّف الكيانات كـ POJOs مع تعليقات JPA التوضيحية (مثل @Entity, @Table, @Id, @Column) وعلاقات مثل @OneToMany, @ManyToOne كما يمكن استخدام الوراثة.
- **طريقة تحويل النتائج:** يقوم بتحويل نتائج الاستعلامات تلقائيًا إلى كائنات Java بناءً على الربط المحدد. يدعم استرجاع أجزاء من الكيانات (Projections).
- **طريقة كتابة الاستعلامات**
 - HQL (Hibernate Query Language)
 - JPQL (Java Persistence Query Language)
 - Criteria API: استعلامات باستخدام كائنات مبرمجة بشكل مسبق.
 - SQL خام
- **سهولة الكتابة وضرورة معرفة SQL:** يوفر مستوى تجريد عالٍ يقلل الحاجة لكتابة SQL مباشرة في البداية. منحى تعلم متوسط إلى مرتفع لاستغلال كافة الميزات. معرفة SQL مفيدة للاستعلامات المعقدة أو تحسين الأداء.
- **سرعة التنفيذ (ملاحظات نظرية):** متوسط كونه يمكن أن يكون هناك حمل زائد (overhead) بسبب التجريد وآلية تتبع التغييرات. التخزين المؤقت (Caching) يمكن أن يحسن الأداء بشكل كبير. الأداء يعتمد بشدة على تصميم الاستعلامات واستخدام نوع التحميل المناسب Lazy/Eager.

```
@Entity
public class Customer implements Serializable {
    @Id Long custId;
    @Column String companyName;
    @Column String contactName;
    @Column String contactTitle;
    @Column String address;
    @Column String city;
    @Column String region;
    @Column String postalCode;
    @Column String country;
    @Column String phone;
    @Column String mobile;
    @Column String email;
    @Column String fax;
```

```
public interface CustomerRepository extends CrudRepository<Customer, Long> {
    @Query("""
        SELECT new com.eliasshallouf.msc.seminar2.domain.model.helpers.CustomerWithTotal(
            c.contactName name,
            sum(od.unitPrice * od.quantity * (1 - od.discount)) total
        )
        FROM Customer c
        join SalesOrder so on so.customer.custId = c.custId
        join OrderDetail od on od.order.orderId = so.orderId
        GROUP by c.custId
        ORDER by total DESC
    """)
    List<CustomerWithTotal> getCustomersWithTotals();

    @Query("""
        SELECT new com.eliasshallouf.msc.seminar2.domain.model.helpers.CustomerLivePlace(
            c.country,
            c.city,
            count(c.custId) as count
        )
        FROM Customer c
        GROUP by c.country, c.city
        ORDER by count desc
    """)
    List<CustomerLivePlace> getCustomersLivePlaces();
}
```

الشكل 3 مثال عن طريقة تعريف الكيانات والاستعلامات في Hibernate

EclipseLink .8.4.2

- نوع النموذج: Data Mapper
- طريقة العمل: تطبيق آخر لمعيار ال JPA، لكن مشابه جدا ل Hibernate لكن يعتبر أحيانا أكثر توافقا مع معيار JPA بشكل صارم مقارنة ب Hibernate الذي يضيف ميزاته الخاصة.
- طريقة كتابة الصفوف والتعامل معها: مطابق ل Hibernate
- طريقة تحويل النتائج: تحويل تلقائي للنتائج إلى كائنات بناءً على الربط.
- طريقة كتابة الاستعلامات: JPQL و Criteria API بشكل أساسي. دعم ل SQL الخام.
- سهولة الكتابة وضرورة معرفة SQL: مشابه ل Hibernate. مستوى تجريد عالٍ، معرفة SQL مفيدة للحالات المتقدمة.
- سرعة التنفيذ (ملاحظات نظرية): مشابه ل Hibernate نظرياً. قد تختلف الأرقام الفعلية في قياسات الأداء المحددة، لكن كلاهما يقع ضمن فئة ORMs كاملة الميزات مع حمل زائد محتمل مقارنة بالوصول المباشر.

```
@Entity
public class Customer implements Serializable {
    @Id Long custId;
    @Column String companyName;
    @Column String contactName;
    @Column String contactTitle;
    @Column String address;
    @Column String city;
    @Column String region;
    @Column String postalCode;
    @Column String country;
    @Column String phone;
    @Column String mobile;
    @Column String email;
    @Column String fax;
```

```
@Repository
public interface CustomerRepository extends JpaRepository<Customer, Long> {
    @Query(value = """
        SELECT c.contactName name, sum(od.unitPrice * od.quantity * (1 - od.discount)) total
        FROM Customer c
        join SalesOrder so on so.customer.custId = c.custId
        join OrderDetail od on od.order.orderId = so.orderId
        GROUP by c.custId
        ORDER by total DESC
    """)
    List<Object[]> getCustomersWithTotals();

    @Query(value = """
        SELECT c.country country, c.city city, count(c.custId) count
        FROM Customer c
        GROUP by c.country, c.city
        ORDER by count desc
    """, nativeQuery = true)
    List<Object[]> getCustomersLivePlaces();
}
```

الشكل 4 مثال عن طريقة تعريف الكيانات والاستعلامات في EclipseLink

Ktorm .8.4.3

- نوع النموذج: يمكن اعتباره DSL-based SQL Builder / Lightweight ORM Mapper (ليس نمط Data Mapper أو Active Record تقليدي صارم).
- طريقة العمل: يوفر DSL (Domain Specific Language) مكتوب بلغة Kotlin لبناء استعلامات SQL بطريقة آمنة النوع (type-safe) وبرمجية. يركز على بناء الاستعلامات وتحويل النتائج، مع تجريد أقل لوحداث العمل من النماذج السابقة.
- طريقة كتابة الصفوف والتعامل معها: يتم تعريف الكيانات عادة كـ data class أو interface في Kotlin. يستخدم Ktorm واجهات لتعريف الجداول وربطها بالكيانات.
- طريقة تحويل النتائج: يوفر آليات لتحويل نتائج الاستعلامات المبنية بالـ DSL إلى كائنات Kotlin.
- طريقة كتابة الاستعلامات: باستخدام الـ DSL الخاص بـ Ktorm في كود Kotlin لبناء استعلامات SELECT, INSERT, UPDATE, DELETE. كما يدعم بناء استعلامات مثل الربط والتجميع والاستعلامات الفرعية.
- سهولة الكتابة وضرورة معرفة SQL: مصمم لمطوري Kotlin، يستفيد من ميزات اللغة. يتطلب فهم كيفية بناء الاستعلامات باستخدام DSL. معرفة الـ SQL مفيدة لفهم ما يتم توليده لكن ليست ضرورية.
- سرعة التنفيذ (ملاحظات نظرية): أدائه جيداً، مع حمل زائد أقل من ORMs كاملة الميزات، ولكنه أعلى من استخدام الـ JDBC المباشر، بسبب بناء الاستعلامات وتحويل النتائج.

```
interface Customer { Entity<Customer> {
    companion object : Entity.Factory<Customer>()

    var id: Int
    var companyName: String
    var contactName: String
    var contactTitle: String
    var address: String
    var city: String
    var region: String
    var postalCode: String
    var country: String
    var phone: String
    var mobile: String
    var email: String
    var fax: String
}

object Customers : Table<Customer>("customer") {
    val id = int("custid").primaryKey().bindTo { it.id }
    val companyName = varchar("companyname").bindTo { it.companyName }
    val contactName = varchar("contactname").bindTo { it.contactName }
    val contactTitle = varchar("contacttitle").bindTo { it.contactTitle }
    val address = varchar("address").bindTo { it.address }
    val city = varchar("city").bindTo { it.city }
    val region = varchar("region").bindTo { it.region }
    val postalCode = varchar("postalcode").bindTo { it.postalCode }
    val country = varchar("country").bindTo { it.country }
    val phone = varchar("phone").bindTo { it.phone }
    val mobile = varchar("mobile").bindTo { it.mobile }
    val email = varchar("email").bindTo { it.email }
    val fax = varchar("fax").bindTo { it.fax }
}
```

```
suspend fun getCustomersWithTotals() = withContext(Dispatchers.IO) {
    val name = Customers.contactName.aliases("name")
    val total = (
        sum(
            OrderDetails.unitPrice
            times
            OrderDetails.quantity
            times
            (1.0 minus OrderDetails.discount)
        )
    ).aliases("total")

    db.from(Customers)
        .leftJoin(SalesOrders, on = it.id.eq(SalesOrders.customer))
        .leftJoin(OrderDetails, on = SalesOrders.id.eq(OrderDetails.order))
        .select(name, total)
        .groupBy(it.id)
        .orderBy(total.desc())
        .map {
            CustomerWithTotal(
                it[name] ?: "",
                it[total] ?: 0.0
            )
        }
}
```

الشكل 5 مثال عن طريقة تعريف الكيانات والاستعلامات في Ktorm

Eloquent .8.4.4

- نوع النموذج: Active Record.
- طريقة العمل: لكل كائن نموذج (Model) يمثل جدولاً في قاعدة البيانات ويحتوي على المنطق اللازم للتعامل مع هذا الجدول بشكل مبني بشكل مسبق (مثل find(), delete(), save()).

- طريقة كتابة الصفوف والتعامل معها: تُعرّف النماذج (Models) كأصناف PHP ترث من *Illuminate\Database\Eloquent\Model*. يتم تعريف العلاقات (Relationships) كدوال داخل النموذج.
- طريقة تحويل النتائج: يتم تحويل النتائج تلقائيًا إلى كائنات من صنف النموذج.
- طريقة كتابة الاستعلامات: من خلال استخدام دوال ستاتيكية مرتبطة بالكائنات على النموذج نفسه (مثل `(User::find(1), User::where('active', 1)->get())`). يوفر واجهة Query Builder. كما يدعم الـ SQL الخام.
- سهولة الكتابة وضرورة معرفة SQL: سهل جداً للبدء والاستخدام خاصة للمهام الشائعة. يتطلب معرفة أقل بـ SQL في البداية. معرفة SQL مفيدة لتحسين الاستعلامات المعقدة.
- سرعة التنفيذ (ملاحظات نظرية): متوسط السرعة بالأخص للعمليات البسيطة.

```
class Customer extends Model
{
    use HasFactory;

    protected $table = 'customer';

    protected $fillable = [
        'custid', 'companyName', 'contactName',
        'contactTitle', 'address', 'city',
        'region', 'postalCode', 'country',
        'phone', 'mobile', 'email', 'fax'
    ];

    public function salesOrder()
    {
        return $this->hasMany(SalesOrder::class, 'custid', 'custid');
    }

    public function customerDemographics()
    {
        return $this->belongsToMany(
            CustomerDemographics::class,
            'custcustdemographics',
            'custid',
            'customertypeid'
        );
    }
}
```

```
$customerPlaces = Customer::select(
    'country',
    'city',
    DB::raw('count(*) as total_customers')
)
->groupBy('country', 'city')
->orderBy('total_customers', 'desc')
->get();
```

الشكل 6 مثال عن طريقة تعريف الكيانات والاستعلامات في Eloquent

Doctrine .8.4.5

- نوع النموذج: Data Mapper.
- طريقة العمل: مشابه لـ Hibernate/EclipseLink. يستخدم EntityManager لإدارة الكيانات. يعتمد على نمط المستودع (Repository pattern) لفصل منطق الاستعلامات. يربط الجداول بـ POPOs (Plain Old PHP Objects) باستخدام التعليقات التوضيحية (Annotations/Attributes) أو XML أو YAML. يدعم التحميل الكسول/النشط والتخزين المؤقت L1/L2.
- طريقة كتابة الصفوف والتعامل معها: تُعرّف الكيانات كـ POPOs مع تعليقات توضيحية أو إعدادات أخرى لتحديد الربط والعلاقات.

- طريقة تحويل النتائج: تحويل تلقائي للنتائج إلى كائنات بناءً على الربط.
- طريقة كتابة الاستعلامات:
 - DQL (Doctrine Query Language)
 - QueryBuilder API
 - SQL خام عبر NativeQuery
- سهولة الكتابة وضرورة معرفة SQL: مستوى تجريد عالٍ. يتطلب فهم مفاهيم مثل EntityManager والمستودعات. معرفة أقل بـ SQL مطلوبة للبدء. معرفة SQL مفيدة للحالات المتقدمة.
- سرعة التنفيذ (ملاحظات نظرية): مشابه لـ Hibernate/EclipseLink. الأداء يعتمد على الاستعلامات والتخزين المؤقت والتحميل. يمكن أن يكون هناك حمل زائد مقارنة بالوصول المباشر.

```
#[ORM\Entity(repositoryClass: CustomerRepository::class)]
#[ORM\Table(name: 'customer')]
class Customer
{
    #[ORM\Id]
    #[ORM\GeneratedValue(strategy: 'AUTO')]
    #[ORM\Column(name: 'custId', type: 'bigint')]
    private ?int $custId = null;

    #[ORM\Column(name: 'companyName', type: 'string', length: 255)]
    private string $companyName;

    #[ORM\Column(name: 'contactName', type: 'string', length: 255, nullable: true)]
    private ?string $contactName = null;

    #[ORM\Column(name: 'contactTitle', type: 'string', length: 255, nullable: true)]
    private ?string $contactTitle = null;

    #[ORM\Column(type: 'string', length: 255, nullable: true)]
    private ?string $address = null;

    #[ORM\Column(type: 'string', length: 255, nullable: true)]
    private ?string $city = null;

    #[ORM\Column(type: 'string', length: 255, nullable: true)]
    private ?string $region = null;

    #[ORM\Column(name: 'postalCode', type: 'string', length: 255, nullable: true)]
    private ?string $postalCode = null;

    #[ORM\Column(type: 'string', length: 255, nullable: true)]
    private ?string $country = null;

    #[ORM\Column(type: 'string', length: 255, nullable: true)]
    private ?string $phone = null;

    #[ORM\Column(type: 'string', length: 255, nullable: true)]
    private ?string $mobile = null;
}

$customerRepository = $this->entityManager->getRepository(Customer::class);

$query = $customerRepository->createQueryBuilder('c')
    ->select('c.country, c.city, COUNT(c.custId) as total_customers')
    ->groupBy('c.country, c.city')
    ->orderBy('total_customers', 'DESC')
    ->getQuery();

$customerPlaces = $query->getResult();
```

الشكل 7 مثال عن طريقة تعريف الكيانات والاستعلامات في Doctrine

Django .8.4.6

- نوع النموذج: يعتبر قريباً من Active Record (النماذج تعرف بياناتها وسلوكها، ولكن الاستعلامات تتم غالباً عبر Manager).
- طريقة العمل: يُعرّف كل نموذج (Model) هيكل جدول قاعدة البيانات. يوفر واجهة برمجة تطبيقات (API) عالية المستوى للاستعلامات (QuerySets) التي تعتبر كسولة (lazy) بطبيعتها (لا يتم تنفيذ الاستعلام إلا عند الحاجة للنتائج).
- طريقة كتابة الصفوف والتعامل معها: تُعرّف النماذج كأصناف Python ترث من `django.db.models.Model`. تُعرّف الحقول والعلاقات كخصائص (attributes) على الصنف.
- طريقة تحويل النتائج: QuerySets تُرجع كائنات من صنف النموذج عند الطلب.

- طريقة كتابة الاستعلامات: باستخدام الـ Manager على النموذج (مثل)، `Model.objects.filter(...)`، `Model.objects.get(...)` لإنشاء QuerySets. دعم لـ SQL الأصلي.
- سهولة الكتابة وضرورة معرفة SQL: سهل التعلم والبدء به خاصة ضمن إطار Django. يقلل الحاجة لكتابة SQL بشكل كبير. معرفة SQL مفيدة لتحسين الأداء وفهم الاستعلامات المعقدة
- سرعة التنفيذ (ملاحظات نظرية): أداء جيد بشكل عام كون طبيعة الـ QuerySets الكسولة تساعد. يجب الانتباه لتجنب مشاكل N+1 عند استخدام `select_related` و `prefetch_related` التي تقوم بتحميل الأغراض المرتبطة بعلاقة مع الجدول الحالي باستخدام المفتاح الأجنبي.

```
from django.db import models

class Customer(models.Model):
    custId = models.AutoField(primary_key=True, db_column="custid")
    companyName = models.CharField(max_length=40, db_column="companyname")
    contactName = models.CharField(max_length=30, db_column="contactname")
    contactTitle = models.CharField(max_length=30, db_column="contacttitle")
    address = models.CharField(max_length=60)
    city = models.CharField(max_length=15)
    region = models.CharField(max_length=15, blank=True, null=True)
    postalCode = models.CharField(max_length=10, db_column="postalcode")
    country = models.CharField(max_length=15)
    phone = models.CharField(max_length=24)
    email = models.EmailField(max_length=225)
    fax = models.CharField(max_length=24, blank=True, null=True)
    mobile = models.CharField(max_length=24)

    class Meta:
        db_table = 'customer'

queryset = Customer.objects.values('country', 'city').annotate(
    count=Count('custId')
).order_by('-count')

results = list(queryset)
```

الشكل 8 مثال عن طريقة تعريف الكيانات والاستعلامات في Django

8.4.7. SQLAlchemy

- نوع النموذج: Data Mapper لكنه يوفر تجربة أقرب لـ SQL Builder.
- طريقة العمل: يُعرّف كل نموذج (Model) هيكل جدول قاعدة البيانات. كما يوفر بنية اتصال مباشر مع قاعدة البيانات تمثل الجلسة Session حيث تستخدم لبناء تعليمات الـ SQL بشكل برمجي.
- طريقة كتابة الصفوف والتعامل معها: تُعرّف النماذج كصفوف Python ترث من الصف `db.Model` حيث `db` هي غرض منشأ باستخدام الباني (`SQLAlchemy()`). تُعرّف الحقول والعلاقات كخصائص (`attributes`) على الصف كـ متحولات من النوع `db.Column`.
- طريقة تحويل النتائج: تُرجع كائنات من صنف النموذج.
- طريقة كتابة الاستعلامات:
 - باستخدام الجلسة لإنشاء تعليمات معقدة مثل الربط `db.session.query(...)`

○ باستخدام صف النموذج بشكل مباشر من أجل عمليات الجلب ذات شروط المقارنة البسيطة

`Model.query.get(id), Model.query.filter_by(column=value...)`

○ SQL خام

- سهولة الكتابة وضرورة معرفة SQL: يتطلب فهم مفاهيم مثل Session ووحدة العمل. معرفة SQL مفيدة لتحسين الأداء وفهم الاستعلامات المعقدة
- سرعة التنفيذ (ملاحظات نظرية): أداء يقترب من الوصول المباشر. أداء النموذج بشكل عام يعتمد على الاستخدام (مشابه لـ Hibernate/Doctrine).

```
db = SQLAlchemy()

class Customer(db.Model):
    __tablename__ = 'Customer'
    custId = db.Column(db.Integer, primary_key=True)
    companyName = db.Column(db.String(40))
    contactName = db.Column(db.String(30))
    contactTitle = db.Column(db.String(30))
    address = db.Column(db.String(60))
    city = db.Column(db.String(15))
    region = db.Column(db.String(15), nullable=True)
    postalCode = db.Column(db.String(10))
    country = db.Column(db.String(15))
    phone = db.Column(db.String(24))
    email = db.Column(db.String(225))
    fax = db.Column(db.String(24), nullable=True)
    mobile = db.Column(db.String(24))
```

```
def count_customers_by_city():
    results = db.session.query(
        Customer.country,
        Customer.city,
        func.count(Customer.custId).label('count')
    ).group_by(Customer.country, Customer.city)
    .order_by(func.count(Customer.custId).desc())
    .all()
```

الشكل 9 مثال عن طريقة تعريف الكيانات والاستعلامات في SQLAlchemy

Sequelize .8.4.8

- نوع النموذج: يعتبر مزيجاً بين Active Record و Data Mapper (النماذج لديها دوال استعمال، ولكن هناك أيضاً تركيز على تعريف النموذج بشكل منفصل).
- طريقة العمل: يُعرّف النماذج (Models) التي تمثل الجداول. يوفر دوال على النماذج لإجراء عمليات CRUD والاستعلامات. يعتمد على الوعود (Promises) للعمليات غير المتزامنة (Async).
- طريقة كتابة الصفوف والتعامل معها: تُعرّف النماذج باستخدام `sequelize.define()` مع تحديد الأعمدة وأنواعها والعلاقات.
- طريقة تحويل النتائج: يحول النتائج إلى كائنات JavaScript (أو نُسخ من النموذج).
- طريقة كتابة الاستعلامات:

○ باستخدام دوال على النموذج (مثل `Model.findAll()`, `Model.create()`) كما يوفر خيارات لتحديد

الشروط والربط بين الجداول والكتابة كـ SQL داخل تعليمات الارجاع برمجيا باستخدام

`.Sequelize.literal()`

○ SQL خام.

- سهولة الكتابة وضرورة معرفة SQL: يتطلب معرفة SQL لبناء الاستعلامات المعقدة كما انه معقد الكتابة نوعياً.
- سرعة التنفيذ (ملاحظات نظرية): أداء مقبول عمومًا، ولكن كانت هناك بعض الانتقادات حول الأداء في بعض الحالات المعقدة أو تحت الحمل العالي مقارنة ببعض البدائل الأحدث.

```
const Customer = (sequelize) => sequelize.define('customer', {
  custId: {
    type: DataTypes.BIGINT,
    primaryKey: true,
    autoIncrement: true,
    field: 'custid'
  },
  companyName: { ... },
  contactName: { ... },
  contactTitle: { ... },
  address: { ... },
  city: { ... },
  region: { ... },
  postalCode: { ... },
  country: { ... },
  phone: { ... },
  mobile: { ... },
  email: { ... },
  fax: { ... }
}, { tableName: 'customer', timestamps: false });
```

```
(await Customer.findAll({
  attributes: [
    'country',
    'city',
    [Sequelize.fn('count', Sequelize.col('custid')), 'count']
  ],
  group: ['country', 'city'],
  order: [[Sequelize.col('count'), 'DESC']]
})).map(r => r.toJSON());
```

الشكل 10 مثال عن طريقة تعريف الكيانات والاستعلامات في Sequelize

Prisma .8.4.9

- نوع النموذج: نوع جديد "Next-generation ORM" (ليس Data Mapper أو Active Record تقليدي).
- طريقة العمل: يعتمد على ملف مخطط (Schema file – schema.prisma) يمثل الهيكل لقاعدة البيانات ونماذج التطبيق. يقوم Prisma بتوليد عميل قاعدة بيانات (Prisma Client) آمن النوع (type-safe) بالكامل بناءً على هذا المخطط.
- طريقة كتابة الصفوف والتعامل معها: يتم تعريف النماذج في ملف schema.prisma. والعميل المولد (Prisma Client) يوفر هذه النماذج ككائنات TypeScript/JavaScript.
- طريقة تحويل النتائج: Prisma Client يحول النتائج تلقائيًا إلى كائنات JavaScript/TypeScript آمنة النوع.
- طريقة كتابة الاستعلامات:

- باستخدام Prisma Client المولّد، والذي يوفر واجهة برمجة تطبيقات (API) واضحة، سلسلة، وأمنة النوع للاستعلامات (مثل `prisma.user.findMany({ where: ... })`)، حيث `prisma` هو غرض منشأ من الصف المولّد `PrismaClient`.

○ SQL خام.

- سهولة الكتابة وضرورة معرفة SQL: يتطلب معرفة SQL لبناء الاستعلامات المعقدة كون الطرق الموجودة فيه محدودة الخيارات كما انه معقد الكتابة بالأخص مرحلة بناء النماذج كونه يعتمد على الكتابة بلغة زائفة خاصة به.
- سرعة التنفيذ (ملاحظات نظرية): يهدف إلى أداء عالٍ من خلال تحسين الاستعلامات المولّدة واستخدام محرك استعلام مكتوب بلغة Rust. غالبًا ما يُعتبر أداؤه جيدًا جدًا ومنافسًا قويًا.

```
generator client {
  provider = "prisma-client-js"
}

datasource db {
  provider = "postgresql"
  url      = env("DATABASE_URL")
}

model Customer {
  custId BigInt @id @default(autoincrement()) @map("custid")
  companyName String @map("companyname")
  contactName String @map("contactname")
  contactTitle String? @map("contacttitle")
  address String? @map("address")
  city String? @map("city")
  region String? @map("region")
  postalCode String? @map("postalcode")
  country String? @map("country")
  phone String? @map("phone")
  mobile String? @map("mobile")
  email String? @map("email")
  fax String? @map("fax")
  SalesOrder SalesOrder[]
  customersDemographics CustomerCustomerDemographics[]

  @@map("customer")
}
```

```
await prisma.customer.groupBy({
  by: ['country', 'city'],
  count: {
    custId: true
  },
  orderBy: {
    count: {
      custId: 'desc'
    }
  }
})
```

الشكل 11 مثال عن طريقة تعريف الكيانات والاستعلامات في Prisma

8.3.10 TypeORM

- نوع النموذج: يدعم كلا النمطين: Data Mapper و Active Record (يمكن الاختيار)
- طريقة العمل: يمكن العمل به بشكل مشابه لـ Hibernate/Doctrine (باستخدام EntityManager والمستودعات - Data Mapper) أو بشكل مشابه لـ Eloquent/Django ORM (حيث تحتوي الكيانات على دوال الحفظ والحذف - Active Record). يستخدم المُزخرفات (Decorators) في TypeScript بشكل مكثف لتعريف الكيانات والربط.
- طريقة كتابة الصفوف والتعامل معها: تُعرّف الكيانات كأصناف TypeScript مع مُزخرفات مثل `@Entity`, `@Column`, `@PrimaryGeneratedColumn`، وعلاقات مثل `OneToMany@`.

- طريقة تحويل النتائج: تحويل تلقائي للنتائج إلى كائنات TypeScript.
- طريقة كتابة الاستعلامات:

- باستخدام QueryBuilder يتم انشاءه باستخدام مستودع النموذج
- دوال معرفة مسبقا على النماذج
- SQL خام

- سهولة الكتابة وضرورة معرفة SQL: مرن جدًا ولكنه قد يكون معقدًا بعض الشيء بسبب دعمه لأنماط متعددة وخيارات الإعدادات الكثيرة. كما ان معرفة SQL ضرورية في حالة استخدام باني الاستعلامات.
- سرعة التنفيذ (ملاحظات نظرية): أداء جيد بشكل عام يعتمد على كيفية كتابة الاستعلامات.

```
@Entity()
export class Customer {
  @PrimaryGeneratedColumn('increment', { type: 'bigint' })
  custId!: number;

  @Column({ type: 'text', nullable: true })
  companyName?: string;

  @Column({ type: 'text', nullable: true })
  contactName?: string;

  @Column({ type: 'text', nullable: true })
  contactTitle?: string;

  @Column({ type: 'text', nullable: true })
  address?: string;

  @Column({ type: 'text', nullable: true })
  city?: string;

  @Column({ type: 'text', nullable: true })
  region?: string;

  @Column({ type: 'text', nullable: true })
  postalCode?: string;

  @Column({ type: 'text', nullable: true })
  country?: string;

  @Column({ type: 'text', nullable: true })
  phone?: string;
}
```

```
let connection = await dataSource.initialize();

const entityManager = dataSource.manager;

const customerRepository = entityManager.getRepository(Customer);

await customerRepository
  .createQueryBuilder('customer')
  .select('customer.country as country')
  .select('customer.city as city')
  .select('COUNT(customer.custId) as count')
  .groupBy('customer.country, customer.city')
  .orderBy('count', 'DESC')
  .getRawAndEntities();
```

الشكل 12 مثال عن طريقة تعريف الكيانات والاستعلامات في TypeORM

8.3.11 Entity Framework Core

- نوع النموذج: Data Mapper.
- طريقة العمل: يستخدم DbContext كوحدة عمل (Unit of Work) ومستودع (Repository). يربط الجداول بأصناف (Plain Old CLR Objects) POCO باستخدام تعليقات توضيحية (Data Annotations) أو واجهة برمجة تطبيقات سلسلة (Fluent API) في OnModelCreating. يتتبع التغييرات تلقائيًا. يدعم التحميل الكسول/النشط.
- طريقة كتابة الصفوف والتعامل معها: تُعرّف الكيانات كـ POCOs. يتم الربط والعلاقات عبر Data Annotations أو Fluent API أي انه مشابه لطريقة تعريف الصفوف باستخدام معيار ال JPA في Java.
- طريقة تحويل النتائج: تحويل تلقائي للنتائج إلى كائنات POCO.

• طريقة كتابة الاستعلامات:

- باستخدام LINQ (Language-Integrated Query) وهي الطريقة الأساسية والمفضلة، توفر استعلامات آمنة النوع ومدمجة باللغة (C# / VB.NET).

○ SQL خام.

- سهولة الكتابة وضرورة معرفة SQL: مستوى تجريد عالٍ جدًا بفضل LINQ. يقلل الحاجة لكتابة SQL بشكل كبير. منحى تعلم متوسط لفهم DbContext و LINQ المتقدم. معرفة SQL مفيدة لتحسين الأداء أو الاستعلامات المعقدة جدًا.

- سرعة التنفيذ (ملاحظات نظرية): يعتبر أدائه الآن جيدًا جدًا ومنافسًا، مع استمرار وجود بعض الحمل الزائد مقارنة بالوصول المباشر.

```
[Table("customer")]
0 references
public class Customer
{
    [Key]
    [DatabaseGenerated(DatabaseGeneratedOption.Identity)]
    0 references
    public int CustId { get; set; }

    0 references
    public string? CompanyName { get; set; }

    0 references
    public string? ContactName { get; set; }

    0 references
    public string? ContactTitle { get; set; }

    0 references
    public string Address { get; set; }

    0 references
    public string City { get; set; }

    0 references
    public string? Region { get; set; }

    0 references
    public string PostalCode { get; set; }

    0 references
    public string Country { get; set; }

    0 references
    public string Phone { get; set; }
}
```

```
public List<CustomerLivePlace> GetCustomersLivePlaces() {
    return [.. context.Customers
        .GroupBy(c => new { Country = c.Country, City = c.City })
        .Select(g => new
        {
            g.Key.Country,
            g.Key.City,
            Count = g.Count()
        })
        .OrderByDescending(x => x.Count)
        .Select(g => new CustomerLivePlace
        {
            Country = g.Country,
            City = g.City,
            Count = g.Count
        })
    ];
}
```

الشكل 13 مثال عن طريقة تعريف الكيانات والاستعلامات في Entity Framework Core

8.3.12 Dapper

- نوع النموذج: Micro ORM / Object Mapper (ليس ORM كامل بالمعنى التقليدي).
- طريقة العمل: مكتبة بسيطة تركز على شيء واحد: تنفيذ استعلامات SQL وتحويل نتائجها إلى كائنات POCO. لا يوفر تتبع تغييرات تلقائي، ولا وحدة عمل، ولا بناء استعلامات بلغة أخرى غير SQL، ولا تحميل كسول.

- طريقة كتابة الصفوف والتعامل معها: تُعرّف الكيانات كـ POCOs بسيطة. لا يتطلب Dapper أي تعليقات توضيحية أو إعدادات ربط معقدة؛ يعتمد على تطابق أسماء الأعمدة مع خصائص (حقول) الكائن (يمكن تخصيصه).
- طريقة تحويل النتائج: يقوم بتحويل النتائج بكفاءة عالية جدًا إلى كائنات POCO باستخدام دوال الإضافة (Extension Methods) على IDbConnection.
- طريقة كتابة الاستعلامات: كتابة SQL خام (Raw SQL) بشكل مباشر كسلاسل نصية، مع دعم ممتاز للمعاملات (Parameters) للحماية من SQL Injection.
- سهولة الكتابة وضرورة معرفة SQL: يتطلب معرفة ممتازة بـ SQL. يوفر تجريّدًا بسيطًا جدًا فوق ADO.NET.
- سرعة التنفيذ (ملاحظات نظرية): سريع جدًا، أداؤه قريب جدًا من ADO.NET المباشر (مثل JDBC في Java). يعتبر من أسرع طرق الوصول للبيانات في .NET بعد ADO.NET الخام بسبب قلة الحمل الزائد.

```
public class Customer
{
    0 references
    public int CustId { get; set; }
    0 references
    public string CompanyName { get; set; }
    0 references
    public string ContactName { get; set; }
    0 references
    public string ContactTitle { get; set; }
    0 references
    public string Address { get; set; }
    0 references
    public string City { get; set; }
    0 references
    public string Region { get; set; }
    0 references
    public string PostalCode { get; set; }
    0 references
    public string Country { get; set; }
    0 references
    public string Phone { get; set; }
    0 references
    public string Mobile { get; set; }
    0 references
    public string Email { get; set; }
    0 references
    public string Fax { get; set; }
}
```

```
public List<CustomerLivePlace> GetCustomersLivePlaces()
{
    return (List<CustomerLivePlace>) connection.Query<CustomerLivePlace>(@"
        SELECT c.country, c.city, count(c.custId) as count
        FROM customer c
        GROUP by c.country, c.city
        ORDER by count desc;
    ");
}
```

الشكل 14 مثال عن طريقة تعريف الكيانات والاستعلامات في Dapper

وبشكل ملخص، وبعد إتمام المقارنة العملية بين سرعة أداء كل من النماذج التي تمت دراستها وذكر التفاصيل النظرية لألية عمل كل منها ونوع النموذج الأساسي فيها ومدى ضرورة الخبرة في لغة الاستعلامات الهيكلية للتعامل معاً، سيتم تلخيص الفروقات بينها من حيث الية التصميم والملاحظات النظرية عنها في الجداول التالية:

النموذج	نوع النموذج	الآلية العمل الأساسية	الآلية الاستعلامات	الآلية تعريف الكيانات
Hibernate	Data Mapper (JPA)	EntityManager, Unit of Work, L1/L2 Cache	JPQL / HQL, Criteria API	POJOs + JPA Annotations/XML
Eclipse Link	Data Mapper (JPA)	EntityManager, Unit of Work, L1/L2 Cache	JPQL, Criteria API	POJOs + JPA Annotations/XML
Ktorm	SQL Builder / Lightweight ORM	Kotlin DSL for SQL يركز على الكتابة الآمنة وأقل تجريد من النماذج الكاملة	Kotlin DSL	Kotlin Data Classes + Table Definitions (Interfaces)
Eloquent	Active Record	نماذج تتضمن منطق الثبات (مثل save()), تعتمد على الاصطلاحات التلقائية وطبقة تجريد قاعدة البيانات DBAL	توابع النموذج QueryBuilder	Models extending base class, Conventions + Properties/Methods
Doctrine	Data Mapper	EntityManager, Unit of Work, L1/L2 Cache	DQL / QueryBuilder	POJOs + JPA Annotations/XML
Django ORM	أقرب ل Active Record	النماذج مع دوال حفظ ومدير النماذج مع خاصية objects للاستعلامات	QuerySet API	النماذج ترث صف أساسي والحقول تمثل كخصائص له
SQLAlchemy	Data Mapper	Session (Unit of Work) مقابلة بشكل تصريحي	Session Query / Core API	صفوف ترث صف النموذج الأساسي أو صف مع ربط بشكل تصريحي
Sequelize	Active Record / Data Mapper	نماذج تتضمن توابع للحفظ وتوابع ثابتة للاستعلامات	توابع النموذج	عن طريق دالة sequelize.define()

ملف تخطيط "schema.prisma" يتم الوصول له من خلال صف العميل المولد	Prisma Client API	مقاد بملف تخطيط قاعدة البيانات ويُنشئ عميلاً آمناً للنوع ، يركز على تجربة المطور	Next-gen ORM / Generator	Prisma
صفوف TypeScript +محددات Decorators (@Entity...)	QueryBuilder توابع النماذج	EntityManager/Repositories BaseEntity المقابلة باستخدام المحددات اتصالات غير متزامنة	Data Mapper / Active Record	TypeORM
POCOs (Plain Old CLR Object)s + Data Annotations/Fluent API Mapping	LINQ (Language Integrated Query)	DbContext (Unit of Work, Repository Gateway)	Data Mapper	EF Core
POCOs	SQL خام	ينفذ استعلامات ال SQL الخام فقط ويقوم بمقابلة النتائج بصفوف برمجية لا يوجد توليد للاستعلامات	Micro- ORM	Dapper

جدول 5 ملخص عن نماذج الربط العلائقي المدروسة

الخاتمة:

في الختام، تجدر الإشارة إلى أن جميع القياسات والتجارب التي أُجريت في هذا البحث تمت على جهاز حاسوبي بمواصفات فوق المتوسطة، واستخدمت اتصالاً مباشراً إلى قاعدة البيانات عبر مسلك رئيسي واحد دون الاعتماد على تقنيات التخزين المؤقت. كما أن بيئة الاختبار استندت إلى قاعدة بيانات صغيرة نسبياً من حيث الحجم، مما يجعل النتائج عرضة للتغير إذا ما تم تطبيق نفس التجارب على قواعد بيانات أكثر تعقيداً وحجماً أكبر.

مع ذلك، يوفر هذا البحث إطارًا واضحًا وموجهًا للمطورين لاختيار نموذج الربط العلائقي Object Relational Mapping (ORM) الأمثل بناءً على متطلبات وأدوات لغات البرمجة الأشهر حتى تاريخ إعداد الدراسة. إن هذه الدراسة تمكن من تقديم مقارنة موضوعية وشاملة بين النماذج المختلفة، مما يسهل على أصحاب المشاريع اتخاذ قرارات مدروسة تعزز جودة الأداء وتناسب بيئات التطوير المختلفة.

كما تفتح هذه الورقة الباب أمام توسيع الدراسات المستقبلية لتشمل بيانات بيئات أضخم وأكثر تعقيدًا، مع تضمين تأثيرات التخزين المؤقت، التوزيع الشبكي، وتعليقات الاستعلامات، إضافة إلى اختبار النماذج في ظروف استخدام واقعية ومتنوعة. مما يساهم في تعزيز فهمنا لأداء نماذج ORM في سيناريوهات عملية متعددة ويؤهل المطورين لاتخاذ قرارات أكثر دقة وفعالية في تطوير برمجياتهم.

المراجع:

- [1] Alexandre Torres, Renata Galantea, Marcelo S. Pimentaa, Alexandre Jonatan B. Martins. (2017) – **Twenty years of object–relational mapping: A survey on patterns, solutions, and their implications on application design**. Information and Software Technology 82 (2017) 1–18.
- [2] Doina Zmaranda, Lucian–Laurentiu Pop–Fele, Cornelia Györfödi, Robert Györfödi, George Pecherle. (2020) – **Performance Comparison of CRUD Methods using NET Object Relational Mappers : A Case Study**. Department of Computer Science and Information Technology, University of Oradea, Oradea, Romania¹.
- [3] Edy Budiman, Muh Jamil, Ummul Hairah, Hario Jati, Rosmasari. (2017) – **Eloquent Object Relational Mapping Models for Biodiversity Information System**. International Conference on Computer Applications and Information Processing Technology (CAIPT).
- [4] Gregory Vial. (2019) – **Lessons in Persisting Object Data using Object–Relational Mapping**. IEEE Software (Volume: 36, Issue: 6, Nov.– Dec. 2019).
- [5] Bernard Che Longho. (2022) – **Room vs. greenDAO for Android : A comparative analysis of the performance of Room and greenDAO**. Mittuniversitetet MID Sweden University.

[6] Magdalena Borys, Beata Panczyk. (2015) – **Improving data processing performance for web applications using object–relational mapping**. Actual problems of economics #2(164).

[7] Bruno Baldez Correa, Yue Wang. (2017) – **Object–relational mapping tools and Hibernate**. University libre de Bruxelles.

[8] Shoaib Mahmood Bhatti, Zahid Hussain Abro, Farzana Rauf Abor. (2013) – **Performance Evaluation of Java Based Object Relational Mapping Tools**. Mehran University Research Journal of Engineering & Technology, Volume 32, No. 2, April.