مجلة جامعة تشرين للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (40) العدد (4) 2018
Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (40) No. (4) 2018

تعزيز الاكتشاف والتجميع الآلي لخدمات الويب باستخدام الانطولوجيا و WordNet

الدكتور يوسف دلاّ*
الدكتور أنور أبو نقطة*
ثراء اصلان**

(تاريخ الإيداع 12 / 2 / 2018. قُبِل للنشر في 2/ 7 / 2018)

□ ملخّص □

ازدادت أهمية خدمات الويب Web Services بعد ظهور الحوسبة السحابية Cloud Computing و انتشار الحوسبة خدمية التوجه (Service Oriented Computing (SOC التي غيرت طريقة تطوير البرمجيات. بسبب التطور الكبير الحاصل في هذا المجال أصبح من الصعب على المستخدمين القيام بتجميع خدمات الويب يدوياً.
يعرض البحث نظاماً مقترحاً للتجميع الآلي لخدمات الويب بالاستفادة من الدلالات المرتبطة به، تطلب التجميع الآلي التعامل مع اكتشاف خدمات الويب، و تجميع الخدمات التي تم اكتشافها. بدايةً تم اقتراح طريقة آلية لاكتشاف خدمات الويب بالاعتماد على الانطولوجيا، حيث عُرفت هرمية للمفاهيم المرتبطة بالمجال ضمن ملف انطولوجيا المجال، ثم استخدمت في توصيف دلالات واجهات الخدمات (الدخل و الخرج) و حساب درجة التشابه فيها بينها، كما عُرفت انطولوجيا تُمثل هرمية التصنيف تحوي مفاهيم مجالات مختلفة، من أجل حل مسألة عدم التجانس و تقليل عدد الخدمات المكتشفة عن طريق حصرها ضمن مجال الطلب نفسه. كما اقتُرحت خوارزمية لتجميع خدمات الويب آلياً، و استُخدمت كلفة تنفيذ الخدمة كمعيار غير وظيفي يعبر عن جودة الخدمات التي سيتم تجميعها، من أجل الحصول على حل أمثلي يرضي المستخدم. كذلك تم العمل على تقليل زمن التجميع عن طريق تحميل ملفات الانطولوجيا عند أول طلب فقط.
تم اختبار النظام من أجل التحقق من دقة النتائج و فعالية الأداء وذلك باستخدام معايير الحساسية Recall و الدقة Precision و F-measure إضافة إلى زمن التجميع.

الكلمات المفتاحية : خدمات الويب، الاكتشاف الآلي، التجميع الآلي، الانطولوجيا، التشابه، WordNet.

* مدرس–قسم البرمجيات ونظم المعلومات– كلية الهندسة المعلوماتية – جامعة تشرين– اللاذقية– سورية
** طالبة دراسات عليا(ماجستير) قسم البرمجيات ونظم المعلومات – كلية الهندسة المعلوماتية –جامعة تشرين– اللاذقية – سورية.
Tharaa87asslan@gmail.com

مجلة جامعة تشرين للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (40) العدد (4) 2018
Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (40) No. (4) 2018

Enhance Automatic Web Service Discovery and Composition Using ontology and WordNet

Dr.Yousef Dalla*
Dr. Anwar Abu Nuktah*
Tharaa Asslan**

(Received 12 / 2 / 2018. Accepted 2 / 7 / 2018)

□ ABSTRACT □

Service Oriented Computing (SOC) is changing the way of developing software systems. Each web service has a specific purpose to serve, so it can not satisfy users' request.
In this paper, we propose a Web services composition method based on OWL ontology, and design an automatic system model for services discovery and composition. This method uses domain ontology and WordNet to calculate matching between input and output parameters and uses Category ontology to solve the problem of semantic heterogeneity in web service description. We use services with single input and single output and cost as QoS criteria. This method can enhance the efficiency and accuracy of service composition, and the experiments are used to validate and analyze the proposed system.

Keywords: Web service, Discovery, composition, ontology, matching, WordNet.

* **Associate professor, Department of Software and Information Systems, Faculty of Informatics Engineering, Tishreen University, Lattakia, Syria.**
Postgraduate student (Master), Department of Software and Information Systems, Faculty of Informatics Engineering, Tishreen University, Lattakia, Syria. Tharaa87asslan@gmail.com

مقدمة :

قامت منظمة المعايير الدولية الرئيسية للشبكة العالمية (W3C) World Wide Web Consortium بتعريف خدمة الويب على أنها " نظام برمجي مصمم لدعم العلاقات البينية من آلة إلى أخرى عبر الشبكة" [1]. تعتبر خدمات الويب مكونات برمجية موصوفة ذاتياً ومستقلة عن المنصات يمكن أن تنشر و يتم استدعاؤها عبر الويب[2]. تكون كل خدمة مخصصة للقيام بوظيفة محددة وغالباً ما تكون غير كافية وحدها لتلبية حاجة المستخدم، لذلك ظهرت الحاجة إلى تجميع عدة خدمات كي تؤدي وظائف معقدة (مسألة تجميع خدمات الويب Web Service Composition (WSC)). فأصبحت تقنية تجميع الخدمات آلياً واحدة من المجالات الرئيسية في عملية تطوير التطبيقات.

تتلخص مزايا تجميع خدمات الويب بزيادة سرعة تطوير التطبيقات، السماح بإعادة استخدام الأكواد، السماح بالوصول إلى خدمات ويب مختلفة [3].

عمل الباحثون على بناء منظومات لتجميع خدمات الويب إما يدوياً(Manual/Static Composition)، أو آلياً (Automatic/Dynamic Composition)[2]، أو بشكل نصف آلي(Semiautomatic/dynamic or static Composition).

تصنف نماذج تجميع خدمات الويب إلى أربعة أنواع [4]. تجميع خدمات الويب بالاعتماد على مخططات سير العمل workflow-based. تجميع خدمات الويب بالاعتماد على الذكاء الصنعي planning-based (AI) . تجميع خدمات الويب بالاعتماد على المفاهيم الدلالية semantic- based . تجميع خدمات الويب بالاعتماد على البيان graph-base [3].

قام الباحثون في[2] بتصميم نظام لتجميع الخدمات آلياً بعد توصيف عملياتها دلالياً بالاعتماد على الانطولوجيا. كذلك قدم الباحثون في [5] طريقة للتجميع الآلي لخدمات الويب باستخدام الانطولوجيا و مرادفات WordNet ، كما حلوا مشكلة عدم التجانس الدلالي عند توصيف الخدمات، و اتبعوا منهجية للتجميع انطلاقاً من خرج الطلب إلى دخله (من اليمين إلى اليسار). لكن البحث لم يهتم بمعايير جودة الخدمة كما لم يعرض أي اختبارات تثبت فاعلية النظام ودقته. اقترح البحث [6] طريقة لتجميع الخدمات بالاستفادة من بنية الأشجار الثنائية و الانطولوجيا لتمثيل العلاقات الدلالية بين واجهات الخدمات. جميع الدراسات السابقة لم تأخذ الخصائص غير الوظيفية للخدمات بعين الاعتبار. في حين عمل الباحثون في [7] على تجميع خدمات الويب اعتماداً على خصائص جودة الخدمة QoS فقط عن طريق حساب semantic similarity بين الخدمات و مقارنته مع عتبة معينة لقبول الخدمة.

تم في البحث اقتراح نظام آلي لتجميع خدمات الويب باستخدام الانطولوجيا. يتيح النظام إمكانية الاستفادة من العلاقات الدلالية بين واجهات الخدمات I/O matching إضافة إلى المرادفات WordNet من أجل اختيار الخدمات و تجميعها. يسمح النظام بالتعامل مع خدمات الويب التي يتم نشرها حديثاً، كما تمت إضافة تصنيف الخدمة كحل لمشكلة عدم التجانس الدلالي. اعتُمدت تكلفة تنفيذ الخدمة كمعيار لجودة الخدمة من أجل زيادة أمثلية الحل. استخدمت خدمات ويب وحيدة الدخل والخرج لاختبار النظام المقترح و تقييمه و اقتراح بعض التحسينات عليه. تم تجميع الخدمات انطلاقاً من دخل الطلب وصولاً إلى خرجه (من اليسار إلى اليمين). أخيراً تم عرض النتائج التي تبين دقة النظام في اكتشاف خدمات الويب، وفاعليته من حيث سرعة التجميع للوصول إلى المطلوب.

أهمية البحث وأهدافه:

مازالت مسألة تجميع خدمات الويب آلياً من أكثر المسائل المعقدة بالرغم من الجهود الكبيرة المبذولة في هذا المجال لعدة أسباب منها الزيادة الضخمة في عدد خدمات الويب الموجودة على الانترنت و تخصص كل منها بوظيفة محددة، إنشاء وتحديث الخدمات في الزمن الحقيقي مما يستدعي امتلاك نظام التجميع القدرة على اكتشاف التعديلات في زمن التنفيذ وإجراء اللازم اعتماداً على المعلومات الجديدة، إضافة إلى تعدد الجهات التي تقوم بنشر الخدمات وعدم وجود لغة موحدة لتوصيفها و تقييمها.

تتركز أهداف البحث في عدة محاور:

1. بناء نظام آلي لاكتشاف و تجميع خدمات الويب بحيث يعيد سلسلة الخدمات التي تلبي حاجة المستخدم انطلاقاً من دخل الطلب وصولاً إلى خرجه (Left To Right من اليسار إلى اليمين).
2. استغلال الويب الدلالي و انطولوجيا المجال من أجل تفسير الروابط بين المفاهيم التي تمثل متحولات دخل و خرج الخدمات لاكتشاف و اختيار الخدمات المتشابهة دلالياً على مستوى الواجهات، و الاستفادة من انطولوجيا التصنيف لحل مشكلة عدم التجانس الدلالي و البحث ضمن الخدمات التي ترتبط بالطلب فقط.
3. استخدام المرادفات WordNet من أجل تعزيز الاكتشاف الآلي لخدمات الويب. إضافة إلى الاستفادة من جودة الخدمة في الوصول إلى تجميع أمثلي. و العمل على تقليل الزمن اللازم للتجميع.

طرائق البحث ومواده:

تم اتباع المنهج التجريبي في البحث، حيث استخدمت لغة جافا مع (Jena , Pellet) لإنجاز النظام المقترح و محرر الانطولوجيا Protégé من أجل بناء ملفي الانطولوجيا. إضافة إلى تعريف قواعد لحساب درجة التشابه الدلالي بين المفاهيم، كما تم اقتراح عدة معايير لتقييم الأداء. كذلك تم تصميم عدة سيناريوهات لاختبار الأداء انطلاقاً من عدة فرضيات لتوصيف الطلب و الخدمات و الانطولوجيا ، حيث تم اختبار قدرة مرادفات WordNet على تعزيز حساسية و دقة النظام بعد الاستفادة من الانطولوجيا في اكتشاف الخدمات آلياً، إضافة إلى سرعة التجميع.

تجميع خدمات الويب بالاعتماد على الانطولوجيا :

"إن الويب الدلالي Semantic Web هو امتداد لشبكة الويب الحالية لكنه يتميز عنها بأن جميع مكوناته لها معاني معرفة بشكل جيد مما يمكن الحواسيب أو برامج الحاسوب من العمل بشكل تفاعلي مع البشر"[1] . شبكة الويب الدلالي عبارة عن مجموعة شبكات متداخلة مع بعضها البعض تضم الكثير من البيانات و المعلومات التي تم تجهيزها و إعدادها بطريقة آلية و أرشفتها بصورة دقيقة من قبل برامج مختصة تعتمد على تقنيات و إمكانيات الذكاء الصنعي. يهدف حسب رؤية منشئه و منشئ WWW إلى تحويل المعلومات الضخمة المتاحة على الشبكة من أصفار و واحدات (نظام ثنائي) إلى بيانات مفهومة من قبل برامج الحاسوب التي تُنشأ خصيصاً لهذا الغرض.

لا تدعم معظم خدمات الويب التوصيف الدلالي لذلك لجأ العديد من الباحثين إلى استخدام الانطولوجيا Ontology من أجل الوصول إلى دلالات المفاهيم المرتبطة بالخدمات. يمكن توصيف الخدمة دلالياًعلى عدة مستويات إما دلالات

[1] **Tim Berners-Lee, James Hendler and Ora Lassila, The Semantic Web, Scientific American, May 2001**

العمليات أي توصيف سلوك الخدمة، أو دلالات الواجهات (الدخل / الخرج) أو دلالات غير وظيفية مثل التوفر والكلفة و زمن التنفيذ.

الانطولوجيا :

الانطولوجيا في الفلسفة فرع أساسي من الميتافيزيقيا تدرس مبادئ و أسباب الوجود، حيث تهدف إلى تحديد الكيانات Entities و أنواعها، كما تشرح طبيعة الموجودات و خصائصها الجوهرية و العلاقات فيما بينها. الانطولوجيا في علم المعلومات هي مفردات بنيوية مرتبة وفق هرمية من أجل توصيف مفاهيم ومعارف تعود إلى مجال معين.[7]

يجب على الانطولوجيا التي توصف مجالاً معيناً أن تقدم فهم مشترك لهذا المجال، إضافة إلى نماذج بصيغ قابلة للمعالجة من قبل الآلة، كما تساعد على تحليل البيانات في هذا المجال من خلال توفير صيغ معينة للمعارف و العلاقات بينها.

تتكون الانطولوجيا من مفردات تستخدم لتوصيف المجال، توصيف واضح للمعنى المقصود من المفردات، معلومات عن التصنيف المستخدم، إضافة إلى قيود على المجال.

إن Ontology Web Language (OWL) هي اللغة الأكثر شيوعاً من أجل توصيف الانطولوجيا [6]، فهي توفر مجموعة كبيرة من التفاصيل و بالتالي تستطيع توصيف علاقات معقدة بين المفاهيم، حيث يوصى بها من قبل W3C. تتيح OWL استخدام العديد من الأدوات من أجل استنتاج الدلالات والحصول على النتائج اعتماداً على قواعد محددة .

معايير تقييم الأداء:

أكثر المعايير الشائعة في تقييم أداء الأنظمة في استرجاع البيانات هي الحساسية Recall التي تعرف بأنها قدرة النظام على عرض جميع الخدمات المرتبطة بالطلب، و الدقة Precision التي تمثل قدرة النظام على عرض الخدمات المرتبطة بالطلب فقط [8]. كان لا بد من تعريف عدة مصطلحات، حيث يمثل True Positive عدد الخدمات المكتشفة و المرتبطة بالطلب. False Positive يمثل عدد الخدمات المكتشفة من قبل النظام و غير مرتبطة بالطلب. False Negative يمثل عدد الخدمات غير المكتشفة من قبل النظام و المرتبطة بالطلب.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (1)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

يمثل F_measure الوسط التوافقي للقيم وهو أكثر ملائمة من المتوسط الحسابي لأنه أكثر استقراراً في التعامل مع القيم المتطرفة. عند وجود فروقات كبيرة بين القيم يكون الوسط التوافقي أقرب إلى القيمة الأصغر و يعرف بأنه مقلوب المتوسط الحسابي لمقلوب كل قيمة.

$$F_measure = 2 * \frac{Recall * Precision}{Recall + Precision} \quad (3)$$

تعاريف أساسية و هرمية النظام المقترح:

فيما يلي بعض التعاريف الأساسية المرتبطة بمشكلة تجميع خدمات الويب و شرح لهرمية النظام المقترح. بما أن التجميع الفعال لخدمات الويب يتطلب توصيف كل خدمة من أجل السماح للمستخدمين والخدمات الأخرى بفهم تفاصيلها والتعامل معه، لذلك تم وضع مجموعة تعاريف و قواعد لخدمات الويب والمفاهيم المرتبطة بها.

مشكلة تجميع خدمات الويب:

التعريف(1): خدمة الويب {I , O, D, C} = WS حيث I دخل الخدمة ، O يمثل خرجها، D تصنيف الخدمة، C يمثل كلفة تنفيذها و هي معيار غير وظيفي يعبر عن الفترة الزمنية بين تلقي الطلب و البدء بتلقي الاستجابة وهي معيار سلبي أي كلما كانت قيمة C أكبر كلما كانت جودة الخدمة أقل. $O \in DO,\ I \in DO,\ D \in CO$

مثال : WS18 ={ ResearchGroup, Research, Study, 58}

التعريف(2): طلب المستخدم{I , O, D} = WSR حيث I دخل الطلب، O خرجه، D التصنيف المطلوب. إن جميع المتحولات السابقة (دخل I / خرج O) هي مفاهيم لها معاني دلالية.

مثال: WSR= {ResearchGroup, Course, PostGraduated }(*)

التعريف(3): تُعرف الانطولوجيا بالشكل Ontology = { Cset , Rset } حيث Cset مجموعة المفاهيم و Rset العلاقات بينها. انطولوجيا التصنيف (CO) و انطولوجيا المجال (DO) لهما نفس التمثيل. تتضمن العلاقات الدلالية بين المفاهيم مايلي: التوافق Equivalent، الصف الأساسي SuperClassOf، الصف الفرعي SubClassOf والفشل Fail [6].

التعريف(4): من أجل مفهومين Cj و Ci في الانطولوجيا إذا عُرّف Ci على أنه صف متوافق EquivalentClass مع Cj عندها نقول عن Ci و Cj أنهما متوافقان دلالياً $Ci \equiv Cj$. أما إذا عُرّف Cj على أنه صف فرعي SubClassof من Ci عندها نقول عن المفهوم Ci أنه يحوي Cj دلالياً أي $Cj \subseteq Ci$.[7]

يتناول البحث ثلاث درجات للتشابه هي التام Exact اذا كان $Ci \equiv Cj$. الاحتواء Subsumes اذا كان $Cj \supseteq Ci$. التوسيع Plugin اذا كان $Cj \subseteq Ci$.

مثال: المفهومين Lecturer ≡ Professor متوافقان و كلاهما صف فرعي من PostDoc

أي $PostDoc \subseteq Professor$ أو $Professor \supseteq PostDoc$.

التعريف(5): تجميع خدمات الويب WSC هو سلسلة من الخدمات [WS0,WS1,…,WSn] = WSC حيث تكون الخدمات مرتبطة مع بعضها بعلاقات دلالية .

مثال: { WS18, WS15, WS8} = WSC سلسلة خدمات تلبي الطلب (*)

يتم إيجاد سلسلة الخدمات الأمثل و تجميعها لنيل قبول المستخدم بالاعتماد على مجموعة قواعد:

القاعدة(1) : إذا كان $WSR.I \equiv WSi.I \vee WSR.I \supseteq WSi.I \vee WSR.I \subseteq WSi.I$ و $WSR.D \subseteq WSi.D$ تتشابه الخدمة WSi مع الطلب WSR دلالياً على مستوى الدخل .

مثال: الخدمة {WS18 ={ ResearchGroup, Research, Study, 58 تتشابه على مستوى الدخل مع الطلب (*).

القاعدة (2) : إذا كان $WSi.O \equiv WSR.O \vee WSi.O \supseteq WSR.O \vee WSi.O \subseteq WSR.O$ و $WSR.D \subseteq WSi.D$ تتشابه الخدمة WSi مع الطلب WSR دلالياً على مستوى الخرج .

مثال: الخدمة {WS8={ Student, Course, Study, 35 تتشابه على مستوى الخرج مع الطلب (*).

القاعدة (3): إذا كانت الخدمة WSi تتشابه مع الطلب WSR دلالياً على مستوى الدخل و الخرج يؤدي WSR و WSi متشابهتان دلالياً.

القاعدة (4): إذا كانت Wsi و WSj خدمتين وكان $WSi.D \subseteq WSR.D$ و $WSj.D \subseteq WSR.D$ و $WSi.O \equiv WSj.I$ نقول أن الخدمتين WSi و WSj متشابهتان دلالياًعلى مستوى الواجهة.

مثال: الخدمة WS18 = { ResearchGroup, Research, Study, 58} تتشابه مع الخدمة

WS15 = {Research, PostGraduated, PostGraduated, 64}

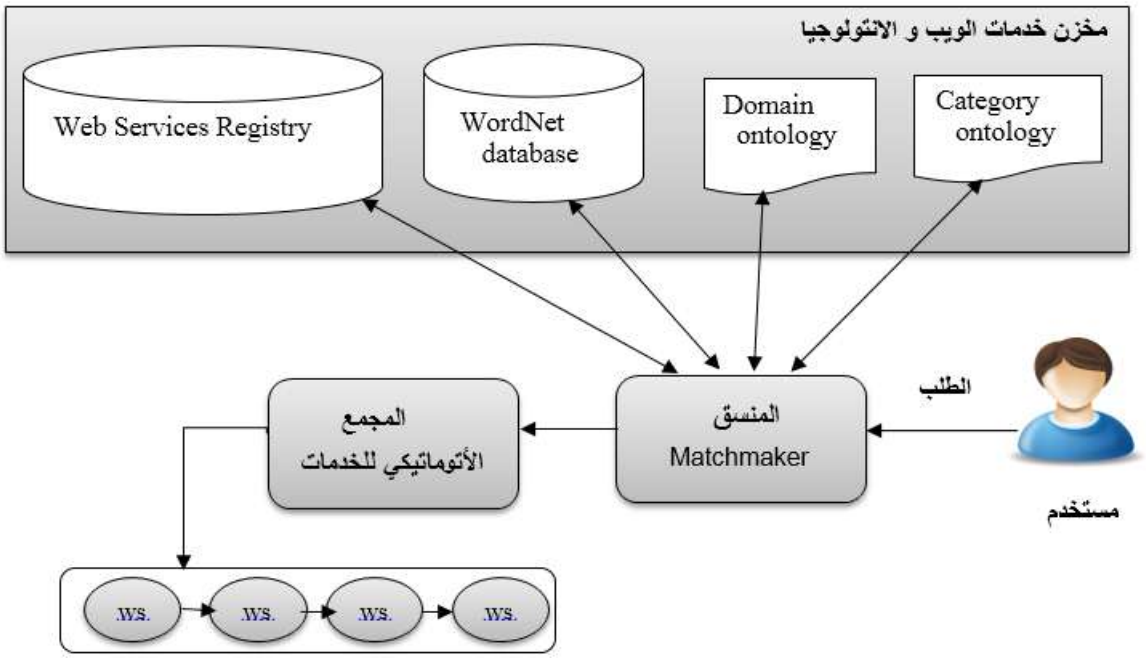
الخدمة WS15 تتشابه مع WS8 = { Student, Course, Study, 35}

هرمية النظام المقترح:

يوضح الشكل (1) مخططاً عالي المستوى للنظام المقترح حيث يتألف من ثلاثة مكونات رئيسية هي مخزن خدمات الويب و ملفات الانطولوجيا و المنسق matchmaker و مجمع خدمات الويب. يحتوي المخزن على توصيف خدمات الويب و انطولوجيا المجال و انطولوجيا التصنيف، توصف الخدمات من حيث الواجهات أي متحولات الدخل و الخرج إضافة إلى كلفة تنفيذ الخدمة كمعيار لجودتها، أما ملفات الانطولوجيا فهي تتضمن انطولوجيا المجال و التصنيف (بصيغة OWL Ontology Web Language) و تستخدم من أجل فهم معاني متحولات الخدمات و استنتاج العلاقات فيما بينها. استُخدم ملف انطولوجيا التصنيف CO الموضح في الشكل(2) من قبل النظام لمعرفة خدمات الويب القابلة للتجميع والتي تتشابه مع تصنيف طلب المستخدم و ذلك لحل مشكلة عدم التجانس الدلالي بين الخدمات فيما بينها و بين الخدمات و الطلب. كما يعرض الشكل(3) انطولوجيا المجال و التي تمثل الجامعة حيث يظهر علاقتي الصف الأب و الصف الفرعي بين المفاهيم. مثلاً PostGraduated صف فرعي من Student و هو بدوره صف فرعي من Person.

يستقبل المنسق الطلب من قبل المستخدم و يقوم بحساب التشابه الدلالي بين واجهات الخدمات باستخدام ملفات الانطولوجيا (انطولوجيا التصنيف Category ontology و انطولوجيا المجال Domain ontology) إضافة إلى استخدام المرادفات عن طريق قاعدة بيانات WordNet حيث يطبق الخوارزمية (1) الموضحة لاحقاً لاكتشاف مجموعة خدمات الويب المرتبطة بالطلب من أجل زيادة فاعلية خوارزمية التجميع الآلي.

ينفذ المجمع الخوارزمية (2) التي تعمل وفق القواعد السابقة لتوليد سلسلة الحل التي تتضمن مجموعة من خدمات الويب، حيث يختار الخدمات التي تكون واجهات كل زوج منها مرتبطة دلالياً، و الأفضل من حيث الكلفة، ثم يعيد الحل الأمثل لتلبية طلب المستخدم.



الشكل (1) : هرمية النظام المقترح

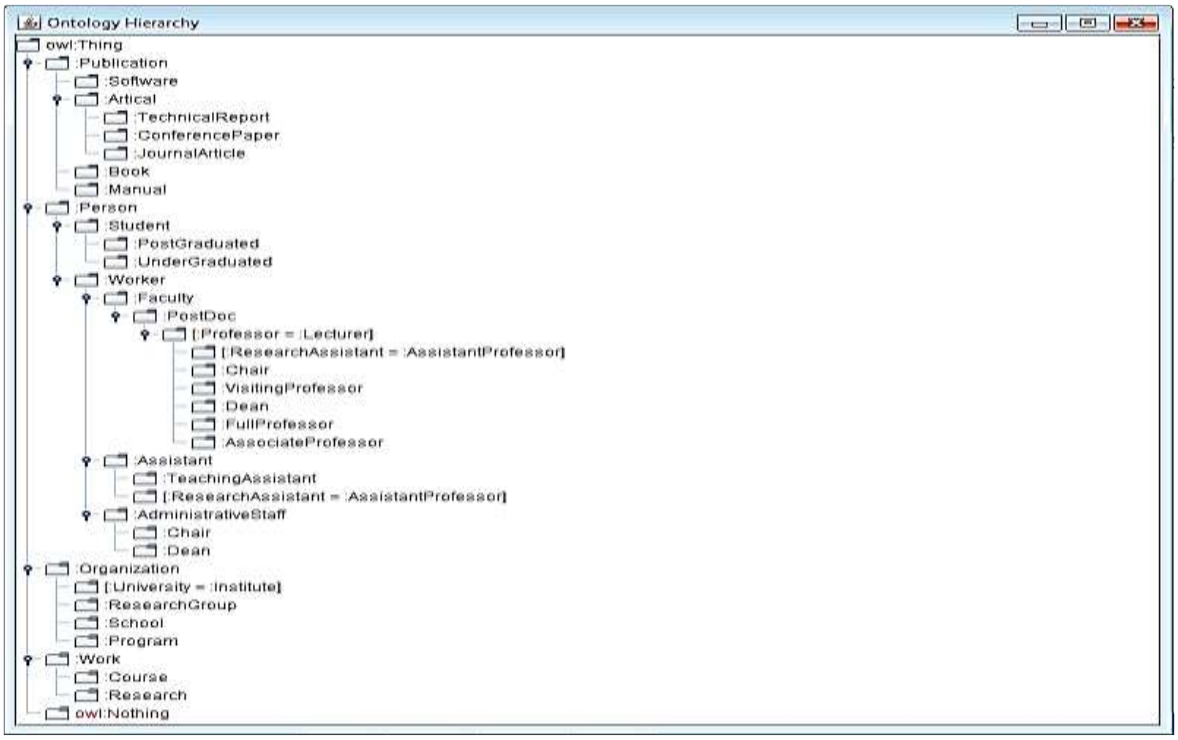
تحقيق الانطولوجيا:

تم إنشاء ملفي الانطولوجيا CO, DO باستخدام Protégé [9] وهو نظام إدارة معارف و محرر انطولوجيا مجاني مفتوح المصدر، يؤمن واجهات رسومية للمستخدم من أجل بناء الانطولوجيا، تم بناؤه باستخدام لغة جافا حيث يسمح بتصدير الانطولوجيا إلى عدة صيغ منها مخططات Extensible Markup Language (XML)، إطار توصيف المصادر Resourse Description Framework (RDF)، لغة انطولوجيا الويب Ontology Web Language(OWL). استُخدم Jena مع محرك استدلال Pellet Reasoner كإطار عمل دلالي مع جافا لاستخلاص دلالات المفاهيم و العلاقات بينها.

يوضح الشكل(2) ملف انطولوجيا التصنيف CO ، استُخدم من قبل النظام لمعرفة خدمات الويب القابلة للتجميع والتي تتشابه مع تصنيف طلب المستخدم و ذلك لحل مشكلة عدم التجانس الدلالي بين الخدمات فيمابينها و بين الخدمات و الطلب.

الشكل (2) : ملف انطولوجيا التصنيف

وفيما يلي هرمية المفاهيم في انطولوجيا المجال DO تبين مفاهيم جامعة و العلاقات بينها، استُخدمت من قبل النظام لمعرفة العلاقات بين واجهات خدمات الويب لاكتشاف و اختيار الخدمات القابلة للتجميع الشكل(3):



الشكل (3) : هرمية المفاهيم في ملف انطولوجيا المجال

خوارزمية اكتشاف خدمات الويب:

إن عملية اكتشاف خدمات الويب هي الخطوة السابقة لعملية التجميع، ولكن لا يمكن تجاهل وجود عدم تجانس بين المفاهيم في الويب الدلالي بسبب اختلاف الاهتمامات والمعارف و الأدوات، إضافة إلى اختلاف مستوى التفاصيل في التوصيف.

بسبب قصور اكتشاف الخدمات بالاعتماد على الكلمات المفتاحية، اقترحت الخوارزمية(1) Matching و هي تختلف عن الخوارزميات السابقة كونها تعتمد في تحسين اكتشاف الخدمات المتشابهة دلالياً على انطولوجيا المجال و انطولوجيا التصنيف للتغلب على عدم التجانس إضافة إلى الاستفادة من المرادفات التي تتيحها قاعدة بيانات WordNet لاكتشاف أكبر عدد من الخدمات التي من الممكن أن تكون جزء من سلسلة الحل.

تأخذ الخوارزمية متحولي دخل C, D ، حيث تعمل على إيجاد جميع المفاهيم الموجودة في الانطولوجيا DO و المتشابهة مع C وفق درجة التشابه المطلوبة (exact, subsumes or plug_in) و تخزينها في مجموعة A، ثم تقوم بإيجاد جميع المفاهيم الموجودة في الانطولوجيا CO و المتشابهة مع D و تخزينها في مجموعة B . كذلك يتم إيجاد جميع مرادفات المفهوم C من قاعدة بيانات WordNet و تخزينها في مجموعة E. أخيراً تقوم الخوارزمية بإيجاد جميع الخدمات التي يكون دخلها هو أحد مفاهيم المجموعتين A,E و تنتمي إلى تصنيف من المجموعة B وتخزينها في المجموعة W و في النهاية تعيد قائمة بجميع الخدمات WS مرتبة تصاعدياً حسب كلفة تنفيذها.

Algorithm 1: Matching (**C, D**)

Input: C input of request or output of web service, **D** request category
Output: WS set of services
{
A = set of concepts from Domain Ontology, which match with **C** based on matching

degree (exact, subsumes or plug_in)
B = set of concepts from Category Ontology, which match with **D**.
E = set of Synenyms of **C** from WordNet.
N = A ∪ E.

WS = ∅
For each concept ***a*** in N do
W = set of web services with input is ***a*** and category ∈ B
WS = WS ∪ W

return WS;
}

خوارزمية تجميع خدمات الويب:

اقترحت الخوارزمية (2) SWSC من أجل توليد سلسلة الخدمات التي تمثل الحل الأمثلي و هي تختلف عن سابقاتها كونها تعمل على إيجاد الحل عودياً مع أخذ كلفة تنفيذ الخدمة كمتطلب غير وظيفي (معيار لجودة الخدمة) بعين الاعتبار. تأخذ الخوارزمية دخلاً يمثل دخل طلب المستخدم WSR. I و الخرج المتوقع WSR. O إضافة إلى التصنيف االمطلوب WSR. D وتقوم أولاً بالبحث عن خدمة واحدة تتشابه دلالياً مع الطلب WSR على مستوى الدخل و الخرج (القاعدة الثالثة) و تعيدها في حال وجودها.

في حالة الفشل في إيجاد خدمة واحدة يتم أولاً إيجاد قائمة خدمات الويب التي تتشابه دلالياً على مستوى الدخل مع دخل الطلب (القاعدة الأولى)، ثم يتم تكرار استدعاء التابع عوديّاً من أجل كل خدمة في القائمة المرتبة تصاعدياً حسب كلفة تنفيذ الخدمات، للبحث عن سلسلة جزئية دخلها هو خرج الخدمة WS. O و خرجها هو خرج الطلب WSR. O اعتماداً على (القاعدة الرابعة)، في النهاية يتم إعادة سلسلة خدمات الويب التي توافق الدخل و الخرج و التصنيف المطلوب.

Algorithm 2: SWSC (WSR. I , WSR. O, WSR. D)

Input: input, output and category of request
Output: WSC
{
If GetService(WSR. I , WSR. O, WSR. D) ≠ ∅
WSC = GetService (WSR. I , WSR. O, WSR. D)
return WSC
else
B = Matching (WSR. I, WSR. D)
for each service WS in B do
C = SWSC (WS. O , WSR. O, WSR. D)
WSC = WSC ∪ C
return WSC
}

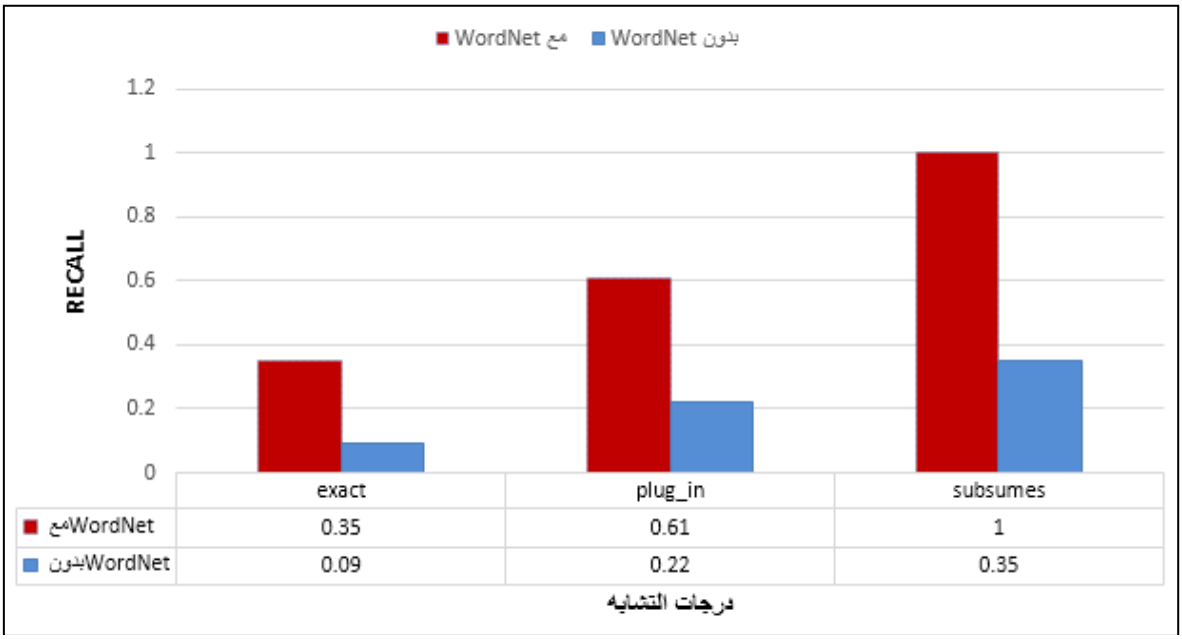
النتائج و المناقشة:

تم إجراء عدة اختبارات من أجل ثلاث مستويات تشابه هي: التطابق exact، الاحتواء subsumes، التوسيع plug_in. حيث تم تنفيذ سيناريوهين من أجل كل درجة تشابه، الأول بدون استخدام المرادفات WordNet، و الثاني مع أخذ المرادفات بعين الاعتبار. و ذلك لاختبار حساسية و دقة النظام المقترح في اكتشاف الخدمات من أجل درجات التشابه المختلفة. كما جرى تقييم فاعلية النظام من حيث الزمن اللازم للتجميع. نُفّذت جميع الاختبارات على نظام تشغيل Windows7 32_bit مع ذواكر 4GB RAM و معالج Intel® Core™2 Duo ، هو معالج متعدد النواة عالي الأداء و السرعة نسبياً، حيث تعالج كل وحدة فيه أربع تعليمات في الثانية مما يعني زيادة السرعة مقارنة بالأجيال السابقة، كما أنه يوفر الطاقة. لكن حالياً توجد عدة أجيال من المعالجات الأسرع التي ظهرت بعده.

السيناريو الأول:

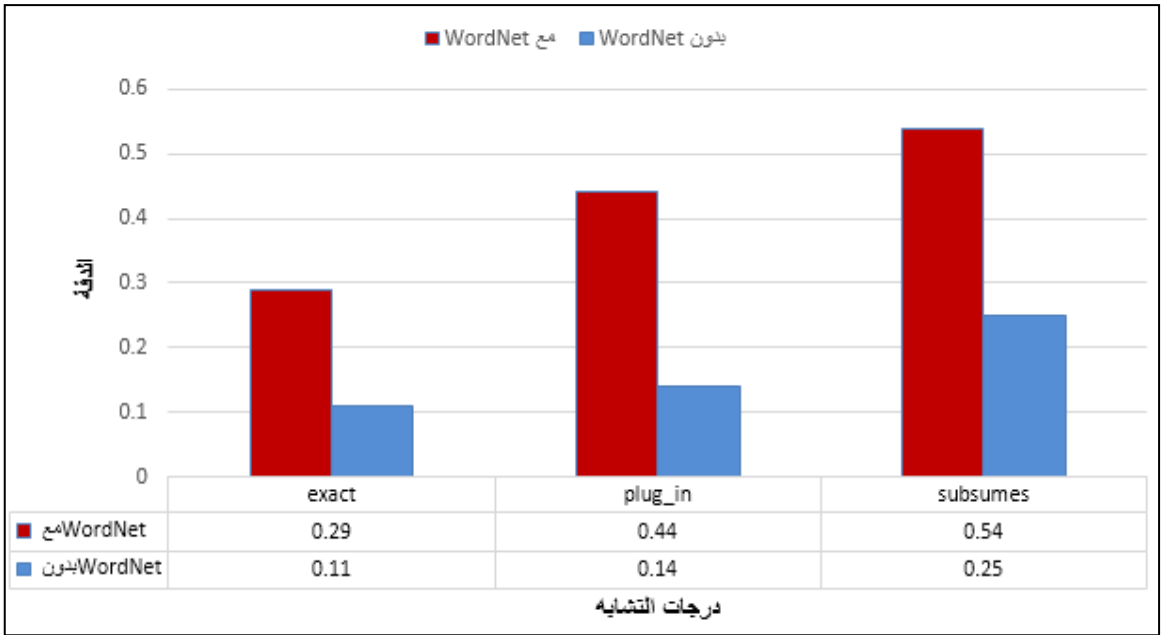
كان الهدف منه اختبار حساسية النظام، تم تكرار الاختبار من أجل ثلاث درجات تشابه (1) التام، (2) الاحتواء، (3) التوسيع.

يظهر الشكل(4) مقارنة بين حساسية النظام في اكتشاف جميع الخدمات المرتبطة بالطلب Recall من أجل حالات الدراسة الثلاث قبل وبعد استخدام المرادفات بالاعتماد على WordNet . تبين النتائج أن الحساسية تزداد بشكل ملحوظ عند أخذ المرادفات بعين الاعتبار من أجل جميع درجات التشابه، كما يحقق النظام أفضل النتائج عند درجة الاحتواء. بينما تكون حساسية النظام أقل ما يمكن عند حالة التطابق، يُفسَّر ذلك بأن عدد الخدمات المناسبة للحل يكون أقل ما يمكن عند التطابق كونها العلاقة الأكثر تقييداً لعملية البحث حيث يتم البحث في ملف الانطولوجيا عن المفاهيم المتطابقة فقط مما يقلل احتمال الحصول على سلسلة خدمات متشابهة دلالياً على مستوى الواجهات، بينما مكّن استخدام المرادفات و علاقة الاحتواء النظام من اكتشاف عدد أكبر من الخدمات التي ترتبط بالطلب مما يزيد قيمة Recall كما تبين المعادلة (1).



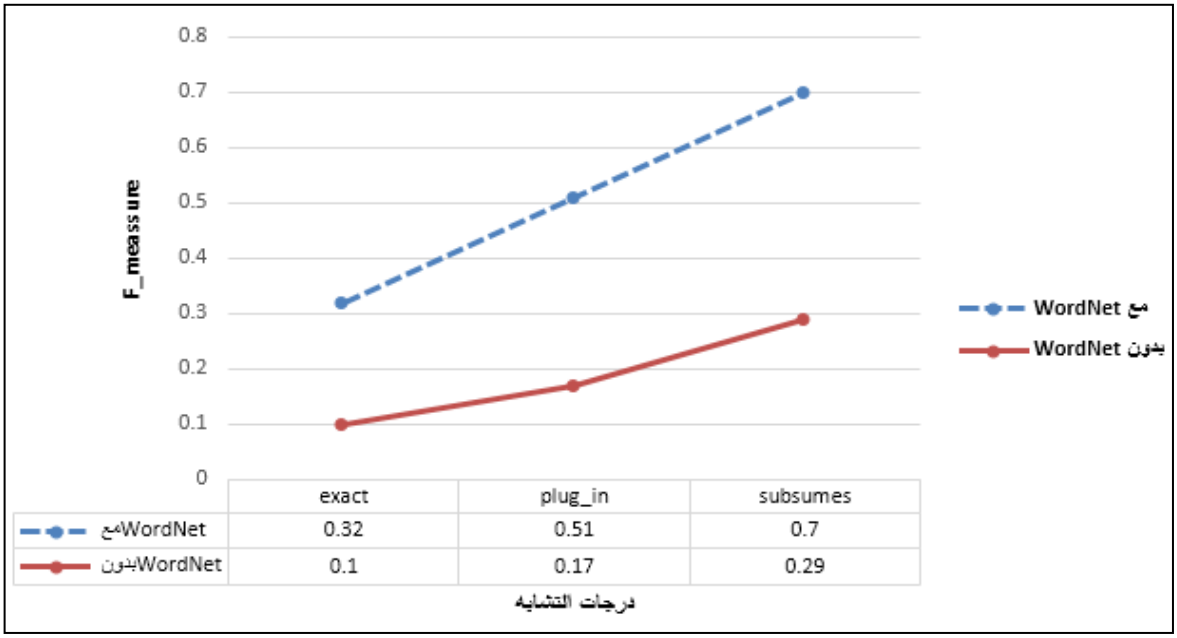
الشكل (4) : مقارنة الحساسية Recall

يظهر الشكل(5) مقارنة بين دقة النظام عند اكتشاف الخدمات المرتبطة بالطلب فقط Precision من أجل حالات الدراسة الثلاث قبل وبعد استخدام مرادفات WordNet، حيث تزداد دقة النظام بشكل ملحوظ من أجل جميع درجات التشابه عند أخذ المرادفات بعين الاعتبار، و يحقق النظام أفضل النتائج عند درجة تشابه الاحتواء، لأن أخذ المرادفات بعين الاعتبار يقلل اكتشاف الخدمات غير المرتبطة بالطلب False Positive والتي هي في مقام المعادلة (2).



الشكل (5) : مقارنة الدقة Precision

يظهر الشكل(6) أن قيم F_measure التي تم حسابها وفق المعادلة (3) تزداد عند الاستفادة من المرادفات في اكتشاف الخدمات المناسبة للتجميع من أجل درجات التشابه الثلاث و تكون أكبر ما يمكن عند درجة الاحتواء، مما يؤكد فاعلية المنهجية المقترحة في الاكتشاف الآلي لخدمات الويب.



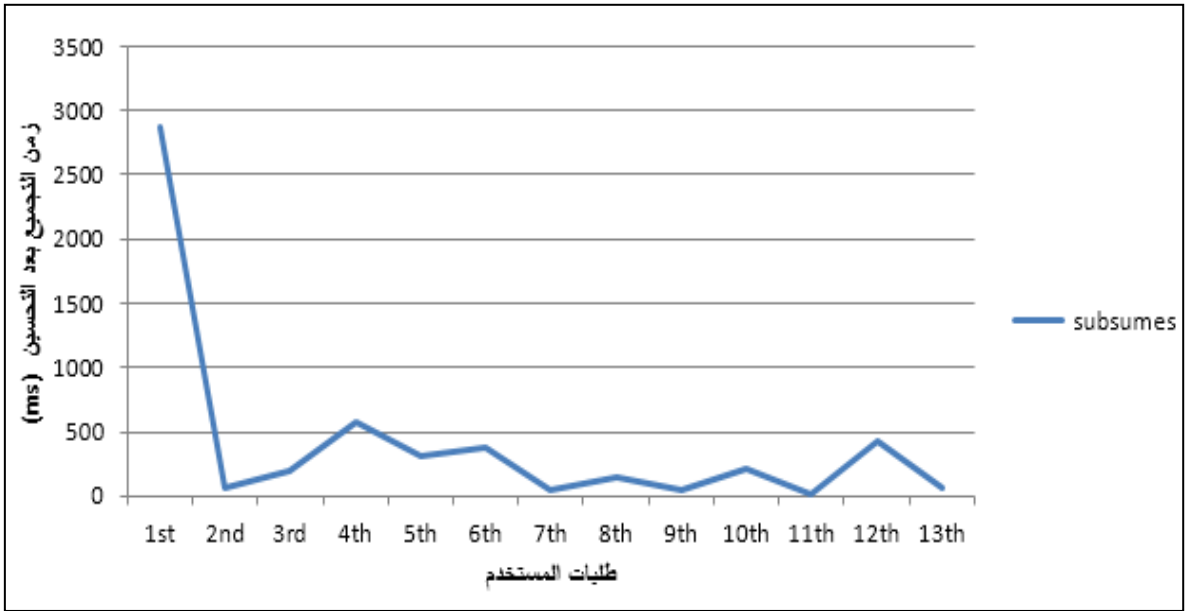
الشكل (6) : مقارنة F_measure

السيناريو الثاني:

يتناول السيناريو الثاني تقييم التجميع الآلي لخدمات الويب باستخدام معيار الزمن اللازم للتجميع composition time، حيث تمت دراسة الزمن في حالتين الأولى عند تحميل ملف الانطولوجيا عند كل طلب، والثانية عند الاكتفاء بتحميل ملف الانطولوجيا عند أول طلب فقط. و كان الهدف منه اختبار التحسينات المقترحة ومدى قدرتها على زيادة فاعلية النظام المقترح حيث تمثل التحسين بتخزين ملفات الانطولوجيا في الذاكرة عند معالجة أول طلب للمستخدم ثم استقبال باقي الطلبات.

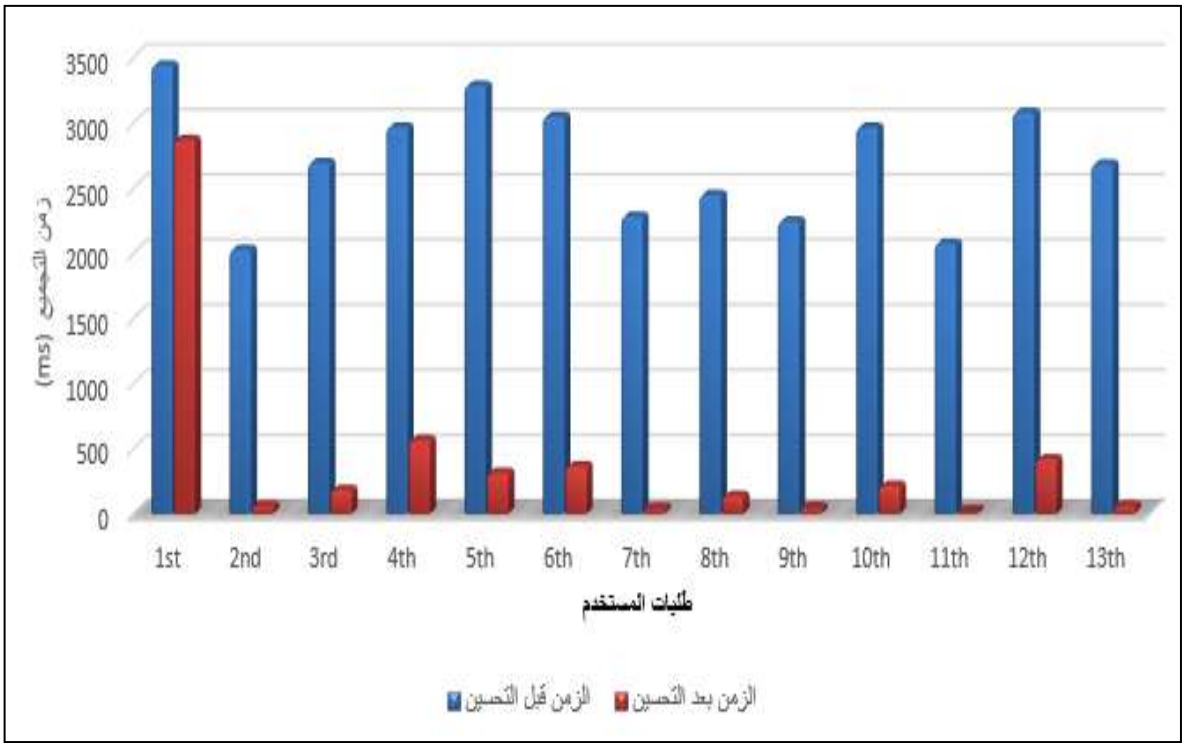
بالعودة إلى خطوات خوارزمية التجميع يتضح أنه من أجل كل خدمة مكتشفة يتم البحث عن سلسلة جزئية لإكمال المطلوب، عندها يتم تحميل ملف الانطولوجيا و قراءته من أجل استخلاص العلاقات بين مفاهيم الدخل و الخرج الخاصة بالخدمات، وذلك لأن النظام يجمع الخدمات آلياً آخذاً بعين الاعتبار التعديلات التي من الممكن أن تطرأ على ملفي الانطولوجيا. يتيح تخزين الانطولوجيا في الذاكرة عند معالجة أول طلب إمكانية توفير الزمن عن طريق تلبية عدة طلبات في نفس الجلسة، كما يسمح للمستخدم بتخصيص كل جلسة لطلب واحد.

اعتمدت علاقة الاحتواء subsumes في الاختبار لأن النظام يعطي أفضل نتائج عندها حسب النتائج السابقة. أظهرت النتائج الموضحة في الشكل(7) أن الزمن عند الطلب الأول هو الأعلى بسبب العبء الناتج عن تحميل الانطولوجيا، بينما يعود و ينخفض بشكل ملحوظ من أجل باقي الطلبات. لا تؤثر زيادة عدد الطلبات في الجلسة الواحدة على زمن التجميع اللازم لتلبية كل طلب.



الشكل (7) : زمن التجميع المحسن

يوضح الشكل(8) مقارنة بين الزمن اللازم للتجميع قبل التحسين و بعده من أجل عدة طلبات، حيث يظهر الأثر السلبي الكبير لعملية الاتصال المتكرر مع ملف الانطولوجيا عند كل طلب.



الشكل (8) : مقارنة زمن التجميع قبل التحسين و بعده

إذا أخذ الفرق بين الزمن اللازم للتجميع قبل عملية التحسين المقترحة و بعده ينتج لدينا الشكل(9) الذي يبين مقدار التوفير في الزمن، حيث تكون الفائدة قليلة نسبياً من أجل الطلب الأول، ثم تزداد بشكل كبير اعتباراً من الطلب الثاني.

الشكل (9) :التوفير في زمن التجميع عند تحميل ملف الانطولوجيا إلى الذاكرة

الاستنتاجات و التوصيات:

1. يعزز استخدام انطولوجيا المجال و انطولوجيا التصنيف و مرادفات WordNet الاكتشاف الآلي لخدمات الويب.
2. يزيد استخدام كلفة تنفيذ الخدمة كمعيار للجودة في اختيار الخدمات من أمثلية التجميع الآلي، لانه يضمن في كل مرحلة اختيار الخدمة الأقل كلفة.
3. استخدام التصنيف يحل مشكلة عدم التجانس الدلالي، كما يقلل عدد الخدمات التي يتم البحث ضمنها و بالتالي يحسن زمن التجميع.
4. الاستفادة من ميزات الانطولوجيا في دعم التجميع الآلي للخدمات كونها تتيح فهم معاني دخل وخرج الخدمات، مما يزيد من الدقة في البحث عن الخدمات، الأمر الذي يسهل تجميع عدة خدمات معاً كونها تقلل المعالجة اليدوية و تدعم المعالجة الآلية.
5. العمل من أجل بناء و توسيع الانطولوجيا لكل مجال عمل فهي تقدم معلومات إضافية عن توصيف الخدمات و هذا يساعد في الوصول إلى طرق أمثلية و فعالة للتجميع.
6. اعتماد استراتيجية تحميل ملف الانطولوجيا إلى الذاكرة عند أول تنفيذ يحسن زمن التجميع بشكل كبير، كونه يقلل من ضياع الزمن الناتج عن التحميل المتكرر للانطولوجيا عند كل طلب و يحصر العبء بالطلب الأول فقط.
7. توسيع قواعد البيانات التي تحوي قواميس المفردات للاستفادة منها بأكبر درجة ممكنة، لأن استخدام المرادفات يزيد دقة النظام في اكتشاف الخدمات الأقرب إلى الطلب.

المراجع:

[1] HAAS, H. and BROWN, A., *Web Services Glossary*, https://www.w3.org/TR/ws-gloss/, January 2018.

[2] GE, J., QIU, Y., YIN, S. *Web Service Composition Method Based on OWL*, IEEE/ International Conference on Computer Science and Software Engineering, Wuhan, Hubei, 2008, pp. 74-77.

[3] GULURU, P. *New Approaches for Service Composition based on Graph* Models, IEEE/ Seventh International Conference on Contemporary Computing (IC3), Noida, 2014, pp. 507-512.

[4] ROSTAMI, N., KHEIRKHAH, E., JALALI, M. *Web Services Composition Methods and Techniques: A Review*, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.3, No. 6, December 2013.

[5] LI, Y., LI, Y., HU, T., LV, Z. *An automatic semantic Web service composition method based on ontology*, IEEE/ ACIS 14th International Conference on Computer and Information Science (ICIS), Las Vegas, NV, 2015, pp. 563-566.

[6] MAO, SH., ZANG, H., NI, B., *Research on Semantic Web Service Composition Based on Binary Tree*, International Journal of Grid Distribution Computing (IJGDC) , 2015, pp.133-142.

[7] WEI, X., JIAN-GUO, C., TAO, H. *Service Composition Algorithm Based on Ontology Semantic*, IEEE/ International Conference on Computational and Information Sciences, Shiyang, 2013, pp. 1886-1889.

[8] POWERS, D., A., *Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation*, Journal of Machine Learning Technologies, 2011, pp-37-63.

[9] HORRIDGE, M. *A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools* Edition 1.3, The University Of Manchester, UK, 2011, 108.