

## موازنة أداء شبكات انترنت الاشياء المعتمدة على بروتوكول LEACH

### باستخدام التعلم المعزز

المهندسة رنيم صالح محمد<sup>1</sup>

إشراف: الدكتور علي الحاتم<sup>2</sup>

#### المخلص

تمثل محدودية الطاقة وموثوقية التسليم تحديين رئيسيين في الحساسات اللاسلكية ضمن شبكات انترنت الاشياء، مما يجعل من التوجيه وإدارة استهلاك الطاقة عاملين حاسمين لضمان استمرار وكفاءة عمل الشبكة. ويعتبر بروتوكول LEACH من أبرز بروتوكولات التجميع الهرمي بفضل قدرته على تجنب استنزاف طاقة العقد من خلال الاختيار الدوري لرؤوس التجمعات (CHs) (Cluster Heads)، إلا أن أدائه ما يزال يتأثر بعشوائية اختيار هذه الرؤوس، وارتفاع كلفة إعادة التجميع عند تنفيذها دورياً. لذا ظهرت العديد من التطويرات على البروتوكول ومن أبرزها بروتوكول LEACH-RLC، الذي جمع بين تحسين تشكيل التجمعات بما فيه اختيار رؤوس التجمعات والعقد المرتبطة بها باستخدام البرمجة الخطية الصحيحة المختلطة (MILP) وبين اتخاذ قرار إعادة التجميع بشكل تكيفي باستخدام التعلم المعزز. لكن يفترض هذا البروتوكول وجود بيئة اتصال مثالية ومعدل تسليم كامل، لذا يعمل هذا البحث على تحسين بروتوكول LEACH-RLC بهدف الحفاظ على أدائه والموازنة بين إدارة الطاقة وموثوقية التسليم في ظل ظروف شبكة غير مثالية، وذلك من خلال تحسين دالة المكافأة وتوسيع متجه حالة الشبكة لدى وكيل التعلم المعزز (DQN) بالإضافة ثمانية ميزات جديدة تعبر عن جودة الاتصال. وأظهرت نتائج الدراسة أن ظروف الاتصال غير المستقرة أدت إلى تدهور واضح في مؤشرات عمر الشبكة ومعدل التسليم عند تشغيل البروتوكول الأصلي، في حين أظهرت النسخة المحسنة المقترحة زيادة في عمر الشبكة ومعدل التسليم وانخفاض في عدد الحزم المفقودة، بالمقابل كان هنالك كلفة تشغيلية تمثلت بزيادة طفيفة في استهلاك الطاقة ناجمة عن الزيادة في عدد حزم التحكم المرتبطة بقرارات إعادة التجميع عند تراجع جودة الاتصال.

<sup>1</sup> طالبة ماجستير - قسم هندسة الشبكات والنظم الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص - سورية

<sup>2</sup> مدرس - قسم هندسة الشبكات والنظم الحاسوبية - كلية الهندسة المعلوماتية - جامعة حمص - سورية

**الكلمات المفتاحية:** بروتوكول LEACH، بروتوكول LEACH-RLC، انترنت الاشياء،

شبكات الحساسات اللاسلكية، الطاقة، موثوقية التسليم، موازنة الأداء، التعلم

## **المعزز . Performance Balancing of LEACH-Based IOT Networks Using Reinforcement Learning**

Eng. Raneem Mohammad<sup>1</sup>

Supervision: Dr. Ali Alhatem<sup>2</sup>

### **Abstract**

Energy constraints and delivery reliability represent two major challenges in wireless sensor networks within Internet of Things (IoT) environments, making routing and energy consumption management critical factors for ensuring network continuity and operational efficiency. The LEACH protocol is one of the most prominent hierarchical clustering protocols due to its ability to reduce node energy depletion through the periodic selection of Cluster Heads (CHs). However, its performance is still affected by the randomness of CH selection and the high cost of reclustering when it is performed periodically. Accordingly, several enhancements to LEACH have been proposed, among which LEACH-RLC is a notable protocol. LEACH-RLC combines optimized cluster formation, including the selection of Cluster Heads and their associated nodes using Mixed-Integer Linear Programming (MILP), with adaptive reclustering decisions based on reinforcement learning. Nevertheless, this protocol assumes an ideal communication environment and a complete packet delivery rate. Therefore, this research aims to improve LEACH-RLC in order to preserve its performance and balance energy management with delivery reliability under non-ideal network conditions. This is achieved by enhancing the reward function and expanding the network state vector of the reinforcement learning agent, based on Deep Q-Networks (DQN), through the addition of eight new features that reflect link quality. The results show that unstable communication conditions caused a clear degradation in network lifetime indicators and packet delivery rate when the original protocol was used. In contrast, the proposed enhanced version

<sup>1</sup> Master's Student – Department of Network and Computer Systems Engineering, Faculty of Informatics Engineering, Homs University, Syria

<sup>2</sup> Lecturer – Department of Network and Computer Systems Engineering, Faculty of Informatics Engineering, Homs University, Syria

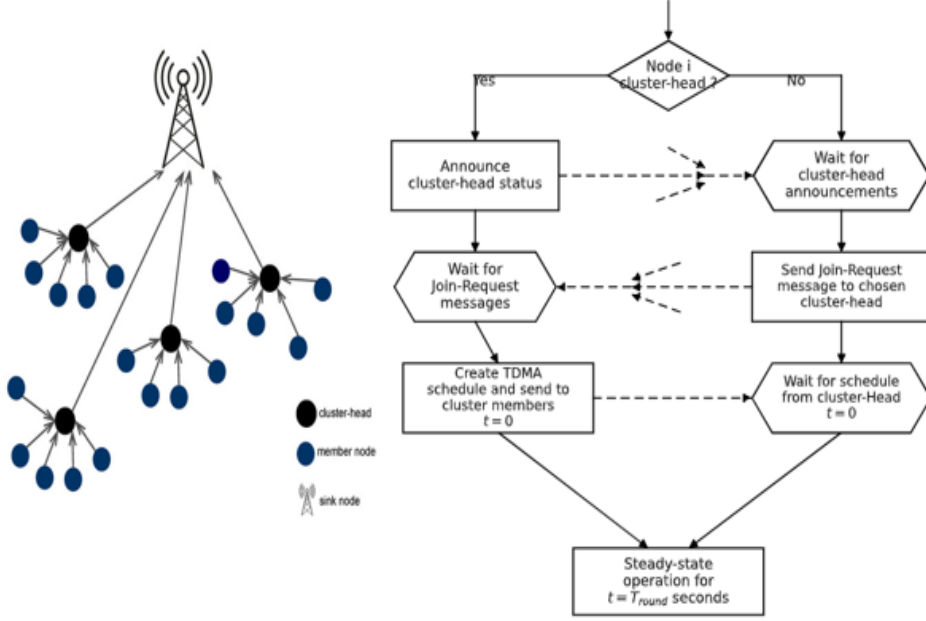
achieved improvements in network lifetime and delivery rate, along with a reduction in the number of lost packets. However, this improvement incurred an operational cost, represented by a slight increase in energy consumption due to the higher number of control packets associated with reclustering decisions when link quality deteriorates.

**Keywords:** LEACH protocol, LEACH-RLC protocol, Internet of Things, Wireless Sensor Networks, energy, delivery reliability, performance balancing, reinforcement learning.

## 1. مقدمة

تؤدي شبكات الحساسات اللاسلكية دوراً محورياً في بنية إنترنت الأشياء، إذ تُستخدم في تطبيقات متنوعة مثل المراقبة البيئية، والأتمتة الصناعية، والمدن الذكية، والرعاية الصحية. وتشكل إدارة الطاقة تحدي جوهري في هذه الشبكات نظراً لاعتماد الحساسات على بطاريات محدودة السعة يصعب استبدالها أو إعادة شحنها في كثير من البيئات العملية [1,2,3]. وفي هذا السياق برزت البروتوكولات الهرمية المعتمدة على التجميع بوصفها من أكثر البروتوكولات فاعلية في إدارة استهلاك الطاقة، لأنها تخفف من كلفة الاتصال المباشر مع المحطة الأساسية من خلال تجميع البيانات محلياً داخل كل تجمع في عقدة تمثل رأس التجمع وهي من تقوم بإرسال البيانات المجمعة إلى المحطة القاعدية.

وجسد بروتوكول LEACH هذا المفهوم بشكل عملي، حيث يعتمد على تقسيم الشبكة إلى تجمعات وتدوير اختيار رأس التجمع بين العقد بهدف منع استنزاف طاقتها، يبين الشكل (1) آلية عمل البروتوكول [1,2].



الشكل (1): آلية عمل بروتوكول LEACH

يعاني LEACH في بنيته التقليدية من عدة محدوديات، أهمها أن اختيار رؤوس التجمعات يتم بطريقة احتمالية لا تراعي بشكل كافٍ الطاقة المتبقية، أو موضع العقد، أو جودة الروابط، مما يؤدي إلى إعادة تشكيل التجمعات بشكل متكرر والذي بدوره يزيد من كلفة حزم التحكم ويؤثر سلباً في الكفاءة العامة للشبكة [4,5,6]. لذا ظهرت العديد من التحسينات على هذا البروتوكول، برز من بينها بروتوكول LEACH-RLC الذي قدّم تطويراً مهماً من خلال الجمع بين تحسين تشكيل التجمعات باستخدام البرمجة الخطية الصحيحة المختلطة (MILP) وبين متحكم يعتمد على التعلم المعزز لاتخاذ قرار إعادة التجميع بشكل تكيفي [7].

لكن ما يزال اتخاذ قرار إعادة التجميع في LEACH-RLC قريباً من افتراضات تشغيل مثالية، في حين تتسم الكثير من بيانات الانترنت الاشياء الواقعية بتذبذب جودة الاتصال، وخسارة الحزم، وتفاوت نجاح الإرسال، وهو ما ينعكس مباشرة على معدل التسليم، وإعادة الإرسال، والطاقة المستهلكة فعلياً [3,8,9,10,11]. ومن هنا تتطوّر مساهمة هذا البحث، بهدف تطوير LEACH-RLC ليصبح أكثر إدراكاً لحالة الشبكة الواقعية عند اتخاذ قرار إعادة التجميع، وأتاح استخدام

التعلم المعزز تحقيق هذه المساهمة عبر مراعاة خصائص الاتصال غير المستقر داخل دالة المكافأة المحسنة وتمثيل المزيد من الميزات في متجه الحالة في DQN، بما يحقق موازنة أفضل لأداء الشبكة في ظل ظروف الشبكة غير المثالية.

## 2. أهداف وأهمية البحث

يهدف هذا البحث إلى تحسين بروتوكول LEACH-RLC ليصبح أكثر قدرة على التعامل مع ظروف الاتصال غير المثالية في شبكات انترنت الأشياء، من خلال جعل قرار إعادة التجميع أكثر ارتباطاً بالحالة الفعلية للشبكة. ويمكن تحديد هدف البحث في النقاط الآتية:

(1) تحسين آلية اتخاذ قرار إعادة التجميع في بروتوكول LEACH-RLC بحيث لا تعتمد فقط على مؤشرات الطاقة أو البنية الحالية للتجمعات، بل تراعي أيضاً جودة الاتصال وحالة الشبكة في البيئات غير المستقرة.

(2) تطوير دالة المكافأة المستخدمة في وكيل التعلم المعزز لتصبح أكثر شمولاً، من خلال إدخال عوامل مرتبطة بمعدل التسليم، والطاقة المستهلكة، وإعادة الإرسال، وحجم حزم التحكم، والتباين في استهلاك الطاقة، وقرار إعادة التجميع.

(3) توسيع متجه الحالة داخل نموذج DQN بإضافة ثماني ميزات جديدة تعبر عن خصائص الاتصال غير المثالي، مثل تذبذب جودة الرابط، واحتمال نجاح الإرسال، ومعدل التسليم، وحجم الحزم المعاد إرسالها.

(4) تمكين وكيل التعلم المعزز من اتخاذ قرار أكثر دقة وتكيفاً عند المفاضلة بين الإبقاء على البنية التجمعية الحالية أو إعادة تشكيلها، وفقاً لحالة الشبكة الفعلية وليس وفق مؤشرات تقليدية محدودة.

(5) تحقيق موازنة أفضل في أداء الشبكة من خلال الجمع بين كفاءة الطاقة، وموثوقية التسليم، وتقليل فقد الحزم، وضبط حمل التحكم في ظل ظروف الاتصال غير المستقرة.

(6) تقييم النسخة المحسنة LEACH-RLC-NI (NI كاختصار لغير مثالي (Non Ideal)) ومقارنتها بالبروتوكول المرجعي LEACH-RLC ضمن ظروف الشبكة نفسها، لقياس أثر دمج خصائص الاتصال الواقعية في قرار إعادة التجميع.

## 3. دراسة مرجعية

شمل تطوير بروتوكول LEACH تحسين بنية التجمعات والتوجيه بينها، والتحكم في توقيت إعادة تشكيل التجمعات بالإضافة إلى تحسين آلية اختيار رؤوس التجمعات [4,3,5,6,12]. وارتبط هذا التطور بالانتقال التدريجي من الأساليب المعتمدة على القواعد الثابتة إلى استراتيجيات أكثر تكيفاً تعتمد على تقنيات الذكاء الصناعي.

قدّم بروتوكول E-FLEACH [13] في هذا السياق استراتيجية تعتمد على المنطق الضبابي لتحسين اختيار رؤوس التجمعات، اعتماداً على مجموعة من العوامل، بينما عملت دراسة NN\_ILEACH [14] على استبدال آلية الاختيار الاحتمالي التقليدية داخل بروتوكول LEACH بآلية تعليمية تعتمد على شبكة عصبية تستفيد من خصائص حالة الشبكة عند اختيار رؤوس التجمعات. وكذلك قدّم بروتوكول LEACH-OCW-IACO [15] في مقترح يجمع بين التحسين الاستدلالي والتوجيه الذكي تطويراً مزدوجاً استهدف كلاً من اختيار رؤوس التجمعات والتوجيه بين التجمعات، من خلال الاعتماد على الأوزان لاختيار الرؤوس وخوارزمية مستعمرة النمل المحسنة لتحديد المسارات. في حين حقق بروتوكول EDK-LEACH [16] خوارزمية K-means لتشكيل تجمعات أكثر انتظاماً واختيار رؤوس التجمعات. ومن الدراسات الحديثة التي قدمت تحسين فريد على بروتوكول LEACH الدراسة [7] والتي عمل الباحثون خلالها على تطوير بروتوكول LEACH-RLC وعالجوا من خلاله مسألة اختيار رؤوس التجمعات إضافة إلى اتخاذ قرار توقيت إعادة التجميع، وجمعت هذه المنهجية بين البرمجة الخطية الصحيحة المختلطة (MILP) من أجل اختيار رؤوس التجمعات وتحديد العقد المرتبطة بها، وبين التعلم المعزز لاتخاذ قرار يقضي إما بالإبقاء على تشكيل التجمع الحالي أو إعادة تشكيله. يلخص الجدول (1) منهجية الدراسات السابقة مع التركيز على التقنيات المستخدمة والعوامل المدروسة ضمن الاستراتيجيات المقترحة مع النتائج.

الجدول (1): ملخص الدراسات المرجعية

| الدراسة       | المنهجية المقترحة                   | التقنية المستخدمة             | العوامل                             | النتائج                                         |
|---------------|-------------------------------------|-------------------------------|-------------------------------------|-------------------------------------------------|
| E-FLEACH [13] | تحسين اختيار رؤوس التجمعات في LEACH | المنطق الضبابي<br>Fuzzy Logic | الطاقة المتبقية، المسافة إلى المحطة | تحسن FND بنسبة 60.38%، و HND بنسبة 32.5%، و LND |

| الدراسة             | المنهجية المقترحة                                                                     | التقنية المستخدمة                                         | العوامل                                                                               | النتائج                                                                                                                                          |
|---------------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
|                     | بالاعتماد على تقييم متعدد العوامل                                                     |                                                           | الأساسية، مركزية العقدة                                                               | بنسبة 17.13% مقارنةً بـ LEACH مع انخفاض في استهلاك الطاقة في الجولات الأولى                                                                      |
| NN_ILEACH [14]      | استبدال الاختيار الاحتمالي التقليدي لرؤوس التجمعات بألية تعلمية تعتمد على حالة الشبكة | شبكة عصبية Neural Network                                 | الطاقة، توزيع العقد، خصائص حالة الشبكة                                                | ارتفاع عمر الشبكة من 505 جولات في LEACH إلى 11361 جولة، وانخفاض استهلاك الطاقة بنحو 40%، وزيادة throughput بنحو 30%، وتحسن معدل التسليم بنحو 25% |
| LEACH-OCW-IACO [15] | تحسين اختيار رؤوس التجمعات والتوجيه بين التجمعات                                      | أوزان اختيار الرؤوس + خوارزمية مستعمرة النمل المحسنة IACO | الطاقة المتبقية، كثافة الجوار، المسافة إلى المحطة الأساسية، عدد القفزات، توازن الطاقة | ارتفاع FND من 141 في LEACH إلى 199 في ILEACH ثم إلى 329 في IACO، وارتفاع LND من 2338 إلى 3487 جولة، مع انخفاض استهلاك الطاقة                     |
| EDK-LEACH [16]      | تشكيل تجمعات أكثر انتظاماً وتحسين                                                     | K-means clustering                                        | إحداثيات العقد، الطاقة المتبقية، القرب من مركز العقدة، المسافة إلى                    | تأخر FND من نحو 1983 إلى 4285 جولة، وارتفاع LND إلى أكثر من 8000 جولة،                                                                           |

موازنة أداء شبكات انترنت الاشياء المعتمدة على بروتوكول LEACH باستخدام التعلم المعزز

| الدراسة       | المنهجية المقترحة                                     | التقنية المستخدمة             | العوامل                                                                                            | النتائج                                                                                                                                                       |
|---------------|-------------------------------------------------------|-------------------------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
|               | اختيار رؤوس التجمعات                                  |                               | المحطة الأساسية                                                                                    | وانخفاض متوسط الطاقة المستهلكة من J 0.052 إلى J0.036                                                                                                          |
| LEACH-RLC [7] | تحسين اختيار رؤوس التجمعات وتحديد توقيت إعادة التجميع | MILP + Reinforcement Learning | الطاقة داخل العنقود، الطاقة بين رأس التجمع والمحطة الأساسية، عدد رؤوس التجمعات، كلفة إعادة التشكيل | تفوق على LEACH و LEACH-C و LEACH-CM في عمر الشبكة وتوازن الطاقة، وبلغ HND نحو 1050 جولة مقابل نحو 950 في LEACH-C و LEACH-CM، مع معدل تسليم أعلى وحمل تحكم أقل |

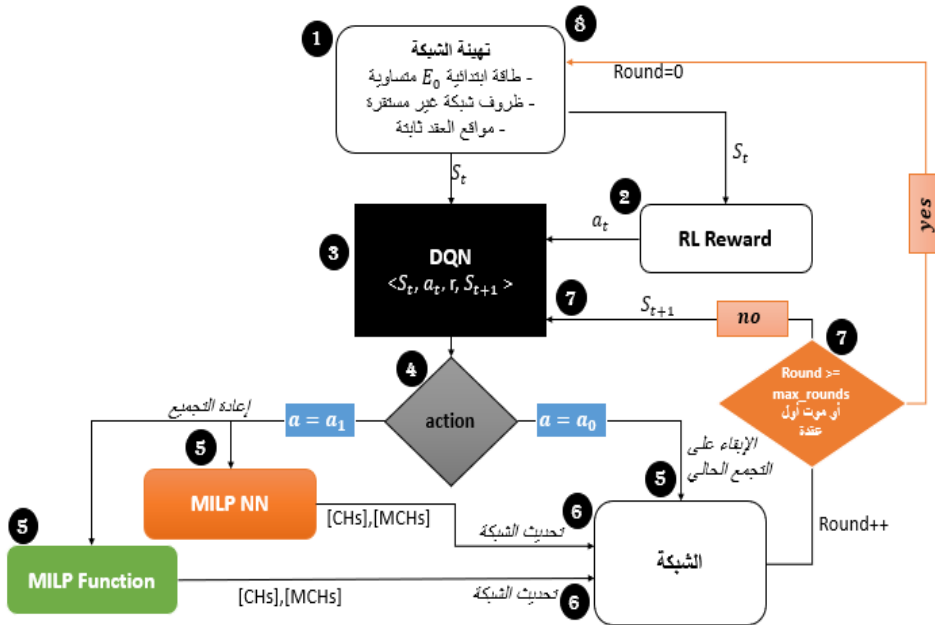
تبيّن من الدراسات المرجعية أن معظم التحسينات على LEACH ركزت على الطاقة المتبقية، والمسافة إلى المحطة الأساسية، وتوازن التجمعات، في حين ما زالت العوامل المرتبطة بجودة الاتصال الواقعية، وموثوقية التسليم، وفقد الحزم، وإعادة الإرسال أقل مراعاة. من هنا تنطلق مساهمة هذا البحث الذي يهدف إلى التطوير على LEACH-RLC ليصبح أكثر إدراكاً للحالة الواقعية للشبكة، من خلال دمج خصائص الاتصال غير المثالي داخل كل من دالة المكافأة ومتجه الحالة، بما يحقق موازنة أفضل بين استهلاك الطاقة وموثوقية التسليم في ظل ظروف الشبكة غير المثالية.

#### 4. مواد وطرائق البحث

##### 1.4. البنية العامة للمنهجية المقترحة

تحافظ المنهجية المقترحة على البنية العامة لبروتوكول LEACH-RLC، حيث تتم على مرحلتين متكاملتين [7]:

- (1) تشكيل التجمعات واختيار رؤوسها باستخدام دالة MILP.
  - (2) اتخاذ قرار إعادة التجميع أو الإبقاء على التجمع الحالي باستخدام التعلم المعزز .  
بينما يكمن الاختلاف بين التطوير المقترح والمنهجية الأصلية في أن وكيل التعلم المعزز لا يعتمد فقط على حالة الطاقة وبنية الشبكة كما في بروتوكول LEACH-RLC، بل يتعلم أيضاً من عوامل مرتبطة بجودة الاتصال، مثل معدل التسليم، وإعادة الإرسال، وحجم حزم التحكم، واحتمال نجاح رابط الاتصال.
- يوضح الشكل (2) البنية العامة للمنهجية المقترحة، حيث نستخدم دالة MILP لتشكيل التجمعات من خلال اختيار رؤوس التجمعات والعقد المرتبطة بكل منها، بينما يتولى DQN اتخاذ القرار بالإبقاء على التجمع الحالي أو إعادة تشكيله وفق حالة الشبكة في كل جولة.



الشكل (2): البنية العامة للمنهجية المقترحة

تبدأ المنهجية بتهيئة شبكة الحساسات وتحديد مواقع العقد ومستوى الطاقة الابتدائي لكل عقدة. بعد ذلك يتم نمذجة حالة الاتصال بين العقد وفق سيناريوهين: روابط مثالية وروابط غير مثالية.

في حالة الروابط غير المثالية، لا يُفترض أن الاتصال بين العقد مضمون، بل يتم تمثيله باحتمالية نجاح تعتمد على جودة الرابط والمسافة وعدد محاولات إعادة الإرسال المسموح بها. بعد بناء نموذج الشبكة، يتم اتخاذ قرار إعادة التجميع من قبل وكيل التعلم المعزز DQN المدرب بناءً على قرار دالة المكافأة وحالة الشبكة، في حال قرار إعادة تجميع يتم اختيار رؤوس العناقيد باستخدام دالة هدف تجمع بين عاملين رئيسيين: الحفاظ على الطاقة وتحسين موثوقية الاتصال. وتهدف هذه الدالة إلى تجنب اختيار عقد منخفضة الطاقة كرؤوس عناقيد، وتقليل تكلفة الإرسال، وتحسين معدل تسليم الحزم. في هذا السياق كما هو مبين في الشكل ، يوجد شبكة عصبونية كمقاربة لدالة MILP تستخدم فقط كبديل أثناء مرحلة تدريب نموذج التعلم المعزز (DQN) بهدف تسريع عملية التدريب وتقليل الكلفة الحسابية الناتجة عن تكرار إعادة التجميع عبر عدد كبير من الجولات (حتى أقصى عدد من الجولات أو حتى موت أول عقدة)، في حين تستخدم دالة MILP الأصلية أثناء التشغيل الفعلي لضمان دقة اختيار رؤوس التجمعات والعقد التابعة لها.

## 2.4. فضاء الأحداث (actions)

يعتمد قرار DQN في هذا العمل على فضاء أحداث مُعرّف كما يأتي:

$$a_t \in \{a_0, a_1\}$$

وتمثل الأحداث كل مما يلي:

- $a_0$ : الإبقاء على التجمع الحالي.
- $a_1$ : إعادة تشكيل التجمعات باستخدام MILP.

حيث يحدد القرار في كل جولة ما إذا كانت البنية الحالية للتجمعات ما تزال ملائمة، أو أنه من الأفضل إعادة تشكيل التجمعات من جديد.

## 3.4. متجه الحالة الأصلي والموسع

يصف متجه الحالة في بروتوكول LEACH-RLC بنية الشبكة من حيث الطاقة والتجمعات والقرار السابق، ويمكن تعريفه كما يأتي [7]:

$$S_t^{base} = [E_t^{total}, E_t, CH_t, N_t^{CH-rep}, M_t, \alpha_{t-1}]$$

تمثل المعلومات الأساسية في هذا المتجه: الطاقة المتبقية الكلية ( $E_t^{total}$ )، والطاقة المتبقية للعقد ( $E_t$ )، ورؤوس التجمعات الحالية ( $CH_t$ )، وعدد مرات تكرار مجموعة الرؤوس والأعضاء داخل التجمعات ( $M_t$ )، والقرار المتخذ في الجولة السابقة ( $\alpha_{t-1}$ ).  
وعملنا في هذا البحث على توسيع متجه الحالة ليصبح أكثر قدرة على تمثيل ظروف الاتصال غير المثالية، وذلك بإضافة ثماني ميزات جديدة مرتبطة بجودة الاتصال، بحيث يصبح:

$$S_t^{NI} = [S_t^{base}, PDR_t, \bar{E}_t^{cons}, OH_t, R_t^{retx}, P_t^{succ}, N_{max}^{retx}, L_t^{dist}, V_t^{ch}]$$

حيث تمثل الحدود المضافة ما يلي:

- $PDR_t$ : معدل التسليم من طرف إلى طرف في الجولة t.
- $\bar{E}_t^{cons}$ : متوسط الطاقة المستهلكة في الجولة.
- $OH_t$ : حجم حزم التحكم في الجولة.
- $R_t^{retx}$ : عدد الحزم المعاد إرسالها.
- $P_t^{succ}$ : احتمال نجاح الاتصال.
- $N_{max}^{retx}$ : العدد الأعظمي لمحاولات إعادة الإرسال.
- $L_t^{dist}$ : تدهور الرابط مع المسافة.
- $V_t^{ch}$ : التذبذب الزمني لقناة الاتصال.

تمنح هذه الميزات المضافة وكيل التعلم المعزز إدراكاً لا يقتصر على بنية الشبكة واستهلاك الطاقة ضمنها فقط، بل يمتد أيضاً إلى خصائص الاتصال الفعلية، وهو ما يجعل قرار إعادة التجميع أكثر ارتباطاً بالسلوك التشغيلي الحقيقي للرابط والشبكة في البيئات غير المستقرة.

#### 4.4. دالة المكافأة المحسنة

تتمثل مهمة دالة المكافأة في توجيه قرار وكيل التعلم المعزز نحو السياسة المطلوبة. بالنظر إلى الدالة الأصلية في LEACH-RLC المبينة في المعادلة (1):

$$R(S_t, a_t) = \begin{cases} 1 & \text{if } a_t = a_2 \text{ and } E_{net} > 0 \text{ for all } n \in N \\ 1.1 & \text{if } a_t = a_1 \text{ and } E_{net} > 0 \text{ for all } n \in N \\ 2 & \text{if } E_{net} \leq 0 \text{ for some } n \in N \end{cases} \quad \text{(المعادلة 1)}$$

حيث أن  $a_1$  يمثل حدث إعادة التجميع و  $a_2$  حدث الإبقاء على التجميع الحالي و  $E_{net}$  طاقة العقد في الشبكة)، وجدنا أنها ركزت على إبقاء العقد جميعها حية قبل موت أول عقدة (FND) لأطول فترة ممكنة مع مراعاة آخر جولة تم فيها إعادة التجميع [7]، بينما يقترح هذا البحث دالة مكافأة أكثر شمولاً تأخذ في الاعتبار إضافة إلى ما تم مراعاته في البروتوكول الأصلي، العوامل المرتبطة بموثوقية التسليم وكلفة الاتصال في البيئات غير المثالية، ويمكن التعبير عنها كما في المعادلة (2):

$$\begin{aligned} R(S_t, a_t) = & w_{alive} \cdot A_t + w_{pdr} \cdot PDR_t - w_E \cdot \bar{E}_t^{cons} \\ & - w_{retx} \cdot R_t^{retx} - w_{oh} \cdot OH_t - w_{\sigma} \cdot \sigma_{E_t} \\ & - w_{recluster} \cdot I(a_t == a_1) \end{aligned} \quad \text{المعادلة (2)}$$

حيث أن:

- $A_t$ : مؤشر بقاء كافة العقد حية في الجولة  $t$ .
- $\sigma_{E_t}$ : التباين في مستويات طاقة العقد.
- $I(a_t == a_1)$ : مؤشر يساوي 1 عند تنفيذ إعادة التجميع و 0 عدا ذلك .
- $w_{alive}, w_{pdr}, w_E, w_{retx}, w_{\sigma}, w_{recluster}$ : أوزان ضبط تمثل عقوبة أو مكافأة وتكون قيمتها حسب أهمية كل عامل، وتم ضبط كل من الأوزان السابقة بالتجريب، حيث تم اختبار عدة قيم لكل وزن ضمن مجال محدود وتم التجريب ضمن مجالات أوزان أعلى [0.7-1.5] للمكونات الأساسية مثل بقاء العقد وموثوقية التسليم، ومجالات أوزان أصغر [0.1-0.5] للعقوبات الخاصة بالمؤشرات مثل حمل التحكم وتكرار إعادة التجميع، حتى لا يصبح الوكيل متردداً في إعادة التجميع عند وجود تدهور حقيقي في جودة الاتصال. ثم تم تقييم كل إعداد باستخدام مؤشرات الأداء الأساسية: FND و PDR ومتوسط الطاقة المستهلكة وكلفة التحكم، وبالاعتماد على نتائج التجريب تم اختيار الإعداد الذي يحقق أفضل توازن بين هذه المؤشرات، مع إعطاء أولوية خاصة لاستقرار الشبكة وموثوقية تسليم الحزم في حالة الروابط غير المثالية.

تزداد المكافأة عندما يتحسن استقرار الشبكة ومعدل التسليم، ويقابلها عقوبة في حال ارتفع معدل أي من استهلاك الطاقة أو حزم التحكم أو إعادة الإرسال أو تزايد تباين الطاقة بين العقد. ويصبح قرار إعادة التجميع يوازن بين الكلفة والمنفعة، بدلاً من ارتباطه فقط ببقاء العقد أو بتوقيت زمني ثابت بين عمليتي تجميع متتاليتين. وتعد هذه الموازنة متوافقة مع ما أشارت إليه الأدبيات الحديثة في هندسة دالة المكافأة (reward engineering)، بأنه يجب على تصميم المكافأة أن يعكس الأهداف الحقيقية متعددة الجوانب للمهمة حتى لا ينحاز الوكيل إلى تحسين جزئي مضر [17].

## 5. النتائج ومناقشتها

حققت المنهجية المقترحة على مرحلتين، في المرحلة الأولى اختبر بروتوكول LEACH-RLC تحت ظروف روابط غير مثالية بهدف إظهار أثر عدم استقرار الاتصال وإثبات الحاجة إلى التحسين عملياً. أما المرحلة الثانية فتمثلت في إدخال التعديلات المقترحة على دالة المكافأة ومتجه الحالة، ثم إعادة تقييم البروتوكول في بنيته الجديدة LEACH-RLC-NI تحت الظروف نفسها، مما يظهر أثر التحسين نفسه على البروتوكول الأصلي عند الانتقال من فرضية مثالية إلى بيئة أكثر واقعية.

### 1.5. بيئة التجريب ومقاييس التقييم

تم تنفيذ التجارب باستخدام محاكي PyNetSim المطور بلغة Python [18]، مع توزيع 100 عقدة داخل حقل استشعار بمساحة  $100 \times 100$  متر، ووضع محطة الأساس عند الإحداثيات (50,175)، بما يطابق إعدادات الدراسة المرجعية الأصلية [7].

أُجريت التجارب عبر خمس تشغليات مستقلة بهدف تقليل أثر العشوائية الناتجة عن توزيع العقد وحالة القناة. وتم تشغيل كل تجربة حتى حد أقصى قدره 1500 جولة، أو حتى تحقق مؤشرات نهاية عمر الشبكة بحسب آلية المحاكاة. وبعد ذلك تم استخراج مؤشرات الأداء لكل seed على حدة، ثم حساب المتوسط، والانحراف المعياري، وفاصل الثقة، واختبار الدلالة الإحصائية للمقارنة بين البروتوكول المرجعي LEACH-RLC [7] والنسخة المقترحة.

تم توليد ظروف القناة غير المستقرة من خلال تغيير معاملات التذبذب الزمني ضمن نموذج قناة مترابط correlated مستلهم من نموذج Gilbert–Elliott [19-21]، حيث يُمثل هذا النموذج حالة جيدة وحالة رديئة للقناة مع احتمالات انتقال بينهما، لتمثيل الفقد المتعاقب أو التذبذب الزمني في جودة الرابط اللاسلكي عبر الجولات. تحدد المعاملات (مثل احتمال أن يبدأ الرابط في الحالة الجيدة ( $p_{init\_good}$ )، واحتمال الانتقال من الحالة الجيدة إلى الحالة الرديئة ( $p_{gb}$ )، واحتمال الانتقال من الحالة الرديئة إلى الحالة الجيدة ( $p_{bg}$ )، ومعامل تحسين جودة الرابط في الحالة الجيدة ( $good\_mult$ )، ومعامل تدهور جودة الرابط في الحالة الرديئة ( $bad\_mult$ )) سلوك الرابط واحتمالات انتقاله بين الحالتين الجيدة والرديئة، وبذلك تعكس التشغيليات الخمسة حالة اتصال أكثر اضطراباً، مع احتمالية أعلى لتدهور الرابط وفقد الحزم، مع إبقاء إعدادات الشبكة الأخرى (حجم الشبكة، عدد العقد، موضع المحطة القاعدية) ثابتة لضمان عدالة المقارنة.

واعتمد التقييم على مجموعة من المقاييس الرئيسية، وهي:

• مقاييس عمر الشبكة:

- FND (First Node Dead): رقم الجولة عند موت أول عقدة .
- LND (Last Node Dead): رقم الجولة عند موت آخر عقدة.
- avg\_energy\_per\_round: متوسط الطاقة المستهلكة في الجولة.
- avg\_overhead\_per\_round: متوسط حمل التحكم في الجولة.
- avg\_pdr\_e2e: معدل التسليم من طرف إلى طرف.
- avg\_pkts\_lost: متوسط عدد الحزم المفقودة.

يبين الجدول (2) إعدادات المحاكاة المعتمدة في الدراسة:

الجدول (2): إعدادات المحاكاة

| الإعداد          | القيمة    |
|------------------|-----------|
| المحاكي          | PyNetSim  |
| عدد العقد الحسية | 100       |
| مساحة الحقل      | 100 * 100 |

| الإعداد                  | القيمة                       |
|--------------------------|------------------------------|
| موضع محطة الأساس         | (50,175)                     |
| مدى الإرسال              | 70 متر                       |
| عدد الجولات الأقصى       | 1500 جولة                    |
| عدد البذور العشوائية     | 5 seeds                      |
| نماذج المقارنة           | NI-RLC-LEACH و RLC-LEACH     |
| طريقة تشكيل التجمعات     | MILP                         |
| متحكم قرار إعادة التجميع | DQN                          |
| نموذج القناة             | correlated non-ideal channel |
| عدد محاولات الإرسال      | 3                            |

## 2.5. نمذجة القناة غير المثالية

تتأثر بيانات شبكات الحساسات اللاسلكية وإنترنت الأشياء بتغير جودة الإشارة مع المسافة، وباحتمال فقد الحزم، والحاجة إلى إعادة الإرسال عند ضعف الرابط [23][22]. لذلك تم تمثيل كل رابط بين العقدتين (i) و (j) باحتمالية نجاح  $(p_{ij})$  بدلاً من افتراض وجود قناة اتصال مثالية لا تسبب أي فقد.

تم اعتماد نموذج احتمالي مبسط مستند إلى المبادئ المستخدمة في نمذجة الروابط غير المثالية، حيث تتدهور جودة الرابط ومعدل استقبال الحزم مع المسافة وظروف القناة [22-25]. تبين المعادلة (3) احتمالية نجاح الرابط كما يلي:

$$p_{ij} = p_{\min} + (p_{\max} - p_{\min}) e^{-\eta \frac{d_{ij}}{d_{\max}}} \quad (3) \text{ المعادلة}$$

حيث تمثل  $(p_{ij})$  احتمالية نجاح الرابط بين العقدتين (i) و (j)، ويمثل  $(p_{\min})$  و  $(p_{\max})$  الحدين الأدنى والأعلى لاحتمالية نجاح الإرسال، استُخدم الحد  $(p_{\min})$  لضمان وجود حد أدنى لاحتمالية نجاح الإرسال حتى في الروابط الضعيفة بدلاً من افتراض أن الرابط يصبح معدوماً تماماً عند تدهور القناة [24][22]، بينما يحدد الفرق  $(p_{\max} - p_{\min})$  مجال التغير الممكن بين أسوأ

وأفضل حالة للرباط بما يسمح بتمثيل تدريجي لجودة الاتصال بدلاً من نموذج ثنائي يفترض أن الرباط إما مثالي أو منقطع [25][23]. أما العامل الأسّي  $e^{-\eta \frac{d_{ij}}{d_{max}}}$  فيمثل تناقص جودة الرباط مع زيادة المسافة، حيث تؤدي الزيادة في المسافة بين العقدتين ( $d_{ij}$ ) إلى انخفاض قيمة هذا العامل تدريجياً ويستند هذا الافتراض إلى نماذج فقد المسار ونمذجة الروابط منخفضة الطاقة، التي تبين أن معدل استقبال الحزم (PRR) وجودة الرباط يتأثران بالمسافة ونسبة الإشارة إلى الضجيج (SNR) [24][22]، بينما تمثل ( $d_{max}$ ) أكبر مسافة مرجعية داخل الشبكة وتستخدم لتطبيع المسافة وجعل أثرها مناسباً لمجال المحاكاة، بينما يتحكم  $\eta$  بدرجة تدهور الرباط مع زيادة المسافة بشكل مشابه لاستخدام معاملات فقد المسار أو معاملات التدهور في نماذج القناة اللاسلكية [24][22].

ولتمثيل أثر إعادة الإرسال، تم حساب احتمالية نجاح الحزمة بعد عدد محدود من المحاولات كما هو مبين في المعادلة (4):

$$P_{ij}^{succ} = 1 - (1 - p_i)^A \quad \text{المعادلة (4)}$$

حيث يمثل  $A$  عدد محاولات إعادة الإرسال الكلية المسموح بها. وتعتمد هذه المعادلة على نموذج المحاولات المستقلة للإرسال، إذ إن احتمال فشل محاولة واحدة يساوي  $(1 - p_i)$ ، واحتمال فشل جميع المحاولات يساوي  $(1 - p_i)^A$ ، وبالتالي تكون احتمالية نجاح الحزمة في محاولة واحدة على الأقل مساوية إلى متمم احتمال فشل جميع المحاولات [26]. ويستخدم هذا النوع من النمذجة في تحليل موثوقية الروابط اللاسلكية في شبكات محدودة الطاقة، لأن عدد محاولات إعادة الإرسال يؤثر مباشرة في احتمالية التسليم وفي الطاقة المستهلكة [27][26]. كما يمكن ربط ما سبق بمفهوم عدد الإرسالات المتوقع (ETX)، الذي يقدر كلفة تسليم الحزمة عبر رباط غير مثالي، حيث تزداد قيمة (ETX) عندما تنخفض احتمالية نجاح التسليم [28]. وبناءً على احتمالية النجاح بعد عدد محدود من المحاولات، يمكن حساب (ETX) وفق المعادلة (5):

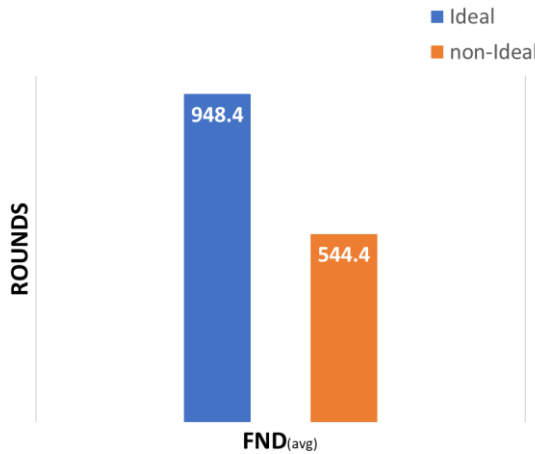
$$ETX_{ij} = \frac{1}{p_{ij}^{succ}} \quad \text{المعادلة (5)}$$

### 3.5. النتائج

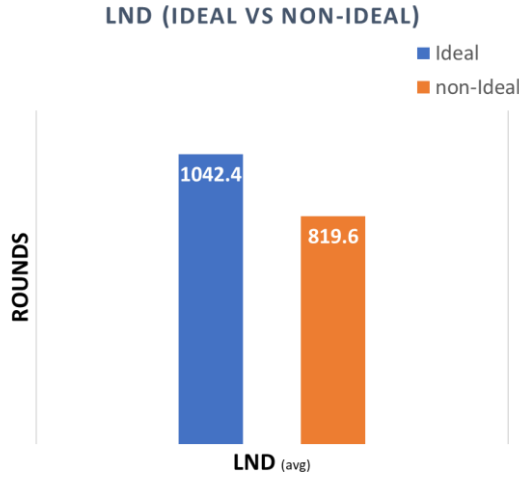
#### 1.3.5 أثر الروابط غير المثالية على أداء بروتوكول LEACH-RLC

أظهرت نتائج اختبار البروتوكول LEACH-RLC تدهور واضح في مؤشرات عمر الشبكة عند الانتقال من ظروف شبكة تشغيلية مثالية إلى ظروف غير مثالية. حيث انخفضت قيمة FND بنسبة (42.6%) من 948.4 في الحالة المثالية إلى 544.4 في الحالة غير المثالية، كما انخفضت قيمة LND من 1042.4 إلى 819.6 بنسبة (21.4%)، كما يوضحه الشكلان (3) و(4). وتشير هذه النتائج إلى تأثير فترة الاستقرار والعمر الكلي للشبكة بوضوح تحت ظروف الاتصال غير المستقرة.

FND (IDEAL VS NON-IDEAL)

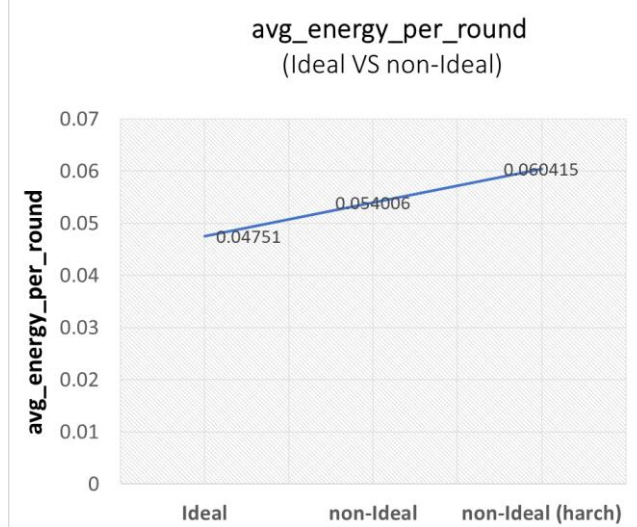


الشكل (3): مقارنة FND بين حالي الروابط المثالية وغير المثالية لبروتوكول LEACH-RLC.



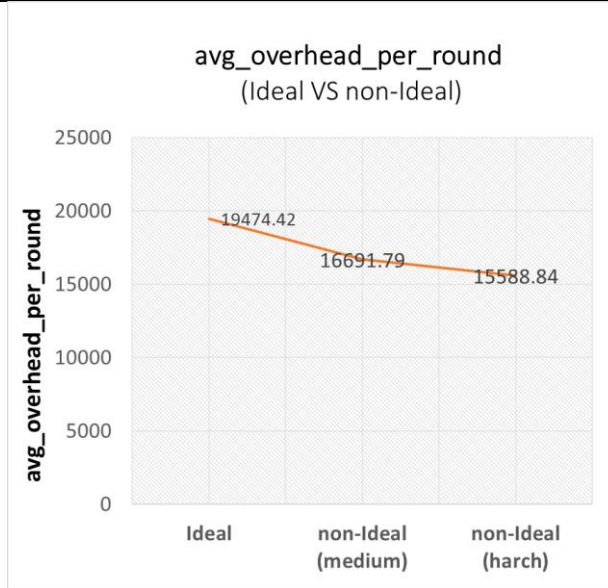
الشكل (4): مقارنة LND بين حالتي الروابط المثالية وغير المثالية لبروتوكول LEACH-RLC.

بينما ارتفع متوسط الطاقة المستهلكة في الجولة من 0.04751 في الحالة المثالية إلى 0.054006، ثم إلى 0.060415 بنسبة (21.4%) تدريجياً مع تدهور الاتصال في الشبكة، كما هو موضح في الشكل (5). ويشير هذا الارتفاع إلى زيادة الكلفة التشغيلية مع ازدياد فشل الإرسال وإعادة الإرسال.



الشكل (5): ارتفاع معدل استهلاك الطاقة في الجولة مع تبدل استقرار الشبكة من مثالية إلى غير مثالية

وفي المقابل انخفض متوسط حمل التحكم من 19474.42 في الحالة المثالية إلى 16691.79، ثم إلى 15588.84 بنسبة (19.95%) تدريجياً عند الانتقال إلى ظروف أسمى غير مثالية، كما هو موضح في الشكل (6). لكن هذا الانخفاض لا يشير بالضرورة إلى كفاءة القرار كونه يوجد ارتفاع في استهلاك الطاقة بسبب إعادة التجميع المتكررة، بل قد يكون السبب في ذلك الموت المبكر للعقد وبالتالي انخفاض عدد حزم التحكم في الشبكة.

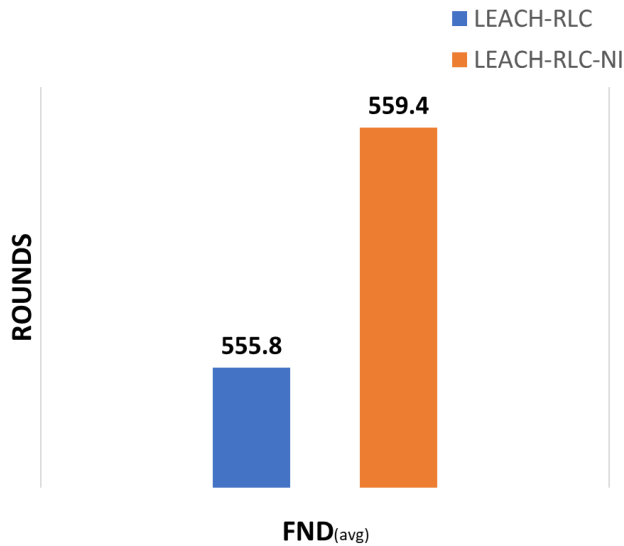


الشكل (6): انخفاض معدل حمل التحكم في الجولة مع تبدل استقرار الشبكة من مثالية إلى غير مثالية

### 2.3.5 نتائج أداء البروتوكول LEACH-RLC-NI تحت ظروف شبكة غير مستقرة

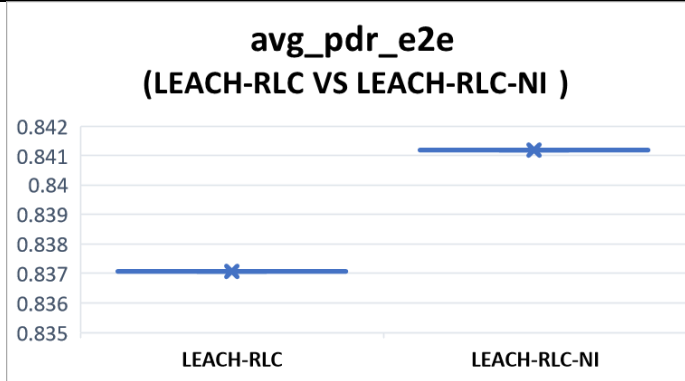
أظهرت النسخة المقترحة LEACH-RLC-NI بعد تطوير دالة المكافأة ومتجه الحالة تحسناً في عدد من مقاييس الأداء مقارنة بالبروتوكول LEACH-RLC تحت الظروف غير المستقرة نفسها. حيث بينت النتائج ازدياد بسيط بمؤشر عمر الشبكة FND بنسبة 0.65% من 555.8 إلى 559.4، كما يوضحه الشكل (7)، وهو ما يشير إلى تحسن محدود في فترة الاستقرار الأولى للشبكة.

### FND (LEACH-RLC VS LEACH-RLC-NI)

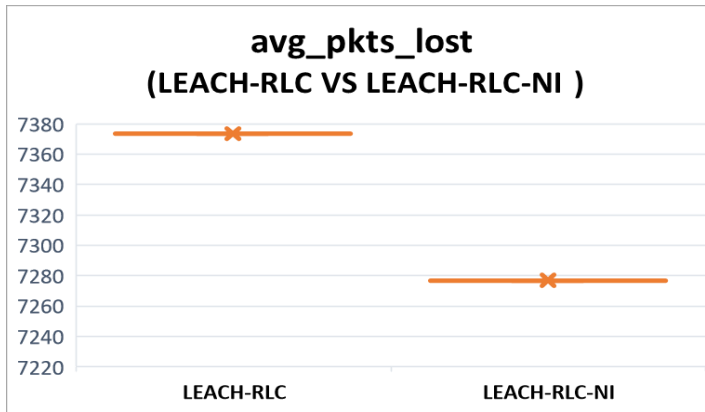


الشكل (7): مقارنة FND بين LEACH-RLC و LEACH-RLC-NI تحت ظروف الشبكة غير المستقرة.

كذلك ارتفع معدل التسليم ( $avg\_pdr\_e2e$ ) بنسبة 0.49% من 0.837069 إلى 0.841195، كما انخفض متوسط عدد الحزم المفقودة بنسبة 1.31% من 7373.4 إلى 7277.0، كما هو موضح في الشكلين (8) و (9). بما يبين أن البروتوكول المقترح حسن موثوقية التسليم تحت ظروف الاتصال غير المستقرة.

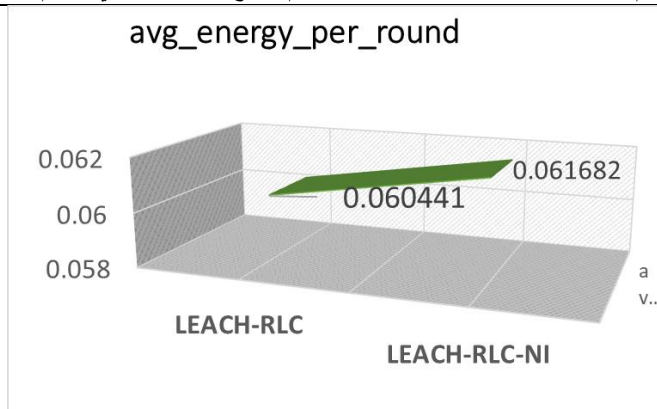


الشكل (8): مقارنة متوسط معدل التسليم بين LEACH-RLC و LEACH-RLC-NI

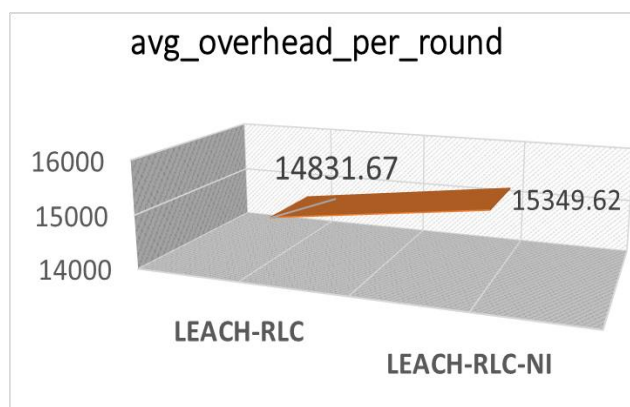


الشكل (9): مقارنة متوسط عدد الحزم المفقودة بين LEACH-RLC و LEACH-RLC-NI

في المقابل كان هناك كلفة تشغيلية طفيفة بسبب قرارات إعادة التجميع عند تدهور جودة الاتصال، مما أدى إلى ارتفاع متوسط الطاقة المستهلكة في الجولة بنسبة 2.05% من 0.060441 إلى 0.061682، وكلفة حزم التحكم من 14831.67 إلى 15349.62، مبين في الشكلين (10) و(11).



الشكل (10): مقارنة متوسط الطاقة المستهلكة في الجولة بين LEACH-RLC و LEACH-RLC-NI.



الشكل (11): مقارنة متوسط حمل التحكم في الجولة بين LEACH-RLC و LEACH-RLC-NI.

يلخص الجدول (3) نتائج تقييم أداء البروتوكول المقترح بالمقارنة مع البروتوكول LEACH-RLC [7] كما يلي:

**الجدول (3):** تقييم أداء البروتوكول المحسن LEACH-RLC-NI مقابل بروتوكول LEACH-RLC في ظروف الشبكة غير المستقرة

موازنة أداء شبكات انترنت الاشياء المعتمدة على بروتوكول LEACH باستخدام التعلم المعزز

| المقياس                | LEACH-RLC | LEACH-RLC-NI | النسبة المئوية | p-value |
|------------------------|-----------|--------------|----------------|---------|
| FND                    | 555.8     | 559.4        | %0.65+         | 0.712   |
| avg_pdr_e2e            | 0.837069  | 0.841195     | %0.49+         | 0.529   |
| avg_pkts_lost          | 7373.4    | 7277.0       | %1.31-         | 0.097   |
| avg_energy_per_round   | 0.060441  | 0.061682     | %2.05+         | 0.189   |
| avg_overhead_per_round | 14831.67  | 15349.62     | %3.49+         | 0.038   |

#### 4.5. مناقشة النتائج

تؤكد النتائج أن الانتقال من دالة مكافأة تركز بصورة رئيسية على مؤشر موت أول عقدة (FND) إلى دالة مكافأة متعددة العوامل أدى إلى نقل قرار التعلم المعزز من منظور يعتمد على الطاقة فقط إلى منظور أوسع يأخذ في الاعتبار موثوقية التسليم وكلفة الاتصال. ففي المنهجية الأصلية يميل الوكيل إلى الإبقاء على البنية التجمعية الحالية ما دامت مؤشرات الطاقة لم تصل إلى مستوى حرج، حتى لو بدأت مؤشرات الاتصال مثل معدل التسليم أو عدد الحزم المفقودة بالتدهور. أما في المنهجية المقترحة فقد أصبح القرار أكثر ارتباطاً بحالة الاتصال الفعلية بفضل إدخال مؤشرات مثل معدل التسليم، وإعادة الإرسال، وحمل التحكم، واحتمالية نجاح الرابط ضمن كل من دالة المكافأة ومتجه الحالة.

وتشير النتائج العددية إلى أن التحسينات المحققة كانت محدودة لكنها ذات دلالة تفسيرية مهمة في سياق الروابط غير المثالية. فقد ارتفعت قيمة (FND) من 555.8 إلى 559.4 جولة، أي بنسبة تحسن بلغت (0.65%)، إلا أن هذا الفرق لم يكن ذا دلالة إحصائية عند مستوى (0.05) حيث بلغت قيمة (p=0.712). كما ارتفع معدل التسليم من طرف إلى طرف من 0.837069 إلى 0.841195 بنسبة (0.49%) مع قيمة (p=0.529)، وهي أيضاً زيادة محدودة وغير دالة

إحصائياً. لذلك لا يتم تفسير التحسن في (FND) أو (PDR) بوصفه دليلاً منفرداً على تفوق حاسم للمنهجية، بل بوصفه اتجاهاً إيجابياً ضمن محاولة موازنة الأداء في بيئة اتصال غير مثالية. بينما أظهرت النتائج انخفاض عدد الحزم المفقودة من 7373.4 إلى 7277.0 بنسبة انخفاض بلغت (1.31%) مع قيمة ( $p=0.097$ ). ورغم أن هذه القيمة لا تصل إلى مستوى الدلالة الإحصائية التقليدي (0.05)، فإنها تعكس اتجاهاً أوضح نحو تحسين موثوقية التسليم مقارنة بمؤشر (FND). وهذا يدعم الفكرة الأساسية للمنهجية المقترحة وهي أن جودة الأداء في شبكات إنترنت الأشياء لا ينبغي أن تُقاس فقط بتأخير موت أول عقدة بل أيضاً بقدرة الشبكة على المحافظة على تسليم البيانات وتقليل الفقد تحت ظروف اتصال غير مستقرة.

لكن مقابل هذا التحسن في الموثوقية كان هناك كلفة تشغيلية محدودة. فقد ارتفع متوسط الطاقة المستهلكة في الجولة من 0.060441 إلى 0.061682 بنسبة (2.05%) مع قيمة ( $p=0.189$ )، كما ارتفع حمل التحكم من 14831.67 إلى 15349.62 بنسبة (3.49%)، وكانت هذه الزيادة دالة إحصائياً حيث بلغت قيمة ( $p=0.038$ ). ويمكن تفسير ذلك بأن الوكيل المحسن أصبح أكثر حساسية لتدهور حالة الاتصال، مما يجعله أكثر ميلاً إلى توليد قرارات أو تغييرات مرتبطة بإعادة التجميع عند ظهور مؤشرات ضعف في الرابط. وعليه فإن زيادة حمل التحكم لا تُقرأ بوصفها فشلاً في المنهجية بل بوصفها الكلفة التشغيلية المباشرة لإدخال موثوقية الاتصال ضمن قرار إعادة التجميع.

وتكشف هذه النتائج أيضاً عن سبب محدودية التحسينات المحققة، فعلى الرغم من أن دالة المكافأة ومتجه الحالة أصبحا أكثر إدراكاً للروابط غير المثالية، فإن عملية تشكيل التجمعات نفسها ما زالت تعتمد على دالة (MILP) الأصلية المطورة أساساً ضمن افتراضات أقرب إلى الاتصال المثالي، حيث تركز على الطاقة والموقع والبنية التجمعية أكثر من تركيزها على احتمالية نجاح الرابط أو كلفة إعادة الإرسال. لذلك فإن وكيل (DQN) يستطيع اتخاذ قرار أفضل حول توقيت إعادة التجميع، لكنه لا يتحكم بالكامل في جودة التشكيل الناتج عن (MILP) إذا لم تكن دالة التشكيل نفسها مدركة لجودة الاتصال. وهذا يفسر أن التحسين ظهر بصورة أوضح في تقليل الفقد وتحسين اتجاه الموثوقية، بينما بقي أثره محدوداً على مؤشرات عمر الشبكة مثل (FND).

بناءً على ذلك لا تتمثل مساهمة هذا العمل في تحقيق تفوق عددي كبير في جميع المؤشرات، بل في تقديم إطار أولي يربط قرار إعادة التجميع بخصائص الروابط غير المثالية داخل التعلم المعزز.

ويُظهر التحليل الإحصائي أن المنهجية المقترحة تتجه نحو تحسين موثوقية التسليم وخفض الفقد، لكنها تحتاج إلى ضبط أدق للأوزان وتطوير إضافي لدالة تشكيل التجمعات حتى تتحول هذه الاتجاهات إلى فروق أقوى وأكثر دلالة إحصائية. وهذا يجعل النتائج الحالية أساساً تجريبياً مناسباً لتطوير لاحق يدمج جودة الرابط داخل (MILP) نفسها، بدلاً من الاكتفاء بإدخالها في دالة المكافأة ومتجه الحالة فقط.

## 6. الاستنتاجات والتوصيات

قدم هذا البحث تطويراً عملياً على بروتوكول LEACH-RLC من خلال مراعاة خصائص الروابط غير المثالية داخل كل من دالة المكافأة ومتجه الحالة في DQN، وذلك بهدف جعل قرار إعادة التجميع أكثر ارتباطاً بالحالة الواقعية للشبكة. وتمثل هذا التطوير في بروتوكول مُحسّن أشرنا إليه باسم LEACH-RLC-NI، يجمع بين الحفاظ على البنية العامة لبروتوكول LEACH-RLC وبين تعزيز رؤية وكيل التعلم المعزز لمؤشرات الاتصال غير المثالي.

وأظهرت نتائج اختبار أداء بروتوكول LEACH-RLC في ظل ظروف اتصال غير مستقرة أنه يتأثر بصورة واضحة عند الانتقال من فرضية الروابط المثالية، حيث تراجعت مؤشرات عمر الشبكة (كل من FND و LND)، وازدادت الطاقة المستهلكة نتيجة فشل الإرسال وإعادة الإرسال وهو ما يؤكد أهمية عدم الاكتفاء بفرضية الاتصال المثالي عند تقييم بروتوكولات التجميع في شبكات إنترنت الأشياء.

وأشارت النتائج أن إدخال مؤشرات جودة الاتصال داخل قرار التعلم المعزز في النسخة المقترحة LEACH-RLC-NI أدى إلى زيادة حساسية الوكيل تجاه تدهور الروابط، لكنه أضاف في الوقت نفسه كلفة تشغيلية مرتبطة بحزم التحكم وإعادة التجميع. لذلك فإن دلالة النتائج لا تكمن في تحقيق تفوق مطلق في جميع المقاييس، بل في إظهار إمكانية توجيه قرار إعادة التجميع نحو موازنة أكثر واقعية بين موثوقية التسليم والكلفة التشغيلية في بيئة اتصال غير مثالية.

ولتحقيق موازنة أفضل نوصي بتوسيع هذا العمل مستقبلاً عبر تطوير دالة (MILP) نفسها لتراعي احتمالية نجاح الرابط، وكلفة إعادة الإرسال، وموثوقية الاتصال عند اختيار رؤوس التجمعات وتحديد العقد التابعة لها. كما نوصي بإجراء ضبط آلي لأوزان دالة المكافأة بدلاً من الاعتماد على ضبط تجريبي محدود لاختيار أوزان تحقق توازناً أفضل بين (FND) ، و (PDR)، والطاقة، وحمل التحكم. ويتطلب ذلك إعادة تدريب وكيل التعلم المعزز على المخرجات الجديدة

## المراجع:

- 167

- [8] Baccour, N., Koubaa, A., Mottola, L., Zúñiga, M. A., Youssef, H., Boano, C. A., & Alves, M. (2012). Radio link quality estimation in wireless sensor networks: A survey. *ACM Transactions on Sensor Networks*, 8(4), Article 34. <https://doi.org/10.1145/2240116.2240123>
- [9] Ding, X., Sun, G., Yang, G., & Shang, X. (2016). Link investigation of IEEE 802.15.4 wireless sensor networks in forests. *Sensors*, 16(7), Article 987. <https://doi.org/10.3390/s16070987>
- [10] Almutairi, H., & Zhang, N. (2024). A survey on routing solutions for low-power and lossy networks: Toward a reliable path-finding approach. *Network*, 4(1), 1–32. <https://doi.org/10.3390/network4010001>
- [11] الهلالي، م.، مرعي، أ.، وعباس، م. (2025). تصميم وتحقيق مقياس جودة وصلة اتصال جديد يحقق توازن الحمل داخل شبكات إنترنت الأشياء التي تعتمد على بروتوكول التوجيه (RPL). مجلة جامعة حمص، سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، 47(6)
- [12] شاهين، خ.، والحمصي، أ. (2024). مقارنة أداء شبكات الحساسات اللاسلكية المعتمدة على طريقة دورة العمل مع تلك المعتمدة على بروتوكولات التوجيه التراتبية. مجلة جامعة حمص، سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية، 46(7).
- [13] Al-Husain, E. A., & Al-Suhail, G. A. (2021). E-FLEACH: An improved fuzzy based clustering protocol for wireless sensor network. *Iraqi Journal for Electrical and Electronic Engineering*, 17(2), 190–197. <https://doi.org/10.37917/ijeec.17.2.21>
- [14] El-Sayed, H. H., Abd-Elgaber, E. M., Zanaty, E. A., Alsubaei, F. S., Almazroi, A. A., & Bakheet, S. S. (2024). An efficient neural network LEACH protocol to extended lifetime of wireless sensor networks. *Scientific Reports*, 14, Article 26943. <https://doi.org/10.1038/s41598-024-75904-1>
- [15] Cheng, X., Xu, C., Liu, X., Li, J., & Zhang, J. (2022). LEACH protocol optimization based on weighting strategy and the improved ant colony algorithm. *Frontiers in Neurorobotics*, 16, Article 840332. <https://doi.org/10.3389/fnbot.2022.840332>
- [16] Lechani, T., Ourari, S., Rahmoune, F., Amari, S., & Termeche, H. (2025). EDK-LEACH: Improving LEACH protocol-based machine learning in wireless sensor networks. *Indonesian Journal of Electrical Engineering and Computer Science*, 37(2), 1251–1261. <https://doi.org/10.11591/ijeecs.v37.i2.pp1251-1261>
- [17] Ibrahim, S., Mostafa, M., Jnadi, A., Salloum, H., & Osinenko, P. (2024). Comprehensive overview of reward engineering and shaping in

- advancing reinforcement learning applications. *IEEE Access*, 12, 175473–175500. <https://doi.org/10.1109/ACCESS.2024.3504735>
- [18] Jurado-Lasso, F. F., Jurado, J. F., & Fafoutis, X. (2025). *PyNetSim* [Computer software]. GitHub. <https://github.com/fdojurado/PyNetSim.git>
- [19] Gilbert, E. N. (1960). Capacity of a burst-noise channel. *Bell System Technical Journal*, 39(5), 1253–1265. <https://doi.org/10.1002/j.1538-7305.1960.tb03959.x>
- [20] Elliott, E. O. (1963). Estimates of error rates for codes on burst-noise channels. *Bell System Technical Journal*, 42(5), 1977–1997. <https://doi.org/10.1002/j.1538-7305.1963.tb00955.x>
- [21] Haßlinger, G., & Hohlfeld, O. (2008). The Gilbert-Elliott model for packet loss in real time services on the Internet. *Proceedings of the 14th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems (MMB)*, 269–286.
- [22] Zuniga, M., & Krishnamachari, B. (2004). Analyzing the transitional region in low power wireless links. *Proceedings of the First IEEE International Conference on Sensor and Ad Hoc Communications and Networks (SECON)*.
- [23] Baccour, N., Koubaa, A., Mottola, L., Zúñiga, M. A., Youssef, H., Boano, C. A., & Alves, M. (2012). Radio link quality estimation in wireless sensor networks: A survey. *ACM Transactions on Sensor Networks*, 8(4), Article 34.
- [24] Mamun, M. M. I., Hasan-Al-Mahmud, T., Kumar, S., & Islam, M. Z. (2010). Analyzing the low power wireless links for wireless sensor networks. *arXiv:1002.4838*. <https://arxiv.org/abs/1002.4838>
- [25] Cerar, G., Yetgin, H., Mohorčič, M., & Fortuna, C. (2018). Machine learning for wireless link quality estimation: A survey. *arXiv:1812.08856*. <https://arxiv.org/abs/1812.08856>
- [26] Mahmood, A., Hossain, M. M. A., & Gidlund, M. (2017). Cross-layer optimization of wireless links under reliability and energy constraints. *arXiv:1705.01774*. <https://arxiv.org/abs/1705.01774>
- [27] Mahmood, A., Hossain, M. M. A., Cavdar, C., & Gidlund, M. (2018). Energy-reliability aware link optimization for battery-powered IoT devices with non-ideal power amplifiers. *arXiv:1810.11902*. <https://arxiv.org/abs/1810.11902>

- [28] De Couto, D. S. J., Aguayo, D., Bicket, J., & Morris, R. (2003). A high-throughput path metric for multi-hop wireless routing. Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom), 134–146. <https://doi.org/10.1145/938985.939000>