

مجلة جامعة البعث

سلسلة العلوم الهندسية الميكانيكية
والكهربائية والمعلوماتية



مجلة علمية محكمة دورية

المجلد 43 . العدد 27

1442 هـ - 2021 م

الأستاذ الدكتور عبد الباسط الخطيب

رئيس جامعة البعث

المدير المسؤول عن المجلة

رئيس هيئة التحرير	أ. د. ناصر سعد الدين
رئيس التحرير	أ. د. درغام سلوم

مديرة مكتب مجلة جامعة البعث

بشرى مصطفى

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

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

الاتصال بالعنوان التالي:

رئيس تحرير مجلة جامعة البعث

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

. هاتف / فاكس : 963 31 2138071 ++

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

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

ISSN: 1022-467X

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

25 دولاراً أمريكياً خارج القطر العربي السوري

قيمة الاشتراك السنوي : 1000 ل.س للعموم

500 ل.س لأعضاء الهيئة التدريسية والطلاب

250 دولاراً أمريكياً خارج القطر العربي السوري

توجه الطلبات الخاصة بالاشتراك في المجلة إلى العنوان المبين أعلاه.

يرسل المبلغ المطلوب من خارج القطر بالدولارات الأمريكية بموجب شيكات

باسم جامعة البعث.

تضاف نسبة 50% إذا كان الاشتراك أكثر من نسخة.

شروط النشر في مجلة جامعة البعث

الأوراق المطلوبة:

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

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

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

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

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

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

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

. بعد الكنية والاسم وسنة النشر يضاف عنوان البحث وتتبعه فاصلة، اسم المجلد ويوضع تحته خط وتتبعه فاصلة . المجلد والعدد (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 البعث

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

المحتوى

الصفحة	اسم الباحث	اسم البحث
48-11	م. جيهان الاسطواني أ.م.د مهلب الداود د. فأتان عجب	التوافق الحيوي لصفائح ترميم العظام (تيتانيوم _ زركون) المسامية
82- 49	الدكتورة زينب راتب خلوف	تصنيفُ التسلّلات إلى الشبكات المحليّة باستخدام الغابات العشوائية و <i>Apache</i> <i>Spark</i>
110-83	د. ايلي خليل	دراسة ومقارنة أداء طرائق تخمين القناة في نظام <i>MIMO-OFDM</i> عند استخدام شفرات <i>STBC</i>
154-111	م. مريان سعدو د. عمار زقزوق	التحكم بسكر الدم لدى مرضى السكري باستخدام متحكم <i>PID</i> وخوارزميات المعاييرة

التوافق الحيوي لصفائح ترميم العظام (تيتانيوم _

زركون) المسامية

م. جيهان الاسطواني (1)

أ.م.د مهلب الداود (2)

د. فاتن عجيب (3)

الملخص

تتطلب الإصابات العظمية زرع مواد حيوية اصطناعية لتقويم وإصلاح العظم المتضرر. تستخدم خلائط التيتانيوم المسامية على نطاق واسع كزرعات في تطبيقات الجراحة التقويمية للعظام، يجب أن تكون الخصائص الميكانيكية لخلائط التيتانيوم المسامية مماثلة للخصائص الميكانيكية للعظم الطبيعي وذلك من أجل الحصول على عملية زرع مرضية وظيفياً، كما يجب أن تتمتع صفائح ترميم العظام المصنعة من خلائط التيتانيوم بمسامية أعلى من (40%). ومع ذلك لا يمكن لهذه الخلائط تحقيق الارتباط المباشر بالعظام و دعم عملية تكوين الأنسجة العظمية الجديدة دون زيادة النشاط الحيوي لأسطح خلائط التيتانيوم حيث يتم إجراء معالجات كيميائية للسطح ثم الغمر في السائل الحيوي (SBF) المحاكي لسائل جسم الإنسان. تم في هذا البحث إعداد صفائح ترميم العظام من خليطة (تيتانيوم _ زركون) وفق النسبة الوزنية (70%wtTi_30%wtZr) باستخدام طريقة ميتالورجيا المساحيق المعدنية، حُددت الخصائص الميكانيكية للعينات الناتجة باستخدام اختبار المقاومة على الانضغاط، تم حساب المسامية للعينات الناتجة للتأكد من الحصول على

(1) معيدة موفدة داخليا في قسم هندسة التصميم الميكانيكي - كلية الهندسة الميكانيكية والكهربائية - جامعة دمشق.

(2) أستاذ مساعد في قسم هندسة التصميم الميكانيكي - كلية الهندسة الميكانيكية والكهربائية - جامعة دمشق.

(3) مدرسة في قسم الهندسة الطبية - كلية الهندسة الميكانيكية والكهربائية - جامعة دمشق.

المسامية المطلوبة لصفائح ترميم العظام، لزيادة النشاط الحيوي لسطح عينة أُجريت عليها المعالجة الكيميائية التالية: (معالجة قلوية بمحلول هيدروكسيد الصوديوم_ معالجة بمحلول كلوريد الكالسيوم_ معالجة حرارية_ معالجة بالماء الساخن) ثم غُمرت العينة المعالجة في السائل الحيوي (SBF) لمدة ثلاثة أيام، ثم حُدّد التركيب الكيميائي لها لمقارنته بالتركيب الكيميائي لعينة غير مغمورة في السائل الحيوي (SBF) باستخدام جهاز تحديد العناصر الكيميائية بطريقة تشتت طاقة الأشعة السينية (EDX). بيّنت النتائج وجود الفوسفور في العينة المغمورة لمدة (3أيام) في السائل الحيوي (SBF)، أُجريت التجارب في المعهد العالي للعلوم التطبيقية والتكنولوجيا.

الكلمات المفتاحية: خلائط التيتانيوم، صفائح ترميم العظام، ميتالورجيا المساحيق المعدنية، التوافق الحيوي.

Biocompatibility of Porous (Ti-Zr) Scaffolds

Eng. Jehan AlEstwani⁽⁴⁾

Mohallab AlDawood⁽⁵⁾

Dr. Faten Ajeeb⁽⁶⁾

Abstract

Bone injuries often require the implantation of biomaterial to heal and recover the damaged bone, Porous Titanium alloys are widely used as implants in orthopedic applications, the mechanical properties of porous Titanium alloys should match those of natural bone in order to obtain a functionally satisfactory implant, and Titanium scaffolds must have a higher porosity from (40%). However, these alloys cannot achieve direct attachment to bone and support the formation of new bone tissue without increasing the bioactivity of the surfaces of titanium alloys, where chemical treatments are carried out on the surfaces followed by immersion in simulated body fluid (SBF). In this research, porous scaffolds were prepared from (70%wt Ti-30%wt Zr) by using Powder Metallurgy method, the mechanical properties of the sintered samples were determined by compressive test, the porosity of sintered samples was calculated to ensure that the required porosity of scaffolds was obtained. To modify the metal surface of a sample to bioactive it, The following chemical treatment was performed on this sample (alkali_ $CaCl_2$ _ heat and hot water) then the sample was immersed

⁽⁴⁾ Teaching Assistance at the Mechanical Engineering Department, Faculty of Mechanical and Electrical Engineering, Damascus University.

⁽⁵⁾ Assistant Professor at the Mechanical Engineering Department, Faculty of Mechanical and Electrical Engineering, Damascus University.

⁽⁶⁾ Teacher at the Biomedical Engineering Department, Faculty of Mechanical and Electrical Engineering, Damascus University.

in (SBF) for three days, then the chemical composition of the treated sample was determined to compare it with the chemical composition of a sample that wasn't immersed in (SBF) by (EDX) device (Energy Dispersive X-ray Fluorescence Spectrometer). The results showed the presence of phosphorous in the sample submerged for (3days) in simulated body fluid (SBF). The experiments were conducted at the Higher Institute of Applied Sciences and Technology.

Key words: Titanium alloys, Scaffolds, Powder Metallurgy, biocompatibility.

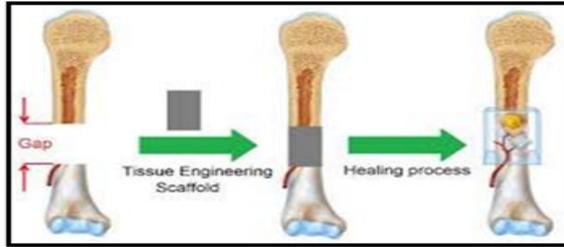
1-المقدمة:

تستخدم خلائط التيتانيوم على نطاق واسع في التطبيقات الطبية الحيوية نظراً لخصائصها الممتازة مثل: المتانة العالية، التوافق الحيوي، المقاومة الجيدة للتآكل، حيث تم استخدامها في تصنيع: مفصل الورك الاصطناعي، مفصل الركبة الاصطناعي، صفائح عظمية وبدائل العظام [1,2].

توسّع استخدام الزركون في التطبيقات الحيوية بسبب خصائصه الفيزيائية والبيولوجية الجيدة، ومقاومة الاهتراء حيث أظهرت بدائل الزركون في عظم الفخذ مقاومة عالية للاهتراء [3].

تعتبر خلائط (Ti-Zr) من المواد الحيوية المعدنية الممتازة النشطة بيولوجياً لأن هذه الخلائط ممكن أن تتشكل على سطحها في الجسم الحي طبقة أباتيت تشبه العظم مما يؤدي إلى حدوث الترابط بين المادة الحيوية المزروعة والعظم السليم [4,5].

حيث يتم تحفيز خلايا العظام الجديدة على النمو وبالتالي شفاء العيب بمساعدة صفائح ترميم العظام (Scaffolds)، يتم تصنيع هذه الصفائح مع بنية مسامية لأنّ المسامات تسمح بنمو الأنسجة العظمية الجديدة وبالتالي تقويم العظم المصاب [6]، ويبيّن الشكل (1) عملية زرع صفيحة ترميم العظام وشفاء العظم المصاب بعد نمو الأنسجة العظمية الجديدة في الهيكل المسامي لصفيحة ترميم العظام [7].



الشكل (1) صفيحة ترميم العظام [7].

وللبنية المسامية أيضاً دور في جعل معامل المرونة لصفيحة الترميم المزروعة قريب من معامل المرونة للأنسجة العظمية المضيفة وبالتالي منع حدوث ارتشاف للعظم والحصول على عملية زرع مرضية وظيفياً [8].

لذلك يجب أن تتمتع صفائح ترميم العظام بمسامية أعلى من (40%) [6] ، وأن تكون خصائصها الميكانيكية قريبة جداً من خصائص العظم الطبيعي:

1- معامل مرونة منخفض (0.1÷30Gpa)

2- مقاومة على الانضغاط (2 ÷ 200Mpa) [9] .

تعتبر تقانة ميتالورجيا المساحيق المعدنية طريقة فعالة لإنتاج المواد الحيوية المسامية، حيث تسمح هذه الطريقة بالحصول على المسامية المطلوبة من خلال التحكم بضغط التدميغ، بالإضافة إلى ذلك فإن استخدام طريقة ميتالورجيا المساحيق المعدنية للحصول على خلائط التيتانيوم تعطي تجانس كيميائي أكبر لأجزاء التيتانيوم المصنوعة من المساحيق [10] .

لزيادة النشاط الحيوي لأسطح خلائط التيتانيوم وبالتالي تحقيق الترابط بين صفيحة الترميم المزروعة والعظم المتضرر نقوم بإجراء معالجات كيميائية لأسطح هذه الخلائط، تشمل الطرق الكيميائية للمعالجة: التتميش بالحمض، الأكسدة باستخدام بيروكسيد الهيدروجين، معالجة (قلوية - حرارية)، معالجة (قلوية - حمضية)، الخ... [11] .

بعد المعالجات الكيميائية والحرارية تصبح أسطح خلائط التيتانيوم نشطة بيولوجياً باستخدام تقنية الطلاء الحيوي المحاكي لسائل الجسم البشري.

تستخدم تقنية الطلاء الحيوي محاليل مائية ذات تركيبة أيونية مماثلة لتلك الموجودة في البلازما البشرية (SBF)، فهي تسمح بترسيب بلورات الأباتيت على طبقة فوسفات الكالسيوم المتشكلة على السطح المعدني بعد إجراء المعالجات الكيميائية والحرارية عليه [12,13].

إنّ المعالجة القلوية بمحلول هيدروكسيد الصوديوم (NaOH) هي الطريقة الأكثر استخداماً لزيادة نشاط السطح المعالج، ومع ذلك أظهرت العديد من الدراسات قدرة أفضل على تشكيل الأباتيت عند إجراء المعالجة الحرارية [14,15]. أثناء المعالجة القلوية طبقة (TiO₂) الموجودة على سطح التيتانيوم تذوب جزئياً في المحلول القلوي بسبب هجوم مجموعات الهيدروكسيل، إضافة إلى ذلك ينتج عن هجوم الهيدروكسيل على (TiO₂) المُمَيَّه هيدرات سالبة الشحنة على السطح المعالج، تتحد هذه الأنواع سالبة الشحنة مع

الأيونات القلوية في المحلول المائي لإنتاج طبقة هيدروجيل تيتانات قلوية ($Na_2 - XH_xTi_3O_7 \cdot nH_2O$) حيث X : عدد ذرات الهيدروجين الممكن أن تحل مكان ذرات الصوديوم في الصيغة) و (n : عدد جزيئات الماء المسببة للهيدروجيل) [16].
أثناء المعالجة الحرارية يتم تجفيف طبقة الهيدروجيل وتكثيفها لتشكيل طبقة تيتانات قلوية مستقرة متبلورة أو غير متبلورة [17]. إن مجموعات ($Ti-OH$) المتشكلة على سطح تيتانات الصوديوم بعد الغمر في السائل المحاكي لسائل الجسم البشري (SBF) تكون سالبة الشحنة، وبالتالي ستتحد بشكل انتقائي مع أيونات (Ca^{2+}) ذات الشحنة الموجبة في المحلول لتشكيل تيتانات الكالسيوم. مع تراكم أيونات الكالسيوم على السطح، يكتسب السطح تدريجياً شحنة موجبة، نتيجة لذلك يتحد السطح الموجب الشحنة مع أيونات الفوسفات سالبة الشحنة لتشكيل فوسفات الكالسيوم غير المتبلور.
يتحول فوسفات الكالسيوم تلقائياً إلى أباتيت في بيئة الجسم لأن الأباتيت هو الطور المستقر [18].

بين الباحث (Takashi Kizuki) وآخرون [19] أن أفضل معالجة للسطح المعدني والتي تتشكل من خلالها طبقة الأباتيت على السطح المعدني بعد غمره لمدة يوم واحد فقط في السائل الحيوي هي (معالجة قلوية بمحلول هيدروكسيد الصوديوم_ معالجة بمحلول كلوريد الكالسيوم_ معالجة حرارية_ معالجة بالماء الساخن)، حيث تُمكن هذه المعالجة من دمج عدد كبير نسبياً من الكالسيوم بسهولة على السطح المعدني المعالج بعد غمره في محلول هيدروكسيد الصوديوم ($NaOH$) ثم غمره في محلول كلوريد الكالسيوم ($CaCl_2$)، وذلك لأن تيتانات هيدروجين الصوديوم التي تشكلت على السطح المعدني خلال المعالجة بمحلول هيدروكسيد الصوديوم ($NaOH$) لها بنية ذات طبقات و يتم استبدال أيونات الصوديوم (Na^+) الموجودة فيها بسهولة بأيونات الكالسيوم (Ca^{2+}) بعد معالجة السطح بمحلول كلوريد الكالسيوم ($CaCl_2$).

في المعالجة الحرارية اللاحقة نحصل على حركية منخفضة لشوارد الكالسيوم (Ca^{2+}) في تيتانات الكالسيوم الناتجة عن تيتانات هيدروجين الكالسيوم. زادت قدرة تشكيل الأباتيت من التيتانيوم المعالج حرارياً بشكل ملحوظ عندما تمت معالجته فيما بعد بالماء

الساخن. لهذا قمنا في هذا البحث باعتماد الطريقة التالية لإجراء المعالجة المسبقة للسطح المعدني: (معالجة قلوية بمحلول هيدروكسيد الصوديوم_ معالجة بمحلول كلوريد الكالسيوم_ معالجة حرارية_ معالجة بالماء الساخن).

السائل الحيوي المحاكي لسائل الجسم البشري {Simulated Body Fluid (SBF)} : هو محلول يحتوي على أيونات غير عضوية بتراكيز مماثلة لتلك الموجودة في بلازما الدم البشري، ولكن لا يحتوي على أي خلايا أو بروتينات [20].

2- هدف البحث:

✓ تحضير صفائح ترميم عظام وفق النسب الوزنية:

(70% wt Ti-30%wt Zr) بطريقة ميتالورجيا المساحيق المعدنية بحيث تكون الخصائص الميكانيكية لهذه الصفائح قريبة جداً من خصائص العظم الطبيعي، ومساميتها أعلى من (40%).

✓ تقييم التوافق الحيوي لصفائح ترميم العظام (70% wt Ti-30%wt Zr) المنتجة بطريقة ميتالورجيا المساحيق المعدنية للتحقق من إمكانية استخدامها في الجراحة التقيومية للعظام.

3- الدراسات المرجعية:

3-1- الدراسة المرجعية الأولى: قام الباحثون:

(Marcio W.D. Mendes, Carola G. Agreda) وآخرون [21] بتوصيف خليطة (Ti-27Nb-13Zr) المُنتجة بتقانة ميتالورجيا المساحيق المعدنية، كما درس الباحثون تأثير زمن الطحن على عملية التليد، وحددوا الخصائص الميكانيكية للعينات (معامل يونغ _ القساوة الميكروية) وقيم الباحثون نتائج تقنية الطلاء الحيوي على العينات للتحقق من إمكانية استخدام هذه الخليطة في تطبيقات الجراحة التقيومية للعظام أو تطبيقات طب الأسنان.

تبيّن أن الأطوار المتشكلة في الخليطة هي (α, β) ويؤثر وقت الطحن على تشكيل هذه الأطوار، وأظهر تقييم تقنية طلاء المحاكاة الحيوية أنه من الممكن تشكل الأباتيت على جميع العينات المغمورة لمدة ثلاثة أيام أو أكثر في السائل الحيوي (SBF)، وبالتالي يمكن استخدام هذه الخليطة في تطبيقات الجراحة التقيومية للعظام أو في تطبيقات طب الأسنان.

3-2- الدراسة المرجعية الثانية: قام الباحثون:

(C.G. Agreda, M.W.D. Mendes) وآخرون [22] بمقارنة قابلية تشكل الأباتيت على عينات تيتانيوم أسطوانية محضرة باستخدام تقانة ميتالورجيا المساحيق المعدنية وعينات أخرى كانت على شكل صفائح من التيتانيوم، وذلك بعد غمر عينات التيتانيوم الصفائحية والعينات المحضرة باستخدام تقانة ميتالورجيا المساحيق المعدنية في السائل الحيوي (SBF). كما قام الباحثون بإجراء عدة أنواع من المعالجات الكيميائية لسطح العينات والمقارنة بين سماكة

طبقة الأباتيت المتشكلة على سطح العينات عند كل نوع من المعالجات الكيميائية المستخدمة. المعالجات الكيميائية المستخدمة: معالجة قلووية، (قلوية- معالجة حرارية)، (حمضية- قلووية)، (قلوية- كلوريد الكالسيوم- معالجة حرارية- معالجة بالماء الساخن). وتبين أن طبقة الأباتيت على عينات التيتانيوم المحضرة باستخدام تقانة ميتالورجيا المساحيق المعدنية تشكلت بعد غمر العينات لمدة يوم واحد فقط في السائل الحيوي (SBF)، بينما تشكلت طبقة الأباتيت على عينات التيتانيوم الصفائحية بعد غمر لمدة تسعة أيام في السائل الحيوي (SBF).

أكبر سماكة لطبقة الأباتيت على عينات التيتانيوم المحضرة باستخدام تقانة ميتالورجيا المساحيق المعدنية كانت باستخدام المعالجة الكيميائية (قلوية- كلوريد الكالسيوم- معالجة حرارية- معالجة بالماء الساخن) بعد غمر ليوم واحد فقط في السائل الحيوي (SBF). أما بالنسبة لعينات التيتانيوم الصفائحية كانت أكبر سماكة لطبقة الأباتيت عند إجراء معالجة (حمضية - قلووية) بعد غمر لمدة تسعة أيام في السائل الحيوي (SBF).

4- المواد والأجهزة المستخدمة:

4-1- المواد المستخدمة:

- ❖ مسحوق التيتانيوم: من إنتاج شركة Merck الألمانية، حبيبات صفيحية الشكل، حجم الحبيبات $(13.77\mu\text{m} \div 14.756\mu\text{m})$ ، النقاوة (98.8%).
- ❖ مسحوق الزركون: من إنتاج شركة Merck الألمانية، شكلها أقرب للكروي، حجم الحبيبات $(9.099\mu\text{m} \div 11.194\mu\text{m})$ ، النقاوة (96.4%).
- ❖ الكواشف الكيميائية التالية:

{NaCl, NaHCO₃, KCl, K₂HPO₄. 3H₂O, MgCl₂. H₂O, 1M HCl, CaCl₂, Na₂SO₄, Tris((CH₂OH)₃ CNH₂)}.

❖ محلول (5M) NaOH، M: تعني مول.

❖ محلول (100 mM) CaCl₂، mM: تعني ميلي مول.

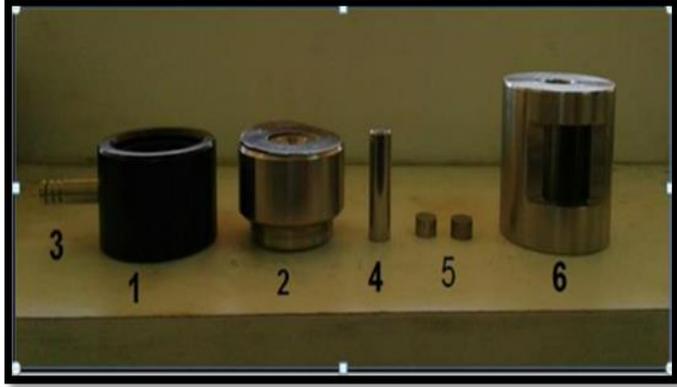
4-2- الأجهزة المستخدمة: تم إجراء هذا البحث في المعهد العالي للعلوم التطبيقية والتكنولوجيا.

❖ خلاط لخلط المساحيق المعدنية المبيّن بالشكل (2)، من تصميم طلاب المعهد العالي للعلوم التطبيقية والتكنولوجيا، يتم تصنيعه بضم أسطوانتين متساويتي الطول على شكل حرف (V) بحيث يمكنه الدوران حول محوره الأفقي كما يبين الشكل (2)، يتم إدخال المواد المراد خلطها إلى داخل حجرة الخلط بحيث لا يُملأ أكثر من النصف للسماح للمساحيق المعدنية بحرية الحركة أثناء الخلط، ثم تتحرك حجرة الخلط كاملة وعندما يكون الخلاط معكوس يسقط المسحوق وينفصل إلى فرعين وفي كل فرع يحصل خلط مضطرب بالهواء المتحرك الصاعد. أقصى سرعة لمحرك الخلاط (90 r.p.m).



الشكل (2) الخلاط الميكانيكي.

❖ قالب معدني لكبس العينات المبين في الشكل (3)، وهو قالب لكبس عينات بقطر (10 mm) وضغط (5 Ton) كحد أقصى، مصنوع من الفولاذ المقاوم للصدأ المقسى لتحمل الضغوط العالية.



الشكل (3) قالب العينات الأسطوانية بقطر (10 mm).

أجزاء القالب: 1- أسطوانة جسم القالب السفلية.

2- أسطوانة جسم القالب العلوية.

3- أنبوب ممكن وصله بأسطوانة هواء لتطبيق ضغط من الطرفين.

4- وتد تطبيق الضغط.

5- قرصا الكبس العلوي والسفلي.

6- اللافت.

❖ مكبس هيدروليكي يدوي المبين في الشكل (4)، تم استخدامه من أجل كبس

العينات الأسطوانية، وهو مكبس مصمم من أجل مجال واسع من تطبيقات الكبس حتى حمل مقداره (15 Ton)، حيث يتم تطبيق الحمل عن طريق تحريك ذراع تطبيق الضغط بشكل يدوي حتى الوصول إلى الحمل المطلوب، ويتم تطبيق قوة الكبس من اتجاه واحد من الأسفل، وللجهاز غطاء حماية بلاستيكي شفاف مع حساس حيث يتوقف الجهاز عن العمل في حال كون الغطاء مفتوحاً،

ويمكن استخدام عدد كبير من القوالب التي تتحمل ضغوط حتى (15 Ton) مع المكبس، دقة المكبس (1Ton).



الشكل (4) مكبس هيدروليكي يدوي.

❖ فرن المعالجة الحرارية الكتيم طراز (HT 1600 M) المبين في الشكل (5)، تم استخدامه لإجراء عملية التليد للعينات، وهو فرن مخصص لمختلف أنواع المعالجات الحرارية ولا سيما درجات الحرارة العالية، حيث يمكن الوصول إلى درجات حرارة حتى (1600°C).

يعتمد الفرن في التسخين على أربع وشائع موزعة على جانبي الفرن اثنتان على اليمين واثنتان على اليسار، ويتم التحكم بسير العمل بواسطة وحدة التحكم الموصولة مع الفرن عن طريق برمجة سرعة التسخين أو التبريد ودرجة الحرارة المطلوب الوصول إليها وزمن التثبيت في كل مرحلة، يعتمد الفرن في العزل على عدة طبقات من أكسيد الألمنيوم (Al_2O_3) تحيط الفرن من الوجوه الستة، ويحتوي باب الفرن على مانعة سيليكونية وعين زجاجية لرؤية العينة وهي داخل الفرن دون الحاجة لفتح الباب، وله مدخلان لإدخال الغازات حيث يتم التحكم بالتدفق عن طريق مقياس تدفق عند مدخل الفرن، مواصفات الفرن الفنية:

- درجة الحرارة العظمى التي يمكن الوصول إليها (1600°C).

- الاستطاعة العظمى للفرن (6 KVA)، الجهد (400 V)، تيار متناوب التردد (50-60 HZ)، ثلاثي الطور (3 phase).
- وزن الفرن مع وحدة التحكم (160 Kg).



الشكل (5) فرن المعالجة الحرارية الكتيم.

❖ جهاز الاختبارات العامة طراز (WDW-50) من شركة (JINAN) الشكل (6) تم استخدامه من أجل اختبار المقاومة على الانضغاط، وهو جهاز يُستخدم من أجل اختبارات الشد والضغط والانحناء ويمكن وصله بملحقات (راسمة، طابعة، حاسب) وهو موصول مع حاسب يمكن التحكم عن طريقه بمتغيرات التجربة المختلفة، يرسم المنحني الناتج ويعطي نتائج التجربة، ويتكون من فك ثابت وفك متحرك، حيث يتم تطبيق القوة بواسطة الفك المتحرك الذي يتحرك بواسطة

منصة متحركة وهذا الانزلاق الحر يؤمن إزاحة تامة للرأس المتحرك مهما يكن الحمل المطبق على العينة.



الشكل (6) جهاز الاختبارات الميكانيكية العامة طراز (WDW-50).

❖ جهاز قياس الأس الهيدروجيني (PH Meter): المبين في الشكل (7)، من شركة (SARTORIUS) نموذج (PB-11).



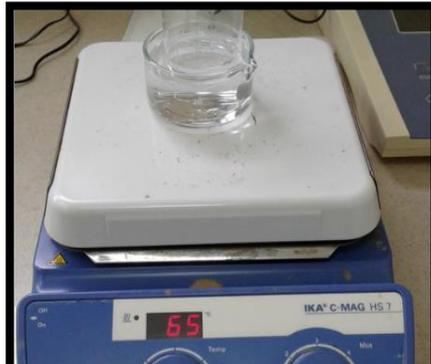
الشكل (7) جهاز قياس الأس الهيدروجيني (PH Meter).

❖ جهاز التنظيف بالأمواف فوق الصوتية: المبين في الشكل (8)، من شركة (ELMA)، نموذج (E 100 H)، تردد الأمواف فوق الصوتية (37 KHZ)، يمكن إجراء عملية التنظيف بزمن يتراوح من (0÷45 min)، كما يمكن أن تتوافق عملية التنظيف مع تسخين بدرجات حرارة تتراوح (0÷85 °C).
مبدأ عمل الجهاز: يتم تحويل الطاقة الكهربائية عالية التردد الناتجة عن مولد الأمواف فوق الصوتية إلى طاقة ميكانيكية بواسطة أنظمة محولات كهروضغطية ثم تنتقل هذه الطاقة الميكانيكية إلى حوض التنظيف. تنتج عن هذه العملية ملايين من الفقاعات الهوائية الصغيرة التي تتفجر بسبب اختلافات الضغط الناتجة عن النشاط الفوق صوتي، ينشأ تدفق عالي لتيارات سائل التنظيف تزيل هذه التيارات الأوساخ من على الأسطح المراد تنظيفها.



الشكل (8) جهاز التنظيف بالأمواف فوق الصوتية.

❖ جهاز تحريك مغناطيسي مع تسخين: المبين في الشكل (9)، من شركة (IKA) طراز (HS7). سرعة التحريك (100÷1500 r.p.m)، أعلى درجة حرارة



للتسخين

الشكل (9) المحرك المغناطيسي مع تسخين.

❖ فرن المعالجة الحرارية طراز (Lenton): المبيّن في الشكل (10)، يعتمد الفرن في التسخين على عشرة وشائع موزعة بانتظام على محيط حجرة التسخين ويتم التحكم بسير العمل بواسطة وحدة التحكم الموصولة مع الفرن. يعتمد الفرن في العزل الحراري على عدة طبقات من أكسيد الألمنيوم (Al_2O_3) تحيط الفرن من الوجوه الستة ويتم رفع العينة وتنزيلها بواسطة مصعد موصول مع سيور وبكرات ويتم التحكم به عن طريق صمام كهربائي كذلك يمكن إنزال المصعد ورؤية العينة المعالجة في أي وقت من التجربة عند درجات حرارة معقولة.

المواصفات الفنية للفرن:

-درجة الحرارة العظمى التي يمكن الوصول إليها ($1700\text{ }^\circ\text{C}$).

-الاستطاعة العظمى للفرن (5000) واط.

-الجهد (240) فولت.

-شدة التيار الكهربائي (21) أمبير.

-أبعاد حجرة التسخين ($15*15*15\text{ Cm}$).



الشكل (10) فرن المعالجة الحرارية.

❖ جهاز تحديد العناصر الكيميائية بطريقة تشتت طاقة الأشعة السينية:
(Energy Dispersive X-ray Fluorescence Spectrometer (EDX)) المبيّن
في الشكل (11) ، من شركة (SHIMADZU) طراز (EDX-8000).
مبدأ عمل الجهاز:

- 1- تعريض العينة المدروسة إلى حزمة عالية من الطاقة.
- 2- تثير هذه الطاقة الكثرونات المدار K من ذرات العناصر المكونة للعينة، يخرج الكثران من المدار الداخلي تاركاً مكانه فجوة الكثرونية.

3- يملأ الكترون من المدار الأعلى هذه الفجوة، ويتم إطلاق الفرق في الطاقة بين المدار الأعلى والمدار الأدنى على شكل أشعة سينية لها طاقة مميزة لهذا العنصر .

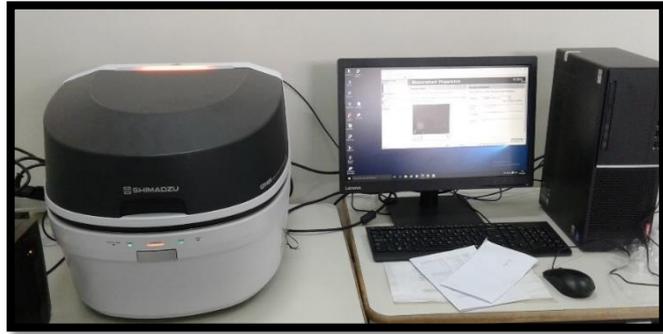
4- يقوم كاشف الجهاز بتفسير الطاقة المنطلقة حسب زمنها وشدتها على شكل عنصر معدني وذلك حسب برمجية خاصة بالاعتماد على أن الأشعة السينية لها طول موجي فريد وطاقة مميزة لكل عنصر .
مواصفات الجهاز:

-يكشف العناصر بدقة عالية من الصوديوم إلى اليورانيوم في الجدول الدوري للعناصر كشفاً دقيقاً .

-الشدّة العليا لطاقة الجهاز (50 Kev) كافية لتحريض العناصر من الصوديوم إلى اليورانيوم .

-يعمل الجهاز لكشف العناصر المعدنية في العينات الصلبة والسائلة والبودرة .

-يعمل الجهاز بجو محيط للعينة هواء أو تخلية .



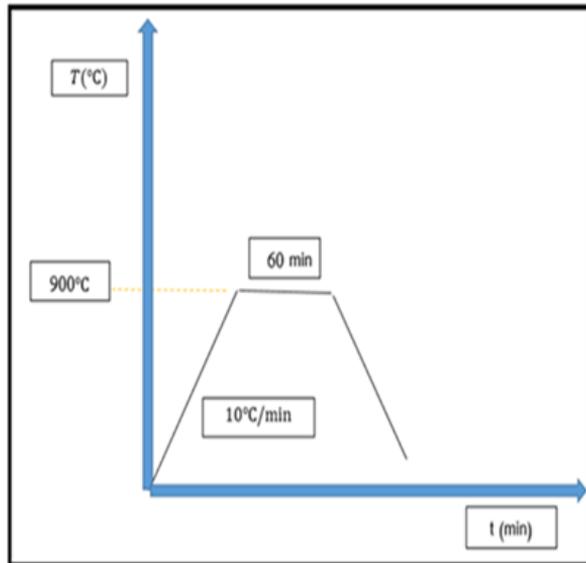
الشكل (11) جهاز تحديد العناصر الكيميائية بطريقة تشتت طاقة الأشعة السينية (EDX).

5- العمل التجريبي:

5-1- تحضير العينات:

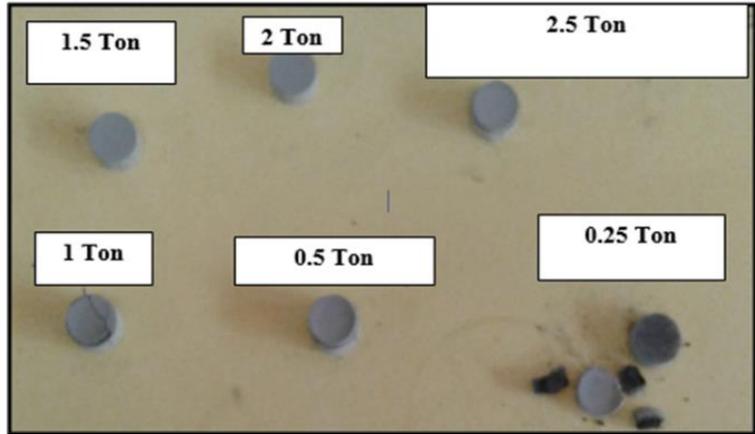
تم تحضير عينات أسطوانية الشكل وفق النسب الوزنية: (70%wt Ti- 30%wt Zr) كما يلي:

- 1- وزن (7.35 g) من بودرة التيتانيوم.
- 2- وزن (3.15 g) من بودرة الزركون.
- 3- خلط الأوزان السابقة باستخدام الخلاط الميكانيكي المبين في الشكل (2) بسرعة دوران (90 r.p.m) ، لمدة ساعة ونصف.
- 4- بعد ذلك تم تشكيل العينات بالكبس على البارد باستخدام قالب العينات الأسطوانية المبين في الشكل (3) ، وباستخدام المكبس الهيدروليكي اليدوي المبين في الشكل (4) وفق الخطوات التالية لكل عينة:
 - وزن (1.5 g) لكل عينة من المساحيق التي تم خلطها.
 - تنظيف القالب، ثم وضع الشحنة السابقة للكبس في القالب.
 - وضع القالب في المكبس الهيدروليكي اليدوي وتطبيق الضغط، عند قيم مختلفة لضغط الكبس:
(0.25, 0.5, 1, 1.5, 2, 2.5 Ton) ، وذلك من أجل تحديد ضغط الكبس المناسب تجريبياً للحصول على المسامية المطلوبة لصفائح ترميم العظام بعد التليد.
 - رفع القالب من المكبس وإخراج العينات.
- 5- تليد العينات عند درجة حرارة (900°C) لمدة (60 min) ، كما هو مبين في البرنامج الحراري للعينات الشكل (12)، باستخدام فرن المعالجة الحرارية الكتيم المبين في الشكل (5) مع استخدام غاز الأرجون كغاز للحماية حيث تم استخدام أسطوانة ارغون مخبرية وضبط التدفق على القيمة (168.9 Liter/hour) على الجهاز الموجود على الفرن، بمعدل تسخين (10°C/min).



الشكل (12) البرنامج الحراري لتلييد للعينات.

العينات الناتجة بعد التلييد مقبولة كما هو مبين في الشكل (13)، باستثناء العينات المدمجة عند قيم منخفضة لضغط الكبس: (0.25, 0.5, 1 Ton) لأن هذه القيم لضغط الكبس لم تعطي متانة خضراء كافية لتماسك العينات.



الشكل (13) العينات (70%wt Ti-30%wt Zr) الملبدة عند (900°C) لمدة (60 min). قمنا بحساب الكثافة والمسامية للعينات الناتجة كما هو موضح في الجدول (1) وذلك لتحديد ضغط الكبس المناسب للحصول على المسامية المطلوبة لصفائح ترميم العظام:

الجدول (1) الكثافة والمسامية لعينات التجربة السابقة بعد التليد:

ملاحظة على العينة	المسامية (%)	الكثافة (g/cm^3)	قطر العينة (mm)	ارتفاع العينة (mm)	وزن العينة (g)	ضغط الكبس (Mpa)	ضغط الكبس (Ton)
مرفوضة لأنها تفتت نتيجة ضغط الكبس المنخفض	62	1.97	11.35	9.33	1.864	31.830	0.25
فيها تشقق	59.2	2.02	11.34	9.12	1.861	63.661	0.5
فيها تشقق	56.07	2.18	11.29	8.455	1.845	127.323	1
مقبولة خالية من أي تشقق	53.14	2.32	11.44	7.703	1.8423	190.985	1.5
مقبولة خالية من أي تشقق	51.18	2.42	11.34	7.015	1.716	254.647	2
فيها تشقق	53.28	2.31	11.51	7.506	1.813	318.309	2.5

من نتائج الجدول (1): تبين أن ضغط الكبس المناسب للحصول على مسامية صفائح ترميم العظام المطلوبة (المسامية < 40%) هو [6] ($1.5\text{Ton}=190.98\text{Mpa}$).
 تُحسب الكثافة التجريبية للعينة من العلاقة:

$$\rho = \frac{m}{v} \text{ g/cm}^3$$

حيث: m: وزن العينة [g].

v: حجم العينة [cm^3].

تحسب المسامية للعينات من العلاقة التالية:

$$P\% = 1 - \frac{\rho}{\rho_{th}} \quad (*)$$

حيث: ρ : كثافة العينة (g/cm^3).

ρ_{th} : الكثافة النظرية للخليطة المدروسة (g/cm^3).

$$\rho_{th} = \frac{1}{\frac{x_{Ti}}{\rho_{Ti}} + \frac{x_{Zr}}{\rho_{Zr}}}$$

حيث: x_{Ti} : النسبة الوزنية للتيتانيوم في الخليطة (70%).

x_{Zr} : النسبة الوزنية للزركون في الخليطة (30%).

ρ_{Ti} : كثافة التيتانيوم في الحالة المصمتة (4.51 g/cm^3).

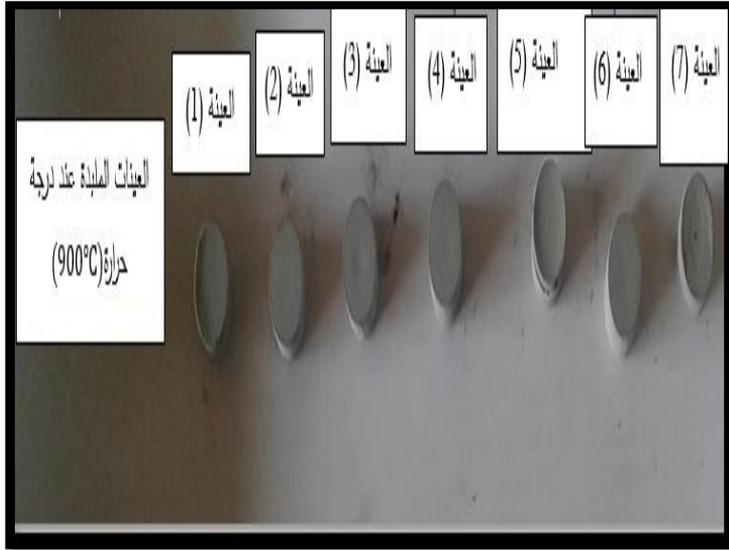
ρ_{Zr} : كثافة الزركون في الحالة المصمتة (6.49 g/cm^3).

بالتعويض نجد: $\rho_{th} = 4.964 \text{ g/cm}^3$

بعد تحديد قيمة ضغط الكبس المناسبة للحصول على مسامية مقبولة لصفائح ترميم

العظام بعد التليد عند درجة حرارة (900°C) لمدة (60 min) قمنا بتحضير (7

عينات)، يوضح الشكل (14) صورة للعينات المحضرة بعد عملية التليد.



الشكل (14) العينات المحضرة بعد عملية التليد.

وتم حساب المسامية لكل عينة باستخدام العلاقة (*) كما هو موضح في الجدول (2):

الجدول (2): مسامية العينات المحضرة باستخدام ضغط كبس (1.5 Ton) والتليد عند (900°C) لمدة (60 min):

رقم العينة	المسامية %
1	45.73
2	45.08
3	45.35
4	44.58

نلاحظ أنّ متوسط المسامية للعينات الناتجة موافقة للمسامية المطلوبة لصفائح ترميم العظام [6].

5-2- اختبار المقاومة على الانضغاط: لتحديد مقاومة العينات على الانضغاط و معامل يونغ للعينات ومقارنة القيم بالقيم المقبولة لصفائح ترميم العظام.
- تم إجراء الاختبار باستخدام جهاز الاختبارات العامة المبين في الشكل (6)، وذلك بعد اختيار ثلاث عينات من المجموعة وهي العينات (2,3,4).
- تم تعيين معامل يونغ والمقاومة على الانضغاط لكل عينة من العينات الثلاثة السابقة بعد وضع العينة بوضع متمركز مع قرص الجهاز (حيث ينطبق محور العينة مع محور تطبيق الضغط) كما هو مبين في الشكل (15)، ثم تطبيق الحمل بشكل تدريجي على العينة وبسرعة للرأس المتحرك (0.5 mm/min) من خلال الحاسب الموصول بالجهاز.



الشكل (15) وضع العينة بوضع متمركز مع قرص الجهاز في اختبار المقاومة على الانضغاط.

تُحسب المقاومة على الانضغاط لكل عينة باستخدام العلاقة التالية:

$$\sigma_{max} = \frac{F_{max}}{A} \quad (Mpa)$$

حيث: F_{max} [N]: القوة العظمى التي تحطمت عندها العينة، نحصل عليها من الجهاز.

A [mm^2]: مساحة مقطع العينة.

سنعوض قيم العينة (4):

القوة التي تحطمت عندها العينة (4) والتي حصلنا عليها من جهاز الاختبارات العامة (6):

$$F_{max} = 1.730 [KN] = 1730[N]$$

مساحة مقطع العينة (4): $A = 97.525 \text{ mm}^2$

وبالتالي تكون المقاومة على الانضغاط للعينة (4) هي: $\sigma_{max} = 17.738 [Mpa]$ يُحسب معامل يونغ من العلاقة:

$$E = \frac{\sigma}{\varepsilon} \quad [Gpa]$$

حيث: σ [Mpa]: إجهاد المقاومة على الانضغاط عند قوة ضغط طبقت على العينة ضمن مجال المرونة.

ε : الانفعال النسبي للعينة المقابل للقوة التي حُسب عندها الاجهاد، يُحسب الانفعال من العلاقة:

$$\varepsilon = \frac{\Delta L}{L_0}$$

حيث: $\Delta L [mm]$: نحصل على قيمتها من الجهاز، يعطي الجهاز عند كل قيمة للقوة المطبقة قيمة لانتقال الرأس المتحرك وهي تمثل التغير الحاصل في طول (ارتفاع) العينة، وذلك عن طريق وجود عداد ضوئي كهربائي عالي الدقة لقياس انتقال الفك المتحرك أثناء تطبيق القوة في جهاز الاختبارات العامة، ($\pm 0.5\% \leq$ دقة قياس الانفعال).
 $L_0 [mm]$: الطول الأصلي للعينة (ارتفاع العينة).

سنحسب معامل يونغ للعينة (4): يُحسب معامل يونغ ضمن مجال المرونة للعينة وبالاعتماد على أن معامل يونغ يمثل ميل مستقيم مجال المرونة على اعتبار أن مجال المرونة للعينة شكله أقرب للخط المستقيم، أي يُحسب معامل يونغ بالاعتماد على قانون ميل المستقيم بين نقطتين وهو (الفرق بين إحداثيات النقطتين على المحور y / الفرق بين إحداثيات النقطتين على المحور x)، سنحسب معامل يونغ للعينة (4) عند القوتان (600N,800N) ضمن مجال المرونة:
أولاً: نحسب الاجهادات عند القوى السابقة:

$$\sigma = \frac{F}{A} \quad , \quad \sigma_1(600N) = \frac{600}{97.525}$$

$$= 6.152246438 [Mpa], \sigma_2(800N) = \frac{800}{97.525}$$

$$= 8.20299525 [Mpa].$$

ثانياً: نحسب الانفعالات عند القوى السابقة:

$$\varepsilon = \frac{\Delta L}{L_0} \quad , \quad \varepsilon_1(600N) = \frac{0.1150}{7.6266} = 0.01507880314, \varepsilon_2(800N)$$

$$= \frac{0.136}{7.626} = 0.01787165972.$$

من الجهاز نعوض قيم (ΔL) : $\Delta L(600N) = 0.1150 \text{ mm}$, $\Delta L(800N) = 0.1363 \text{ mm}$

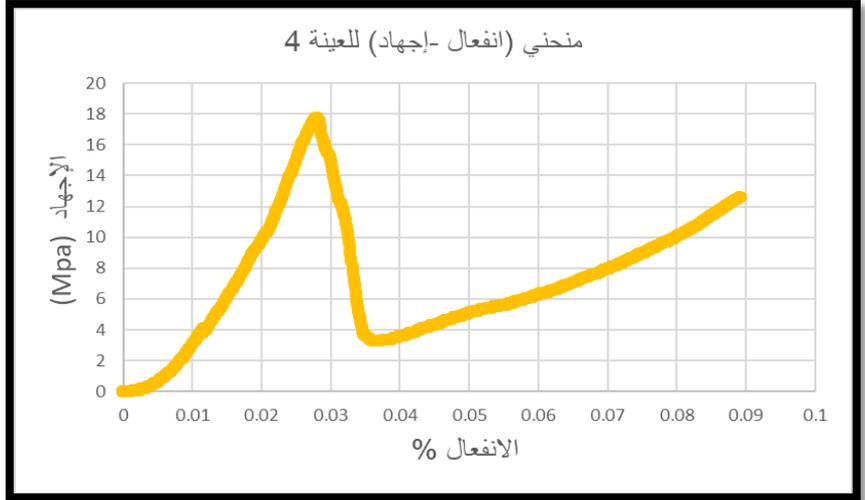
الطول الأصلي للعينة: $L_0 = 7.626 \text{ mm}$

ومنه نجد أن معامل يونغ للعينة (4):

$$E = \frac{\Delta \sigma}{\Delta \varepsilon} = \frac{\sigma_2 - \sigma_1}{\varepsilon_2 - \varepsilon_1} = \frac{8.20299525 - 6.152246438}{0.01787165972 - 0.01507880314}$$

$$= 734.283 [Mpa] = 0.734 [Gpa].$$

يبين الشكل (16) منحنى (انفعال-إجهاد) للعينة (4).



الشكل (16) منحنى (انفعال-إجهاد) للعينة (4).

بنفس الطريقة السابقة تم حساب المقاومة على الانضغاط ومعامل يونغ للعينات (2,3) ثم تم حساب المتوسط لقيم المقاومة على الانضغاط، ومعامل يونغ. حصلنا على النتائج التالية:

متوسط معامل يونغ للعينات الثلاثة [Gpa]	متوسط المقاومة على الانضغاط للعينات الثلاثة [Mpa]
0.71	14.03

نلاحظ أن قيمة المقاومة على الانضغاط (المتانة على الانضغاط)، وقيمة معامل يونغ للعينات ضمن المجال المطلوب لصفائح ترميم العظام المذكور في المرجع [9].

5-3- التوافق الحيوي للعينات:

5-3-1: إجراء المعالجة الكيميائية لزيادة النشاط الحيوي لسطح العينة قبل غمرها في السائل الحيوي (SBF):

قمنا بإجراء المعالجة الكيميائية الوارد ذكرها في المرجع [22] كأفضل معالجة كيميائية مقارنة بطرق كيميائية أخرى لزيادة النشاط الحيوي لسطح العينة قبل غمرها في السائل الحيوي (SBF):

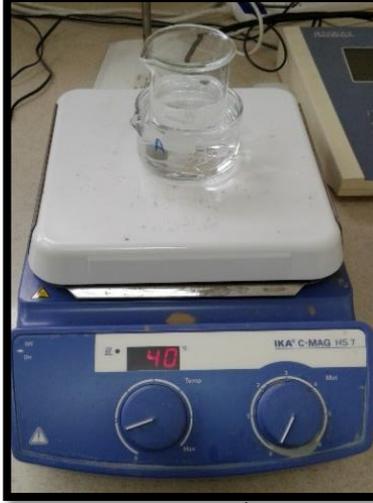
1- غسل العينة (6) من العينات في الشكل (14) التي سيتم غمرها في السائل الحيوي وغسل عينة مرجعية لن يتم غمرها في السائل الحيوي (العينة 5 من العينات في الشكل 14) بالأسيتون وماء عالي النقاوة باستخدام جهاز التنظيف بالأمواج فوق الصوتية المبيّن في الشكل (8) لمدة (30 min).

2- غمر العينة (6) في محلول (NaOH 5M) عند (60 °C) باستخدام جهاز التحريك المغناطيسي المبيّن في الشكل (9) لضبط درجة الحرارة عند (60 °C)، لمدة (24 h)، كما يوضح الشكل (17).



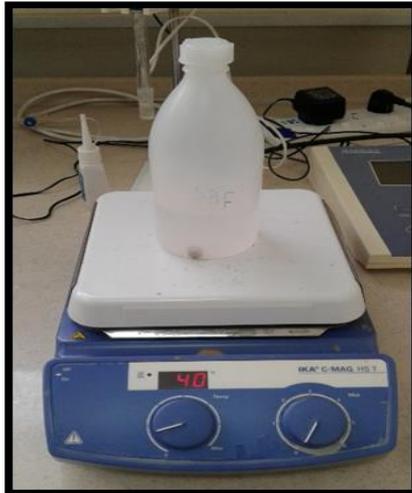
الشكل (17) غمر العينة (6) بمحلول (NaOH 5M).

3- غمر العينة (6) في محلول (Ca Cl₂ 100mM) عند درجة حرارة (40 °C) باستخدام جهاز التحريك المغناطيسي المبيّن في الشكل (9) لضبط درجة الحرارة عند (40 °C)، لمدة (43 h)، كما هو مبيّن في الشكل (18).



الشكل (18) غمر العينة بمحلول (Ca Cl₂ 100mM).

- 4- إجراء المعالجة الحرارية للعينة (6) بفرن المعالجة الحرارية المبيّن في الشكل (10) ، حيث تم تسخين العينة لدرجة حرارة (600 °C) بمعدل تسخين (5 °C/min) مع تثبيت عند درجة الحرارة (600 °C) لمدة (60 min).
- 5- غمر العينة (6) بماء عالي النقاوة عند درجة حرارة (60 °C) لمدة (24 h) ، باستخدام فرن المعالجة الحرارية المبيّن في الشكل (10).
- 6- غسل العينة (6) بشكل سريع بماء عالي النقاوة وتجفيفها عند درجة حرارة (40 °C) باستخدام فرن المعالجة الحرارية المبيّن في الشكل (10).
- 7- بعد الانتهاء من المعالجة الكيميائية السابقة تم غمر العينة (6) في السائل الحيوي (SBF) لمدة (3 أيام) عند درجة حرارة (40 °C) باستخدام جهاز التحريك المغناطيسي المبيّن في الشكل (9) لضبط درجة الحرارة عند (40 °C) ، لتكون درجة حرارة غمر العينة قريبة من درجة حرارة الجسم البشري، كما هو مبيّن في الشكل (19).



الشكل (19) غمر العينة في السائل الحيوي (SBF).

8- غسل العينة (6) بشكل سريع بماء عالي النقاوة وتجفيفها عند درجة حرارة (40 °C) باستخدام فرن المعالجة الحرارية المبين في الشكل (10) قبل إجراء الاختبار للكشف عن الفوسفات باستخدام جهاز تحديد العناصر الكيميائية بطريقة تشتت طاقة الأشعة السينية (EDX) المبين في الشكل (11).

5-3-2: طريقة تحضير السائل الحيوي (SBF):

يُحضّر السائل الحيوي المحاكي لسائل الجسم البشري {Simulated Body Fluid (SBF)} عن طريق الذوبان في الماء عالي النقاوة للكواشف التحليلية التالية المبينة في الجدول (3) كما وردت في المرجع [23].

الجدول (3): التركيب الكيميائي للسائل الحيوي (SBF):

الكمية بالغرام في (1000 mL)	الكاشف المستخدم
8.035	NaCl
0.355	NaHCO ₃
0.225	KCl
0.231	K ₂ HPO ₄ . 3H ₂ O
0.311	MgCl ₂ . H ₂ O
40 mL	1 M HCl
0.292	CaCl ₂
0.072	Na ₂ SO ₄
6.118	Tris ((CH ₂ OH) ₃ CNH ₂)
الكمية المناسبة لضبط الأس الهيدروجيني (PH 7.4)	1 M HCl

المحضّر وهي مطابقة
 قمنا بقياس قيمة الأس

الهيدروجيني باستخدام جهاز قياس الأس الهيدروجيني (PH Meter) المبين في الشكل (7):



الشكل (20) قيمة الأس الهيدروجيني للسائل الحيوي المحضّر.

يوضع السائل الحيوي بعد الانتهاء من تحضيره في وعاء مغلق من البولي إيثيلين.

3-3-5: اختبار تحديد العناصر الكيميائية بطريقة تشتت طاقة الأشعة السينية

باستخدام جهاز (EDX) المبين في الشكل (11):

قمنا بتحديد التركيب الكيميائي للعينة (5) غير المغمورة في السائل الحيوي.

- قمنا بتحديد التركيب الكيميائي للعينة (6) المغمورة في السائل الحيوي لمدة (3 أيام)،

يبين الشكل (21) مكان وضع العينات في جهاز (EDX):



الشكل (21) مكان توضع العينة في جهاز (EDX).

حصلنا على النتائج التالية للعينتان (5,6):

الجدول (4) نتائج التركيب الكيميائي للعيينة (5):

العنصر الكيميائي	النسبة المئوية
Ti	69.552%
Zr	22.836%
Cl	6.823%
Hf	0.584%
Fe	0.145%
Y	0.061%

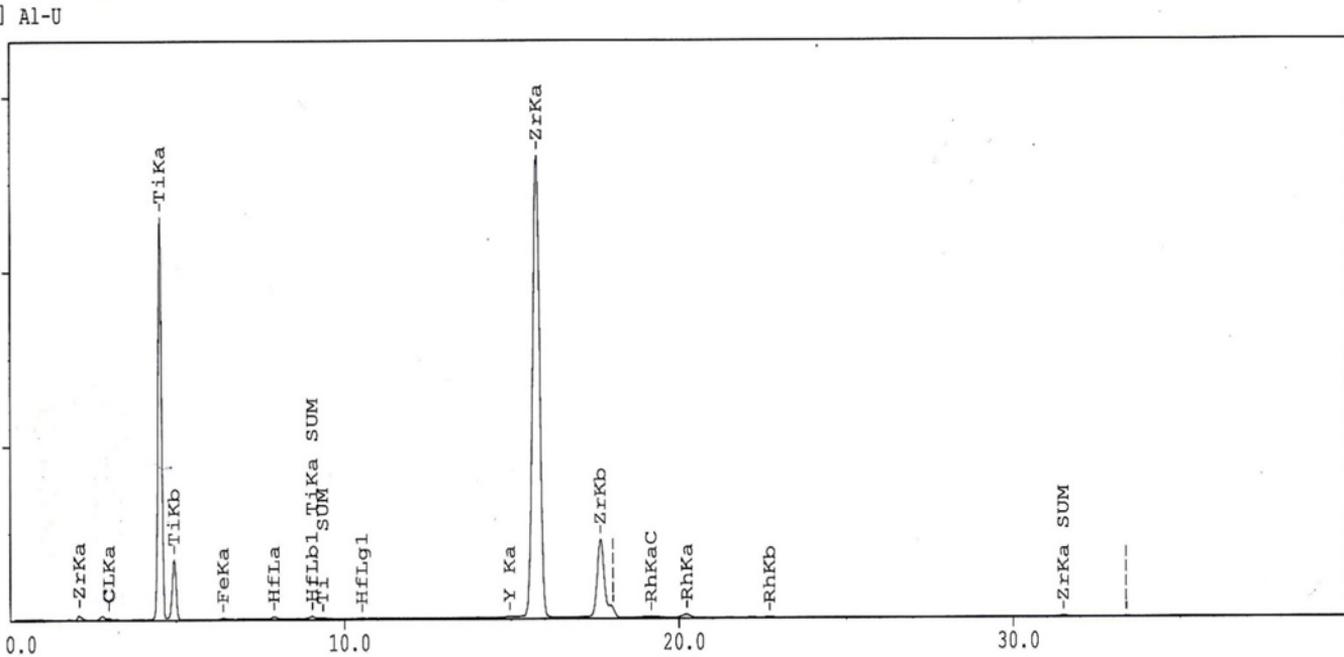
الجدول (5) نتائج التركيب الكيميائي للعيينة (6) المغمورة في السائل الحيوي:

العنصر الكيميائي	النسبة المئوية
Ti	68.803%
Zr	20.827%
P	9.129%
Hf	0.539%
S	0.203%

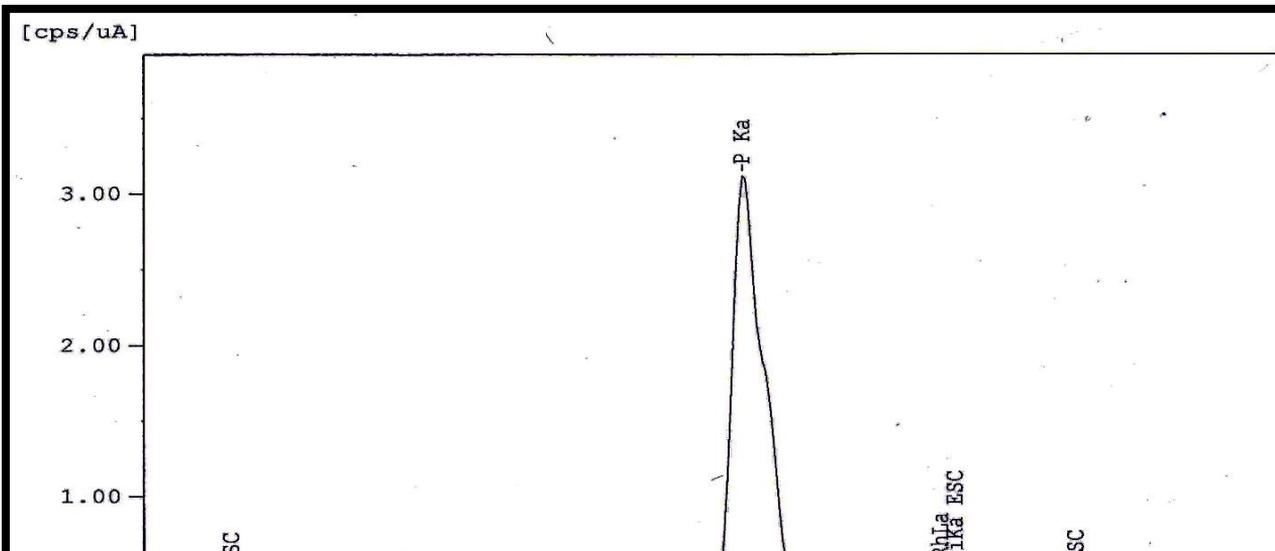
التوافق الحيوي لصفائح ترميم العظام (تيتانيوم _ زركون) المسامية

Ca	0.170%
Al	0.112%
Ag	0.085%
Fe	0.077%
Y	0.055%

تبيّن الأشكال (22)، (23) الطيف الناتج عن العينتان (5)، (6) على التوالي:



الشكل (22) طيف العينة (5) الغير مغمورة في السائل الحيوي.



الشكل (23) طيف العينة (6) المغمورة في السائل الحيوي.

نلاحظ من الشكل (22) عدم ظهور أي قيمة لعنصر الفوسفور في العينة (5) الغير مغمورة في السائل الحيوي، بينما نلاحظ من الشكل (23) ظهور قمتين لعنصر الفوسفور في العينة (6) التي تمت معالجتها كيميائياً ثم غمرها في السائل الحيوي. ظهرت القمة الأولى للفوسفور عند (0.30 Kev)، أما القمة الثانية للفوسفور ظهرت عند (2.04 Kev).

يدل وجود عنصر الفوسفور في العينة (6) المغمورة في السائل الحيوي كما لاحظنا من الجدول (5) والشكل (23) على تشكل فوسفات الكالسيوم على سطح العينة (6)، وبالتالي العينة متوافقة حيويًا وسيتحول فوسفات الكالسيوم المتكون على سطحها إلى أبائيت في بيئة الجسم كما ورد المرجع [19].

6- المناقشة:

تبيّن من التجارب السابقة أنه من الممكن الحصول على المسامية المطلوبة لصفائح ترميم العظام (< 40%) باستخدام ضغط كبس قيمته (190.98 Mpa = 1.5 Ton)، حيث تم تحديد قيمة ضغط الكبس تجريبياً كما ورد في فقرة تحضير العينات، وكانت العينات الناتجة بعد التليد عند (900 °C) لمدة (60 min) خالية من أي تشقق ومتوسط المسامية للعينات السبعة (من الجدول (2) 45.58 %) وهذه القيمة متوافقة مع المسامية المطلوبة لصفائح ترميم العظام (< 40%) كما ذكر في المرجع [6].

وبإجراء اختبار المقاومة على الانضغاط كانت القيمة الوسطية لمعامل يونغ للعينات الثلاثة المختبرة هي (0.71 Gpa)، والقيمة الوسطية للمقاومة على الانضغاط للعينات الثلاثة المختبرة هي (14.03 Mpa) وهي قيم متوافقة مع قيم الخصائص الميكانيكية للعظم الطبيعي كما ذكر في المرجع [9].

وبعد إجراء المعالجة الكيميائية التالية: (معالجة قلوية بمحلول هيدروكسيد الصوديوم_ معالجة بمحلول كلوريد الكالسيوم_ معالجة حرارية_ معالجة بالماء الساخن) الموصى بها في المرجع [22] باعتبارها أفضل معالجة كيميائية للسطح المعدني لتشكيل طبقة الأباتيت على العينة (6) وغمر العينة في السائل الحيوي (SBF) المحضّر لمدة (3 أيام) عند درجة حرارة (40 °C) أقرب درجة حرارة للجسم البشري يعطيها جهاز التحريك المغناطيسي المبيّن في الشكل (9)، تبيّن وجود عنصر الفوسفور في العينة (6) المغمورة في السائل الحيوي (SBF) بعد إجراء اختبار تحديد العناصر الكيميائية بطريقة تشتت طاقة الأشعة السينية باستخدام جهاز (EDX) المبيّن في الشكل (11). بلغت نسبة الفوسفور في العينة (6) (9.129 %) المبيّنة في الجدول (5)، ولم يلاحظ وجود عنصر الفوسفور في العينة (5) الغير مغمورة في السائل الحيوي (SBF) كما هو مبيّن في الجدول (4)، كما ظهرت قمتان للفوسفور في العينة (6) كما هو مبيّن في الشكل (23) وعدم وجود أي قمة للفوسفور في نتائج العينة (5) المبيّنة في الشكل (22). أي إن العينة (6) تتشكّل على سطحها فوسفات الكالسيوم.

وهذا يتوافق مع ما وجدته الباحثة (Marcio W.D. Mendes) وآخرون [21] حيث أكد الباحثون وجود قمم (PO_4^{-3}) لعينات (Ti-27-Nb-13Zr) تمت معالجتها معالجة قلوية فقط بمحلول هيدروكسيد الصوديوم ثم غمرها لمدة (15 , 3 يوم) في السائل الحيوي (SBF) باستخدام اختبار مطيافية الأشعة تحت الحمراء بعكسية الانتشار الذي يسمى ب:

(Diffuse Reflectance Infrared Fourier Transform spectroscopy):
(DRIFTS).

وبسبب عدم توفر جهاز يعمل بتقنية (DRIFTS) في كلية الهمة والمعهد العالي للعلوم التطبيقية والتكنولوجيا تم استخدام جهاز (EDX) المبيّن في الشكل (11) للكشف عن وجود عنصر الفوسفور في العينة المعالجة والمغمورة.

كما توافقت نتائج وجود عنصر الفوسفور في العينة (6) مع ما وجدته الباحث (C.G. Agreda) وآخرون [22] حيث بيّن الباحثون وجود قمع (PO_4^{-3}) في عينات التيتانيوم المحضّرة باستخدام تقانة ميتالورجيا المساحيق المعدنية بعد معالجتها كيميائياً كما يلي (معالجة قلووية بمحلول هيدروكسيد الصوديوم_ معالجة بمحلول كلوريد الكالسيوم_ معالجة حرارية_ معالجة بالماء الساخن) وغمر العينات لمدة يوم واحد في السائل الحيوي (SBF)، تم الكشف عن قمع (PO_4^{-3}) باستخدام تقنية (DRIFT).

7- الاستنتاجات:

1- حصلنا على صفائح ترميم عظام من خليطة:

(70%wt Ti-30%wt Zr) باستخدام تقانة ميتالورجيا المساحيق المعدنية بمسامية وخصائص ميكانيكية متوافقة مع القيم المطلوبة لصفائح الترميم باستخدام ضغط كبس (1.5 Ton) وبالتلبيد عند درجة حرارة (900 °C) لمدة (60 min) في جو من الأرجون بتدفق (168.9 Liter/hour).

2- أثبتت تجربة السائل الحيوي (SBF) أن العينة (6) المحضّرة والمعالجة كيميائياً والمغمورة في السائل الحيوي متوافقة حيوياً ولها القدرة على تشكيل الأباتيت عند زرعها في الجسم البشري.

8- التوصيات:

1- دراسة إمكانية الحصول على المسامية المطلوبة لصفائح ترميم العظام باستخدام مواد تساعد على تشكيل المسامات.

2- إجراء معالجات كيميائية مختلفة على العينات الملبّدة وتحديد أفضل معالجة كيميائية تعطي أكبر سماكة لطبقة الأباتيت الناتجة عن الغمر في السائل الحيوي (SBF) باستخدام مجهر المسح الالكتروني.

3- إجراء عملية الغمر في السائل الحيوي (SBF) بفترات زمنية مختلفة وتحديد أفضل فترة تعطي أكبر سماكة لطبقة الأباتيت الناتجة بعد الغمر في السائل الحيوي (SBF) باستخدام مجهر المسح الإلكتروني.

9- المراجع (References):

- 1-Geetha, M., Singh, A.K., Asokamani, R., Gogia, A.K. (2009, May). "Ti based biomaterials, the ultimate choice for orthopedic implants". Progress in Materials Science (Vol. 54, pp.397-425). India: Vellore Institute of Technology.
- 2-Niinomi, M., Nakai, M., Hieda, J. (2012, November). "Development of new metallic alloys for biomedical applications". Acta Biomaterialia (Vol. 8, pp. 3888-3903). Japan: Institute for Materials Research, Tohoku University.
- 3-Mehjabeen, A., song, T., Xu, W., Tang, H.P., Qian, M. (2018, June,22). "Zirconium alloys for orthopedic and dental applications". Advanced Engineering Materials (Vol. 20). Australia: Australian Research Council (ARC).
- 4- Uchida, M., Kim, H.M., Miyaji, F., Kokubo, T., Nakamura, T. (2002, January). "Apatite formation on zirconium metal treated with aqueous NaOH". Journal of Biomaterials (Vol. 23, pp. 313-317).
- 5- Kokubo, T., Kim, H.M., Kawashita, M. (2003, June). "Novel bioactive materials with different mechanical properties". Journal of Biomaterials (Vol. 24, pp.2161- 2175).

- 6-Yilmaz, E., Gokce, A., Findik, F., Gulsoy, H.O., Iyibilgin, O. (2018, November). "Mechanical properties and electrochemical behavior of porous Ti-Nb biomaterials". Journal of Mechanical Behavior of Biomedical Materials (Vol. 87, pp.59-67).
- 7-Barati, D. (2016). "Biodegradable hybrid tissue engineering scaffolds for reconstruction of large bone defects". University of South Carolina, South Carolina.
- 8-Arifvianto, B., Zhou, J. (2014)." Fabrication of metallic biomedical scaffolds with the spacer holder method". Delft University of Technology, Netherland.
- 9-Evans, G. (1969). "The mechanical properties of bone", The University of Michigan.
- 10-Bratu, V., Vibu, R. (2017). "Porous metallic biomaterials processing (Review) part1: compaction, sintering behavior, properties and medical applications". The scientific Bulletin of Valahia university materials and mechanics (Vol. 15).
- 11-Feng, B., Chen, J Y., (2002, January, 23). "Carbonate apatite coating on titanium induced rapidly by precalcification". Journal of Biomaterials.
- 12-Kokubo, T., Takadama, H. (2006, May). "How useful in SBF in predicting in vivo bone bioactivity?". Journal of Biomaterials (Vol. 27, Issue 15, pp. 2907-2915).
- 13-Bohner, M., Lemaitre, J. (2009, April). "Can bioactivity be tested in vitro with SBF solution?". Journal of Biomaterials.
- 14- Nishiguchi, S., Kato, H. (2001, September, 1)." Titanium metals form direct bonding to bone after alkali and heat treatments". Journal of Biomaterials.
- 15- Miyazaki, T., Miyaji, F., (2002). "Mechanism of bonelike apatite formation on Bioactive tantalum metal in a simulated body fluid". Journal of Biomaterials (Vol. 23, pp. 827-832).
- 16- Chen, Y., Zheng, X. (2007, December, 5). "Effect of Ti-OH formation on bioactivity of vacuum plasma sprayed titanium coating after chemical treatment". Journal of Surface and coatings Technology (Vol 202, pp 494-498).
- 17- Kokubo, T., (2005, April). "Design of bioactive bone substitutes based on bio mineralization process". Journal of Materials Science and Engineering.

- 18- Lotsari, A., Mats, H. (2018, October, 9). "Transformation of amorphous calcium phosphate to bone- like apatite". Journal of nature communications.
- 19- Kizuki, T., Takadama, H. (2010). "Preparation of bioactive Ti metal surface enriched with calcium ions by chemical treatment". Journal of Acta Biomaterialia.
- 20- Kamitakahara, M., Ohtsuki, C. (2007, November,2). "Coating of bone- like appetite for development of bioactive materials for bone reconstruction". Journal of Biomedical Materials.
- 21- W.D. Mendes, M., G. Agreda, C., (2016, June,1)." A new titanium based alloy Ti-27Nb-13Zr produced by powder metallurgy with biomimetic coating for use as a biomaterial". Journal of Materials Science and Engineering (Vol. 63, pp. 671-677).
- 22- Agreda, C.G., Mendes, M.W.D, (2013). "Apatite coating on titanium samples obtained by powder metallurgy". Journal of Advances in science and Technology (Vol. 86, pp. 28-33).
- 23- Sasikumar, Y., Ebenso, E., (2017, January). "Effect of surface treatment on the bioactivity and electrochemical behavior of magnesium alloys in simulated body fluid: Effect of surface treatment on the bioactivity". Journal of Materials and corrosion.

تصنيفُ التسلّاتِ إلى الشبكاتِ المحليّة باستخدام

الغابات العشوائية و Apache Spark

الدكتوراة زينب راتب خلوف*

الملخص

مع الاعتماد المتزايد على النظم المعلوماتية و الشبكات الحاسوبية جذب موضوع حماية البيانات و النظم اهتماماً واسعاً ، وخاصة مع تطوّر قدرات المهاجمين وزيادة حالات الاختراق. لكن من المعروف أنّ بناء نظام قادر على تجاوز جميع الهجمات غير مُمكن في وقتنا الحالي ومن هنا تأتي أهمية نُظم كشف التسلّات بأنواعها. يستندُ قسمٌ من نظم كشف التسلّل إلى تعلّم الآلة الخاضع للإشراف بحيث يُدرّب النظام على بيانات مصنّفة في فئات وبناءً على مرحلة التعلّم هذه يصنف سجلات الاتصالات القادمة لتمييز التسلّات إلى الشبكة. نحقق في هذا البحث عدّة نماذج تصنيف لاكتشاف التسلّات إلى الشبكات المحليّة باستخدام خوارزمية الغابات العشوائية والمُضمّنة في مكتبة تعلّم الآلة من أبانتشي سبارك ، كما نقيّم هذه النماذج بناءً على مجموعة معايير لتقييم الأداء.

الكلمات المفتاحية: أمن النظم و الشبكات الحاسوبية ، نُظم كشف التسلّات ، تعلّم الآلة ، الغابات العشوائية ، Apache Spark ، لغة سكال ، مجموعة البيانات NSL-KDD.

* أستاذ مساعد-قسم هندسة الشبكات و النظم الحاسوبية- كلية الهندسة المعلوماتية-جامعة البعث

Classification of Intrusions in Local Area Networks (LANs) using Random Forests and Apache Spark

Dr. Zainab Khallouf*

Data and systems security has attracted increasing attention as computer systems and networks influence every aspect of modern life.

It is known that an "Attack Proof" system doesn't exist, but it is possible to detect suspicious network activities using intrusions detection systems (IDSs). IDSs that utilize supervised machine learning techniques such as classification to learn from existing data can classify new connections to detect abnormal behavior.

In this paper, we implement and evaluate several models to detect intrusions in local area networks (LANs) using the random forest algorithm provided by the machine learning library in Apache Spark. In order to evaluate the performance of the proposed framework, several performance matrices are used.

Keywords: Computer systems and networks security, intrusions detection systems, machine learning, random forest, Apache Spark, Scala, NSL-KDD data set.

* Associate Professor, department of systems and computer networks engineering, faculty of informatics engineering, Al-Baath University, Homs, Syria.

1. مقدمة والهدف من البحث:

يُعرّف كشف التسلّات بأنه عملية مُراقبة و تحليل الأحداث (Events) في الشبكة أو ضمن النظام المعلوماتي بهدف كشف أي إشارة لحدوث اختراق للنظام. تأتي هذه التسلّات إمّا من مهاجمين ، أو من مستخدمين يحاولون الحصول على سماحيات أعلى ، أو يسيئون استخدام السماحيات الممنوحة لهم [19].

تقسّم نُظْم كشف التسلّات (IDSs) لقسمين رئيسيين [1]:

أولاً : (Anomaly Detection Systems) نُظْم كشف الشذوذ: يعتمدُ هذا النموذج على تعريف سلوك طبيعي للنظام المعلوماتي بطرق إحصائية ، و من ثمّ اعتبار أي انحراف عن هذا السلوك كدليل على وجود اختراق.

ثانياً : (Signature Based Detection Systems) نُظْم معتمدة على مجموعة من الأنماط المسبقة التعريف: يعتمدُ هذا النموذج على مجموعة من الأنماط المُسبقة التعريف للهجمات الممكنة على النظام ، بالإضافة لمجموعة من القواعد الواجب تطبيقها عند اكتشاف أحد هذه الأنماط في الرزم الملتقطة.

يُمكن استخدام التصنيف في نُظْم كشف التسلّات من النوع الأول (نظم كشف الشذوذ) التي تتعلم من مجموعة بيانات (Dataset) تتضمن اتصالات طبيعية "Normal" و اتصالات تابعة لمتسللين "Attack" ، وبناءً على مرحلة التعلم هذه يتمّ تصنيف سجلات الاتصالات القادمة لتميز التسلّات إلى الشبكة [1].

نحقّق في هذا البحث عدة نماذج تصنيف لاكتشاف التسلّات إلى الشبكات المحليّة باستخدام الغابات العشوائية و مكتبة تعلّم الآلة من أباتشي سبارك (Apache Spark) وهو نظام مفتوح المصدر لمعالجة موزعة للبيانات الكبيرة. يتميزُ سبارك بكونه متسامح مع الأخطاء ، وذو أداء مرتفع نظراً لإمكانية معالجة البيانات في الذاكرة ، بالإضافة لقابلية التوسع [2]. نقيم النموذج المحقّق اعتماداً على عدة معايير لتقييم أداء نماذج تعلّم الآلة.

يقدمُ البحث في قسمه الأول لمحة نظريّة عن التقنيات المستخدمة و يبدأ بللمحة عن تقنيات تعلّم الآلة و خوارزمية أشجار القرار التي تعدّ المُكون الأساس للغابات العشوائية ، مروراً بأباتشي سبارك ، و انتهاءً بمجموعة البيانات المُستخدمة في البحث NSL-KDD. أمّا في قسمه الثاني فيلخّص عدد من الأعمال البحثية المُتعلقة بموضوع البحث ، و في قسمه الثالث

يوضّح النماذج المحقّقة و يعرضُ تقيماً للأداء ، يليه مجموعة من التوصيات و الأعمال المستقبلية في خاتمة البحث.

2. لمحة عن تقنيات تعلم الآلة [3]:

يهدفُ تعلمُ الآلة لتعليم الحواسيب كيفية التعرف على أنماط و استنباط رؤى من البيانات و يمكن استخدام هذا العلم لحل العديد من المسائل في مجالات شتى مثل التعرف على الصور ، التعرف على الكلام ، معالجة اللغات الطبيعية ، اكتشاف الاحتيال ، النظم الناصحة ، الروبوتية و غيرها. عند استخدام تقنية لتعلم الآلة يمكن أن نميّز بين المفاهيم الآتية:

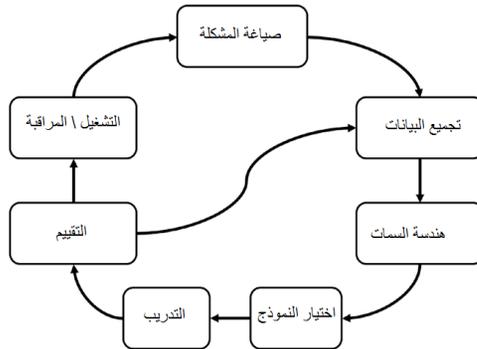
- خوارزمية تعلم الآلة: مجموعة خطوات تُنفَّذ بشكل تكراري بهدف استنباط أنماط من مجموعة بيانات لتعلم كيفية توليد تصنيف أو تنبؤ من أجل دخل معيّن. يوجد العديد من خوارزميات تعلم الآلة لكن التحدي يكمنُ في اختيار الخوارزمية الأنسب لحل مسألة معينة.
- الملاحظة (Observation): و تمثّل مثال من الكائن المستخدم للتعلم.
- التسمية (Label): قيمة تُستخدم لعنونة الملاحظة (على سبيل المثال: اتصال "طبيعي" أو اتصال "هجمة").
- السمات (Features): تمثّل مجموعة الخصائص الهامة للملاحظات و التي لها التأثير الأهم في خرج التنبؤ ؛ مثلاً معدل الإرسال من عنوان IP معيّن.
- بيانات التدريب (Training data): قسم من البيانات يستخدم لتدريب خوارزمية تعلم الآلة للحصول على النموذج.
- بيانات التحقق (Validation): قسم من البيانات يستخدم لتقييم أداء نموذج تعلم الآلة خلال عملية ضبط النموذج.
- بيانات الاختبار (Test data): تُستخدم لتقييم أداء نموذج تعلم الآلة بعد انتهاء عملية ضبط النموذج.

- النموذج (Model): ينتج عن تدريب خوارزمية تعلم الآلة و يُستخدم للتنبؤ أو لاتخاذ قرارات مُتعلقة ببيانات جديدة. الهدف هو انتاج نموذج مُعمم يُنفذ جَيداً حتى على بيانات لم يُدرب عليها.

يوجد ثلاثة أنواع رئيسية لخوارزميات تعلم الآلة:

- التعلم بالإشراف (Supervised Learning): في هذا النوع من التعلم يتم تدريب الخوارزمية على بيانات معنونة (Labeled). من أمثله: التصنيف (Classification) و هو التقنية المستخدمة في هذا البحث و الإنحدار (Regression).
- التعلم الغير خاضع للإشراف (Unsupervised Learning): يتم تدريب الخوارزمية على بيانات غير معنونة ، صُمم هذا النوع من التعلم بهدف إيجاد بنية أو أنماط مخفية في البيانات وأحد أمثله التجميع (Clustering).
- التعلم المعزز (Reinforcement Learning): في هذا النوع من التعلم لاتتعلم الخوارزمية من البيانات و إنما من التفاعل مع البيئة من خلال سلسلة من الأفعال (Actions) ومن خلال معلومات التغذية الراجعة تقوم بتعديلات (Adjustments) بهدف تعظيم قيمة ربح ما.

تتكون عملية تعلم الآلة من الخطوات الآتية الموضحة في الشكل 1:



الشكل 1: خطوات تعلم الآلة (مترجم عن [17])

- 1) فهم و صياغة المسألة التي ننتظر من تقنية تعلّم الآلة حلها و من الجيد في هذه المرحلة دراسة امكانية استخدام حلول بديلة لتقييم الكلفة بالإضافة للميزات و للسليبيات. عند بناء ثقة قوية بأن تعلّم الآلة هو الخيار الأفضل للوصول للهدف عندئذ ننتقل إلى الخطوة التالّية.
- 2) جمع البيانات: حيث أنّ نوعيّة و كمّيّة البيانات المجمعة سيكون لها تأثير مباشر على أداء النموذج المدرب و من المهم أن تكون البيانات ممثّلة قدر الإمكان للمشكلة التي نحاول حلها.
- 3) هندسة السمات (Feature engineering): وهي واحدة من أكثر المراحل أهميّة و تطلباً للوقت. في هذه المرحلة تتمّ تنقيّة البيانات (Data cleaning) و الاستنارة بمعرفة المجال (Domain knowledge) لتحديد خصائص أساسيّة أو سمات مهمّة لتعلّم الآلة. كذلك يمكن أن يتمّ تحويل قيم السمات المحرفية لأدلة (String Indexing) وأيضاً ترميز قيم السمات الفئوية العديدة باستخدام رموزات مثل ترميز واحد حار (One hot encoding) لإلغاء أي علاقة ترتيب قد تعتبرها خوارزمية تعلّم الآلة موجودة بينها و في هذا البحث سنستخدم هذه التقنيات.
- 4) اختيار خوارزمية تعلم الآلة وتدريبها مما يتطلب الفهم الجيد للمشكلة المطلوب حلها ، مع معرفة عمليّة بالخصائص المختلفة لكل خوارزمية وامتلاك خبرة تطبيق هذه الخوارزميات على سيناريوهات مماثلة في السابق.
- 5) تقييم النموذج ولاتهدف هذه المرحلة إلى تحديد أداء النموذج فحسب و إنما أيضاً لتحديد متى يتوقف ضبط النموذج (Model tuning) عندما يصل أدائه إلى جودة مقيّمه باستخدام معايير محدّدة. يُمكن أن تؤدي هذه المرحلة إمّا لنموذج قابل للاستخدام أو للعودة إلى تجميع البيانات و المرور على المراحل من جديد عندما يكون أداء النموذج المولّد غير كاف.

3. أهم المقاييس لتقييم أداء نماذج التصنيف [17]:

يوجدُ عدد من مقاييس تقييم أداء نماذج التصنيف التي تهدفُ لضمان تقييم "غير متحيز" لأداء النموذج و لعل أكثرها شهرة المقاييس الآتية:

- مصفوفة الارتباك (Confusion matrix) المُوضحة في الجدول 1 ، و تعدّ أساساً لحساب قيمة عدد من المقاييس المعبّرة عن أداء نموذج التصنيف.

		التنبؤ (Predicted)	
		طبيعي (Normal)	هجمة (Anomalous)
الواقع (Actual)	طبيعي (Normal)	TN	FP
	هجمة (Anomalous)	FN	TP

الجدول 1: مصفوفة الارتباك

في الجدول 1 نميّز بين المفاهيم التالية:

- إيجابية حقيقية (True Positive (TP)): عدد الملاحظات التي تتمثل هجمة و التي صُنفت بشكل صحيح على أنها هجمة.
- سلبية حقيقية (True Negative (TN)): و تتمثل عدد الملاحظات الطبيعية (أي التي لاتتمثل هجمة) و التي صُنفت بشكل صحيح على أنها لاتتمثل هجمة.
- إيجابية خاطئة (False Positive (FP)): عدد الملاحظات الطبيعية لكنها صُنفت بشكل خاطيء على أنها هجمة (يسمى أيضاً خطأ من النوع الأول).
- سلبية خاطئة (False Negative (FN)): و تتمثل عدد الملاحظات التي تتمثل هجمة و صُنفت بشكل خاطيء على أنها طبيعية (يسمى أيضاً خطأ من النوع الثاني).

تصنيفُ التسلّات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

- الصّحة (Accuracy)؛ النسبة بين عدد الملاحظات التي تمّ تصنيفها بشكل صحيح و العدد الكلي للملاحظات أي:

$$Accuracy = \left(\frac{TP + TN}{TP + TN + FP + FN} \right)$$

تعدّ الصحة مقياساً جيداً في الفئات المتوازنة (أي أنّ أعداد الملاحظات في كل فئة متقاربة) وعندما تكون كلفة الخطأ الأول و الثاني متقاربة و إلا يمكن أن لا يكون هذا المقياس معبراً ، فلو فرضاً تضمنت مجموعة بيانات 1000 ملاحظة من فئة 1 و 10 ملاحظات من فئة 2 عندئذ صحة مصنف يصنّف جميع الملاحظات على أنها من الفئة 1 تعادل:

$$accuracy = (1000 + 0)/1010 = 99\%$$

من الواضح أنّ هذه القيمة لاتعكس أداء جيداً للمصنف في هذه الحالة وخصوصاً في الحالات الحرجة مثل اكتشاف وجود تسلل.

- الإحكام (Precision)؛ النسبة بين عدد التنبؤات الإيجابية الصحيحة ومجموع عدد التنبؤات الإيجابية الخاطئة و الإيجابية الصحيحة؛ أي مثلاً: من بين كل الاتصالات التي تنبأت بها الخوارزمية بأنها تمثل هجمة ، كم منها تمثل هجمة فعلياً؟ كلما كان عدد التنبؤات الإيجابية الخاطئة أقل كلما كانت قيمة الإحكام مرتفعة.

$$Precision = \frac{TP}{TP + FP}$$

- (الحساسية/الاستدعاء/معدل التنبؤات الإيجابية الصحيحة)(TPR/Recall/Sensitivity)؛ النسبة بين عدد التنبؤات الإيجابية الصحيحة ومجموع عدد التنبؤات الإيجابية الصحيحة و عدد التنبؤات السلبية الخاطئة. السؤال الذي تُجيب عنه الحساسية هو: من بين كل الاتصالات التي تمثل هجمة فعلياً ، كم عدد الاتصالات التي تنبأت بها الخوارزمية على أنها هجمة؟

$$Recall = \frac{TP}{P} = \frac{TP}{TP + FN}$$

- (النوعية/معدل التنبؤات السلبية الصحيحة)(TNR/Specificity)؛ النسبة بين عدد التنبؤات السلبية الصحيحة ومجموع عدد التنبؤات السلبية الصحيحة و عدد التنبؤات الإيجابية الخاطئة. السؤال الذي تُجيب عنه النوعية هو: من بين كل الاتصالات الطبيعية (أي ليست هجمة) ، كم عدد الاتصالات التي تنبأت بها الخوارزمية على أنها طبيعية؟

$$Specificity = \frac{TN}{N} = \frac{TN}{TN + FP}$$

- معدّل التصنيف الإيجابي الخطأ أو معدّل الإنذارات الخاطئة (FPR) ويمثّل النسبة بين عدد الملاحظات الطبيعية التي صنّفت بشكل خاطئ ، والعدد الكلي للملاحظات الطبيعية. يُعبّر عنه بالعلاقة:

$$FPR = \frac{FP}{FP + TN}$$

- النتيجة F1 (F1-score)؛ المتوسط الحسابي الموزون لقيمة الإحكام والحساسية.

$$F1 - score = 2 \frac{Recall * Precision}{Recall + Precision}$$

- المنطقة تحت منحنى خصائص تشغيل المستقبل (AUC-ROC) ؛ تقيس مدى قدرة النموذج على الفصل بين الفئات و كلما كانت قيمة AUC أكبر كلما كان النموذج أفضل. أما المنحنى ROC فيعبّر عن تغيّر العلاقة بين TPR و FPR.
- المنطقة تحت منحنى الاستدعاء و الإحكام (AUC-PR) ؛ بما أنّ كلّ من الإحكام و الاستدعاء لايتضمنان في حسابهما (عدد التنبؤات السلبية الحقيقية TN) و لذلك في الحالات التي يكون فيها عدد الملاحظات السلبية أكبر بكثير من الملاحظات الإيجابية (سمة أساسية للفئات الغير متوازنة العدد) يتم استخدام هذا المقياس [4].

تصنيف التسلسلات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

تقدّم مكتبة تعلم الآلة في سبارك عدد مهم من هذه المقاييس و نميز بين مقاييس تُستخدم للتصنيف الثنائي (Binary classification) و للتصنيف المتعدد الفئات (Multiclass classification) الذي يعدّ تعميماً للتصنيف مُتعدد الفئات و يصنّف الملاحظات بين عدة فئات ممكنة.

4. ضبط نموذج التصنيف (Classification Model Tuning):

يهدف ضبط النموذج إلى تدريب النموذج باستخدام المجموعة الصحيحة من الوسائط لتحقيق أفضل أداء في تحقيق الهدف المطلوب و الذي حُد في الخطوة الأولى من مراحل تعلم الآلة (الشكل 1). غالباً ما تكون هذه المرحلة تكرارية تتطلب جهداً و وقتاً و ربما تتضمن تدريب عدة خوارزميات تعلم آلة أو عدة مجموعات من الوسائط. في هذا السياق نميز بين المصطلحين الآتيين:

- الوسائط الفائقة للنموذج (Model hyperparameters): و تمثل الوسائط التي تُحدّد قبل بدء مرحلة التدريب من قبل محلّل البيانات وتستخدم للتحكم بعملية تدريب خوارزمية تعلم الآلة. تعدّ هذه الوسائط خارجية عن النموذج و لا يمكن تعلمها من بيانات التدريب كما يمكن ضبطها من خلال عملية تكرارية. كمثال على هذا النوع من الوسائط عدد الأشجار في الغابة العشوائية.
- وسائط النموذج (Model parameters): و تمثل الوسائط التي لا تُضبط من قبل محلّل البيانات و إنّما يتمّ تعلمها و أمثلتها خلال عملية التدريب ، مثلا السمة التي ستشكل الجذر في شجرة القرار.

تقدّم مكتبة تعلم الآلة (ML) في سبارك صفتين لضبط النموذج: الأول TrainValidationSplit و الثاني CrossValidator و كلا الصفتين يتطلبان العناصر التالية كدخل:

- المكون المراد ضبطه: نموذج ، خوارزمية ، أو تسلسل عمليات (Pipeline).
- مجموعة الوسائط التي ستستخدم في عملية الضبط ، و تسمى شبكة الوسائط (grid Parameter).

- أحد مقاييس التقييم التي تستخدم لتقييم أداء النموذج اعتماداً على مجموعة بيانات اختبار (مثلاً أحد المقاييس التي وضحت سابقاً: الصحة ، النوعية ، الاستدعاء...).

يقسم الصف الأول TrainValidationSplit مجموعة البيانات إلى مجموعتين: مجموعة تدريب (Training dataset) ومجموعة للتحقق من الصحة (Validation Dataset) و بعد ذلك يتم تدريب و تقييم مجموعتي البيانات على كل توافق (Combination) ممكن من الوسطاء التي عرفت في شبكة الوسطاء. على سبيل المثال ، لو تضمنت شبكة الوسطاء ست توافقيات ممكنة من الوسطاء عندئذ يتم تدريب وتقييم النموذج ست مرات من أجل كل مجموعة ممكنة. أما الصف الثاني CrossValidator فهو تحقيق لتقنية التحقق المتقاطع. في هذه الطريقة يتم تقسيم الملاحظات إلى k مجموعة أو مايسمى طية (Fold) لها تقريباً نفس الحجم ؛ الطية الأولى تستخدم لعملية التحقق من الصحة و بقية الطيات للتدريب و يتم تدريب النموذج k مرة مع تدوير مجموعات التدريب و التحقق كما يوضح الشكل 2.

Experiment 1					<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; width: 20px; height: 10px; background-color: white;"></div> Training </div> <div style="border: 1px solid black; width: 20px; height: 10px; background-color: #cccccc;"></div> Validation
Experiment 2					
Experiment 3					
Experiment 4					

الشكل 2: التحقق من الصحة المتقاطع مع $k=4$ [3]

5. أشجار القرار (Decision Trees) و الغابات العشوائية (Random Forests) [5]:

تعد عائلة شجرة القرار من خوارزميات تعلم الآلة ذات الاستخدام الواسع في تطبيقات التصنيف (Classification) كما يمكن أن تستخدم للانحدار (Regression). يمكن أن يتم بناء الشجرة أو عدد من الأشجار على التوازي و تقبل الخوارزمية السمات العددية (Numeric) و كذلك الفئوية (Categorical) كما أنها تقبل بيانات من أنماط مختلفة ومن مقاييس مختلفة دون أن يكون هناك حاجة لمعالجة مسبقة لهذه البيانات أو لتقييسها (Normalization). تتميز بأنها متينة أمام الشذوذات (Outliers) في البيانات أي أنّ وجود عدد قليل من نقاط البيانات ذات القيم القصوى وعلى الأغلب الغير صحيحة لا يؤثر على الخوارزمية [6].

تصنيفُ التسَلَّاتِ إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

تعتمدُ الخوارزمية على قيم السمات (Features) لتقسيم فضاء القرار (Decision space) وبشكل تكراري إلى فضاءات جزئية (Subspaces) أصغر بدءاً من اختيار العقدة الجذر ، ولتنمو تدريجياً إلى التصنيف النهائي الممثل بالأوراق.

من حيث المبدأ يوجد عدّة طرق لإنشاء الشجرة وجميعها تؤدي إلى نفس النتائج النهائية للتصنيف ، لكن يُمكن مع ازدياد عدد الخصائص و مجال القيم العددية أو عدد القيم الفئوية أن تصبح غاية في التعقيد و أن لا تتبأ الأشجار بنفس النتيجة في هذه الحالة ، وخصوصاً عندما تتضمن قيم بعض الخصائص ضجيج أو قيم مفقودة [7].

لاختيار السمة التي سيتم تقسيم الملاحظات وفقها عند كل عقدة بدءاً من العقدة الجذر يوجد معايير مثل معيار ربح المعلومات (Information Gain) المُعتمد في الخوارزمية الأساسية لبناء أشجار القرار: ID3. يتم تقسيم البيانات في خوارزمية ID3 حسب السمات التي تجعل قيمة ربح المعلومات أعظمية ، مع تكرار هذه العملية عند كل عقدة حتى نصل لأوراق نقيّة (Pure) تكون الملاحظات عندها من نفس الفئة. يُمكن أن تؤدي هذه العملية التكرارية إلى شجرة ذات عدد كبير من العقد و بالتالي إلى نموذج ذو تكيف زائد (Overfitting) مع بيانات التدريب ؛ أي أنّ النموذج يعكس مجموعة بيانات التدريب لكن لا يمكنه التنبؤ بفئة بيانات لم يدرّب عليها و لذلك نقوم بعملية تقليم (Prune) من خلال وضع حد لعمق الشجرة. يُعرّف ربح المعلومات بالعلاقة:

$$IG(D_p, f) = I(D_p) - \sum_{j=1}^m \frac{N_j}{N_p} I(D_j)$$

حيث أنّ: f تمثل السمة التي سيتم التقسيم عندها، D_p و D_j تمثّلان مجموعتي البيانات للأب و للعقدة البنت j ، I مقياس عدم النقاوة أو عدم تجانس البيانات (Impurity) ، N_p العدد الكلي للعينات عند عقدة الأب و N_j عدد العينات عند العقدة البنت j . أي أن ربح المعلومات يمثّل الفرق بين عدم النقاوة عند العقدة الأب ومجموع عدم النقاوة عند العقد الابناء ، و كلما كان عدم النقاوة عند الابناء أقل كلما كانت المعلومات المكتسبة أكبر . لكن للتبسيط و لتقليل مجال البحث فإن معظم المكتبات تحقّق أشجار قرار ثنائية أي أنّ كل عقدة أب تقسم إلى

عقدتين أبناء D_{left} و D_{right} .

يوجد ثلاثة معايير شائعة لعدم النقاوة:

1. الانتروبي (Entropy (I_H)) و تُعطى بالعلاقة الآتية من أجل جميع الفئات الغير خالية:

$$I_H(t) = - \sum_{i=1}^c p(i|t) \log_2 p(i|t), p(i|t) \neq 0$$

حيث أن $p(i|t)$ تمثل نسبة العينات المنتمية للفئة c من أجل عقدة محددة t ، وبالتالي الانتروبي تساوي صفر إن كانت جميع العينات عند عقدة منتمية لنفس الفئة، في حين تأخذ قيمة أعظمية عندما يكون توزع العينات منتظم (Uniform class distribution).

2. معامل جيني (Gini impurity (I_G)) و يُعطى بالعلاقة:

$$I_G(t) = \sum_{i=1}^c p(i|t)(1 - p(i|t)) = 1 - \sum_{i=1}^c p(i|t)^2$$

يُنظر لمعامل جيني كمعيار لتقليل احتمال خطأ التصنيف، كما تكون قيمة معامل جيني أعظمية عندما تكون العينات من فئات مختلفة وفقاً لتوزيع منتظم. مثلاً في حالة فئتين ($c = 2$):

$$I_G(t) = 1 - \sum_{i=1}^2 0.5^2 = 0.5$$

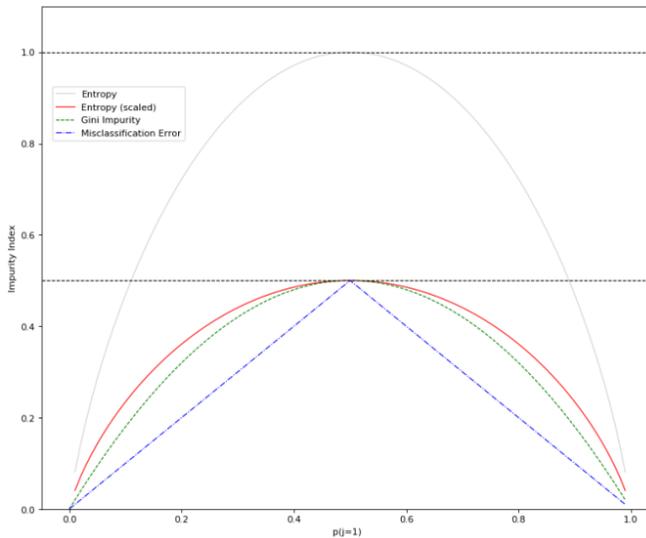
عملياً كل من المعيارين جيني و الانتروبي يعطيان نتائج متشابهة.

3. خطأ التصنيف (Error classification (I_E)):

$$I_E = 1 - \max\{p(i|t)\}$$

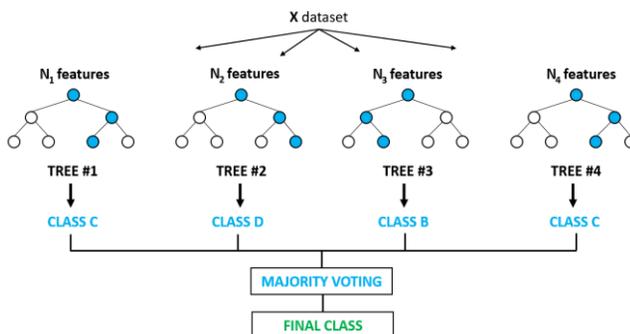
ويُنصح بهذا المعيار في حالة تقليم الشجرة لكن ليس لبناء الشجرة لأنه أقل حساسية لتغير احتمالات الفئات.

يوضح الشكل 3 هذه المعايير من أجل مجال احتمال $[0, 1]$ للفئة $c = 1$:



الشكل 3: المعايير الشائعة لعدم النقاوة [5]

تعدّ الغابات العشوائية تقنية لتعلم الآلة بالإشراف (Supervised Learning) ، من نوع تجميع (Ensemble) لمجموعة من أشجار القرار بحيث تعطي النتيجة من خلال تصويت بالأغلبية (Majority vote) على نتائج الأشجار المكونة للغابة كما يوضح الشكل 4. يمكن أن تُستخدم للتصنيف (Classification) أو للانحدار (Regression) و تهدف لتشكيل نماذج ذات أداء أفضل و أقل عُرضة للتكيف الزائد من شجرة القرار .



الشكل 4: التصويت في الغابة العشوائية [18]

يُمكن تلخيص خوارزمية الغابات العشوائية بالخطوات الأربعة التالية:

1. اختيار مجموعة تمهيدية (Bootstrap) عشوائية من مجموعة التدريب (Training set) مكونة من n ملاحظة بحيث يتم الاختيار مع استبدال (With replacement).
2. بناء شجرة قرار من المجموعة التمهيدية ، وعند كل عقدة:
 - (a) اختيار d سمة بشكل عشوائي بدون استبدال.
 - (b) تقسيم الملاحظات عند العقدة باستخدام السمة التي تعطي أفضل قيمة لربح المعلومات (أو لتابع هدف مختلف Objective function).
3. تكرار الخطوتين الأولى و الثانية k مرة.
4. تجميع نتائج التنبؤ من كل شجرة لتحديد فئة من خلال تصويت أغلبية.

تكمُن الميزة الأساسية للغابات العشوائية في أنها لا تتطلب جهود كبيرة لاختيار الوسائط الفائقة (Hyperparameters) والتي تُحدّد قبل مرحلة التعلّم ، كما لا تتطلب تقليم لأن النماذج المعتمدة على تجميع نماذج بسيطة تكون مُقاومة لضجيج النماذج الجزئية. عادة الوسيط الوحيد الذي يفترض تحديده هو عدد الأشجار k (الخطوة 3) ، وكلما ازداد عدد الأشجار تحسّن أداء الغابة على حساب زيادة الكلفة الحسابية.

من الوسائط الفائقة التي يمكن تحديدها أيضاً عدد العناصر n في المجموعة التمهيدية (الخطوة الأولى) ، وكذلك عدد السمات d التي يتم اختيارها بشكل عشوائي عند كل تقسيم (الخطوتين 1، 2). يمكن من خلال حجم مجموعة البيانات التمهيدية التحكم بالموازنة بين الانحياز و التشتت (Bias-variance tradeoff) في الغابات العشوائية ، لأنّ إنقاص حجم مجموعة البيانات التمهيدية يزيد التنوع (Diversity) بين أشجار القرار المكونة للغابة نظراً لأنّ احتمال أن تكون ملاحظة معينة مضمّنة في المجموعة التمهيدية أقل ، وبالتالي تقلص حجم المجموعة الابتدائية يمكن أن يحسّن العشوائية في الغابة العشوائية و يقلّل تأثير التكيف الزائد لكنه يعطي أداء كلي أقل ، كما أنّ زيادة حجم العينة التمهيدية يُمكن أن يزيد التكيف الزائد نظراً لأنّ الأشجار الجزئية تصبح أكثر تشابهاً و تتعلم من بيانات التدريب الأصلية بشكل مشابه. في معظم التحقيقات يتم اختيار عدد العناصر في المجموعة التمهيدية مساوياً لعدد الملاحظات في مجموعة التدريب الأصلية ، أمّا بالنسبة لعدد السمات d عند كل تقسيم يتم اختيار عدد أقل من عدد السمات

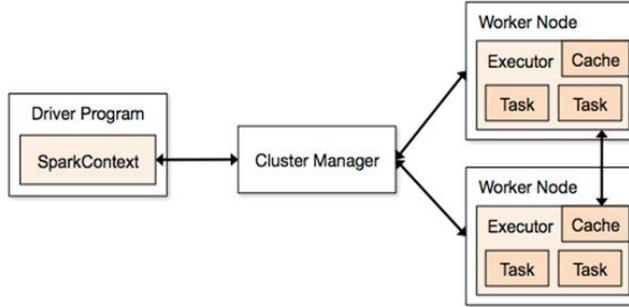
تصنيفُ التسلّلات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

الكلية في مجموعة بيانات التدريب و عادة بشكل افتراضي يتم اختيار $d = \sqrt{m}$ ، حيث أن m تمثل عدد السمات في مجموعة التدريب.

6. أباتشي سبارك (Apache Spark) [3]:

نظام مفتوح المصدر لمعالجة موزعة للبيانات الكبيرة يتميز بكونه متسامح مع الأخطاء ، و ذو أداء مرتفع نظراً لإمكانية معالجة البيانات في الذاكرة. حُقّق سبارك بلغة سكالالا لكن يمكن كتابة تطبيقاته بأي من R ، بايثون ، جافا أو سكالالا.

يتم عادة تشغيل سبارك على مجموعة من الأجهزة والتي تُعرف باسم عنقود سبارك (Spark cluster) ، علماً أن أكبر عنقود تمّ التصريح عنه يضمّ أكثر من 8000 جهاز. يتولى إدارة مجموعة الأجهزة المكوّنة للعنقود بشكل فعال و ذكي نظام لإدارة الموارد مثل Apache YARN أو Apache Mesos مكونّ عادةً من مدير عنقود (Cluster manager) و عمال (Workers) كما يوضّح الشكل 5 . يعرفُ المدير مكان توضع العامل ، كمية الذاكرة التي يمتلكها ، وعدد نوى الـ CPU ، كما أنّ واحدة من المهام الأساسية للمدير تنظيم العمل من خلال توزيعه على كل عامل.

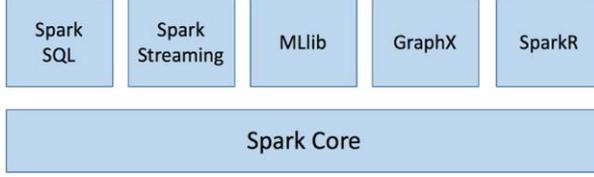


الشكل 5: التفاعل بين تطبيق سبارك و مدير العنقود [3]

تعدّ سواقة سبارك (Spark driver) المنسق المركزي لتطبيق سبارك و تتفاعل مع مدير العنقود لمعرفة الأجهزة التي سيتم تنفيذ منطق معالجة البيانات عليها. على كلٍ من هذه الأجهزة تطلب سواقة سبارك من مدير العنقود تشغيل عملية تدعى المنقذ (Executor) ، كما تجمع نتائج التنفيذ إن لزم من المنقذات وتعرض النتيجة للمستخدم. تتمثل نقطة البداية في تنفيذ تطبيق سبارك في الصف SparkSession و الذي يُمكن أن يُستخدم لتحديد مجموعة من التشكيلات (Configurations) و واجهات برمجة التطبيقات (APIs) اللازمة للتطبيق.

بُني سبارك بشكل مكس كما يوضّح الشكل 6 ، وفي طبقتة الأولى المُسمّاة نواة سبارك (Spark Core) تمّ تضمين الوظائف الأساسية اللازمة لإدارة و تنفيذ التطبيقات مثل: الجدولة ،

التنسيق ، و التسامح مع الأخطاء بالإضافة لبنى معطيات تُسمى مجموعات البيانات الموزعة و المتينة (*Resilient distributed datasets (RDDs)*) كتجريد برمجي لتمثيل ومعالجة البيانات. تتوضع فوق النواة مجموعة مكونات خُصص كلٌ منها لوظيفة محددة مُعلقة بمعالجة البيانات.



الشكل 6: مكدس سبارك [3]

حالياً يقدم سبارك مكتبتين لتعلم الآلة: Spark MLlib و Spark ML لهما واجهات APIs مختلفة لكن خوارزميات متشابهة. المكتبة الأولى و الأقدم Spark MLlib دخلت حالياً مرحلة الصيانة و التصحيحات (Bug-fix) ، أما المكتبة الثانية Spark ML و المُستخدمة في هذا البحث فهي الأحدث و مستوحاة من مكتبة تعلم الآلة في بايثون scikit-learn ، كما أنها تقدم وظائف لإدارة و تبسيط العديد من مهام بناء نماذج تعلم الآلة مثل هندسة السمات (Features Engineering)، سلسلة العمليات (Pipeline) لبناء النماذج و تقييمها و ضبط أدائها ، بالإضافة لتخزين النماذج بحيث تستخدم في التطبيقات و النظم بشكل فعلي.

7. توصيف مجموعة البيانات NSL-KDD [8]:

كانت KDD Cup مسابقة سنوية تُنظّم من قبل مجموعة مختصة من مؤسسة ACM ، و في عام 1999 خصصت المسابقة لمسألة كشف التسلات في الشبكات باستخدام تعلم الآلة. مازالت مجموعة البيانات المتعلقة KDD'99 متاحة حتى الآن ، كما تُعتبر إحدى أشهر مجموعات البيانات المستخدمة لتقييم أداء نظم كشف الاختراقات.

لتكوين مجموعة بيانات التدريب (Training data) في KDD'99 تمّ تجميع حوالي 4 gigabytes من البيانات الشبكية خلال سبع أسابيع من شبكة محلية صُممت لتحاكي شبكة واقعية ، و تمّت معالجة تلك البيانات في خمسة ملايين سجل ، كما تمّ تجميع بيانات الاختبار (Test data) خلال أسبوعين إضافيين لتشكّل حوالي 2 مليون سجل. يختلف التوزيع الاحتمالي لمجموعة بيانات الاختبار عن التوزيع الاحتمالي لبيانات التدريب ، كما تتضمن بيانات الاختبار سجلات غير موجودة ضمن بيانات التدريب لجعل عملية التصنيف أكثر واقعية.

تصنيفُ التسلّات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

عُرّف الاتصال و المكون من 100 بايت على أنه سلسلة من الرزم التي تبدأ و تنتهي في نقاط زمنية محددة ، من عنوان IP مصدر إلى عنوان هدف و تابعة لبروتوكول محدد و تمّت عنونته على أنه اتصال "طبيعي" ، أو "هجمة" مع تحديد نوع الهجمة. تمّ تصنيف أربع فئات (Categories) أساسية للهجمات:

- منع الخدمة (Denial of Service (DoS)): مثل (SYN flood).
- بعيد لمحلي (Remote to Local (R2L)): محاولة الوصول من جهاز بعيد إلى الجهاز المحلي ، مثل (محاولة تخمين كلمة السر).
- مستخدم لجذر (User to Root): وصول غير مسموح به لميزات (الجذر) ، مثل "Buffer Overflow".
- الجس (Probing): مثل مسح المنافذ (Port scanning).

عرّف (Stolfo et al.) [9] مجموعة من السمات التي تسمح بتمييز الاتصالات

الطبيعية من الهجمات كالآتي:

- السمات المرتبطة بنفس المضيف "Same host" و تفحص فقط الاتصالات التي لها نفس الهدف مثل الاتصال الحالي في الثانيتين السابقتين.
- السمات المتعلقة بنفس الخدمة "Same service" و تفحص فقط الاتصالات في الثانيتين السابقتين المرسله لنفس الخدمة مثل الاتصال الحالي.

كل من سمات "نفس المضيف" و "نفس الخدمة" تدعى السمات المعتمدة على الوقت (Time-based).

- بعض هجمات الجس (Some probing) تسمح الأجهزة (أو المنافذ) باستخدام فترات زمنية أكبر من ثانيتين ، مثلاً مرّة كل دقيقة. لذلك تمّ ترتيب سجلات الاتصال وفقاً للمضيف الهدف ، وتمّ بناء السمات باستخدام نافذة من مئة اتصال إلى نفس المضيف بدلاً عن النافذة الزمنية لتنتج مجموعة من السمات المعتمدة على المضيف (Host-based).

- بعض الهجمات مثل R2L و U2R تكون متضمنة في قسم البيانات من الرزمة و تتطلب اتصال وحيد ، لذلك تمت إضافة سمات تُعنى بالسلوك المشكوك به في أقسام البيانات مثل عدد محاولات التسجيل الفاشلة وسميت هذه السمات: سمات المحتوى (Content features).

وبناء عليه فإن كل سجل يتكون من 41 سمة (Features) و تسمية (Label).

في هذا البحث استخدمنا مجموعة البيانات NSL-KDD [8] والتي تحل بعض مشاكل الكامنة في مجموعة البيانات KDD'99 [10] ، وبالرغم من أن مجموعة البيانات هذه مازالت تتضمن بعض المحدوديات ومن الممكن أن لا تشكل تمثيل مثالي لشبكات فعلية إلا أنها مازالت مُستخدمة بشكل كبير من قبل الباحثين كمجموعة بيانات فعّالة لتقييم ومقارنة أداء نظم كشف الاختراقات.

يوضح السطر الآتي مثال عن سجل من مجموعة البيانات NSL-KDD:

```
0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,0.00,0.00,0.00,1.00,0.00,0.00,150,25,0.17,0.03,0.17,0.00,0.00,0.00,0.05,0.00,normal,20
```

من ضمن الـ 41 سمة في السجل تتوزع السمات كالتالي: تستخدم السمات 1-9 لتمثيل السمات الأساسية للرمز مثل مدة الاتصال ، نوع البروتوكول والخدمة ، 10-22 لسمات المحتوى، -23 31 للسمات المعتمدة على نافذة زمنية من ثانييتين ، و 32-41 للسمات المعتمدة على المضيف ، كما توضح الجداول 2 ، 3 ، 4 ، 5 الآتية:

رقم السمة	اسم السمة	توصيف	مثال
1	Duration	المدة الزمنية للاتصال	0
2	Protocol_type	نوع البروتوكول	TCP
3	Service	الخدمة	FTP
4	Flag	حالة الاتصال: صحيح أم خطأ	SF
5	Src_bytes	عدد البايتات المنقولة من المصدر إلى الهدف ضمن اتصال	491
6	Dst_bytes	عدد البايتات المنقولة من المصدر ضمن اتصال	0
7	Land	1 إذا كان الاتصال من نفس/إلى نفس المضيف/المنفذ، و 0 وإلا تأخذ	0
8	Wrong_fragment	العدد الكلي للأجزاء الخاطئة في هذا الاتصال	0
9	Urgent	العدد الكلي للرمز المستعجلة التي تحمل البت urgent مفعلاً	0

الجدول 2: السمات الأساسية للاتصال

تصنيفُ التسلّات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

رقم السمة	اسم السمة	توصيف	مثال
10	Hot	عدد مرات الوصول لمجلدات النظام و انشاء و تنفيذ البرامج	0
11	Num_failed_logins	عدد مرّات فشل تسجيل الدخول	0
12	Logged_in	حالة تسجيل الدخول: 1 نجاح و 0 إفاشل	0
13	Num_compromised	عدد الشرط "Compromised"	0
14	Root_shell	1 إذا تمّ الوصول لـ Root shell ، و 0 إفا	0
15	Su_attempted	1 إذا في حال محاول تنفيذ الأمر "su root" ، و 0 إفا	0
16	Num_root	عدد مرات الوصول بصلاحيات "Root" أو عدد العمليات التي نفذت بصلاحيات "Root" في الاتصال	0
17	Num_file_creations	عدد مرات انشاء ملف في الاتصال	0
18	Num_shells	عدد الـ shell prompts	0
19	Num_access_files	عدد العمليات على ملفات التحكم بالوصول	0
20	Num_outbound_cmds	عدد الأوامر الصادرة (Outbound) في جلسة FTP	0
21	Is_hot_login	1 إن كان تسجيل الدخول ينتمي لقائمة "hot" مثل Admin أو Root	0
22	Is_guest_login	1 إن كان تسجيل الدخول كـ "guest" و 0 إفا	0

الجدول 3: السمات المتعلقة بالمحتوى

رقم السمة	اسم السمة	توصيف	مثال
23	count	عدد الاتصالات إلى نفس الهدف مثل الاتصال الحالي في الثائنتين السابقتين	2
24	Srv_count	عدد الاتصالات إلى نفس الخدمة (رقم المنفذ) مثل الاتصال الحالي في الثائنتين السابقتين	2
25	Serror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) s0 ، s1 ، s2 or s3 بين الاتصالات المجمعة في count (وسمة رقم 23)	0
26	Srv_serror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) s0 ، s1 ، s2 or s3 بين الاتصالات المجمعة في srv_count (وسمة رقم 24)	0
27	Rerror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) REJ بين الاتصالات المجمعة في count (وسمة رقم 23)	0
28	Srv_rerror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) REJ بين الاتصالات المجمعة في srv_count (وسمة رقم 24)	0
29	Same_srv_rate	نسبة الاتصالات إلى نفس الخدمة من ضمن مجموعة الاتصالات المجمعة في count (وسمة رقم 23)	1
30	Diff_srv_rate	نسبة الاتصالات إلى خدمات مختلفة من ضمن مجموعة الاتصالات المجمعة في count (وسمة رقم 23)	0

31	Srv_diff_host_rate	نسبة الاتصالات إلى أجهزة مختلفة من بين الاتصالات المجمعة في srv_count (وسمة رقم 24)	0
----	--------------------	---	---

الجدول 4: السمات المتعلقة بالزمن

رقم السمة	اسم السمة	توصيف	مثال
32	Dst_host_count	عدد الاتصالات التي لها نفس عنوان الانترنت (IP) الهدف	150
33	Dst_host_srv_count	عدد الاتصالات التي لها نفس المنفذ	25
34	Dst_host_same_srv_rate	نسبة الاتصالات لنفس الخدمة من بين الاتصالات المجمعة في dst_host_count (الوسمة 32)	0.17
35	Dst_host_diff_srv_rate	نسبة الاتصالات إلى خدمات مختلفة من بين الاتصالات المجمعة في dst_host_count (الوسمة 32)	0.03
36	Dst_host_same_src_port_rate	نسبة الاتصالات إلى نفس المنفذ من بين الاتصالات المجمعة في dst_host_srv_count (الوسمة 33)	0.17
37	Dst_host_srv_diff_host_rate	نسبة الاتصالات إلى أجهزة مختلفة المنفذ من بين الاتصالات المجمعة في dst_host_srv_count (الوسمة 33)	0
38	Dst_host_serror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) s0 ، s1 ، s2 or s3 بين dst_host_count (الوسمة 32)	0
39	Dst_host_srv_serror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) s0 ، s1 ، s2 or s3 بين الاتصالات المجمعة في dst_host_srv_count (الوسمة 33)	0
40	Dst_host_rerror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) REJ بين الاتصالات المجمعة في dst_host_count (الوسمة 32)	0.05
41	Dst_host_srv_rerror_rate	نسبة الاتصالات التي فعلت الراية (وسمة رقم 4) REJ بين الاتصالات المجمعة في dst_host_srv_count (الوسمة 33)	0

الجدول 5: السمات المعتمدة على المضيف

8. الأبحاث المتعلقة:

يتمحور عدد كبير من الأعمال البحثية أو التقنية حول موضوع تحليل مجموعة البيانات KDD'99 ونظم كشف الاختراقات ، لكن عدد قليل نسبياً ركز على موضوع تحليل البيانات الكبيرة المتعلقة بالحماية ، و من هذه الأبحاث اخترنا تلك الأقرب إلى موضوع هذا البحث.

تصنيف التسلات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

في عام 2016 ، قدّم الباحثان (M. A. Jabbar و Nabila Farnaaz) [11] نموذج يعتمد على الغابات العشوائية لكشف الاختراقات في الشبكات الحاسوبية. تمّ تقييم أداء النموذج المقترح في [11] باستخدام مجموعة البيانات NSL-KDD بعد اختيار السمات وفقاً لمعيار عدم الشك التناظري (Symmetrical uncertainty). استُخدم النموذج لتصنيف أنواع مختلفة من الهجمات وتمّ تقييمه من خلال تحديد أربعة معايير: الصحة (Accuracy) ، معدّل الكشف (Detection rate) و يمثل النسبة بين العدد الكلي للهجمات التي تمّ اكتشافها من قبل النظام الى العدد الكلي للملاحظات التي تمّ تصنيفها بشكل صحيح أي $Dr = TP/(TP + TN)$ ، معدل الانذارات الخاطئة (FAR) ، و معامل ماتيو للترابط (MCC). بيّنت النتائج بأن للنموذج معدّل كشف مرتفع و معدّل إنذارات خاطئة منخفض ، كما أنّ النموذج صنّف هجمات منع الخدمة بصحة 99.67% وهذه القيمة أعلى بـ7% من خوارزمية شجرة القرار J48، كما بيّنت النتائج أيضاً أن اختيار السمات له دور في تحسّن أداء نموذج التصنيف. جميع الاختبارات في هذا البحث تمت باستخدام الأداة WEKA.

في عام 2016 ، قدّم الباحثان (Jonathan Graham و Tonya Fields) [12] مُصنّف مُعتمد على الغابات العشوائية لبناء نظام كشف اختراقات شبكي (NIDS) لاكتشاف الشذوذات (Anomaly). تمّ استخدام مجموعة البيانات KDD'99 و أيضاً الأداة WEKA وتمحورت الدراسة حول كيفية تغيير الزمن اللازم لبناء النموذج مع تغيير عدد الأشجار في الغابة ، وكذلك كيفية تحديد عدد الاشجار الأمثلي ، ودراسة دور عدد السمات بالنسبة لعدد الاشجار في تحقيق الدقة (Accuracy). أظهرت النتائج أنه يمكن الحصول على نتائج مناسبة عند تقليص مجموعة السمات بما يحسّن الكلفة الحسابية و يحسن صحة التنبؤ.

في عام 2017 ، قدّم الباحثان (Bayu Adhi Tama و Rifkie Primartha) [13] مقارنة لأداء نظام كشف اختراقات معتمد على مصنف الأشجار العشوائية من خلال ثلاث مجموعات بيانات متاحة على الانترنت: NSLKDD ، UNSW-NB15، and GPRS ، ومعيارين: الصحة (Accuracy) و معدل الانذارات الخاطئة ، وذلك باستخدام تحقيق الغابات العشوائية المضمّنة في حزمة H2O. تم تغيير أحجام الاشجار و البحث عن الوسائط الانسب باستخدام شبكة البحث (Grid search) لتظهر النتائج تفوق الاشجار العشوائية على مصنفات مجمعة (Ensembles) أخرى.

وفي عام 2018 ، قيم الباحثون (Mustapha Belouch et al.) [14] أداء أربع خوارزميات تصنيف ألاهي: شعاع الدعم الآلي ((Support Vector Machines (SVM))، بايز (Naïve Bayes)، أشجار القرار، و الغابات العشوائية باستخدام أباتشي سبارك و مجموعة البيانات UNSW-NB15 المفتوحة و المخصصة لاكتشاف الاختراقات. تمت المقارنة من خلال تقييم خمسة معاملات للأداء: الصحة (Accuracy) ، الحساسية (Sensitivity) وتقيس نسبة الهجمات التي تم تصنيفها بشكل صحيح (TP/TP+FN) ، النوعية (Specificity) ، زمن بناء النموذج (Building time) و زمن التنبؤ (Prediction time). أظهرت هذه الدراسة تفوق مصنف الغابات العشوائية على بقية المصنفات من حيث الصحة و الحساسية ، وتقارب بين الغابات العشوائية و أشجار القرار من حيث النوعية. في حين كانت خوارزمية بايز هي الأسرع من حيث التدريب، وكانت الأشجار العشوائية الأسرع باكتشاف الهجمات.

في عام 2018 ، قدم الباحثان (Priyanka Dahiya و Devesh Kumar Srivastava) [15] إطار لكشف التسلات باستخدام سبارك يحذف بداية السمات الأقل أهمية باستخدام تقنيتين:

LDA (Linear Discriminant Analysis) و CCA (Canonical Correlation Analysis) ، ومن ثم تطبيق سبع تقنيات تعلم خاضع للإشراف ومنها الغابات العشوائية لكشف الاختراقات في مجموعة البيانات UNSW-NB15. تمت المقارنة بين التقنيات المطبقة من خلال مجموعة من المعايير لتظهر الدراسة أن خوارزمية الشجرة العشوائية حققت الأداء الأفضل و بأن خوارزمية LDA كانت افضل تقنية لتقليص السمات.

في عام 2018 ، قدم الباحثان (M.Kranthi Kumar و K. Raja Kumar) [16] نظام لكشف الاختراقات في التطبيقات ذات البيانات الكبيرة باستخدام سبارك. أيضاً في هذا البحث تمت المقارنة بين عدد من المصنفات من ضمنها الغابات العشوائية ، واستخدمت مجموعتي البيانات KDD Cup99 و NSL KDD. قبل تطبيق التصنيف تم تقليص السمات باستخدام ثلاث خوارزميات: ربح المعلومات (Information Gain) ، نسبة الربح (Gain Ratio) و الترابط (Correlation). بينت النتائج أنه في حالة كل من مجموعتي البيانات حقق كل من الغابات العشوائية و مصنف الانحدار اللوجستي أفضل أداء من حيث دقة التصنيف و كانت الغابات العشوائية الأفضل من حيث زمن التنبؤ.

نحقق في هذا البحث نموذج لاكتشاف التسلات إلى الشبكات المحلية باستخدام خوارزمية الغابات العشوائية المضمنة في مكتبة تعلم الآلة (ML) من سبارك ولغة سكالادون

تصنيف التسلسلات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

أعدنا بعد ذلك تقسيم مجموعة البيانات لمجموعة تدريب مكونة من 80% من السجلات و مجموعة اختبار تتضمن 20% من السجلات كما استقنا من ميزة تخبئة للبيانات يُقدمها سبارك بهدف تسريع الحسابات (الشكل 11).

```
scala> val Array(trainingData, testData) = data.randomSplit(Array(0.8, 0.2), 11L)
trainingData: org.apache.spark.sql.Dataset[org.apache.spark.sql.Row] = [duration: int, protocol_type: string ... 40 more fields]
testData: org.apache.spark.sql.Dataset[org.apache.spark.sql.Row] = [duration: int, protocol_type: string ... 40 more fields]

scala> trainingData.cache()
res5: trainingData.type = [duration: int, protocol_type: string ... 40 more fields]

scala> testData.cache()
res6: testData.type = [duration: int, protocol_type: string ... 40 more fields]
```

الشكل 11: تقسيم مجموعة البيانات و التخبئة

يبين الشكل 12 عدد الملاحظات ضمن كل فئة في البيانات الكلية.

```
scala> KDDTrainIndexed.groupBy("label").count().orderBy("count").show(23)
[Stage 15:=====> (1 + 3) / 4]
[Stage 16:=====> (154 + 4) / 200]
```

label	count
spy1	21
perl1	31
phf1	41
multihop1	71
ftp_writel	81
loadmodule1	91
rootkit1	101
imap1	111
land1	181
warezmaster1	201
buffer_overflow1	301
guess_passwd1	531
pod1	2011
warezclient1	8901
teardrop1	8921
back1	9561
nmap1	14931
smurf1	26461
portsweep1	29311
ipsweep1	35991
satan1	36331
neptune1	412141
normal	1673431

الشكل 12: عدد الملاحظات ضمن كل فئة في البيانات الكلية

ثانياً: تعريف تحويلات لمقابلة السمات الفئوية المحرفية (البرتوكول ، الخدمة ، الراية) بأدلة رقمية وذلك باستخدام أحد صفوف تحويل البيانات: StringIndexer والذي يرمز القيم بأدلة رقمية حسب تكرار القيمة بحيث يتم ترميز القيمة الأكثر تكراراً ب 0 و هكذا.

ثالثاً: تعريف تحويل لتشكيل أشعة السمات (Features Vectors) من خلال تجميع جميع أعمدة السمات التي ستستخدم في نموذج تعلم الآلة في شعاع واحد.

رابعاً: بناء النموذج باستخدام خوارزمية الغابات العشوائية (الشكل 13).

```
scala> val rf = new RandomForestClassifier().
  setSeed(5043).
  setNumTrees(50).
  setMaxDepth(10).
  setMaxBins(100).
  setFeatureSubsetStrategy("auto").
  setLabelCol("indexedLabel").
  setFeaturesCol("features").
  setPredictionCol("prediction")
rf: org.apache.spark.ml.classification.RandomForestClassifier = rfc_168917bbb0b6
```

الشكل 13: بناء نموذج الغابات العشوائية

ومن ثمّ تعريف تسلسل عمليات (Pipeline) يتضمن الخطوات الثلاث السابقة ليؤدي في النتيجة لتوليد نموذج تمّ تدريبه على مجموعة بيانات التدريب (الشكل 14).

```
scala> val model = pipeline.fit(trainingData)
[Stage 11:> (0 + 4) / 5]
[Stage 11:===== (1 + 4) / 5]
[Stage 11:===== (2 + 3) / 5]
[Stage 11:===== (4 + 1) / 5]
```

الشكل 14: تدريب النموذج

يوضّح الشكل 13 إعدادات المصنّف حيث تمّ بداية تحديد رقم عشوائي في `setSeed` ، ومن ثمّ عدد المجالات التي ستستخدم لتقطيع السمات ذات القيم المستمرة (`setMaxBins(100)`) ، أما الخاصية `featureSubsetStrategy` فأعطيت القيمة `auto` حتى تختار الخوارزمية مجموعة من السمات الأفضل بشكل أوتوماتيكي ، بالإضافة لعمود العنونة ("indexedLabel") و اسم العمود الجديد الذي يُضيفه المصنّف للقيم التي يتنبأ بها ("prediction"). بعد بناء النموذج حددنا أهمية السمات والتي يقصد بها مدى مساهمة كل سمة في التنبؤات الصحيحة وهذه ميزة تقدمها خوارزمية أشجار القرار و الغابات العشوائية [6]. يبيّن الشكل 15 ترتيب أهمية السمات في الحالة الأولى:

تصنيف التسلات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

```
scala> Model.featureImportances.toArray.zip(combined).sorted.reverse.foreach(println)
(0.11365616216748639,src_bytes)
(0.10862363062893887,diff_srv_rate)
(0.10819379056857317,same_srv_rate)
(0.056594790771059146,src_serror_rate)
(0.056314539264287866,dst_host_serror_rate)
(0.054712752338939345,flag_indexed)
(0.049258123128110336,dst_host_srv_serror_rate)
(0.04257625997879035,dst_bytes)
(0.04209478844805279,service_indexed)
(0.039523537542342536,count)
(0.03775590236702081,protocol_type_indexed)
(0.03423642982038563,dst_host_diff_srv_rate)
(0.033784553469266176,error_rate)
(0.0292125598688617,dst_host_same_srv_rate)
(0.023064329676120873,dst_host_same_src_port_rate)
(0.02209348622595998,dst_host_srv_diff_host_rate)
(0.020939659802086816,src_count)
(0.020075253884009835,dst_host_srv_count)
(0.017434475953492023,dst_host_srv_rerror_rate)
(0.017321388355693704,dst_host_rerror_rate)
(0.016061194584460012,dst_host_count)
(0.010911834243300939,src_rerror_rate)
(0.010353265909217764,rerror_rate)
(0.006630880204496068,wrong_fragment)
(0.006595301677786092,logged_in)
(0.00632266164005416,num_compromised)
(0.006092532934831828,src_diff_host_rate)
(0.00543889188836696,hot)
(0.00293031973156353,duration)
(6.331569438613598E-4,is_guest_login)
(2.5926693827047197E-4,num_failed_logins)
(7.522258722246172E-5,num_root)
(7.38193247078914E-5,root_shell)
(6.268638454945078E-5,land)
(5.006819275451807E-5,num_file_creations)
(2.3635611465897796E-5,num_shells)
(1.682438269282028E-5,num_access_files)
(2.0225609194435472E-6,su_attempted)
(0.0,urgent)
(0.0,num_outbound_cmds)
(0.0,is_host_login)
```

الشكل 15: أهمية السمات مرتبة تنازلياً

خامساً: تقييم أداء النموذج

في هذه المرحلة طبقنا النموذج على مجموعة بيانات الاختبار (الشكل 16).

```
scala> val predictions = model.transform(testDF1)
predictions: org.apache.spark.sql.DataFrame = [duration: int, protocol_type: str
ing ... 47 more fields]
```

الشكل 16: تطبيق النموذج على بيانات الاختبار

يوضح الشكل 17 صحة النموذج :

```
scala> val testErr = predictionAndLabels.filter(r => r._1 != r._2).count.toDouble / testData.count()
[Stage 88:> (0 + 4) / 4]
[Stage 88:=====> (1 + 3) / 4]
[Stage 88:=====> (2 + 2) / 4]

testErr: Double = 0.0043118794256101906

scala> println("Accuracy = " + (1-testErr) * 100 + "%")
Accuracy = 99.56881205743898 %
```

الشكل 17: صحة النموذج في الحالة الأولى

ونظراً لعدم تقارب عدد الملاحظات في الفئات قيماً الإحكام ، الاستدعاء و معدل الانذارات الخاطئة (FPR) لكل فئة. يوضح الجدول 6 النتائج:

Precision	Recall	FPR
Precision(normal)=99.32 %	Recall(normal)=99.96 %	FPR(normal)=0.79 %
Precision(neptune)=100.00 %	Recall(neptune)=99.98 %	FPR(neptune)=0.00 %
Precision(satan)=100.00 %	Recall(satan)=97.17 %	FPR(satan)=0.00 %
Precision(ipsweep)=98.77 %	Recall(ipsweep)=99.17 %	FPR(ipsweep)=0.04 %
Precision(portsweep)=99.64 %	Recall(portsweep)=98.76 %	FPR(portsweep)=0.01 %
Precision(smurf)=100.00 %	Recall(smurf)=99.63 %	FPR(smurf)=0.00 %
Precision(nmap)=99.18 %	Recall(nmap)=93.08 %	FPR(nmap)=0.01 %
Precision(back)=100.00 %	Recall(back)=100.00 %	FPR(back)=0.00 %
Precision(teardrop)=100.00 %	Recall(teardrop)=100.00 %	FPR(teardrop)=0.00 %
Precision(warezclient)=100.00 %	Recall(warezclient)=81.67 %	FPR(warezclient)=0.00 %
Precision(pod)=100.00 %	Recall(pod)=100.00 %	FPR(pod)=0.00 %
Precision(guess_passwd)=100.00 %	Recall(guess_passwd)=100.00 %	FPR(guess_passwd)=0.00 %
Precision(buffer_overflow)=100.00 %	Recall(buffer_overflow)=50.00 %	FPR(buffer_overflow)=0.00 %
Precision(warezmaster)=100.00 %	Recall(warezmaster)=50.00 %	FPR(warezmaster)=0.00 %
Precision(land)=50.00 %	Recall(land)=100.00 %	FPR(land)=0.01 %
Precision(imap)=100.00 %	Recall(imap)=100.00 %	FPR(imap)=0.00 %
Precision(rootkit)=0.00 %	Recall(rootkit)=0.00 %	FPR(rootkit)=0.00 %
Precision(loadmodule)=0.00 %	Recall(loadmodule)=0.00 %	FPR(loadmodule)=0.00 %
Precision(ftp_write)=0.00 %	Recall(ftp_write)=0.00 %	FPR(ftp_write)=0.00 %
Precision(multihop)=0.00 %	Recall(multihop)=0.00 %	FPR(multihop)=0.00 %
Precision(phf)=0.00%	Recall(phf)=0.00%	FPR(phf)=0.00 %

الجدول 6: نتائج تقييم الإحكام ، الاستدعاء و معدل الإنذارات الخاطئة للنموذج الأول تشير النتائج إلى أن المصنف حقق أداء مرتفع بالنسبة لمعظم الفئات ، لكن يوجد فئات مثل ftp_write لم تظهر أي نتيجة لها غير "صفر" و من الممكن أن ذلك يعود لقلّة عدد الملاحظات في هذه الفئات بالنسبة لبقية الفئات.

نشير إلى أنه نظراً لتثبيت رقم ضمن setSeed في المصنف فإن تكرار التجربة سيؤدي لنفس النتائج السابقة.

سادساً: ضبط الوسائط الفائقة (Hyperparameters)

في هذا الاختبار استخدمنا الصف ParamGridBuilder و تم البحث من خلال ثلاث وسائط يأخذ كل منها قيمتان مما يعني أنه سيتم تقييم 8 وسطاء (الشكل 18):

```
scala>
| val paramGrid = new ParamGridBuilder().
| addGrid(rf.impurity, Seq("gini", "entropy")).
| addGrid(rf.maxDepth, Seq(1, 20)).
| addGrid(rf.maxBins, Seq(100, 300))
paramGrid: org.apache.spark.ml.tuning.ParamGridBuilder = org.apache.spark.ml.tuning.ParamGridBuilder@56fe3ec2
```

الشكل 18: تحديد ثلاث وسائط فائقة للبحث عن النموذج الأفضل

ثم طبقنا التحقق المتقاطع مع تحديد أن البحث عن النموذج الأفضل سيتم بناء على معيار الصحة. نتيجة لضبط النموذج وجدنا تحسن في الصحة لتصبح (99.82%) (الشكل 19).

تصنيف التسلات إلى الشبكات المحلية باستخدام الغابات العشوائية و Apache Spark

```
scala> val accuracy = multiclassEval.evaluate(predictionscv)
[Stage 672:>] (0 + 4) / 4]
[Stage 672:=====>] (1 + 3) / 4]
[Stage 672:=====>] (2 + 2) / 4]

[Stage 674:=====>] (1 + 3) / 4]

accuracy: Double = 0.9981803077653388
```

الشكل 19: صحة النموذج الأفضل بعد عملية البحث

كما وجدنا أن الوسائط الفائقة لهذا النموذج كما يوضح الشكل 20:

```
scala> bestModel.asInstanceOf[PipelineModel].stages.last.extractParamMap
res7: org.apache.spark.ml.param.ParamMap =
{
  rfc_388209a1ddb-cacheNodeIds: false,
  rfc_388209a1ddb-checkpointInterval: 10,
  rfc_388209a1ddb-featureSubsetStrategy: auto,
  rfc_388209a1ddb-featuresCol: features,
  rfc_388209a1ddb-impurity: gini,
  rfc_388209a1ddb-labelCol: indexedLabel,
  rfc_388209a1ddb-maxBins: 300,
  rfc_388209a1ddb-maxDepth: 20,
  rfc_388209a1ddb-maxMemoryInMB: 256,
  rfc_388209a1ddb-minInfoGain: 0.0,
  rfc_388209a1ddb-minInstancesPerNode: 1,
  rfc_388209a1ddb-numTrees: 50,
  rfc_388209a1ddb-predictionCol: prediction,
  rfc_388209a1ddb-probabilityCol: probability,
  rfc_388209a1ddb-rawPredictionCol: rawPrediction,
  rfc_388209a1ddb-seed: 5043,
  rfc_388209a1ddb-subsamplingRate: 1.0
}
```

الشكل 20: وسائط النموذج الأفضل في الحالة الأولى

9.2 الحالة الثانية: تصنيف مُتعدد الفئات مع ترميز واحد حار:

في الحالة الثانية قمنا بترميز قيم السمات الفئوية الرقمية قبل عملية التصنيف باستخدام ترميز واحد حار بهدف دراسة إن كان هذا الترميز سيؤدي لتحقيق أداء أفضل للمصنف. لكن صحة النموذج الناتج (99.26%) لم تكن أفضل من سابقه مما يعني أن استخدام الترميز واحد حار لم يؤدي لتحقيق مصنف ذو أداء أفضل.

9.3 الحالة الثالثة: تصنيف ثنائي بدون ترميز واحد حار:

تم اعتبار الاتصالات الطبيعية فئة أولى (0) و البقية كفئة ثانية (1) و لم يتم ترميز قيم السمات الفئوية الرقمية. يوضح الشكل 21 تابع تحويل العنونة إلى صفر أو واحد حسب إن كان الاتصال طبيعي أم غير ذلك.

```
scala> val func = udf( (i:String) => if(i == "normal") 0.0 else 1.0 )
func: org.apache.spark.sql.expressions.UserDefinedFunction = UserDefinedFunction
(<function1>,DoubleType,Some(List(StringType)))
```

الشكل 21: تابع تحويل العنونة إلى صفر أو واحد

تم اتباع نفس خطوات المرحلة الأولى لنجد بداية أن ترتيب أهمية السمات قد تغير و كانت السمات العشر الأكثر أهمية كالتالي:

```
(0.18560993211330665,dst_bytes)
(0.1517893921246571,src_bytes)
(0.08985512205020213,same_srv_rate)
(0.0851156044180343,diff_srv_rate)
(0.06677556272995512,dst_host_srv_count)
(0.06562562869258275,dst_host_same_srv_rate)
(0.052777321343388105,service_indexed)
(0.04115571036797827,protocol_type_indexed)
(0.033677374186416066,flag_indexed)
(0.027352062422142383,logged_in)
(0.0270730898054533,dst_host_diff_srv_rate)...
```

وكذلك وجدنا تحسن ملحوظ في أداء النموذج ، حيث كانت النتائج كما يوضح الجدول 7:

Area under PR	Area under ROC Area under ROC	F1	الصحة
99.99%	99.74%	99.75 %	99.75 %

الجدول 7: نتائج تقييم الأداء للنموذج الثالث

طبقتنا البحث عن الوسائط في هذه الحالة أيضاً لنلاحظ تحسن في قيمة معيار الصحة ليصبح

(99.88%) ، كما أن وسائط النموذج الأفضل موضحة في الشكل 22:

```
scala> bestModel.asInstanceOf[PipelineModel].stages.last.extractParamMap
res7: org.apache.spark.ml.param.ParamMap =
{
  rfc_ea1ab87d22c7-cacheNodeIds: false,
  rfc_ea1ab87d22c7-checkpointInterval: 10,
  rfc_ea1ab87d22c7-featureSubsetStrategy: auto,
  rfc_ea1ab87d22c7-featuresCol: features,
  rfc_ea1ab87d22c7-impurity: gini,
  rfc_ea1ab87d22c7-labelCol: label,
  rfc_ea1ab87d22c7-maxBins: 100,
  rfc_ea1ab87d22c7-maxDepth: 20,
  rfc_ea1ab87d22c7-maxMemoryInMB: 256,
  rfc_ea1ab87d22c7-minInfoGain: 0.0,
  rfc_ea1ab87d22c7-minInstancesPerNode: 1,
  rfc_ea1ab87d22c7-numTrees: 50,
  rfc_ea1ab87d22c7-predictionCol: prediction,
  rfc_ea1ab87d22c7-probabilityCol: probability,
  rfc_ea1ab87d22c7-rawPredictionCol: rawPrediction,
  rfc_ea1ab87d22c7-seed: 5043,
  rfc_ea1ab87d22c7-subsamplingRate: 1.0
}
```

الشكل 22: وسائط النموذج الأفضل في الحالة الثالثة

10. الخاتمة و التوصيات المستقبلية:

حققنا في هذا البحث عدة نماذج لاكتشاف التسلسلات إلى الشبكات المحلية باستخدام الغابات العشوائية و مكتبة تعلم الآلة من أباتشي سبارك وهو نظام مفتوح المصدر لمعالجة موزعة

تصنيفُ التسلّلات إلى الشبكات المحليّة باستخدام الغابات العشوائية و Apache Spark

للبيانات الكبيرة ، كما قيّمنا أداء هذه النماذج اعتماداً على عدة معايير. بيّنت النتائج دقة عالية في اكتشاف التسلّلات ، كما بيّنت أن التصنيف الثنائي يحقق نتائج أفضل من التصنيف المتعدد الفئات و بأن ترميز السمات الفئوية الرقمية بمرمّز واحد حار لم يكن له تحسين ملحوظ على أداء المصنّف.

يمكن في الأعمال المستقبلية استخدام مجموعات بيانات مختلفة لمقارنة أداء التصنيف ، كما يمكن دراسة أداء المصنّف بعد انتقاء مجموعة من السمات الأكثر الأهمية.

11. المراجع:

- [1]. د. زينب خلوف ، دراسة ومقارنة تقنيات تحسين أداء نظم كشف الاختراقات في الشبكات الحاسوبية. بحث مسجل قُبل للنشر في مجلة جامعة البعث ، المجلد 32 لعام 2010.
- [2]. Apache Spark Lightning-fast unified analytics engine.
URL: <https://spark.apache.org/>, last retrieved: 29-9-2019
- [3]. H. Luu, Beginning Apache Spark 2 With Resilient Distributed Datasets, Spark SQL, Structured Streaming and Spark Machine Learning library. Apress; 1st edition (2018).
- [4]. Differences between Receiver Operating Characteristic AUC (ROC AUC) and Precision Recall AUC (PR AUC), URL: <http://www.chioka.in/differences-between-roc-auc-and-pr-auc/>, last retrieved: 29-9-2019.
- [5]. S. Raschka and V. Mirjalili, Python Machine Learning. Birmingham, UK: Packt Publishing (2017).
- [6]. J. Wills, S. Ryza, S. Owen, and . Laserson, Advanced Analytics with Spark: Patterns for Learning from Data at Scale. O'Reilly Media; 2nd edition (2017).
- [7]. X-She Yang, Introduction to Algorithms for Data Mining and Machine Learning. Academic Press (2019).
- [8]. NSL-KDD, <https://www.unb.ca/cic/datasets/nsl.html>. Last retrieved: 29-9-2019.
- [9]. S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, Cost based modeling for fraud and intrusion detection: Results from the jam project, discex, vol. 02, p. 1130, 2000.
- [10]. KDD99, <http://kdd.ics.uci.edu/databases/kddcup99/task.html>.
- [11]. N. Farnaaz, M. A. Jabbar, "Random forest modeling for network intrusion detection system", Procedia Comput. Sci., vol. 89, pp. 213-217, Jan. 2016.
- [12]. T. Fields and J. Graham. Classifying network attack data using random forest. Computers and Their Applications (CATA), Dec 2016.
- [13]. R. Primartha and B. Tama. Anomaly Detection Using Random Forest: A Performance Revisited. International Conference on Data and Software Engineering (ICoDSE), Palembang Sumatra Selatan, Indonesia, pp. 1--6, IEEE, 2017.
- [14]. Belouch M, El Hadaj S, Idhammad M. Performance evaluation of intrusion detection based on machine learning using Apache Spark. Procedia Computer Science. 2018.
- [15]. Dahiya P, Srivastava DK. Network intrusion detection in big dataset using Spark. Procedia Comput Sci. 2018.

- [16]. M. Kranthi Kumar, K. Raja Kumar, Intrusion Detection System For Big Data Driven Application Using Apache Spark, International Journal of Management, Technology And Engineering, Volume 8, Issue XI, 2018.
- [17]. Ofir Shalev, Recall, Precision, F1, ROC, AUC, and everything, May 27, 2019,
URL: <https://medium.com/swlh/recall-precision-f1-roc-auc-and-everything-542aedef322b9>, last retrieved: 24-5-21
- [18]. Random forests or random decision forests,
URL: <https://www.kaggle.com/getting-started/176257>, last retrieved: 24-5-21
- [19]. K. Scarfone, P. Mell, Special Publication 800-94: Guide to Intrusion Detection and Prevention Systems (IDPS), National Institute of Standards and Technology (NIST) (2007).

دراسة ومقارنة أداء طرائق تخمين القناة في نظام *MIMO-OFDM* عند استخدام شفرات *STBC*

م. ايلي خليل

ملخص:

ترسل البيانات في نظام التجميع بتقسيم التردد المتعامد orthogonal frequency division multiplexing على عدد كبير من الحوامل المتعامدة، وتعرض هذه الحوامل لأشكال مختلفة من قنوات الخفوت التي تؤثر على البيانات المرسلة.

تتم عملية كشف البيانات عبر معرفة استجابة القناة عند كل حامل، وذلك عبر ارسال اشارات دليلية (PILOT) معروفة مسبقاً لدى المستقبل واعتماد طرائق تخمين القناة .

يتم دمج نظام OFDM مع نظام متعدد الدخل- متعدد الخرج (multiple input and multiple output) (MIMO)، عبر ارسال واستقبال اشارات OFDM على أكثر من هوائي ويتم ذلك عبر تشفير هذه الاشارات بأحد انماط التشفير المتبعة في نظام MIMO.

ندرس في هذا البحث الشفرات البلوكية المفصولة زمنياً (Spaced-Time Block codes)، وذلك في حالتين عند اعتماد هوائي ارسال وهوائي استقبال واحد، أو عند اعتماد هوائي ارسال واستقبال .

ونبين تأثير هذه الشفرات على أداء عدد من طرائق تخمين القناة، ونرسم مخطط خطأ BER من أجل حالتين من حالات الخفوت عند زمن تأخير الانتشار العالي ومن أجل تردد دوبلر المرتفع.

الكلمات المفتاحية : OFDM ، MIMO ، STBC، تخمين القناة .

Comparison Between The Performance of Channel Estimation in MIMO-OFDM When We Use STBC

Abstract

In OFDM system we transmit data symbols on a large number of subcarriers, and these data symbols exposed to changes in amplitude and phase, caused of fading Channel.

To get the transmitted data at the receiver we must cope with these changes, so we need to know the channel response at every subcarrier and every symbol, and for that reason we use PILOTS and we use many methods to estimate the channel response depend on pilots.

We merge OFDM System with MIMO System By Transmit and receive OFDM signal over multiple antennas, and using type of codes to transmit data .

We will discuss in these issue the Space time block code, at two case:

2*1: system with tow transmitter and one receiver antennas.

2*2: system with tow transmitter and tow receiver antennas.

And its effect at the performance of many channel estimation we will study the BER and MSE in both case of frequency selective fading and the case of high Doppler shift .

Key Word: MIMO, OFDM ,STBC, Channel Estimation .

1- مقدمة:

يعد نظام (OFDM) من أكثر الأنظمة شيوعاً في الاتصالات اللاسلكية ؛ حيث اعتمد في العديد من المعايير اللاسلكية، وفي الجيل الرابع من أنظمة الاتصال الخلوية 4G ، وذلك لكفاءته الطيفية ومعدلات نقل البيانات العالية التي يؤمنها [1].

إنّ إحدى الطرائق لتحسين أداء أي نظام اتصال في قنوات الخفوت هي تغيير مواصفاته الإحصائية، ويتم ذلك بشكلٍ عام بوساطة تقنية تدعى التعدد، حيث تخفف تأثيرات الخفوت عملياً في أنظمة التعدد بوساطة خوارزميات معالجة الإشارة المصممة بشكلٍ مناسب في كل من المرسل و المستقبل .

يعرّف التعدد على أنّه قنوات مستقلة متعددة بين المرسلات و المستقبلات ؛ لذلك تحدث تعددية الهوائيات أو التمييز القطبي عندما يكون للمسارات المستقلة حيّز مكاني في الطبيعة . حيث يمكن القول أنّه يوجد فراغ كافٍ بين عناصر الهوائي في المرسلات و المستقبلات بحيث أنّه لا يوجد ارتباط، أو يوجد ارتباط قليل جداً بين إشاراتها المتوقعة. تستخدم تعددية الهوائيات إما من أجل تحسين أداء الوصلة للإشارة، أو لزيادة معالجة البيانات [2] .

تم استخدام نظام OFDM مع عدة هوائيات في كل من المرسل والمستقبل من أجل تحسين أداء النظام في قنوات الخفوت المتعددة، وزيادة معدل نقل البيانات كما ذكرنا سابقاً، ويطلق على هذه الهوائيات المتعددة في كل من الخرج والدخل مصطلح (multiple input and multiple output) الأنظمة متعددة الدخل- متعددة الخرج (MIMO).

وقد حسنت أنظمة MIMO من أداء أنظمة الاتصالات اللاسلكية بالنسبة للعديد من العوامل، مثل تحسين السعة و الحصول على معدل معطيات أعلى و معدل خطأ بت (BER) أخفض [3]، كما أعطت أنظمة MIMO إنتاجية جيدة وكفاءة طيفية عالية مع

نفس الاستطاعة المرسل، وعرض حزمة المتطلبية في حالة الأنظمة التي لا تعتمد التعدد
(Single input-Single output) (SISO) [4].

تطبق أنظمة MIMO وفق عدة نماذج فمن الممكن أن نرسل نفس الإشارة على كل الهوائيات، مما يعني اننا سنحسن من أداء النظام تجاه قنوات الخفوت، ولكننا لن نزيد من سعة البيانات، ويسمى هذا النموذج بتقنية التباين الفراغي (Spatial Diversity)؛ حيث يتم توليد N نسخة بديلة عن الإشارات، وهذه الإشارات ستكون مرتبطة بهوائيات الإرسال، ويتم تخفيض تأثير خفوت الإشارة بشكل كبير في هذه الحالة.

أو يتم تقسيم قطار البتات للدخل إلى عدد من القطارات الفرعية وعددها N ، حيث N يحدد العدد الكلي لهوائيات الإرسال، ويتم تجميع الإشارة القادمة في طرف الاستقبال باستخدام تقنيات التجميع المختلفة، وتدعى هذه التقنية بتقنية التجميع الفراغي (Spatial diversity) وتزيد هذه التقنية من معدل المعطيات للإرسال وذلك بالإرسال التفرعي الذي يتم تحقيقه [4].

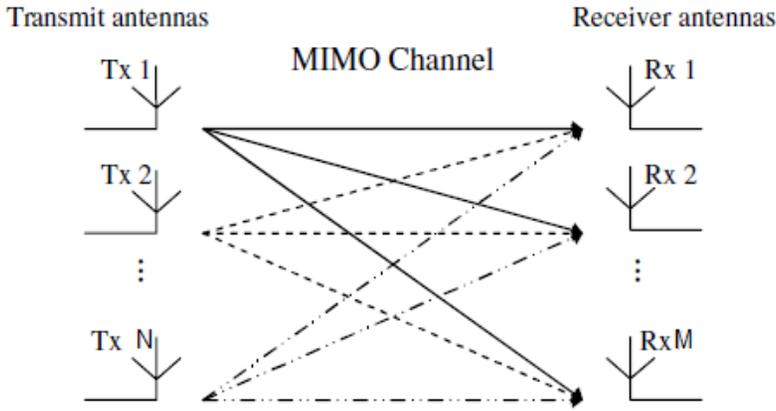
يتم تشفير البيانات في مشفر MIMO قبل أن ترسل على الهوائيات المتعددة وفق عدة أنواع من الشفرات كالشفرات الطبقيّة المفصولة زمنياً (layered Spaced-Time codes)، أو الشفرات التشعبية المفصولة زمنياً (Trellis Spaced-Time codes)، أو الشفرات البلوكية المفصولة زمنياً (Spaced-Time Block codes).

2-هدف البحث:

يهدف البحث الى دراسة استخدام الشفرات البلوكية المفصولة زمنياً STBC في نظام MIMO-OFDM، في حالتين عند اعتماد هوائي ارسال وهوائي استقبال واحد، أو عند استخدام هوائي ارسال و هوائي استقبال، ودراسة تأثيرها على أداء عدد من طرائق التخمين المستخدمة في نظام OFDM، وذلك من أجل حالتين من حالات الخفوت عند زمن تأخير الانتشار العالي، ومن أجل تردد دوبلر المرتفع.

3- نموذج القناة لأنظمة MIMO (دخل متعدد / خرج متعدد) :

لنعتبر لدينا وصلة اتصالات تتضمن N هوائي إرسال و M هوائي استقبال المبينة في الشكل (1)، و التي تعمل في قناة MIMO . يستجيب كل هوائي مستقبل لكل هوائي إرسال من خلال مسار خفوت مستقل إحصائياً . إنَّ الإشارات المستقبلية تكون محزفة بوساطة ضجيج إضافي و الذي يكون مستقل إحصائياً بين M هوائي للمستقبل و أدوار الإرسال .



الشكل (1) نظام MIMO يستعمل N هوائي إرسال و M هوائي استقبال

تعطى الإشارة المستقبلية عند هوائي الاستقبال K بالعلاقة التالية [2]:

$$r_k = \sum_{i=1}^N H_{ik} * S_i + n_i \quad 1 < k < M \quad (1)$$

حيث: r_k هي الإشارة المستقبلية عند الهوائي K

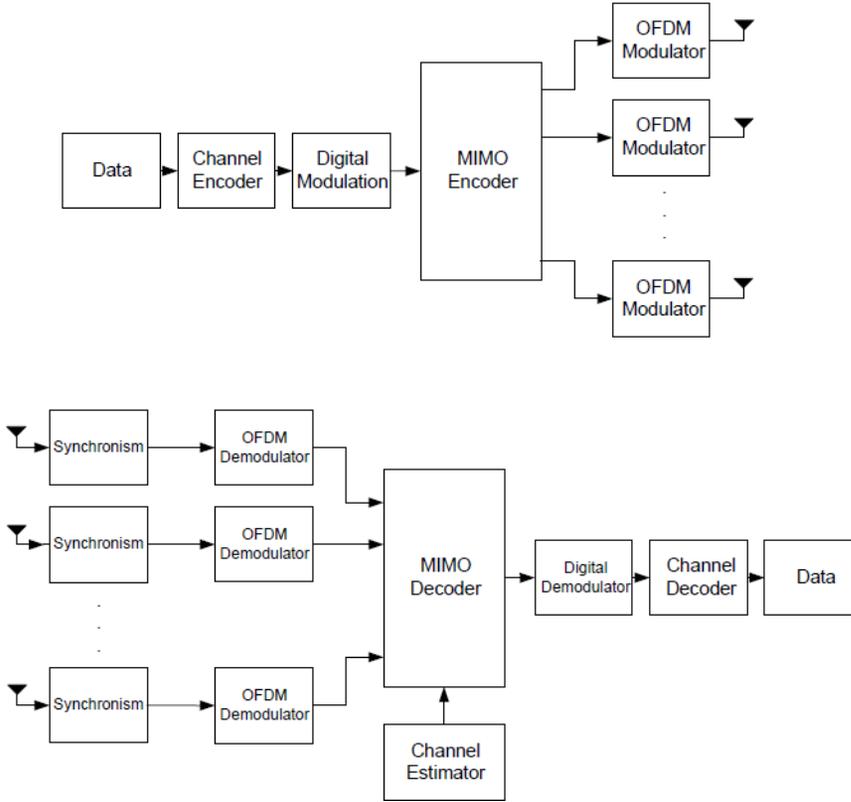
H_{ik} هي مسار الخفوت بين هوائي المرسل i و هوائي الاستقبال K

S_i هي الإشارة المرسلة عند الهوائي i

n_i ضجيج غاوصي مستقل عشوائياً بتباين معياري قيمته (1)

4-بنية النظام الهجين OFDM-MIMO :

يتم دمج نظام OFDM مع نظام MIMO عبر ارسال اشارة OFDM عند كل هوائي ارسال كما يبين الشكل (2) .



الشكل (2) مخطط صندوقي لنظام OFDM-MIMO

تشفر البيانات أولاً باستخدام شيفرة القناة لحمايتها من العيوب الناجمة عن القناة، ثم تمرر الى مشفر MIMO (MIMO encoder) الذي يقسم انسياب البيانات المشفر إلى N انسياب متوازي.

يتم تشفير البيانات في مشفر MIMO وفق عدد من الشفرات، كشفرة التأخير (Delay) التي تعد أبسط الشفرات، وتعتمد الى ارسال نفس البيانات على كل هوائي

بتأخير زمني، أو الشفرات الطبقيّة المفصولة زمنياً (layered Spaced-Time codes)، والتي تفترض أن يكون عدد هوائيات الإرسال نفس عدد هوائيات الاستقبال، وتعالج الإشارة المرسلّة من الهوائي المرسل الأول عند المستقبل كإشارة مرغوبة، بينما تعامل الإشارات المرسلّة من الهوائيات الأخرى عند هذا المستقبل كإشارات ضجيج.

أو الشفرات التشعبية المفصولة زمنياً (Trellis Spaced-Time codes)، التي تقسم البيانات إلى مجموعات، وترسل على كل هوائي وتقدم معدلات نقل بيانات عالية.

أو الشفرات البلوكية المفصولة زمنياً (Spaced-Time Block codes)، والتي سندرس أداءها بشكل تفصيلي في هذا البحث.

يعدل كل انسياب وفق نظام OFDM، ويرسل كل منها باستخدام هوائيات منفصلة.

يتم توليد إشارة OFDM عند كل هوائي وتضمين البيانات المرسلّة على حوامل فرعية متعامدة تشكل مجملها إشارة OFDM، و تتم عملية تشكيل إشارة OFDM باستخدام تحويل فورييه المعكوس IFFT [5].

يتم ادراج مجموعة من الاشارات الدليلية المعروفة مسبقاً لدى المستقبل (PILOTS)، من أجل تخمين القناة عند كل حامل فرعي من حوامل إشارة OFDM، وذلك من أجل معرفة استجابة القناة عند المسار بين هوائي المرسل i و هوائي الاستقبال K (H_{ik}).

يتم بعد ذلك إضافة فترة الحماية، حيث يمدد رمز OFDM بشكل دوري، وذلك من أجل المحافظة على التعامدية بين الإشارات قبل أن تتم عملية إرسال الإشارة.

يتم عند الاستقبال التعامل أولاً مع كل إشارة OFDM على حدة، واستخلاص استجابة القناة عند كل مسار، قبل أن تتم عملية الاستفادة من تعدد الاشارات الواصلة إلى كل هوائي كما سنبين لاحقاً.

إن معالجة الإشارة الرقمية تبدأ بتهيئة الطور لتحديد توقيت الرمز، و انزياح التردد. يستخدم FFT لفك تعديل كل الحوامل الفرعية [5].

تتم معرفة استجابة القناة عند كل حامل فرعي في النظام وعند كامل الرموز بالاعتماد على الاشارات الدليلية المرسله ، وذلك بالاعتماد على طرائق تخمين القناة مثل طريقة LS (Least Square) أو طريقة linear mean minimum square error (LMMSE) [5].

بعد معرفة استجابة القناة عند كل وصلة والبيانات المرسله على كل حامل ترسل الى فاك تشفير MIMO (MIMO decoder) ليتم التعامل معها حسب التقنية المعتمدة عند الارسال.

5- الشفرات البلوكية المفصولة زمنياً (Spaced-Time Block codes) :

اقترحت هذه الشفرة من قبل العالم Alamouti بسبب تعقيد فك التشفير في مخططات التشعب المستخدمة في شفرات التشعب المفصولة زمنياً، وقد طبق Alamouti هذا النمط من أجل هوائي إرسال واحد و M هوائي استقبال.

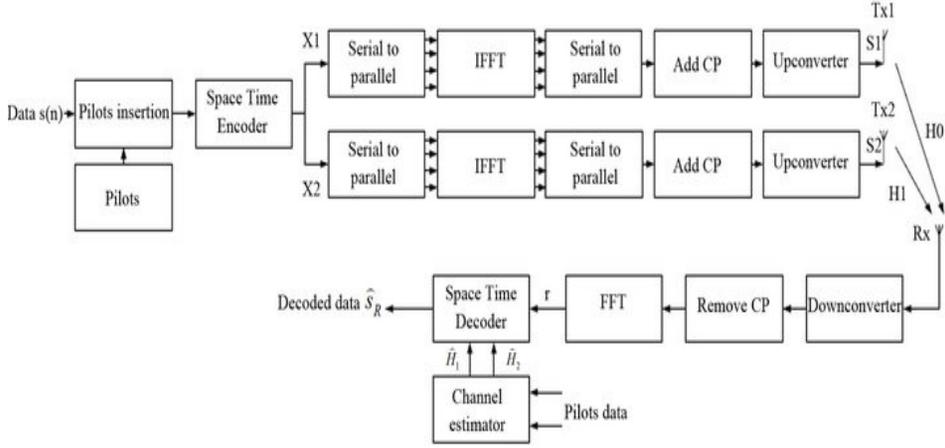
ويعتبر STBC نمط تشفير قوي جداً ويملاً الفجوة بين التجميع الفراغي والتباين الفراغي أي يقدم حلاً وسطاً بين معدل نقل البيانات العالي والأداء الجيد ويبين الشكل مخطط ارسال واستقبال لنظام MIMO-OFDM مع اعتماد شفرات STBC.

يمكن تلخيص الخواص الرئيسية لصندوق الشيفرات المفصولة زمنياً-block space time codes بما يلي:

1- لا يوجد ضياع في عرض الحزمة لأن الخوارزمية الخاصة به تعطي أعلى معدل نقل ممكن عند استخدام full diversity.

2-الفعالية الطيفية لهذا النظام أقل من الفعالية الطيفية لشيفرات التشعب المفصولة زمنياً STC trellis.

3-يتم استخدام خوارزمية بسيطة جدا تستخدم جوامع خطية فقط عند المستقبل لفك تشفير الاحتمال الأعظمي [2].



الشكل (3) مخطط صندوقي لنظام OFDM-MIMO

و تأتي بساطة الخوارزمية من تصميمها بحيث تكون أعمدها متعامدة فيما بينها.

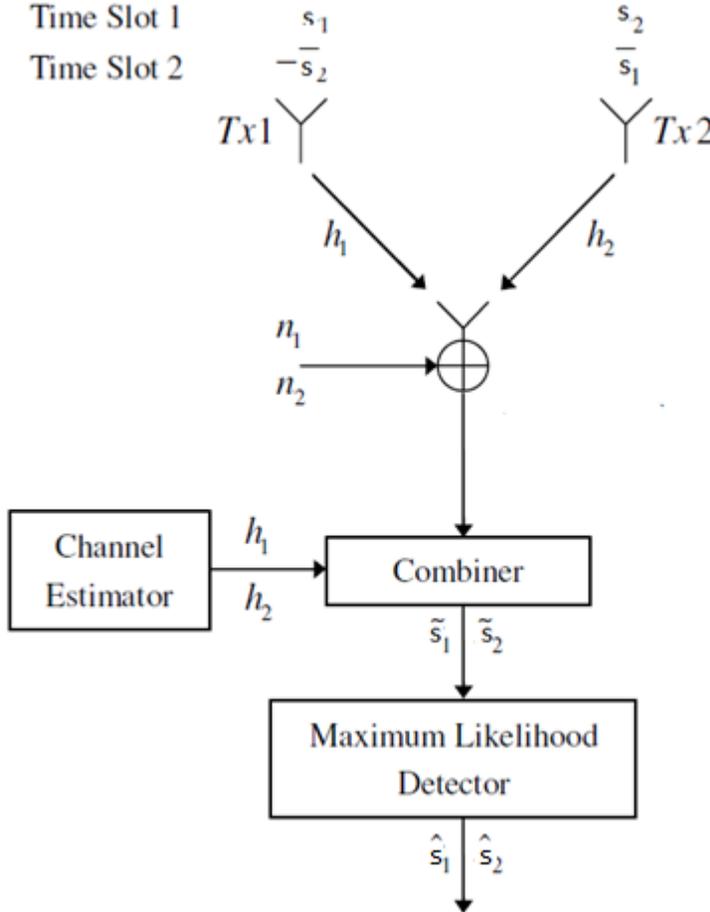
وتعطى مصفوفة STBC من أجل هوائي إرسال كما يلي [6]:

$$C2 = \begin{pmatrix} S_1 & -S_2^* \\ S_2 & S_1^* \end{pmatrix}$$

حيث يرسل الرمز S_1 من الهوائي (1) خلال الحيز الزمني t_1 ، ويرسل الرمز S_2 من الهوائي (2) خلال الحيز الزمني t_1 ، بينما يرسل الرمز $-S_2^*$ من الهوائي (1) خلال الحيز الزمني t_2 ، أما الرمز S_1^* يرسل من الهوائي (2) خلال الحيز الزمني t_2 كما هو مبين في الشكل (3).

وتكون العلاقة بين الزمنيين $t_2 = t_1 + \tau$ و S_1^*, S_2^* ، هما المرافقان العقديان للرمزين S_2, S_1 على الترتيب .

يرسل الرمزين ضمن حيزين زمنيين اثنين لذلك فإن معدل المعطيات هو $2/2=1$ ،
والذي يعتبر معدل معطيات كامل و رتبة التباين هي $2M$ حيث M يحدد عدد هوائيات
الاستقبال .



الشكل (4) مخطط صندوق لآلية عمل الشفرات البلوكية المفصولة زمنياً

تكون الإشارة المستقبلية عند طرف الاستقبال كما هو مبين في المعادلة التالية [7] :

$$r_1(t_1) = h_{1,1}S_1 + h_{1,2}S_2 + n \quad (2)$$

$$r_1(t_2) = -h_{1,1} S_2^* + h_{1,2}S_1^* + n \quad (3)$$

حيث $r_1(t_1)$ الاشارة المستقبلية في الحيز الزمني 1

$r_1(t_2)$ الاشارة المستقبلية في الحيز الزمني 2

نحصل على استجابة القناة $h_{1,1}$ ، $h_{1,2}$ من مخزن نظام OFDM بالاعتماد على الاشارات الدليلة المرسله و طرائق تخمين القناة.

يمكن تعميم المعادلتين السابقتين على أكثر من مستقبل وفق المعادلة التالية [7] :

$$r_j(t_1) = \sum_{i=1}^N h_{i,j} S_i + n_j \quad (4)$$

حيث $r_j(t_1)$ هي الاشارة المستقبلية عند هوائي الاستقبال J وخلال الحيز الزمني (t_1)

و N عدد الهوائيات المرسله وهي في حالتنا تساوي 2 .

يمكن استعادة عينات البيانات المرسله بعد معرفة استجابة القناة بوساطة جامع خطي linear combining عند اللحظة (t_1) و اللحظة (t_2) بالترتيب كما هو مبين في المعادلتين التاليتين [2]:

$$\tilde{S}_1 = h_1^* r_1(t_1) + h_2 r_2^*(t_1) + n_1 \quad (5)$$

$$\tilde{S}_2 = h_2^* r_1(t_1) - h_1 r_2^*(t_1) + n \quad (6)$$

حيث \tilde{S}_1 ، \tilde{S}_2 هي الاشارات المستخرجه مع الضجيج.

بالتعويض عن $r_1(t_1)$ ، $r_2(t_1)$ نجد :

$$\tilde{S}_1 = (|h_1|^2 + |h_2|^2) S_1 + h_1^* n_1 + h_2 n_2^* \quad (7)$$

$$\tilde{S}_2 = (|h_1|^2 + |h_2|^2)S_2 + h_2^*n_1 + h_1n_2^* \quad (8)$$

إن المعادلتين (7)،(8) تعطي تخمين أفضل من حالة عدم التعدد لأن احتمال خفوت كلا من h_1 و h_2 معاً في وقت واحد أقل من احتمال خفوت h_1 في المعادلة 3. والميزة الأكبر أنه في حال تعرضت h_1 و h_2 لفقد ريليه، فإن $|h_1|^2$ تتعرض لفقد Chi-square من الدرجة الثانية (two degrees of freedom) أما $|h_1|^2 + |h_2|^2$ تتعرض لفقد Chi-square من الدرجة الرابعة Chi-square distribution with four degree of freedom [2].

بالتالي احتمال أقل للخفوت العميق deep fade. هذا يبين الفرق بين نظام اللا تعدد والنظام الذي يستخدم التعدد.

توجه الآن القيم المستخلصة \tilde{S}_1 ، \tilde{S}_2 الى مخمن الأرجحية العظمى Maximum likelihood Decoder الذي يحدد الرموز المرسله الأكثر احتمالاً وذلك باختيار الرموز التي تجعل المسافة الاقليدية (Euclidean Distance) عن الاشارات المستقبله أقل مايمكن وفق المعادلة التالية [7]:

$$\sum_{t=1}^p \sum_{j=1}^M \left| r_j(t) - \sum_{i=1}^N h_{i,j} s_t^i \right|^2 \quad (9)$$

حيث p هي عدد الفواصل الزمنية التي تتم فيها عملية الارسال.

ومن أجل حالتنا التي تعتمد هوائيي ارسال ووشفرة STBC نحصل على الخوارزمية الاتية من أجل تحديد S_1 [8] :

$$\left[\sum_{i=1}^M r_i(t_1)h_{i,1}^* + r_i^*(t_2)h_{i,2}^* \right] - S_1 \left| -1 + 2 \sum_{i=1}^M \sum_{j=1}^2 |h_{i,j}|^2 \right| S_1 \left| \right|^2 \quad (10)$$

ومنها نختار S_1 التي تجعل القيمة أصغر ما يمكن

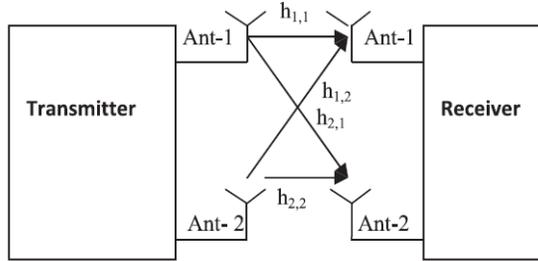
و من أجل تحدي S_2 لدينا المعادلة (11) :

$$\left[\sum_{i=1}^M r_i(t_1)h_{i,2}^* - r_i^*(t_2)h_{i,1}^* \right] - S_2 \left| -1 + 2 \sum_{i=1}^M \sum_{j=1}^2 |h_{i,j}|^2 \right| S_2 \left| \right|^2 \quad (11)$$

ومنها نختار S_2 التي تجعل القيمة اقل مايمكن.

6- حالة هوائي ارسال و هوائي استقبال (2*2):

نستخدم بهذه الحالة نفس مصفوفة الترميز المستعملة في الحالة السابقة .



الشكل (5) نموذج نظام 2*2

هنا لدينا إشارتين مستقبليتين وزمنييين مختلفين كما هو محدد في المعادلات التالية [6]:

$$r_1(t_1) = h_{1,1}S_1 + h_{1,2}S_2 + n_1 \quad (12)$$

$$r_1(t_2) = -h_{1,1} S_2^* + h_{1,2}S_1^* + n_2 \quad (13)$$

$$r_2(t_1) = h_{2,1}S_1 + h_{2,2}S_2 + n_3 \quad (14)$$

$$r_2(t_2) = -h_{2,1} S_2^* + h_{2,2}S_1^* + n_4 \quad (15)$$

وهذه المخارج المستقبلية تعطى كمدخل إلى كاشف ML ليتم كشف الرموز S_2 ، S_1 ، وتتم عملية الكشف بنفس الطريقة المبينة في الحالة السابقة و يتم إيجاد المسافات الأصغرية لكل القيم الممكنة للرموز .

نقوم الآن بإجراء محاكاة باستخدام الماتلاب لنظام MIMO-OFDM ، يستخدم شفرات STBC في كلا الحالتين : عند استخدام هوائي ارسال وهوائي استقبال واحد (2^*1)، أو عند استخدام هوائي ارسال و هوائي استقبال (2^*2) ، وذلك من أجل عدة طرائق تخمين للقناة في نظام OFDM، و من أجل حالتين من حالات الخفوت عند زمن تأخير الانتشار العالي ومن أجل تردد دوبلر المرتفع.

1- عند اعتماد طريقة (Least Square) LS لتخمين القناة:

تعتمد طريقة LS على جعل مربع المسافة بين الإشارة المرسله ، والإشارة المستقبلية أصغر ما يمكن . كما يلي [9] :

$$\min_{H^H} J(H) = \min_{h^H} \{|Y - XH|^2\} = \min_{h^H} \{(Y - XH)^H (Y - XH)\} \quad (16)$$

حيث H^* تعني المرافق العقدي للمنقول

$$\frac{\partial}{\partial H^H} J(H) = 0 \quad \text{يكون هذا الفرق أصغر ما يمكن عندما يكون المشتق صفراً:}$$

و بالتالي تكون استجابة القناة تساوي:

$$\hat{H}_{LS} = \left[\frac{Y_1}{X_1} \frac{Y_2}{X_2} \dots \frac{Y_{NP}}{X_{NP}} \right]^T \quad (17)$$

سنقوم الآن بدراسة أداء هذه الطريقة عند استخدام شفرات STBC في نظام MIMO من أجل حالتين (2*1) و (2*2) في كلا الحالتين التاليتين :

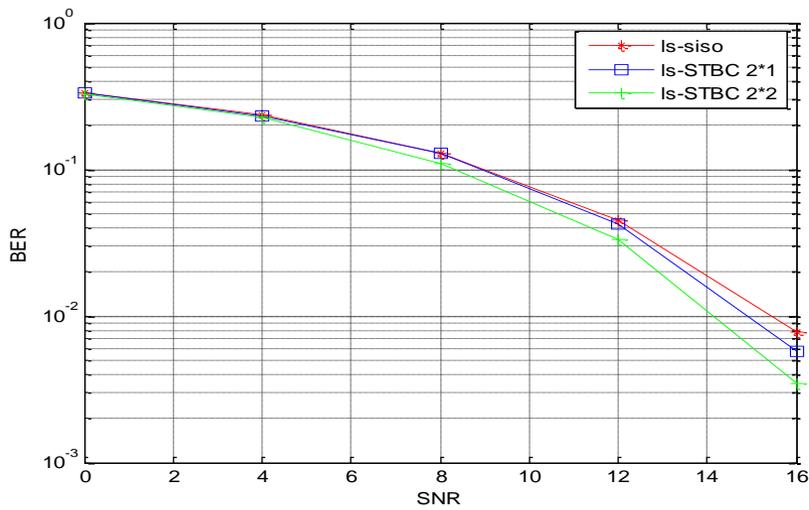
أولاً: من أجل بيئة خفوت تتسم بزمن تأخير انتشار عالي :

نرسم منحنى الخطأ BER عند استخدام هذه الطريقة باستخدام برنامج Matlab و سنعتمد في دراستنا على النمط البلوكي لتوزيع الاشارات الدليلية بحيث نضع لكل 9 رموز إشارة دليلية (Pilot) وذلك في حال كانت رموز المعطيات هي رموز أنظمة التعديل 16QAM أو QPSK.

علماً أن بارامترات نظام OFDM المستخدمة هي: عرض حزمة كلي B=16MHz و زمن الحماية $T_g = 8\mu s$ وزمن الرمز الكلي $T_s = 64\mu s$ وعدد الحوامل N=64 وعدد حوامل فترة الحماية $N_g = 8$ فيكون عدد الحوامل الكلي $N_{OFDM} = 72$.

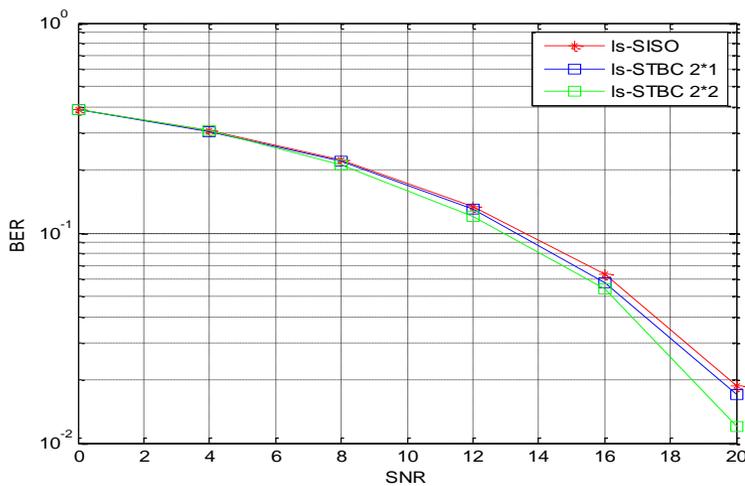
و تمت الدراسة عند تأخير انتشار $\tau_{rms} = 8\mu s$

A. من أجل رموز معطيات بتعديل QPSK:



الشكل (6) خطأ BER عند استعمال طريقة LS عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين 2×2 و 1×2 وذلك من أجل تعديل QPSK

B. من أجل رموز معطيات بتعديل 16QAM:



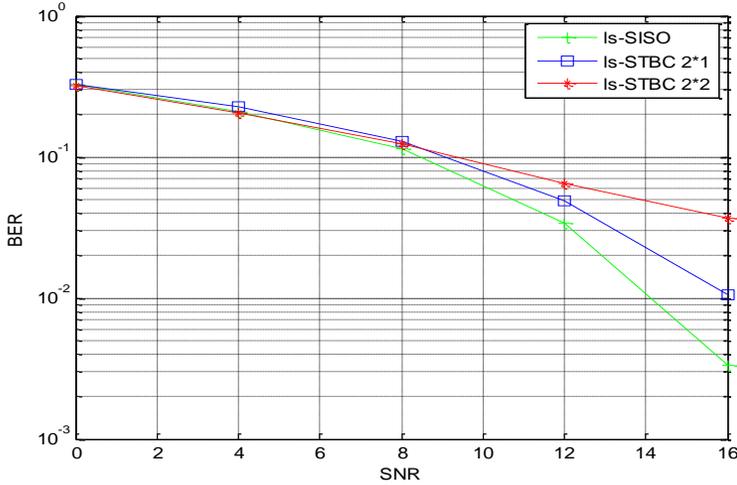
الشكل (7) خطأ BER عند استعمال طريقة LS عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين 2×2 و 1×2 وذلك من أجل تعديل 16QAM

نلاحظ من الشكلين (5,6) أن استخدام التعدد مع شفرات STBC يحسن من أداء طريقة LS بشكل بسيط من أجل زمن تأخير انتشار مرتفع.

ثانياً: من أجل بيئة خفوت تتسم بتردد دوبلر مرتفع:

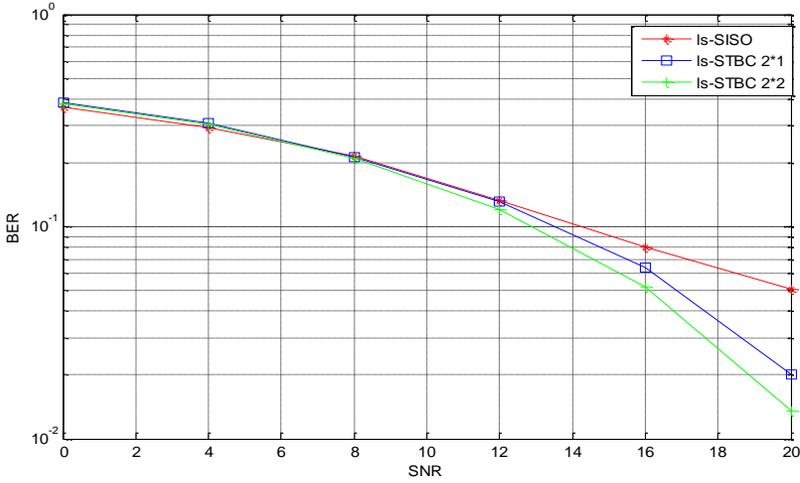
سنقوم الآن بأجراء نفس عملية المحاكاة السابقة ورسم منحنى الخطأ BER عند استخدام طريقة تخمين LS من أجل أنماط توزع الاشارات الدليلية الثلاث وسنعمد نفس البارامترات السابقة مع تغيير بارامترات بيئة الخفوت حيث سندرس حالة تردد دوبلر مرتفع $f_d = 35Hz$.

A. من أجل رموز معطيات بتعديل QPSK:



الشكل (8) خطأ BER عند استعمال طريقة LS عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين 1*2 و 2*2 وذلك من أجل تعديل QPSK

B. من أجل رموز معطيات بتعديل 16QAM:



الشكل (9) خطأ BER عند استعمال طريقة LS عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين 1*2 و 2*2 وذلك من أجل تعديل QAM16

يبين الشكلين (7,8) أن استخدام التعدد مع شفرات STBC يحسن بدرجة كبيرة أداء طريقة LS عند تردد دوبلر العالي.

2- طريقة التخمين المعتمدة على تحويل فورييه DFT :

إن استجابة القناة المخمنة بطريقة LS Least Square (h_{ls}) تعطى بالعلاقة :

$$\hat{h}_{ls} = Y X^{-1} \quad (18)$$

حيث X هي الإشارة المرسلة و Y هي الإشارة المستقبلية.

بتطبيق تحويل فورييه المعكوس IDFT، يمكن الحصول على الاستجابة النبضية للقناة Channel Impulse Response CIR من الاستجابة الترددية للقناة Channel Frequency Response CFR كما يلي :

$$\hat{H}_{ls} = \text{IDFT}[\hat{h}_{ls}] \quad (19)$$

تمتلك عدة رموز من CIR طاقة منخفضة في قنوات خفوت متعدد المسارات، بالتالي فإننا نأخذ فقط أول L رمز ذي الطاقة الأعلى، و نحصل بالتالي على [10]:

$$\hat{H}_{ls} = \begin{cases} IDFT[\hat{h}_{ls}] & 0 \leq n \leq L-1 \\ 0 & otherwise \end{cases} \quad (20)$$

بتطبيق تحويل فورييه نحصل على :

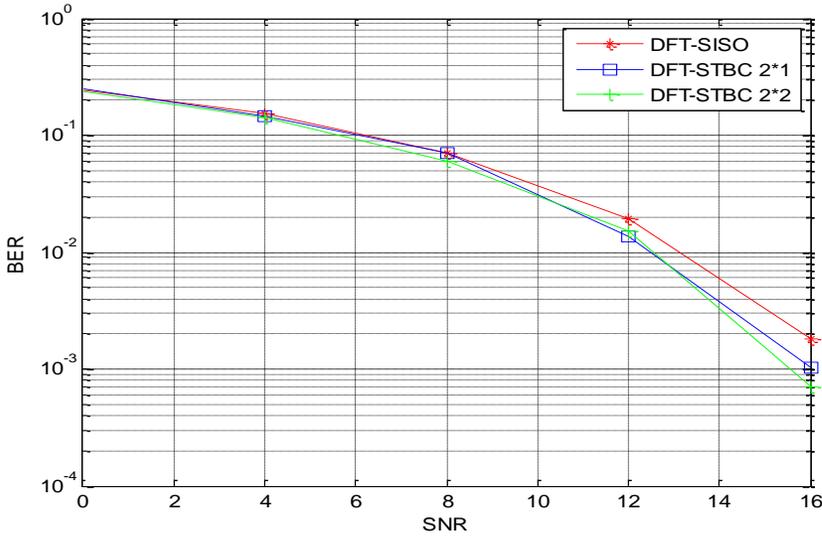
$$\hat{h}_{DFT-LS} = DFT[\hat{H}_{ls}] \quad (21)$$

سنقوم الآن بدراسة أداء هذه الطريقة عند استخدام شفرات STBC في نظام MIMO من أجل حالتين $(2*1)$ و $(2*2)$.

أولاً: من أجل بيئة خفوت تتسم بزمن تأخير انتشار عالي:

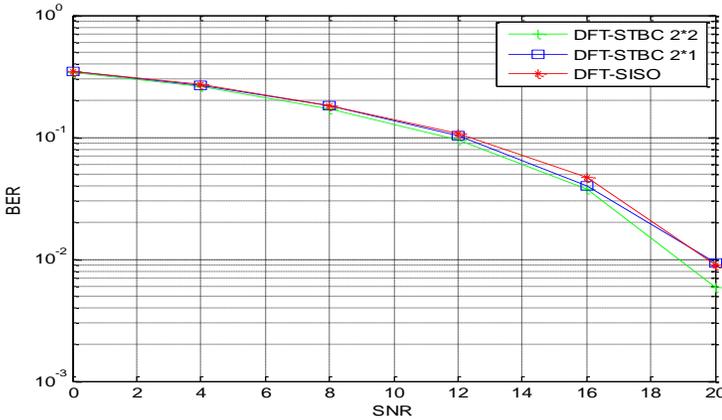
نقوم الآن بأجراء محاكاة باستخدام برنامج الماتلاب ورسم منحنى الخطأ BER عند استخدام هذه الطريقة وسنعمد نفس البارامترات السابقة ومن أجل زمن تأخير انتشار $\tau_{rms} = 8\mu s$.

A. من أجل رموز معطيات بتعديل QPSK :



الشكل (10) خطأ BER عند استعمال طريقة DFT عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين $1*2$ و $2*2$ وذلك من أجل تعديل QPSK

B. من أجل رموز معطيات بتعديل 16QAM:

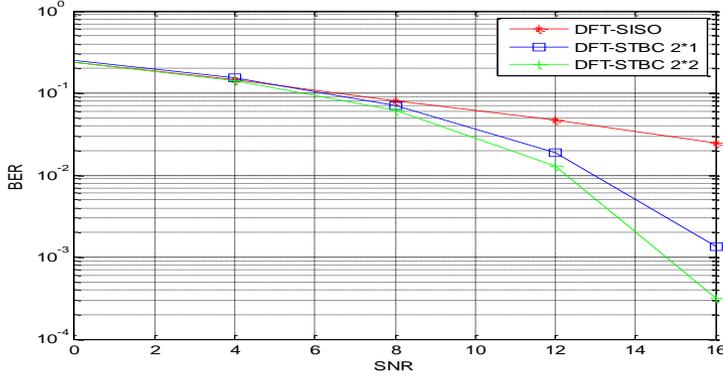


الشكل (11) خطأ BER باعتماد طريقة DFT عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين $1*2$ و $2*2$ وذلك من أجل تعديل 16QAM

ثانياً: من أجل بيئة خفوت تتسم بتردد دوبلر مرتفع

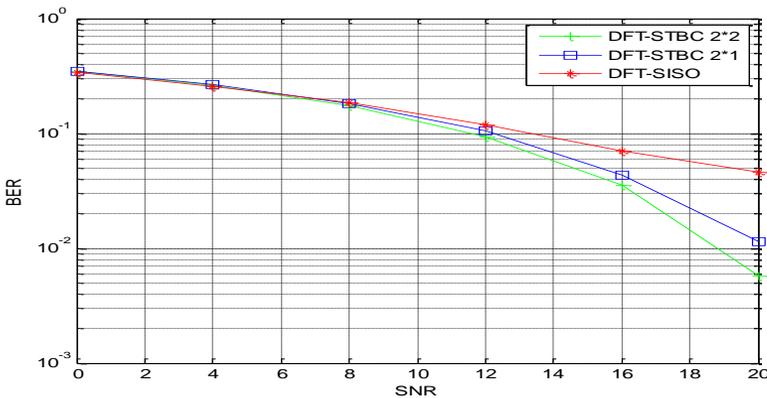
باعتقاد نفس البارامترات السابقة مع تردد دوبلر مرتفع $f_d = 35\text{Hz}$.

A. من أجل رموز معطيات بتعديل QPSK:



الشكل (12) خطأ BER عند استعمال طريقة DFT عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين $1*2$ و $2*2$ وذلك من أجل تعديل QPSK

B. من أجل رموز معطيات بتعديل 16QAM:



الشكل (13) خطأ BER باعتماد طريقة DFT عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين $1*2$ و $2*2$ وذلك من أجل تعديل 16QAM

نستنتج من المخططات (9,10,11,12) أن شفرة STBC تحسن من أداء طريقة DFT من أجل تردد دوبلر مرتفع بشكل واضح بينما يبقى التحسين محدوداً من أجل زمن تأخير انتشار عالي.

3-طريقة (LMMSE) Liner minimum mean square error لتخمين القناة (Wiener Filtering):

تستخدم هذه الطريقة خصائص القناة (الترابط الزمني والترابط الترددي) لتقدير تابع استجابة القناة غير المعلومة عند الحوامل الفرعية التي لا تحوي الإشارات الدليلية انطلاقاً من الترابط بين هذه الحوامل والحوامل التي تحمل الإشارات الدليلية [11].

أي أن هذه الخصائص يجب أن تكون معروفة عند المستقبل ليتم التخمين بشكل مثالي،

تعطى استجابة القناة باستخدام طريقة LMMSE (H_{lmmse}) بدلالة H_{LS} [12]:

$$\hat{H}_{lmmse} = \{R_h (R_h + \sigma^2 (XX^H)^{-1})\}^{-1} \hat{H}_{ls} \quad (22)$$

حيث \hat{H}_{ls} هي استجابة القناة المخمّنة بطريقة (LS) Least Square.

R_h هي مصفوفة الترابط الذاتي بين الترددات التي تحمل الإشارات الدليلية.

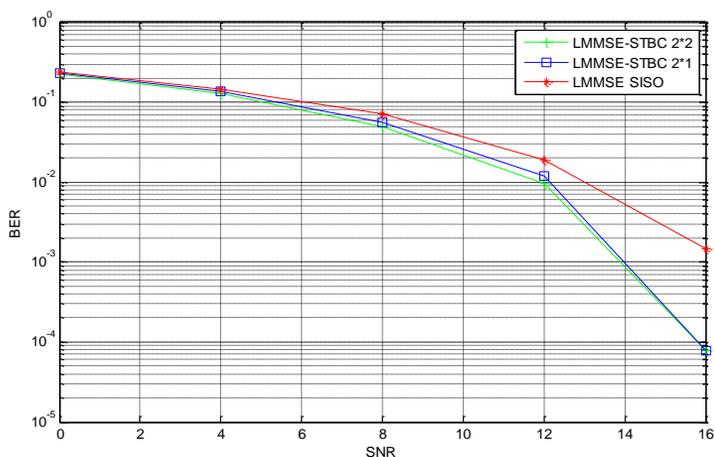
σ^2 هي تشتت الضجيج من أجل كل حامل فرعي.

أولاً: من أجل بيئة خفوت تتسم بزمن تأخير انتشار عالي :

سنقوم الآن بأجراء محاكاة باستخدام برنامج الماتلاب ورسم منحنى الخطأ BER عند استخدام طريقة تخمين هذه وسنعمد في دراستنا نفس البارامترات السابقة ومن أجل زمن

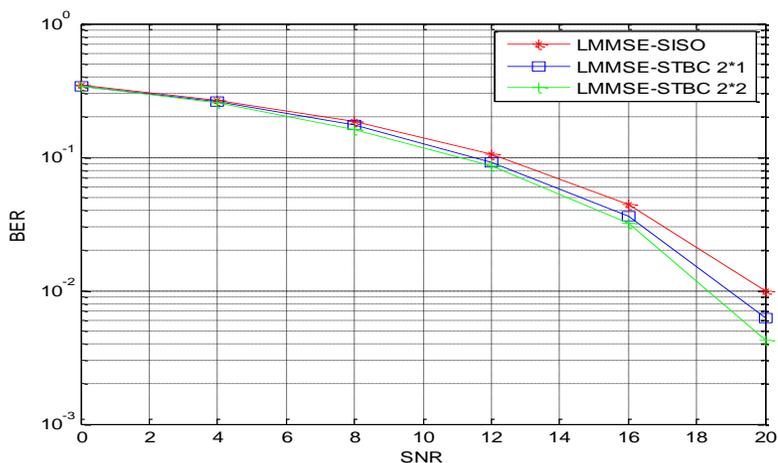
$$\tau_{rms} = 8\mu s$$

A. من أجل رموز معطيات بتعديل QPSK:



الشكل (14) خطأ BER باعتماد طريقة LMMSE عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين $1*2$ و $2*2$ وذلك من أجل تعديل QPSK

B. من أجل رموز معطيات بتعديل 16QAM



الشكل (15) خطأ BER عند استعمال طريقة LMMSE عند عدم وجود تعدد (SISO) وعند استعمال STBC في كلا الحالتين $1*2$ و $2*2$ وذلك من أجل تعديل 16 QAM

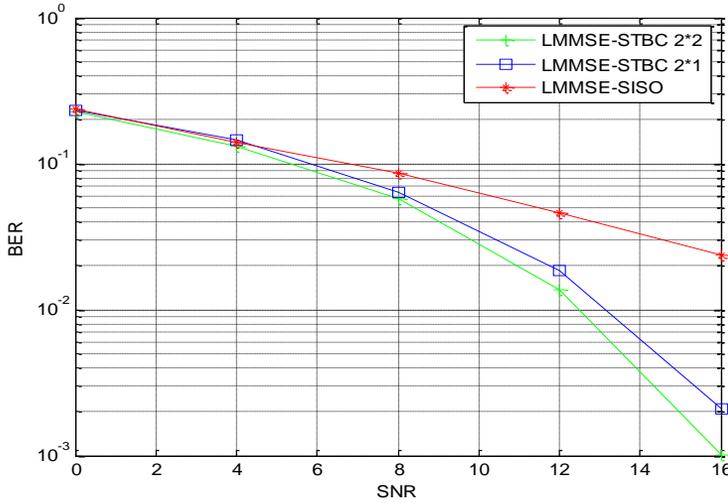
نستنتج من الأشكال (13,14) أن التعدد مع شفرات STBC تحسن من أداء طريقة LMMSE بشكل أكبر من طريقة LS و DFT عند زمن تأخير انتشار مرتفع

ثانياً: من أجل بيئة خفوت تتسم بتردد دوبلر مرتفع

باعتداف نفس البارامترات السابقة مع تغير بارامترات بيئة الخفوت؛ حيث سندرس حالة

$$f_d = 35 \text{ Hz} .$$

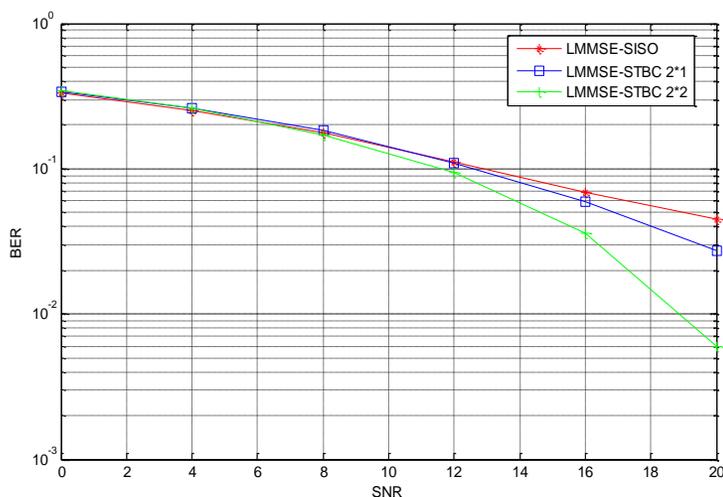
A. من أجل رموز معطيات بتعديل QPSK:



الشكل (16) خطأ BER باعتماد طريقة LMMSE عند عدم وجود تعدد (SISO) وعند استعمال

STBC في كلا الحالتين 2*2 و 2*2 وذلك من أجل تعديل QPSK

B. من أجل رموز معطيات بتعديل 16QAM:



الشكل (17) خطأ BER باعتماد LMMSE عند عدم وجود تعدد (SISO) وعند استعمال

STBC في كلا الحالتين 1×2 و 2×2 وذلك من أجل تعديل 16 QAM

يبين الشكلين (15,16) أن استخدام التعدد مع شفرات STBC يحسن أيضاً بدرجة كبيرة أداء طريقة LMMSE عند تردد دوبلر العالي.

8- النتائج والمقترحات:

تم دراسة أداء نظام MIMO-OFDM عند استعمال شفرات STBC ، وإجراء محاكاة للنظام عند اعتماد عدد من طرائق تخمين القناة، وتبين أن استخدام شفرة STBC يحسن الأداء عند استخدام طريقة التخمين LS و طريقة التخمين المعتمدة على DFT بشكل بسيط من أجل زمن تأخير انتشار مرتفع، ويزداد هذا التحسن عند اعتماد طريقة LMMSE لتأخير القناة.

يكون التحسن في الأداء أكبر عند استعمال هوائيين في الاستقبال، ولكن مع زيادة التعقيد والكلفة عند المستقبل.

كما تحسن شفرة STBC من أداء طرائق التخمين المدروسة كلها بشكل كبير من أجل تردد دوبلر مرتفع، أي يمكن اعتمادها في بيئات الخفوت السريع التي تكون فيها سرعة دوبلر B_d أكبر من عرض حزمة القناة W ، أي عندما يكون الزمن المتماسك T_c أصغر من زمن الرمز للإشارة T_s .

نقترح البحث في الشفرات التي تحسن الأداء بشكل أكبر عند زمن تأخير انتشار عالي كالشفرات البلوكية المفصولة ترددياً، كما نقترح البحث بالشفرات التشعبية ودراسة طرائق تخفيض تعقيدها.

9- المراجع:

- [1] Anibal Luis Intini ,2000-Orthogonal Frequency Division Multiplexing For Wireless Network Standard IEEE 802.11a,University of California.
- [2] Heiskala . j , Terry . j , - OFDM Wireless LANs – A Theoretical and Practicl Gide
- [3]A Molisch ,Wireless Communications. Wiley–IEEE Press,2005 .
- [4] T.Kaiser ,''Smart antenna–State of art'', Hindawi Publishing Corp–2005
- [5] R . v .Nee ,R .Parsad ,2000- OFDM for wireless multimedia communication , Artich House ,London.
- [6] Luis Miguel Cortes-Pena, "MIMO Space time block coding (STBC):Simulation and Results" ,Cortes-Pena EC6604 Personal and Mobile Communications presented to Dr Jordan stuber, April-2009 .
- [7] Wang.L , Akhtman.j , Jiang.M, 2011- MIMO-OFDM for LTE, Wi-Fi and WiMAX.
- [8] D.Gesbert,M.Shafi,D.Shiu,P.Smith,and A. Naguib," From Theory to Practice : An Overview of MIMO Space-Time Coded Wireless Systms",IEEE Journal on selected areas in Communications,vol.21,no.3,pp.281-302,2003 .
- [9]Marc Engels, 2002- Wireless OFDM Systems: How to make them work?. Kindle Edition.
- [10]Jun-heeJang,Se-bin Im,Jeog-soon Park,Hyung-jin Choi,14th 2008, DFT-based decision directed channel estimation for OFDM systems in very large delay spread channels.

[11]Jie Ma, HuaYa, Shouyinliu, 2009- The MMSE Channel Estimation Based on DFT for OFDMSystem .

[12]Samuel.c.young,2010-OFDMA System Analysis and Design, ,Artech House, Boston, London.

التحكم بسكر الدم لدى مرضى السكري باستخدام متحكم PID وخوارزميات المعايرة

م. مريان سعدو اشراف : د. عمار زقزوق

طالبة ماجستير - كلية الهندسة الميكانيكية والكهربائية - جامعة البعث

ملخص البحث:

مرض السكري هو مرض واسع الانتشار في العالم اليوم. يعمل العديد من الباحثين على طرق لتشخيص مرض السكري وعلاجه، ومن الأدوات المستخدمة لهذا الغرض نماذج رياضية لغلوكوز الدم والأنسولين. تم في هذا البحث عرض دراسة تحليلية لبعض خوارزميات الذكاء الصناعي (خوارزميات المعايرة)، والتي استخدمت مع نظام التحكم هذا وساعدت على إيجاد الحل الأمثل، وعرض النتائج التي تم الوصول إليها عند استخدام كل خوارزمية. تم استخدام الخوارزمية الجينية وخوارزمية النحل وخوارزمية محاكاة التلدين وخوارزمية سرب الطيور في تحديد القيم المثلى للمتحكم في النموذج الرياضي لنظام الغلوكوز والأنسولين، مع عرض زمن التنفيذ لكل خوارزمية عن طريق الواجهات الرسومية في بيئة MATLAB وتحديد الخوارزمية الأفضل.

الكلمات المفتاحية: النموذج الرياضي لنظام الغلوكوز والأنسولين، المعايرة، الخوارزمية الجينية، خوارزمية النحل، خوارزمية محاكاة التلدين.

Controlling Blood Glucose in Diabetics using PID Controller and Calibration Algorithms

Eng. Marian Saado

MSc. Student - Mechanical and Electrical Engineering Faculty - Albaath University

Abstract:

Diabetes is a widespread disease in the world today. Many researchers are working on ways to diagnose and treat diabetes. Among the tools used for this purpose are mathematical models of blood glucose and insulin kinetics. In this research, an analytical study of some artificial intelligence algorithms was presented.

In this research, an analytical study was presented for some artificial intelligence algorithms (Calibration Algorithms) that were used with this control system and they helped to find the optimal solution and displayed that were reached when using each algorithm. In this research, we focus of the simulated annealing algorithm, artificial Bee Colony algorithm, particle swarm optimization, genetic algorithm in determining the optimal values for the controller in the proposed mathematical model for the patient. Showing the execution time for each algorithm via graphical interfaces in the MATLAB environment and determining the best algorithm.

Keywords: Mathematical Model of Glucose System, Calibration, Genetic Algorithm, Simulated Annealing Algorithm Artificial Bee Colony.

1- مقدمة:

مرض السكري من الأمراض المنتشرة بشكل كبير وهو مرض يؤدي إلى تلف بعض الأعضاء وحتى الوفاة إذا لم تتم مداراته بشكل صحيح في عصرنا الراهن تطبيقات الذكاء الصناعي تجتاح المجال الطبي بكافة نواحيه، وفي هذا البحث سوف يتم التدقيق في خوارزميات الذكاء الصناعي التي استخدمت لهدف التحكم بنسبة السكر في الدم وتطبيقها من خلال نموذج رياضي باستخدام الماتلاب لاختيار الخوارزمية الأفضل التي يمكن استخدامها مع البنكرياس الصناعي، وسوف يتم استخدام خوارزميات الحل الأمثل (المعايرة) لمعرفة القيم الأفضل لضبط معدل ضخ الأنسولين والمقارنة بين الخوارزميات المستخدمة وتحديد الأفضل.

سيتناول القسم الثاني من هذه المقالة الهدف من هذا البحث، بينما يحوي القسم الثالث أهم الدراسات المرجعية، ويتضمن القسم الرابع دراسة خوارزميات الحل الأمثل (المعايرة) (GA,SA,ABC,PSO) مع النموذج الرياضي المقترح، بينما يعرض القسم الخامس التنفيذ العملي وجدول مقارنة الخوارزميات مع آلية عمل كل منها وفق النموذج المقترح، والقسم السادس يحتوي على أهم الاستنتاجات والتوصيات.

2- الهدف من البحث:

نسعى من خلال الدمج ما بين النموذج الرياضي لمريض السكري مع خوارزميات الذكاء الصناعي إلى تحقيق الأهداف التالية:

- A. دراسة داء السكري كونه من الامراض البيولوجية الواسعة الإنتشار.
- B. استخدام تقنيات الذكاء الصناعي للوصول للقيم المثلى للمتحكم المستخدم في النموذج الرياضي لمريض السكري.

3- الدراسات المرجعية:

- عام 2015 قامت الباحثة شذا شحادة في بحثها (دراسة وتصميم نظام تحكم استدلالي يعتمد على الرواصد في فضاء الحالة). باستخدام طريقة الاستدلال المعتمد على

الرواصد لتخمين المتغيرات القابلة للقياس لنظام الغلوكوز-أنسولين في الجسم البشري الحي ومن ثم إجراء عملية التحكم في فضاء الحالة للحفاظ على مستوى تركيز الغلوكوز عند القيمة المرجعية في الدم.

- في عام 2018 قام الباحثان امير راشد صبكان وأمجد عبود عبد الجليل في بحثهم (نموذج رياضي للكشف عن داء السكر في الدم)، باستخدام النماذج الرياضية التي تساعد على تحديد وتشخيص الأشخاص المصابين بداء السكري من خلال توصيف المعادلات الرياضية في تحديد نسبة الغلوكوز في الدم وكذلك جرعة الأنسولين التي يحتاجها هؤلاء الأشخاص.

- كما استخدام الباحثون مضخة أنسولين لضخ الأنسولين اللازم بشكل مستمر إلى الدم عبر الجلد وتم استخدام متحكم ضبابي نوع Mamdani واستخدام الخوارزمية الجينية لاختيار المدخل الأفضل إلى نظام التحكم [1].

وقاموا بدراسة البنكرياس الصناعي وتطبيق خوارزميات التحكم الثلاث المعتمدة على PID, FLC, MPC لكن دون التطرق لأي نوع من خوارزميات الذكاء الصناعي الممكن استخدامها مع هذا النظام [2].

- بالإضافة إلى استخدام نظام الحلقة المغلقة الذي يعتمد على المتحكم الضبابي نوع Mamdani للوصول لمستوى سكر الدم لدى الشخص السليم بالاعتماد على وحدة التحكم المقترحة، وخوارزمية التحسين ABC [3].

- تم تصميم جهاز تحكم الحلقة المغلقة لمرضى السكري من النوع الأول للحفاظ على مستوى غلوكوز الدم عند الحد الطبيعي باستخدام متحكم PD وخوارزمية PSO [4].

- أيضاً التنبؤ بسكر الدم لدى مرضى السكري من النوع الاول وتوقع نسبة السكر بشكل دقيق لخدمة الأجهزة الجديدة القابلة للارتداء والتي تتطلب المزيد من المعلومات الدقيقة. تستخدم هذه الدراسة خوارزمية تحديد المدخلات المتسلسلة (SISAL) والتي تم إعدادها خصيصاً لبيانات السلاسل الزمنية. هنا يمكن استخدام الأفكار كاستراتيجيات لاختيار البيانات والتنبؤ بسكر الدم والعمل على توفر القوة الحسابية والسرعة المطلوبة والدقة

المطلوبة. مع العلم أن التنبؤ بسكر الدم هو الخطوة الأولى نحو البنكرياس الصناعي [5].

- تم اقتراح نظام لعلاج مرض السكري من النوع الثاني حيث تم تقديم علاج لمرض السكري من النوع الثاني بالاعتماد على الخوارزميات الجينية (GA)، وإنشاء مجموعة من الكروموسومات التي يتكون كل منها من شجرة القرار الثنائية وقد أثبت هذا النظام كفاءته بنسبة جيدة وتمكن من مساعدة الأطباء في العثور على علاج لمرض السكري من خلال الاختبارات لتحديد العلاج المناسب [6].

4- الطرائق المستخدمة:

في البحث المطروح سوف نقدم خوارزميات عدة مستخدمة في إيجاد الحل الأمثل لبارامترات المتحكم، وهذه الخوارزميات هي الخوارزمية الجينية، خوارزمية النحل، خوارزمية محاكاة التلدين، خوارزمية سرب الطيور. أما بالنسبة للنموذج الرياضي المستخدم فهو النموذج الأصغري لبييرغمان.

4-1 خوارزميات الحل الأمثل (المعايرة):

4-1-1 الخوارزمية الجينية (GA):

هي خوارزمية أمثلة عديدة مستوحاة من الاختيار الطبيعي وعلم الجينات، وتعتبر هذه الخوارزمية عامة أي يمكن تطبيقها على مجال واسع من المشكلات.

تتميز الخوارزمية الجينية عن الطرائق التقليدية بالآتي:

1. تتعامل مع البارامتر ككود وليس مع البارامتر نفسه.
2. تبحث عن مجموعة نقاط وليس عن نقطة واحدة.
3. تستخدم في إيجاد حلولها على توابع أهداف موضوعية وليس على معلومات إضافية.

من فوائد الخوارزمية الجينية مايلي:

1. لا داعي لمعرفة قواعد المشكلة لأنها تتعامل مع قواعد داخلية خاصة بها.

2. المقترحات السلبية لا تؤثر على الحل النهائي بشكل سلبي لأنها ببساطة يتم تجاهلها.
3. لها القدرة على التعامل مع المشاكل المتعددة لإيجاد حلول مثالية.

4-1-2 خوارزمية محاكاة التلدين (Simulated Annealing):

هي طريقة لحل مشاكل إيجاد الحل الأمثل المقيدة وغير المقيدة، هذه الطريقة تحاكي العملية الفيزيائية لتسخين المعادن وبعدها تخفيض درجة حرارتها ببطء من أجل تقليل العيوب، وبالتالي تقليل ضياعات النظام. تعمل الخوارزمية حتى تصبح قيمة التغير المتوسط في تابع الهدف أقل من قيمة التابع (tolerance function). تتوقف الخوارزمية عندما تصبح أفضل قيمة لتابع الهدف أقل أو تساوي قيمة الهدف (Objective Limit). تتوقف الخوارزمية عندما يتجاوز عدد التكرارات العدد الأعظمي من التكرارات (maximum number of iterations)[8] .

4-1-3 خوارزمية النحل (ABC):

خوارزمية مستوحاة من سلوك النحل في البحث عن الطعام، إذ يعمل النحل بشكل تعاوني لجمع أكبر كمية من الطعام، يتم تنفيذ هذه العملية بشكل غريزي بدون آلية تحكم مركزية. يتم تصنيف النحل إلى مجموعات:

- النحل الكشاف بمهمة إيجاد مصدر غذاء جديد.
- النحل العامل يأخذ معلومات من النحل الكشاف حول جودة موقع الغذاء ويقوم هذا النوع بمجموعة تجارب ويختار مصدر الغذاء الأفضل.

4-1-4 خوارزمية سرب الطيور (PSO):

عبارة عن خوارزمية حديثة مستوحاة من السلوكيات في الطبيعة مثل تدفق الطيور، على

عكس الخوارزمية الجينية لا يوجد بها طفرات وهي أقل عرضة للوقوع في شرك الحدود الدنيا المحلية ولها كفاءة حسابية جيدة. تبدأ الخوارزمية بمجموعة أولية من الحلول العشوائية تسمى الطيور. تتوافق الطيور مع الكروموسومات في GA، كل جسيم له موقع وسرعة بمثابة حل مرشح للمشكلة يتم الاحتفاظ بأفضل وضع سابق، أفضل جسيم من بين جميع الجسيمات في المجتمع هو أفضل قيمة إجمالية [9].

5- نموذج بيرغمان المستخدم:

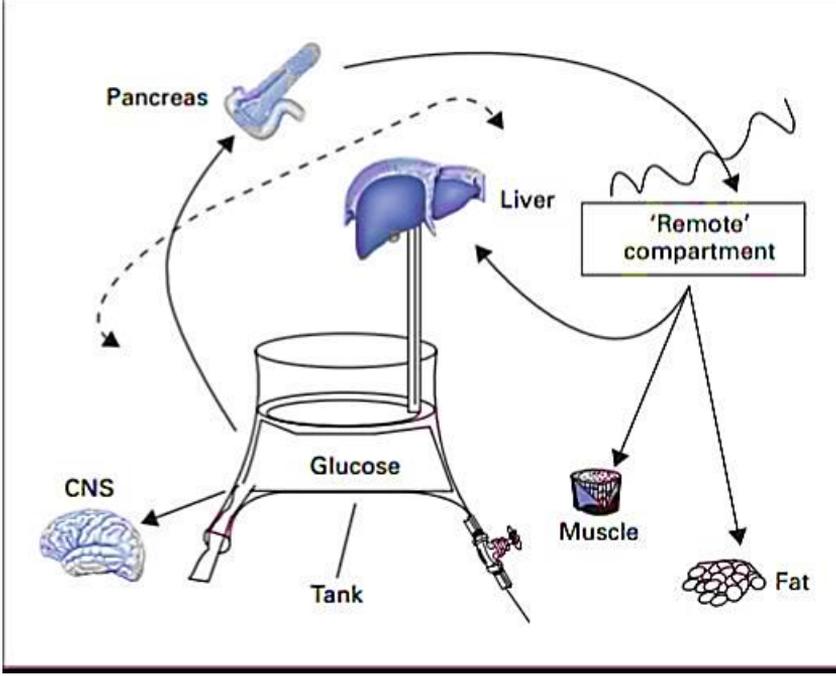
يوصف نموذج بيرغمان البسيط الحيز الذي يتضمن التركيز الأساسي للأنسولين. ومستوى الغلوكوز في الجسم.

وهو يتألف من نموذجين بسيطين:

الأول يصف حركية الغلوكوز أي كيفية تفاعل الغلوكوز مع الأنسولين في الدم، بينما الثاني يصف حركية الأنسولين أي كيفية تفاعل تركيز الأنسولين مع غلوكوز الدم عن طريق أخذ بيانات الأنسولين والغلوكوز كمدخلات. يرتفع مستوى الغلوكوز بعد تناول الطعام ويتم تحديد هذا المستوى من خلال الموازنة بين معدل الاستهلاك ومعدل الإنتاج. وبشكل أكثر تحديداً يتم تمثيل إفرازات البنكرياس من الانسولين الذي بدوره يقوم بتنظيم الغلوكوز في الدم.

المشكلة الرئيسية لمريض السكري هي أنه ليس لديه القدرة على إنتاج الأنسولين وخاصة مريض السكري من النوع الأول، لذلك يتعين علينا المراقبة والحقن مرات عدة في اليوم للمساعدة في التحكم بمستوى الغلوكوز لديه.

مريض سكر الدم يحتاج إلى مراقبة مستمرة لأنه على المدى الطويل وإذا حدث إهمال يمكن أن يصاب بالعمى أو بعض مشكلات القلب والأوعية الدموية، كما يمكن أن يسبب ارتفاع السكر في الدم مشكلات أخرى مثل غيبوبة السكري أو الإغماء.



الشكل (1): نموذج بيرغمان الأصغري. [4]

بسبب هذه المشاكل، توصل المهندسون إلى مزيد من التكنولوجيا المتقدمة للتحكم بالجلوكوز في نظام الحلقة المغلقة، حيث تحتوي التكنولوجيا الحالية على مضخة خارجية مع خزان يحتوي على الأنسولين ويحقن الأنسولين بشكل مستمر اعتماداً على قراءات الجهاز، هذه التقنية مناسبة للمريض الذي يرغب في حقن الأنسولين عدة مرات في اليوم. والعمل مستمر لتطوير مستشعر الجلوكوز ومضخة الأنسولين القابلة للزرع.

يقوم هذا المستشعر بقياس مستوى الجلوكوز وإرساله إلى وحدة التحكم التي تكون في حجم ساعة المعصم، وحدة التحكم ترسل إشارة إلى مضخة حقن الأنسولين، تعمل هذه الحلقة المغلقة كبنكرياس صناعي لمرضى السكري.

يتم اقتتران النموذج الأصغري لبيرغمان مع بعض التعديلات والإضافات. ويشار إلى هذا النموذج المعدل بالنموذج الأساسي ويتم تمثيله بالمعادلات التفاضلية الآتية:

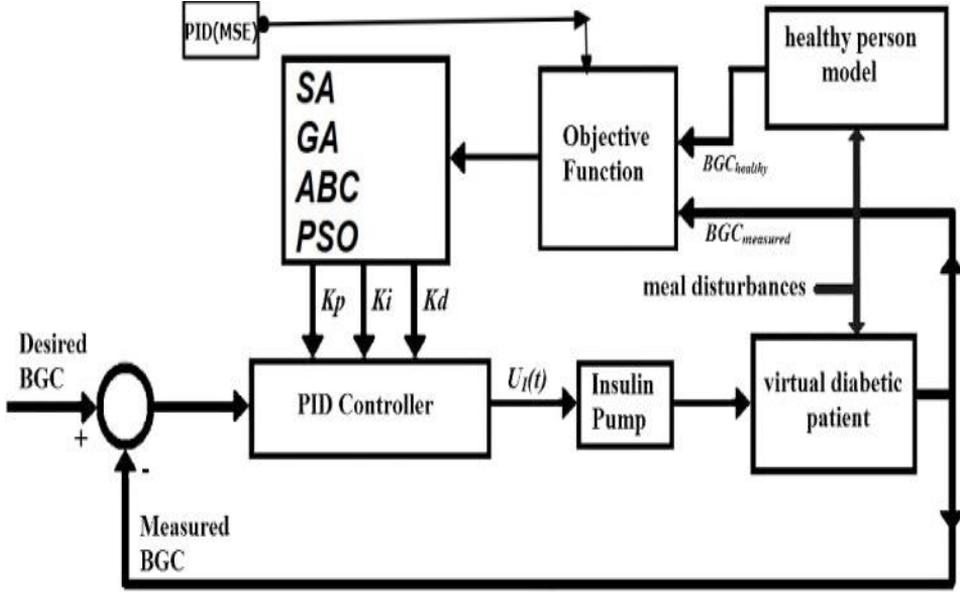
$$\begin{aligned} \frac{dG(t)}{dt} &= -(p_1 + X(t))G(t) + p_1 G_b + D(t) & G(0) &= G_0 \\ \frac{dX(t)}{dt} &= -p_2 X(t) + p_3 (I(t) - I_b) & X(0) &= X_0 \end{aligned}$$

$$\begin{aligned} \frac{dI(t)}{d(t)} &= -p_4 I(t) + \frac{U(t)}{V_I} & I(0) &= I_0 \\ \frac{dD(t)}{d(t)} &= -d \text{rate} \cdot D(t) & D(0) &= D_0 \\ \frac{dG_{sc}(t)}{d(t)} &= \frac{G(t) - G_{sc}(t)}{5} - R_{utln} & G_{sc}(0) & \\ &= G(0) - 5 \cdot R_{utln} & & \end{aligned}$$

الجدول (1): بارامترات النموذج الرياضي الأساسي.

الرمز	الواحدة	الوصف
$G(t)$	[mg/dL]	تركيز الجلوكوز في الدم
$X(t)$	[1/min]	تأثير الأنسولين النشط
$I(t)$	[mU/L]	تركيز الأنسولين في الدم
$D(t)$	[mg/dL/min]	وظيفة اضطراب الوجبة
$G_{sc}(t)$	[mg/dL]	تركيز الجلوكوز تحت الجلد
$U(t)$	[mu/min]	الأنسولين الخارجي
G_b	[mg/dL]	تركيز جلوكوز الدم القاعدي
I_b	[mU/L]	تركيز الأنسولين القاعدي
V_I	[L]	حجم تجمع توزيع الأنسولين
p_1	[1/min]	معدل تصفية الجلوكوز بشكل مستقل عن الأنسولين
p_2	[1/min]	معدل تصفية الأنسولين النشط
p_3	[L/(min ² mU)]	زيادة في قدرة الامتصاص ناتجة عن الأنسولين
p_4	[1/min]	معدل استهلاك الأنسولين في الدم
D_{rate}	[1/min]	معدل الاستهلاك لاضطراب الوجبة
R_{utln}	[1/min]	معدل الاستخدام (الفرق بين التركيزين تركيز الجلوكوز في الدم وتركيز الجلوكوز تحت الجلد في الحالة المستقرة)

5-1 المخطط الصندوقي للخوارزميات مع المتحكم PID:



الشكل (2): المخطط الصندوقي للخوارزميات مع المتحكم.

5-2 اختيار تابع التقييم أو الهدف (objective function):

هدفنا الرئيس أثناء تصميم وحدة التحكم (PID) هو تقليل الخطأ في هذه الوحدة بالاعتماد على النموذج الرياضي لمرضى السكري.

يمكن استخدام العديد من توابع التقييم أو الهدف لتحسين المجال الزمني لوحدة التحكم (PID) في هذا البحث سوف يتم استخدام متوسط الخطأ التربيعي.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

حيث Y_i هي القيمة الحقيقية.

و \hat{Y}_i هي القيمة المقاسة.

وهذا يحقق نتائج جيدة رياضياً للمتحكم، لكن فيزيائياً نحن بحاجة لتابع تقييم إضافي من أجل إعطاء أفضل قيمة لبارامترات المتحكم من خلال مقارنة تركيز الجلوكوز بين المريض الافتراضي وشخص سليم، وهو متوسط الانحراف المعياري:

$$MAPE = \frac{1}{n} \sum_{t=0}^n \left| \frac{G_h - G_{measured}}{G_h} \right|$$

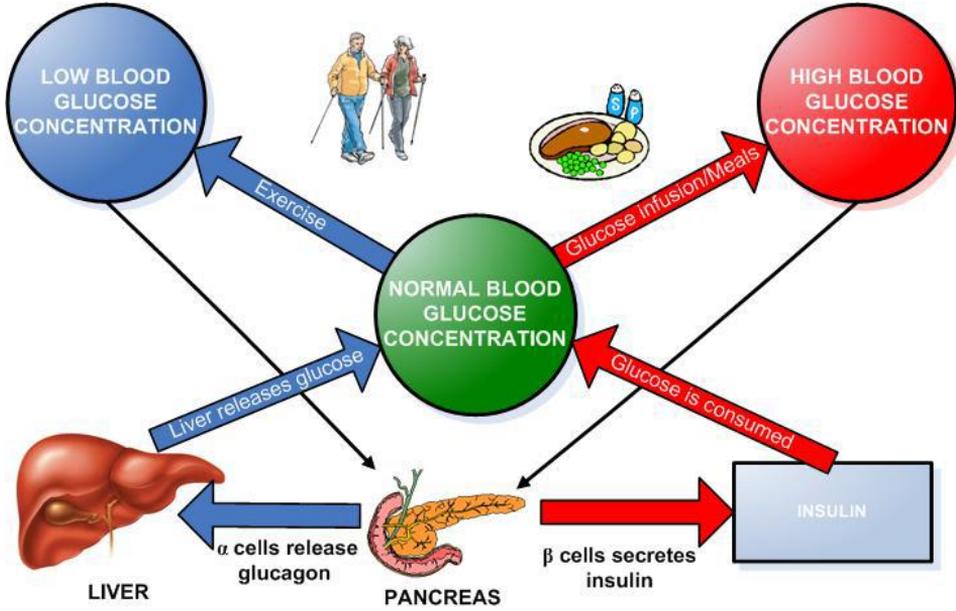
حيث G_h تركيز الغلوكوز للشخص السليم.

و $G_{measured}$ تركيز الغلوكوز للشخص المصاب بالسكري (المريض الافتراضي). والفرق بين الاثنين هو أنه الأول يحسب الخطأ للمتحكم نفسه أما الثاني فليس له علاقة بالمتحكم عندما تصل النسبة المئوية للخطأ إلى أدنى قيمة تتوقف الخوارزمية ويتم تثبيت قيم المتحكم.

6- نظام الغلوكوز والأنسولين في الدم:

يعد نظام الغلوكوز والأنسولين مثالاً على النظام الفسيولوجي ذي الحلقة المغلقة. عند الشخص السليم عادة ما يكون تركيز الغلوكوز في الدم عنده 70-110 ملغ/دسليتر. يساعدنا نظام الغلوكوز والأنسولين في الحفاظ على هذه الحالة المستقرة. في الشل التالي يتم وصف بسيط لنظام الغلوكوز والأنسولين. الشخص السليم في المنطقة الخضراء، إذ يكون تركيز الغلوكوز في الدم طبيعياً إذا تناول الشخص بعد ذلك غلوكوز إضافياً، على سبيل المثال عن طريق إحدى الوجبات، فإن الشخص ينتقل إلى المنطقة الحمراء، مع ارتفاع تركيز الغلوكوز في الدم.

مع ارتفاع تركيز الغلوكوز في الدم. وعندما يحدث هذا يتم إرسال إشارة إلى البنكرياس، حيث تتفاعل الخلايا نوع بيتا عن طريق إفرازها هرمون الأنسولين. يزيد هذا الأنسولين من امتصاص الخلايا للغلوكوز، وذلك يعيد الشخص إلى المنطقة الخضراء. إذا كان غلوكوز الدم التركيز أقل من المستوى الطبيعي، يكون الشخص في المنطقة الزرقاء. هذا وصف بسيط للغاية لنظام أكثر تعقيداً. لكن هذه هي الطريقة المبسطة لشرح التمثيل الغذائي، والذي سيتم تقديمه في نموذج رياضي في هذه الأطروحة.



الشكل (3): نموذج الغلوكوز والأنسولين.

7- التطبيق العملي:

تمت محاكاة نظام الغلوكوز والأنسولين بعد تطبيق خوارزميات الذكاء الصناعي المختارة، وتحديد بارامترات المتحكم $(T_d \ T_i \ K_c)$ بالشكل الأمثل، إذ تم استخدام كل من الخوارزمية الجينية وخوارزمية النحل وخوارزمية محاكاة التلدين وخوارزمية سرب الطيور في إيجاد القيم الأفضل لبارامترات المتحكم وفق النموذج المقترح لنظام الغلوكوز والأنسولين.

يوضح الجدول (2) القيم المختلفة لبارامترات الخوارزميات الأربعة.

الجدول (2): القيم المختلفة لبارامترات الخوارزميات المستخدمة.

GA		PSO		ABC		SA	
Parameters	Value	Parameters	Value	Parameters	Value	Parameters	Value
Generation Size	100	Maximum Iteration	60	Maximum Cycle	60	Maximum Iteration	60

Population Size	1000	Size of the swarm	1000	Size of the colony	1000	Initial Temperature	1000
Selection Method	Roulette Wheel	Inertia weight	[0.9-0.4]	Selection Method	Roulette Wheel	-	-

مقارنة النتائج:

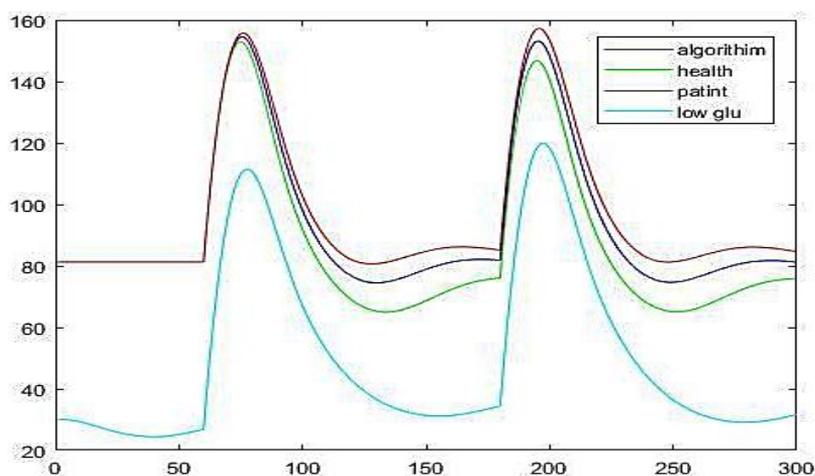
تمت مقارنة أداء الخوارزميات المذكورة على نظام الجلوكوز والأنسولين، وتوضيح النتائج عن طريق الجداول والرسوم البيانية الآتية.

نتائج المحاكاة باستخدام خوارزمية محاكاة التلدين:

الجدول (3): مقارنة بين القيم الصغرى والعظمى لجلوكوز الدم عند استخدام

خوارزمية محاكاة التلدين.

	minimum BGC (mg/dL)	maximum BGC (mg/dL)
Healthy person	64.9600	152.9458
Patient-PID	80.6431	157.3126
SA-PID	72.28	155.59

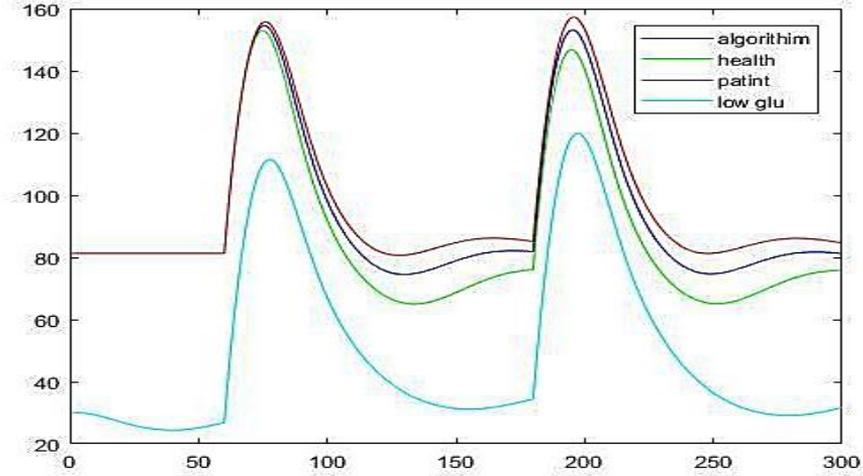


الشكل (3): المحاكاة عند استخدام خوارزمية محاكاة التلدين.

نتائج المحاكاة باستخدام خوارزمية سرب الطيور:

الجدول (4): مقارنة بين القيم الصغرى والقيم العظمى لغلوكوز الدم عند تطبيق خوارزمية سرب الطيور.

	minimum BGC (mg/dL)	maximum BGC (mg/dL)
Healthy person	64.9600	152.9458
Patient-PID	80.6431	157.3126
PSO-PID	73.97	156.13

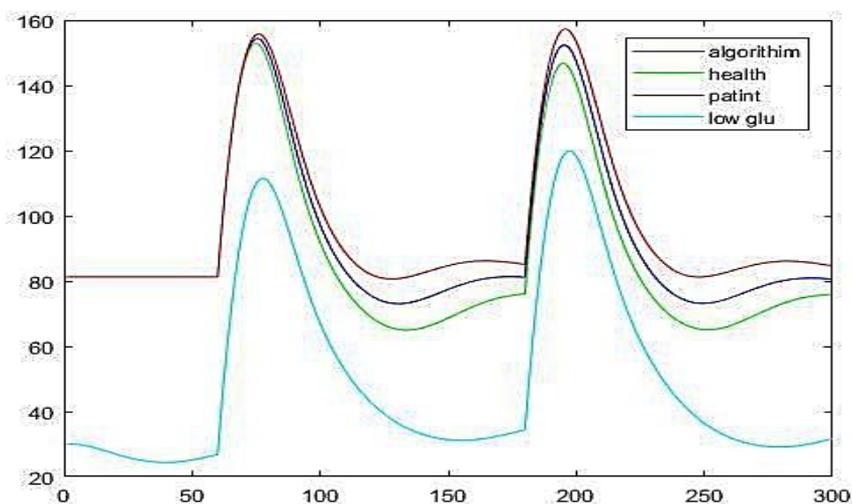


الشكل (4): المحاكاة عند استخدام خوارزمية سرب الطيور.

نتائج المحاكاة باستخدام الخوارزمية الجينية:

الجدول (5): مقارنة بين القيم الصغرى والعظمى لغلوكوز الدم عند تطبيق الخوارزمية الجينية.

	Minimum BGC (mg/dL)	maximum BGC (mg/dL)
Healthy person	64.9600	152.9458
Patient-PID	80.6431	157.3126
GA-PID	74.72	156.13

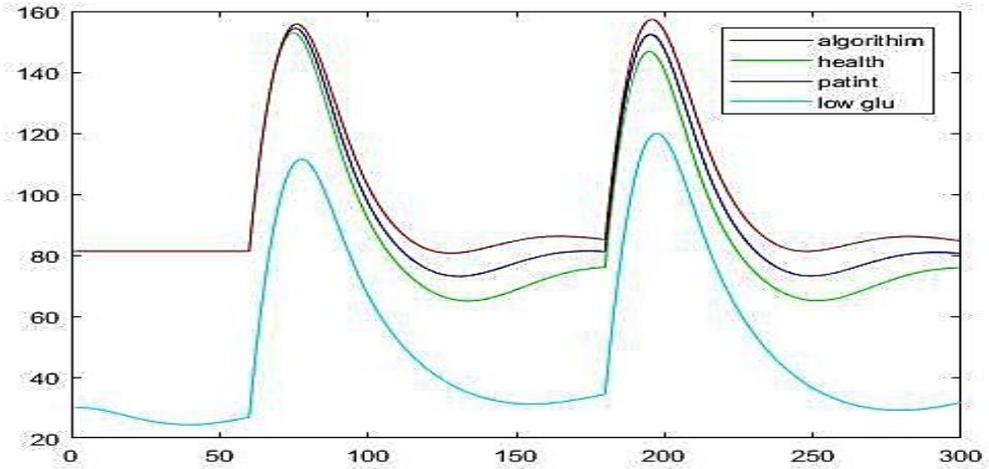


الشكل (5): المحاكاة بعد تطبيق الخوارزمية الجينية.

نتائج المحاكاة باستخدام خوارزمية النحل:

الجدول (6): مقارنة بين القيم الصغرى والعظمى للغلوكوز الدم عند تطبيق خوارزمية النحل.

	minimum BGC (mg/dL)	maximum BGC (mg/dL)
Healthy person	64.9600	152.9458
Patient-PID	80.6431	157.3126
ABC-PID	72.32	155.65



الشكل (6): المحاكاة بعد تطبيق خوارزمية النحل.

يبين الجدول (7) مقارنة بين القيم العظمى والصغرى للغلوكوز لشخص سليم مع شخص مصاب بمرض السكري ومقارنتهم مع نتائج الخوارزميات المستخدمة في هذا البحث.

الجدول (7) مقارنة بين الخوارزميات.

	minimum BGC (mg/dL)	maximum BGC (mg/dL)
Healthy person	60	180
Patient	115.46	313.02
GA-PID	74.72	156.13
PSO-PID	73.97	156.01
ABC-PID	72.32	155.65
SA-PID	72.28	155.59

نلاحظ تفوق خوارزمية محاكاة التلدين (SA) في المحافظة على القيم الصغرى والعظمى لتركيز الجلوكوز في الدم يليها خوارزمية النحل (ABC)، ثم خوارزمية سرب الطيور (PSO)، وأخيرا الخوارزمية الجينية (GA).

الجدول (8): الزمن اللازم لكل خوارزمية لتنفيذ المهمة المطلوبة.

PSO	GA	SA	ABC	Iteration
25.3sec	52.5sec	8.75sec	37.6sec	100

نلاحظ من الجدول (8) أن خوارزمية محاكاة التلدين (SA) هي الأسرع من حيث زمن التنفيذ.

7- الاستنتاجات والتوصيات:

- 1- تسهل هذه الدراسة التحليلية استخدام خوارزميات الذكاء الصناعي مع وحدة التحكم PID والنموذج الرياضي لمريض السكري، ويكون عمل الباحثين في اختيار الخوارزمية الأنسب، أو الدمج بين خوارزميتين بهدف التحسين من خلال ميزات كل خوارزمية.
- 2- تمتاز خوارزمية محاكاة التلدين (SA) وخوارزمية النحل (ABC) في المحافظة على القيم الصغرى والعظمى لتركيز الجلوكوز في الدم.
- 3- إن خوارزمية SA أسرع من خوارزمية GA و ABC و PSO.

8- الخاتمة:

قمنا في هذا البحث بضبط بارامترات المتحكم المستخدم مع النموذج الرياضي لمرضى السكري، وذلك من خلال الاستفادة من خصائص خوارزميات الذكاء الصناعي (المعايرة) لتوليد متحكم يمتلك أهم خصائص الاستقرار، وذلك وفقاً لكل خوارزمية من الخوارزميات المستخدمة (ABC, PSO, SA, GA) وحساب الزمن اللازم لكل منها أثناء المحاكاة.

9- الآفاق المستقبلية:

يمكن تطوير هذا البحث والعمل عليه من خلال الدمج بين اثنتين من الخوارزميات أو أكثر للاستفادة من الإيجابيات في كل خوارزمية. فالخوارزميات تساعد في الحصول على أصغر نسبة للخطأ عن طريق ضبط المتحكم، وبالتالي تحسين النظام ككل والحصول على نتائج أكثر دقة وموثوقية. كما نلاحظ خوارزمية محاكاة التلدين التي طبقناها لأول مرة على هذا النظام امتازت بدقتها، على أمل الاستفادة منها أو دمجها مع خوارزمية أخرى وتطبيقها مع البنكرياس الصناعي مستقبلاً. كما يمكن التركيز على نقاط أخرى لتحسينها في النظام سواء في المتحكمات أو مضخة الأنسولين، على أمل انتشار الجهاز وتطبيقه في وقت ليس ببعيد وحصول مريض السكري على حياة طبيعية دون الخوف من مضاعفات هذا المرض، إذ سيتمكن الجهاز من المراقبة المستمرة وضبط سكر الدم بطريقة سهلة ودقيقة مدى حياة المريض.

10- المراجع:

- [1] Basher, A. S. (2017). design fuzzy control system for blood glucose level for type-1 diabetes mellitus patients using GA a simulation study.
- [2] Mehmood, S., Ahmad, I., Arif, H., Ammara, U. E., & Majeed, A. (2020). Artificial Pancreas Control Strategies Used for Type 1 Diabetes Control and Treatment: A Comprehensive Analysis. *Applied System Innovation*, 3(3), 31.
- [3] Soylu, S., & DANIŞMAN, K. (2018). Blood glucose control using an ABC algorithm-based fuzzy-PID controller. *Turkish Journal of Electrical Engineering & Computer Sciences*, 26(1), 172-183.
- [4] Malik, S. F. J. M. (2018). Particle swarm optimization based blood glucose control in type1 diabetes (Doctoral dissertation, Kauno technologijos universitetas).
- [5] Rodríguez-Rodríguez, I., Rodríguez, J. V., González-Vidal, A., & Zamora, M. Á. (2019). Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL). *Symmetry*, 11(9), 1164.
- [6] Al Switi, M., Alshraideh, B., Alshraideh, A., Massad, A., & Alshraideh, M. (2019). Treatment of Diabetes Type II Using Genetic Algorithm.
- [7] F. Aguiree, et al., *IDF Diabetes Atlas: sixth edition*. Switzerland: International Diabetes Federation, 2013.
- [8] Varty, Z. (2017). Simulated annealing overview.
- [9] Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: an overview. *Soft Computing*, 22(2), 387-408.

