

# ارائه روشی جهت کشف وب سرویس معنایی بر مبنای مفاهیم و ویژگی های مفاهیم

محمد رستمی\*

کارشناس ارشد گروه مهندسی نرم افزار موسسه آموزش عالی جهاد دانشگاهی اهواز، خوزستان

پست الکترونیکی: [Mohamad.rostami10@yahoo.com](mailto:Mohamad.rostami10@yahoo.com)

ایمان عطارزاده

عضو باشگاه پژوهشگران و نخبگان جوان دانشگاه آزاد اسلامی واحد دهاقان، اصفهان

استادیار گروه مهندسی نرم افزار دانشگاه آزاد اسلامی واحد دزفول، خوزستان

پست الکترونیکی: [attarzadeh.std@gmail.com](mailto:attarzadeh.std@gmail.com)

حسین بیگی هرچگانی

مریی گروه مهندسی نرم افزار، موسسه آموزش عالی جهاد دانشگاهی اهواز، خوزستان

پست الکترونیکی: [hossein\\_beigi\\_h@yahoo.com](mailto:hossein_beigi_h@yahoo.com)

## چکیده

افزایش روز افزون تعداد وب سرویس ها و رشد انفجاری دانش در دسترس WWW باعث شده است که راهکارهای فراوانی در جهت کشف وب سرویس که یکی از مهمترین چالش های اصلی حوزه وب سرویس می باشد مطرح گردد. راهکارهایی که تاکنون در محافل تحقیقاتی مطرح گردیده است به علت محدود بودن پارامترهایی که در جهت کشف وب سرویس در نظر گرفته می شوند مانند راهکارهایی که تنها به مقایسه ورودی و خروجی می پردازند دارای دقت کمی می باشند که این امر باعث می گردد در خواست های بازگردانده شده به کاربر نتوانند به طور شایسته نیازهای کاربر را پوشش دهند.

الگوریتم ارائه شده در این پژوهش علاوه بر در نظر گرفتن ورودی و خروجی پارامترهای دیگر مانند مفاهیم و ویژگی های مفاهیم را نیز در نظر می گیرند که این امر باعث می گردد که فرآیند کشف سرویس با جزئیات بیشتر و در نتیجه دقت بیشتری همراه باشد و برای اثبات آن نیز با سه الگوریتم دیگر مقایسه می گردد. همچنین در این پژوهش سعی شده است که علاوه بر درجات تطبیق دیگری که در کارهای پیشین مطرح گردیده درجه تطبیق substitute که می تواند سرویس هایی را به کاربر پیشنهاد کند که حتی اگر شباهت زیادی به درخواست کاربر نداشت ولیکن به عنوان مکمل درخواست کاربر باشد به کاربر پیشنهاد گردد.

کارایی و میزان دقت الگوریتم ارائه شده به وسیله نسخه سوم مجموعه داده OWLS-TC ارزیابی و با سه الگوریتم کشف سرویس مقایسه شده است که دارای دقت بالایی می باشد. الگوریتم ارائه شده در این پژوهش توانسته است با ترکیب مقایسه میان مفاهیم و ویژگی های آن ها و همچنین امتیازدهی به سرویس ها و رتبه بندی آن ها در مجموعه های درجات تطبیق، گامی در راستای افزایش میزان دقت در کشف سرویس ها و همچنین با ارائه درجه تطبیق جدیدی با نام جایگزین گامی در راستای بهبود فرآیند کشف برداشته باشد.

---

\* نویسنده مسئول

## 1- مقدمه

فناوری‌های فراهم شده به وسیله وب معنایی در محیطی کار می‌کنند که قابل تفسیر توسط ماشین است. وب سرویس‌ها نیز در محیطی کار می‌کنند که سازمان‌ها می‌توانند بخشی از توانایی‌های خود را از طریق اینترنت در دسترس قرار دهند. این کار با بسته بندی یا در کنار هم قرار دادن توانایی محاسباتی با واسط وب سرویس، و اجازه دادن به دیگر سازمان‌ها برای کشف آن (به وسیله UDDI<sup>1</sup>) و تعامل با آن (به وسیله WSDL<sup>2</sup>) امکان پذیر شده است.

به علت موجود بودن تعداد زیادی از وب سرویس‌ها، یافتن وب سرویس متناسب با نیازهای کاری چالشی عظیم است. این چالش نیاز به برقراری یک فرآیند مورد اعتماد و مؤثر کشف وب سرویس را تضمین می‌کند. امروزه تحقیقات بسیاری برای طراحی روش‌هایی به منظور افزایش میزان دقت کشف وب سرویس انجام می‌گیرد که بهترین سرویس را تطبیق دهد. در نظر گرفتن رابطه معنایی میان کلمات استفاده شده در توصیف سرویس و همچنین پارامترهای ورودی و خروجی می‌تواند منجر به دقت در کشف وب سرویس شود. علاوه بر پارامترهای ورودی و خروجی، پارامترهای دیگری (پیش شرط، اثر) نیز در توصیف سرویس وجود دارند که استفاده از آن‌ها منجر به بالا بردن دقت کشف وب سرویس‌ها می‌شود. در بیشتر روش‌های کشف وب سرویس، پنج درجه تطبیق در نظر گرفته می‌شود که عبارتند از: [13]

Exact: سرویس ارائه شده A و درخواست R دو مفهوم یکسان هستند.

Plug-In: درخواست R زیر مفهوم سرویس ارائه شده A می‌باشد.

Subsume: درخواست R ابر مفهوم سرویس ارائه شده A می‌باشد.

Intersection: سرویس ارائه شده A و درخواست R به روابط عمومی‌تر و بخشی از (part of) تعلق ندارند اما اشتراک آن‌ها تهی نیست.

Fail: سرویس ارائه شده A و درخواست R دو مفهوم متفاوت هستند.

در این مقاله روشی برای کشف وب سرویس‌ها بر اساس مقایسه میان پارامترهای ورودی و خروجی، پیش شرط و اثر درخواست و سرویس، با زبان OWL ارائه می‌شود و الگوریتمی برای آن طراحی می‌گردد، اقدام به تعریف درجه تطبیق جدیدی می‌شود که سرویس‌هایی را کشف می‌نماید که گرچه نیاز کاربر را بر آورده نمی‌نمایند اما جایگزینی نزدیک به نیاز کاربر را ارائه می‌دهد.

## 2- راهکارهای پیشین

جستجوی سرویسی که متناسب با درخواست کاربر باشد یکی از چالش‌های بزرگ در زمینه وب سرویس‌ها است. این به خاطر عواملی از قبیل (1) تعداد زیاد سرویس‌های موجود بر روی اینترنت (2) جستجو بر اساس کلمات کلیدی و در نظر نگرفتن معانی است.

تحقیقات زیادی که در زمینه کشف وب سرویس معنایی صورت گرفته است را می‌توان به زمینه‌های مختلفی تقسیم نمود. در ادامه به معرفی چند روش کشف وب سرویس می‌پردازیم.

### - روش معنایی

بیشتر تحقیقاتی که اخیراً در زمینه کشف وب سرویس صورت گرفته، به منظور افزایش دقت کشف وب سرویس، بر روی تطبیق معنایی تمرکز داشته است. فرآیند تطبیق معنایی سرویس، پیاده‌سازی عملیات متفاوت میان سرویس‌های ارائه شده و سرویس‌های

<sup>1</sup> universal description, discovery, and integration

<sup>2</sup> Web Service Description Language

درخواست شده می‌باشد. اگر وب سرویس‌ها با هستان شناسی مدل شده باشند، تطبیق معنایی می‌تواند به وسیله مفاهیم و روابط آن‌ها با یکدیگر صورت گیرد [44].

در [29] راه‌حلی با استفاده از ماشین حالت قالبی برای کشف سرویس پیشنهاد شده است. قالب پیشنهادی سه عمل انجام می‌دهد: پیدا نمودن هدف، کشف وب سرویس معنایی، و درنهایت انتخاب سرویس. در مرحله کشف هدف، پرس‌وجوی کاربر مورد بررسی قرار می‌گیرد و اهداف کاربر مشخص می‌شوند. تطبیق معنایی توصیف هدف در مرحله کشف وب سرویس معنایی به منظور پیدا نمودن مجموعه سرویس‌های مرتبط انجام می‌پذیرد. در مرحله انتخاب سرویس، سرویس انتخاب شده از نظر سازگاری با پروتکل یا پردازش‌های کاربر مورد بررسی قرار می‌گیرد.

با افزایش تعداد وب سرویس‌های موجود، نیاز به خودکار نمودن فرآیند کشف سرویس و ایجاد ساختار فراخوانی مناسب وجود دارد. توصیف معنایی سرویس‌ها برای تسهیل کشف معنایی مهم است و براساس معنایی مفاهیم به‌کار رفته در توصیف سرویس و ساخت هستان شناسی خاص، این روش می‌تواند سرویس‌های مناسب را به‌وسیله استدلال منطقی بر روی توصیف سرویس‌ها بیابد. توصیف سرویس‌ها باید انعکاس دهنده عملکرد آن‌ها باشد.

[45] یک سیستم تطبیقی ارائه نموده است که بر پایه استفاده از UDDI برای نگهداری از توصیف سرویس‌ها استوار است. الگوریتم تطبیق سعی می‌کند که پارامترهای ورودی و خروجی سرویس ارائه شده را با پارامترهای نظیر آن در سرویس درخواست تطبیق نماید. یکی از ایده‌های مهم در الگوریتم ارائه شده توسط پائولوچی و دیگران این است که هم مشتری و هم فراهم‌کننده از یک هستان شناسی یکسان استفاده می‌کنند تا بتوانند مقادیر معنایی یکسانی به هر یک از سرویس‌ها بدهند. در این روش، به دلیل بارگذاری یک هستان شناسی برای درخواست‌کننده و ارائه دهنده وظیفه موتور استنتاج یا استدلال‌گر تسهیل شده است که از طریق این هستان شناسی درجه شباهت میان پارامترهای ورودی و خروجی تعریف شده را اندازه‌گیری کند. الگوریتم تطابق آن‌ها، میان چهار درجه تطبیق تمایز قایل شده است: Fail, Subsume, Plug-in, Exact. آن‌ها فقط روابط میان رده‌های والد و پدر هستان شناسی را بررسی می‌کند که منجر به recall پایین می‌شود. راه حل آن‌ها بر مبنای DAML-S<sup>3</sup> طراحی شده است که قادر به نمایش کارکرد وب سرویس‌ها می‌باشد. متأسفانه، برخی از اطلاعات مهم از قبیل پیش‌شرط‌ها، اثرها، و طبقه‌بندی سرویس‌ها در این روش حذف شده یا در نظر گرفته نشده‌اند که این موجب بالا رفتن False Negative (منفی غلط) می‌شود در نتیجه سرویس‌هایی که توسط الگوریتم شناسایی می‌شوند نیازهای کاربر را به‌طور کامل برآورده نمی‌نمایند.

[46] قالبی برای کشف خودکار سرویس ارائه نموده است. در این قالب، کشف سرویس در سه مرحله انجام می‌گیرد: بررسی نام، تطبیق پروفایل، و مقایسه معنایی پارامترهای ورودی/خروجی وب سرویس. دو پایگاه دانش در این قالب به‌کار گرفته شده‌اند، یکی برای اطلاعات پایه که در آن از OWL استفاده شده است و دیگری برای ذخیره توصیف وب سرویس‌هایی که در OWL-S<sup>4</sup> نوشته شده‌اند. فرآیند تطبیق، سرویس‌های بالقوه را براساس مفاهیم به‌کار رفته در ورودی‌های آن‌ها انتخاب می‌کند. در این قالب مفهوم جدیدی به نام برنامه‌ریزی سرویس<sup>4</sup> ارائه شده است. برنامه‌ریزی سرویس ترتیبی از فراخوانی سرویس را برای ارضای نیاز کاربر مشخص می‌کند. برای نایل شدن به این هدف، درخت‌هایی از سرویس‌ها خلق شده است و در هنگام فراخوانی درخت‌ها به روش پس‌ترتیب<sup>5</sup> پیمایش می‌شوند.

کاوامورا و دیگران [21] روشی ارائه نموده‌اند که با استفاده از فضای نام، متن، نوع ورودی/خروجی، و فیلتر محدودیت، سرویس‌ها را کشف می‌نماید. درحالت تطبیق نوع ورودی/خروجی، شباهت میان سرویس‌ها براساس مدل شمول محاسبه می‌شود.

تعداد زیادی از قالب‌های تطبیق بر مبنای DL وجود دارند. این قالب‌ها بر روی سرویس‌هایی عمل می‌کند که در RDF، DAML+OIL یا DAML-S<sup>6</sup> نوشته شده‌اند. یک استدلال‌گر DL نقش کلیدی در محاسبه درجه تطبیق میان دو توصیف

<sup>3</sup> semantic DARPA Agent Markup Language

<sup>4</sup>service planning

<sup>5</sup> postorder

<sup>6</sup>coverage

سرویس بازی می‌کند. در این روش‌ها، توصیف سرویس به‌عنوان یک موجودیت در نظر گرفته می‌شود یعنی مفاهیم در DLها، و تبلیغات در استدلال‌گر بر اساس روابط شمولی طبقه‌بندی می‌شوند. فرآیند استدلال باید پاسخ مناسبی را به‌عنوان نتیجه به درخواست بازگرداند.

تلاش‌های زیادی در عرصه کشف وب سرویس معنایی، به ویژه در OWL-S، صورت گرفته است. OWL-S در هستان شناسی OWL برای مدل کردن ویژگی‌های مختلف وب سرویس می‌باشد [47]. در نتیجه OWL-S استدلال بر روی ویژگی‌های وب سرویس با کمک استدلال‌گر DL را فراهم می‌سازد. OWL-S ساختارهایی برای مدل کردن پیش‌شرط و اثر وب سرویس دارد. با این‌که OWL-S هیچ منطق خاصی را برای مشخص کردن شرایط ارائه نمی‌کند اما SWRL را به‌عنوان نامزد این کار پیشنهاد می‌کند. اما SWRL<sup>7</sup> در ذات خود غیر قابل تصمیم‌گیری است.

OWL-S پیشنهاد می‌کند که پیش شرط‌ها و اثرهای یک وب سرویس به‌صورت ویژگی‌های آن مدل شوند. در نتیجه حتی اگر زیرمجموعه‌ای قابل تصمیم‌گیری از SWRL استفاده شود، نمی‌توان بر روی پیش شرط و اثر سرویس با یک استدلال‌گر DL به‌طور مستقیم استدلال نمود. علت این است که DL قادر به اخذ معانی قوانینی که فقط به وسیله برخی ویژگی‌ها قابل دسترس می‌باشند، نیست. متعاقباً با این‌که OWL-S اصولی را برای مدل کردن پیش شرط و اثر فراهم می‌آورد اما الگوریتم تطبیق بر مبنای OWL-S فقط انواع پارامترهای ورودی و خروجی وب سرویس را در نظر می‌گیرد. بیشتر تطبیق دهنده‌هایی که برای OWL-S پیشنهاد شده‌اند بر پایه تطبیق ورودی‌ها و خروجی‌ها هستند و پیش شرط و اثرهای سرویس را در نظر نمی‌گیرند. مقایسه میان پیش شرط و اثر در [48] انجام گرفته که مواردی از استدلال‌گر برای استدلال این‌که پیش شرط‌ها و اثرها در چه وضعیتی نسبت به یکدیگر هستند، استفاده نموده و از SWRL برای تعریف پیش شرط و اثر استفاده نموده است. روش ارائه شده برای تطبیق معنایی از روش‌های مشابهت لغوی نیز بهره گرفته و از ماژول مشابهت WordNet برای این کار استفاده نموده است و برای تعیین مشابهت معنایی از روابط استنتاج استفاده نموده است. استفاده از روش‌های لغوی بر مبنای WordNet به‌دلیل محدودیت این هستان شناسی جوابگوی تمامی سرویس‌ها نیست و از این رو الگوریتم کشف به نوعی دچار دقت پایین می‌گردد.

[43] نیز روش مطابقت معنایی‌ای را ارائه کرده است که از OWL-S بهره برده اما رتبه‌بندی سرویس برای سادگی قوانین تطبیق در نظر گرفته نشده است و این حوزه مانع دقیق و مناسب بودن سرویس‌های باز گرداننده شده به‌عنوان نتیجه می‌باشد. روش ارائه شده از استدلال‌گر RACER DL استفاده نموده است و بر مبنای عامل می‌باشد.

حاصل تطبیق می‌تواند شامل سرویس‌های مختلفی باشد و بنابراین نیاز به مشخص کردن این امر است که کدامیک از سرویس‌ها نیازهای مشتریان را برآورده می‌نماید. گاهی اوقات دخالت کاربر به منظور انتخاب این‌که کدامیک مناسب‌ترین سرویس می‌باشد، مورد نیاز است. هیچ یک از روش‌های مذکور از روش‌های رتبه‌بندی برای سرویس‌های منطبق یا تطبیق داده شده استفاده نمی‌کنند. اما کارهایی نیز هستند که بر روی رتبه‌بندی سرویس تمرکز کرده‌اند. لو [49] کشف وب سرویس معنایی و رتبه‌بندی را بر اساس ویژگی‌های کارکردی ارائه می‌کند که در آن هر چه سرویس از نظر معنایی نزدیک‌تر و شبیه‌تر باشد امتیاز رتبه‌بندی بیشتر خواهد بود.

روش‌هایی برای تطبیق ارائه شده‌اند که از منطق فازی برای چکیده کردن داده‌های اصلی یا پایه وب سرویس‌ها در فرآیند تطابق بهره می‌گیرند. [50,56] به معرفی قالبی پرداخته است که از منطق فازی به منظور چکیده نمودن و دسته‌بندی داده پایه وب سرویس‌ها که به‌صورت عبارات و قوانین فازی هستند، استفاده نموده است. این روش شناسایی ابعاد مخفی وب سرویس‌ها یعنی داده‌های آن‌ها و نمایش چکیده داده‌های وب سرویس‌ها در تئوری مجموعه فازی را فراهم آورده است.

کارهای پیشین نگاهت بین مفاهیم را بر اساس رده‌ها مشخص می‌نمایند و اگر رده‌های آن‌ها یکسان باشد آن‌ها را با یکدیگر تطبیق می‌دهند، درحالی‌که رابطه تطبیق میان مفاهیم ممکن است تنها محدود به روابط شمولی نباشد بلکه به روابط میان ویژگی‌ها نیز

---

<sup>7</sup> Semantic Web Rule Language

بستگی داشته باشد. اما اگر ویژگی‌های درخواست کاربر متمایز از ویژگی‌های سرویس ارائه شده باشد، تطبیق پایه نادرست در نظر گرفته می‌شود.

کارهای پیشین روابط تطبیق میان مفاهیم رابراساس رابطه شمولی محاسبه می‌کنند و روابط تطبیق چند به چند<sup>۸</sup>، یعنی در نظر گرفتن پارامترهایی غیر از رده‌های مفاهیم، بین پارامترهای ورودی/خروجی را در نظر نمی‌گیرند.

#### - روش بر مبنای بازیابی اطلاعات

روش‌های بازیابی اطلاعات بیشتر بر روی واژگان و اهمیت آن‌ها در متون تاکید دارند. برای انتخاب واژگان مهم از مجموعه بزرگی از واژگان روش‌های آماری مانند فرکانس واژه<sup>۹</sup>، inverse-documentfrequency و یا ترکیبی از این دو استفاده می‌شود.

فرکانس واژه t در متن D میزان اهمیت واژه در متن را مشخص می‌کند. این امر به عنوان فرکانس واژه شناخته می‌شود. inverse-documentfrequency میزان اختصاصی بودن واژه در مجموعه متون را اندازه‌گیری می‌کند. این اندازه به وسیله لگاریتم مجموع تعداد متون بخش بر تعداد متونی که واژه حداقل یکبار در آن‌ها آمده باشد محاسبه می‌شود. ترکیب دو روش فرکانس واژه و inverse-document frequency معیار بسیار مناسبی برای یافتن واژگان مهم در مجموعه‌ای از متون است. این روش‌ها برای کشف بهترین سرویس‌ها کافی نیستند زیرا متون موجود در وب سرویس‌ها بسیار خلاصه و چکیده شده هستند و اطلاعات ساختاری که به اخذ معانی عملیات کمک می‌کند را نادیده می‌گیرند.

از دیگر تکنیک‌های بازیابی اطلاعات که کاربرد گسترده‌ای نیز پیدا کرده است، تکنیک LSI می‌باشد. این تکنیک از SVD<sup>۱۰</sup> برای شناسایی مفاهیم استفاده می‌کند. LSI<sup>۱۱</sup> فرض می‌کند که ساختاری در استفاده از کلمات وجود دارد که به وسیله تنوع در انتخاب لغات پنهان شده است. [52] با استفاده از تکنیک LSI و توصیف سرویس با OWL-S روش تطبیق بر مبنای نحو برای کشف وب سرویس ارائه نموده است. در این روش ابتدا مشابهت معنایی پارامترهای ورودی و خروجی سرویس و درخواست محاسبه شده و سپس میزان مشابهت نحوی آن‌ها به وسیله LSI تعیین شده است. مجموع میزان مشابهت نحوی و معنایی درجه تطبیق سرویس می‌باشد که برای رتبه‌بندی سرویس استفاده می‌شود. رتبه‌بندی به درخواست کننده در انتخاب مرتبط‌ترین سرویس در میان تمامی سرویس‌های مرتبط کمک شایانی می‌نماید.

[53] نیز الگوریتمی برای کشف ارائه نموده است که از اطلاعات ساختاری و متنی سرویس برای کشف آن استفاده می‌نماید. این الگوریتم مفهوم جدیدی به نام درجه ترجیح<sup>۱۲</sup> را معرفی کرده است که نشان‌دهنده اولویت یک سرویس برای کاربر می‌باشد. این مفهوم معیاری برای رتبه‌بندی سرویس می‌باشد. همچنین ویژگی‌های serviceimportance و servicerelevance را برای محاسبه درجه ترجیح تعریف نموده است. در این الگوریتم مفهومی به نام serviceconnectivity تعریف شده که معیاری برای محاسبه میزان اهمیت سرویس می‌باشد.

#### - روش بر مبنای کیفیت سرویس

اخیرا کیفیت سرویس<sup>۱۳</sup> اهمیت زیادی در کشف و ترکیب وب سرویس‌ها یافته است. به منظور در نظر گرفتن کیفیت سرویس، هر وب سرویس به وسیله یک چندتایی<sup>۱۴</sup> نشان داده می‌شود که شامل توصیف سرویس، ویژگی‌های کیفیت سرویس و هزینه می‌شود. درخواست کنندگان سرویس انتخاب‌های بسیاری از مجموعه سرویس‌های با عملکرد مشابه و کیفیت‌های متفاوت در اختیار دارند.

<sup>8</sup>many-to-many

<sup>9</sup>Term-Frequency

<sup>10</sup>Single Vector Decomposition

<sup>11</sup>Latent Semantic Indexing

<sup>12</sup>preference degree

<sup>13</sup>QOS

<sup>14</sup>Tuple

روش کشف بر مبنای کیفیت سرویس با ویژگی‌های غیرعملکردی سرویس از قبیل کیفیت سرویس، سیاست‌های امنیتی، اطلاعات قیمت و دیگر قراردادهای مابین وب سرویس‌ها مرتبط است.

کشف سرویس بر مبنای هستان شناسی OWL-QOS شامل سه لایه است. لایه پروفایل QOS که وظیفه تطبیق را برعهده دارد. لایه تعریف ویژگی QOS، ساخت دامنه و تعریف ویژگی‌های آن را بر عهده دارد. لایه استانداردها<sup>15</sup> که استانداردها و معیارها را تعریف می‌نماید.

Cardoso, sheth [22] سرویس‌های متناسب را با استفاده از شباهت‌های نحوی، عملیاتی و معنایی کشف می‌نمایند. شباهت نحوی به‌وسیله مقایسه نام‌ها و مشخصات سرویس‌ها محاسبه می‌شود. شباهت معنایی و عملیاتی به ترتیب به‌وسیله اطلاعات کیفیت سرویس و مفاهیم ورودی /خروجی محاسبه می‌شود. کیفیت سرویس قبلاً به منظور ارتقا یا بهبود انتخاب سرویس استفاده می‌شد. [54] روشی بر مبنای ماشین حالت ارائه می‌کند که در این روش ابتدا پرس‌وجو یا همان درخواست کاربر به عنوان هدف در نظر گرفته شده و سپس پیش‌پالایش بر روی آن انجام می‌شود و سپس تطابق معنایی انجام گرفته و پس از آن تطابق کیفیت سرویس انجام می‌گیرد و پس از این مرحله سرویس کشف شده مشخص می‌شود. این روش از راهبردهای مختلف تطبیق استفاده می‌کند.

### - روش بر مبنای داده کاوی

به‌کارگیری روش‌های داده کاوی در کشف وب سرویس می‌تواند دقت در کشف وب سرویس را افزایش دهد. عدم وجود توصیف معنایی در وب سرویس‌ها نتایج کمی در جست‌وجو به‌دست می‌دهد [54] بسیاری از نتایج حاصل از یک پرس‌وجو بخشی از نیازهای کاربر را ارضا می‌نماید. دسته‌بندی وب سرویس معنایی از نمایش مناسبی برای گروه‌بندی وب سرویس‌های مشابه استفاده می‌کند. برای محاسبه مشابهت میان وب سرویس‌ها به منظور گروه‌بندی آن‌ها، [54] از ضریب همبستگی jaccard استفاده نموده است. به منظور گروه‌بندی وب سرویس‌های مشابه از تکنیک‌های دسته‌بندی استفاده شده است. میزان مشابهت سرویس با در نظر گرفتن توصیف سرویس و ویژگی‌های OWL-S محاسبه می‌شود. [34] woogle الگوریتم‌های دسته‌بندی را برای گروه‌بندی نام‌های پارامترهای عملیات وب سرویس در مفاهیم معنی‌دار معنایی ارائه نموده است. این الگوریتم بیش از حد بروی نام‌ها و پارامترها تکیه کرده است.

روش‌های برمبنای داده کاوی روش‌هایی داده - محور هستند. مشکل این است که با استفاده از این روش‌ها امکان بررسی رفتار کاربر وجود ندارد. بنابراین استفاده از این روش‌ها منجر به تمرکز بررسی وب سرویس‌های موجود به جای بررسی نحوه استفاده از آن‌ها شده است.

## 3- روش پیشنهادی

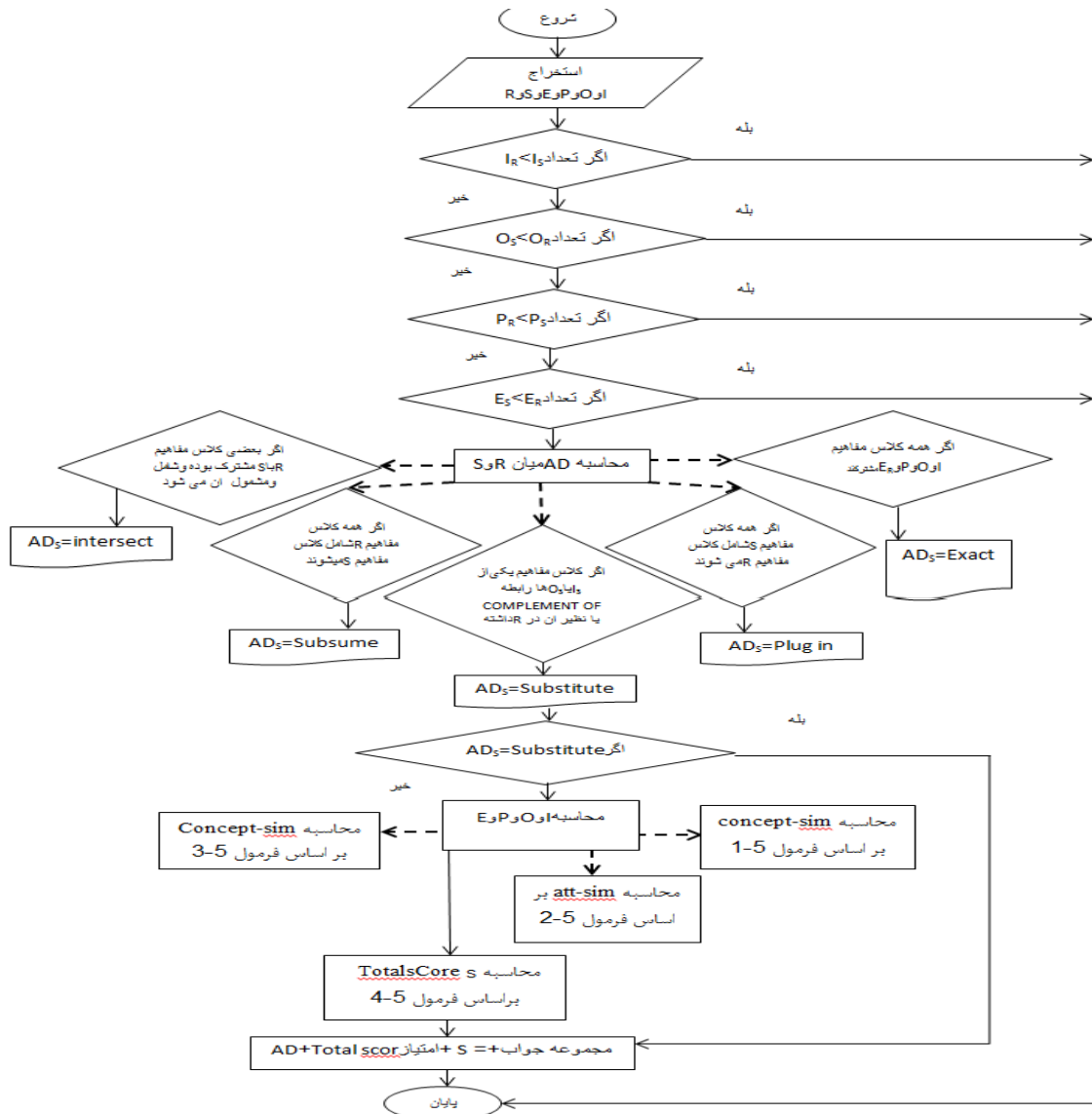
کشف وب سرویس معنایی شامل چهار مرحله است. بدیهی است که الگوریتم ارائه شده نیز از همین سیستم تبعیت نموده است. مرحله انتشار و تبلیغ سرویس با استفاده از ویژگی‌های پروفایل OWL-S اقدام به ثبت مشخصات تبلیغ سرویس غنی شده با معانی موجود در هستان شناسی در مخزن سرویس می‌نماید. در مرحله دوم درخواست غنی شده با معانی کاربر برای کشف سرویس مرتبط به سیستم تحویل داده می‌شود. مرحله بعدی که مهم‌ترین مرحله است تطبیق ورودی، خروجی، پیش‌شرط و اثر سرویس با نظایر آن‌ها در درخواست است که این مرحله در موتور تطبیق انجام می‌پذیرد. در نهایت سرویس‌های مرتبط تشخیص داده شده توسط موتور تطبیق، به درخواست کننده بازگردانده می‌شود.

### 1-3 الگوریتم روش پیشنهادی

الگوریتم ارائه شده بدین صورت می‌باشد.

<sup>15</sup>Metrics

- ورودی: سرویس پیشنهادی S، سرویس درخواست R، هستان شناسی، وزن مفاهیم W1، وزن مفاهیم W2 خروجی: مجموعه سرویس‌های مرتب شده براساس درجه تطبیق و امتیاز
- گام 1: استخراج ورودی، خروجی، پیش شرط و اثر سرویس پیشنهادی و سرویس درخواست
- گام 2: اگر تعداد ورودی‌های سرویس پیشنهادی بیشتر از ورودی‌های درخواست است به گام 11 و در غیر این صورت به گام سه برو.
- گام 3: اگر تعداد خروجی‌های درخواست بیشتر از سرویس پیشنهادی است به گام 11 و در غیر این صورت به گام 4 برو.
- گام 4: اگر تعداد پیش شرط‌های سرویس پیشنهادی بیشتر از درخواست است به گام 11 و در غیر این صورت به گام 5 برو.
- گام 5: اگر تعداد اثرهای درخواست بیشتر از سرویس پیشنهادی است به گام 11 و در غیر این صورت به گام 6 برو.
- گام 6: محاسبه درجه تطبیق میان درخواست و سرویس پیشنهادی
- 6-1: اگر تمامی رده‌های مفاهیم ورودی، خروجی، پیش شرط و اثر درخواست مشترک هستند درجه تطبیق سرویس پیشنهادی Exact است.
- 6-2: اگر تمامی رده‌های مفاهیم سرویس پیشنهادی شامل رده‌های مفاهیم سرویس درخواست می‌شوند درجه تطبیق سرویس پیشنهادی plug-in است.
- 6-3: اگر تمامی رده‌های مفاهیم سرویس درخواست شامل رده‌های مفاهیم سرویس پیشنهادی می‌شوند درجه تطبیق سرویس پیشنهادی subsume است.
- 6-4: اگر بعضی از رده‌های مفاهیم سرویس درخواست با سرویس پیشنهادی مشترک بوده یا شامل و یا مضمول آن‌ها می‌شوند درجه تطبیق سرویس پیشنهادی intersect است.
- 6-5: اگر رده مفاهیم یکی از ورودی‌ها یا خروجی‌های سرویس پیشنهادی رابطه complementOf با نظیر آن در سرویس درخواست داشته باشد درجه تطبیق سرویس پیشنهادی substitute می‌شود.
- گام 7: اگر درجه تطبیق substitute است به گام 10 برو و در غیر این صورت به گام 8 برو.
- گام 8: محاسبه subscore برای هر عنصر ورودی، خروجی، پیش شرط و اثر
- 8-1: محاسبه concept-sim براساس فرمول 1
- 8-2: محاسبه att-sim براساس فرمول 2
- 8-3: محاسبه Concept-sim براساس فرمول 3
- گام 9: محاسبه TotalsCore سرویس پیشنهادی براساس فرمول 4
- گام 10: اضافه نمودن سرویس پیشنهادی با درجه تطبیق و امتیاز TotalScore به مجموعه جواب.
- گام 11: پایان



شکل 1: روندنمای روش پیشنهادی

## 2-3 تطبیق ورودی، خروجی، پیش شرط و اثرهای درخواست و سرویس نامزد

### 1-2-3 تطبیق

روش‌های بسیاری برای تطبیق میان عناصر درخواست و سرویس وجود دارد. اکثر این روش‌ها از درخت طبقه‌بندی هستان‌شناسی استفاده می‌کنند. این روش بدین صورت است که با در نظر گرفتن سلسله مراتب هستان‌شناسی رابطه میان مفاهیم را می‌یابد. روابطی که در هستان‌شناسی میان مفاهیم وجود دارد به صورت *is-a* و *part-of* می‌باشد و هستان‌شناسی بر مبنای همین روابط شکل می‌گیرد. بدیهی است که هرچه به سمت بالای سلسله مراتب هستان‌شناسی برویم مفاهیم کلی‌تر شده و بالعکس با حرکت به سمت پایین مفاهیم جزئی‌تر می‌شوند. این روابط سلسله‌مراتبی را تشکیل می‌دهند که می‌توان بر روی آن استدلال نمود و روابط موجود میان مفاهیم را دریافت. برای استدلال بر روی مفاهیم هستان‌شناسی از استدلال‌گر استفاده می‌شود. استدلال‌گرهای مختلفی برای استدلال بر روی هستان‌شناسی موجود هستند اما هر یک از این استدلال‌گرها از الگوریتم استدلال خاصی برای استدلال استفاده می‌کنند. در الگوریتم ارائه شده برای استدلال بر روی هستان‌شناسی از استدلال‌گر معنایی *pellet* به دلیل پشتیبانی از قوانین *SWRL* و همچنین *OWL-DL* استفاده شده است. هدف از استدلال بر روی هستان‌شناسی کشف روابط



میان مفاهیم است. روابط میان مفاهیم بیشتر به معنی رابطه شمولی میان مفاهیم است. این رابطه مهم‌ترین و اصلی‌ترین رابطه‌ای است که برای تشخیص مشابهت میان دو مفهوم به کار می‌رود. مسئله اصلی برای تشخیص مشابهت میان عناصر وب سرویس معنایی در واقع همین مسئله است یعنی این که مهم نیست که این عناصر دارای نام یکسان یا مشابه یا مترادف باشند بلکه باید مفهومی که در هستان شناسی به آن اشاره می‌کنند مشابه یا یکسان باشد. در این صورت است که عناصر مشابه یکدیگر تشخیص داده می‌شوند. به همین منظور در الگوریتم ارائه شده پاسخ به این سوال که آیا مفهومی که عنصر وب سرویس نامزد به آن اشاره می‌کند آبرمفهوم یا زیرمفهوم مفهومی است که عنصر نظیر در درخواست به آن اشاره می‌کند، بررسی شده است. با تشخیص وضعیت مفاهیمی که عناصر درخواست و سرویس نامزد نسبت به یکدیگر دارند می‌توان مشابهت آن‌ها را دریافت. بر اساس رابطه شمولی میان مفاهیم می‌توان درجه تطبیقی بسته به این که کلیه عناصر مورد مقایسه سرویس و درخواست (منظور همان IOPE است) نسبت به یکدیگر چه وضعیتی دارند، به سرویس منتسب نمود. این درجه تطبیق نشان دهنده رابطه میان درخواست و سرویس است.

### 2-2-3 تطبیق IOPEها

ورودی و خروجی سرویس از نوع کلمه هستند یعنی از یک کلمه تشکیل شده‌اند اما پیش شرط و اثر از نوع عبارت هستند و بنابراین از مجموعه‌ای از کلمات تشکیل شده‌اند. بنابراین مقایسه میان ورودی و خروجی و نحوه تعیین درجه تطبیق سرویس با نظیر همین عمل در پیش شرط و اثر متفاوت است. به بیان دیگر، چون برای تعیین میزان مشابهت از رابطه شمولی استفاده می‌شود، تعیین این که یک ورودی / خروجی در درخواست نسبت به عنصر نظیر آن در سرویس چه وضعیتی دارد با تعیین این که دو پیش شرط یا اثر نسبت به یکدیگر در چه وضعیتی قرار دارند متفاوت است. از اینرو انجام تطبیق به دو بخش تقسیم می‌شود. بخش اول تطبیق ورودی و خروجی و بخش دوم تطبیق پیش شرط و اثر.

#### 1-2-2-3 تطبیق ورودی و خروجی درخواست با سرویس نامزد

برای تطبیق ورودی و خروجی درخواست و سرویس نامزد با توجه به مطالبی که مطرح شد و در نظر گرفتن رابطه شمولی میان مفاهیم مورد اشاره عناصر موجود در درخواست و سرویس، مقایسه میان آن‌ها به صورت زیر انجام می‌پذیرد. ابتدا تمامی رده‌های مفاهیمی (منظور همان مفهوم است در طبقه‌بندی هستان شناسی) که ورودی‌ها / خروجی‌های درخواست به آن‌ها اشاره می‌کنند، به صورت یک آرایه در نظر گرفته می‌شوند. سپس رده‌های مفاهیم ورودی‌ها / خروجی‌های سرویس از مخزن استخراج شده و آن‌ها نیز در یک آرایه قرار می‌گیرند. لازم به ذکر است که در هنگام ذخیره عناصر سرویس‌های تبلیغ در مخزن نام رده مفاهیمی که به آن اشاره می‌کنند نیز در مخزن ذخیره می‌شود. پس از این مرحله هر یک از عناصر آرایه رده مفاهیم ورودی/خروجی درخواست با تمامی عناصر موجود در آرایه رده مفاهیم ورودی/خروجی سرویس نامزد مورد مقایسه قرار می‌گیرند. مقایسه نیز بدین صورت است که کلاس‌های مفاهیم از نظر این که آیا با یکدیگر مساوی یا آبرمفهوم یا زیرمفهوم یکدیگر هستند مورد بررسی قرار می‌گیرند. تا اینجا آنچه که از مقایسه ورودی‌ها و خروجی‌ها گفتیم مربوط به انتساب درجات تطبیق Exact, plug-in, subsume, intersect بود. در این الگوریتم درجه تطبیق جدیدی به نام substitute نیز ارائه شده است که برای سرویس‌هایی است که هیچ یک از درجات تطبیق بالا را نسبت به درخواست ارائه شده دارا نمی‌باشند. این درجه تطبیق به دنبال سرویس‌هایی می‌باشد که گرچه نیاز کاربر را برطرف نمی‌کنند اما تا حدودی معادل آن را ارائه می‌دهند. این بدین معنی است که سرویس نامزدی با درجه تطبیق substitute نیاز کاربر را برطرف نمی‌کند بلکه جایگزین آن را ارائه می‌کند. به طور مثال درخواستی را در نظر بگیرید که در آن کاربر تقاضای رزرو هتل در تاریخ معین را نموده است. اگر هیچ یک از سرویس‌های نامزد موجود قادر به انجام چنین کاری نبودند و سرویسی موجود بود که امکان اقامت در کمپ را برای کاربر فراهم می‌آورد به عنوان سرویس جایگزین انتخاب می‌شود و درجه تطبیق آن substitute خواهد بود. به جهت نایل شدن به این درجه تطبیق، ورودی و خروجی سرویس و درخواست از نظر این که رده‌های مفاهیم آنان مکمل یکدیگر باشند، مورد بررسی قرار می‌گیرند.

#### 2-2-2-3 تطبیق پیش شرط و اثر درخواست با سرویس نامزد

پیش شرط و اثر به صورت عبارت هستند و بنابراین نحوه انتساب درجه تطبیق (مشخص نمودن درجه تطبیق) به آنان با ورودی و خروجی متفاوت است. با استفاده از قوانین SWRL می توان پیش شرط و اثر را توصیف نمود. این قوانین متشکل از یک اتم و یک گزاره و یا از دو اتم و یک گزاره تشکیل شده اند. هر یک از این اتم ها به مفهومی در هستان شناسی اشاره می کنند. اما گزاره ها بر دو نوع هستند، یک نوع که با classpredicate مشخص می شوند و گزاره های رده هستند مانند اتم ها به مفاهیم هستان شناسی اشاره می کنند. اما نوع دیگر propertypredicateها هستند که به ویژگی های اشیا در هستان شناسی اشاره می کنند.

برای تطبیق عبارات پیش شرط و اثر باید رده های مفاهیم اتم های درخواست و سرویس با یکدیگر و گزاره های درخواست و سرویس نیز با یکدیگر مقایسه شود. بنابراین برای این کار به یک آرایه برای مفاهیم اتم ها و یک آرایه برای گزاره هایی از نوع ClassPredicate و گزاره هایی از نوع PropertyPredicate نیاز است. اتم ها و گزاره های عبارات را جدا نموده و پس از پیدا نمودن مفهوم مورد اشاره، در آرایه قرار داده می شود. سپس عناصر هر آرایه با آرایه نظیر آن مقایسه می شود. این مقایسه نیز از جنس مقایسه های میان ورودی ها و خروجی ها می باشد. راه حل ارائه شده در این الگوریتم برای حل این مسئله بدین صورت است: هر عبارت متشکل از دو یا سه جزء می باشد. بنابراین می توان گفت وضعیت یک عبارت نسبت به دیگری به هر دو یا سه جزء آن وابسته است. یعنی اگر هر سه جز یک عبارت در وضعیت یکسانی نسبت به اجزای متناظر عبارت دیگر بودند. آنگاه کل عبارت دارای همان وضعیت نسبت به عبارت دیگر است.

از دیگر ویژگی های که این الگوریتم را از دیگر الگوریتم های تطبیق متمایز می سازد، در نظر گرفتن و مقایسه میان ویژگی های مفاهیم است.

### 3-3 رتبه بندی سرویس های تبلیغ

در بخش قبل نحوه مقایسه IOPE سرویس نامزد با درخواست بررسی شد. نتیجه این بخش تعیین درجه تطبیق سرویس نامزد می باشد. اما با تعیین درجه تطبیق یک سرویس نمی توان مناسب ترین سرویس را از میان سرویس های نامزد موجود در مخزن انتخاب نمود زیرا اگر چند سرویس با درجه تطبیق یکسان موجود باشد چگونه می توان تشخیص داد که کدامیک از این سرویس ها به سرویس درخواست شده نزدیکتر است. به بیان دیگر، کدامیک از سرویس ها نیازهای کاربر را بیشتر برطرف می نماید. اینجاست که بحث رتبه بندی سرویس مطرح می شود.

برای رتبه بندی سرویس نیاز به منتسب کردن امتیاز به سرویس است. در اینجا سوالی مطرح می شود که امتیازدهی به سرویس بر چه معیاری باید باشد. پاسخ به این سوال منجر به انتخاب بهترین سرویس خواهد شد. معیارهای مختلفی برای امتیازدهی به سرویس وجود دارد. برخی از این معیارها عبارتند از زمان پاسخدهی سرویس، کیفیت سرویس، میزان شباهت عناصر سرویس به یکدیگر و غیره. معیارهای دیگری نیز برای امتیازدهی وجود دارند اما می توان سه معیار بالا را از اصلی ترین معیارهای امتیازدهی به سرویس برشمرد. معیارهای دیگر بیشتر به کاربر درخواست کننده باز می گردد از قبیل این که مثلا سرویس توسط فراهم کننده خاصی که کاربر اعتماد بیشتری به آن دارد، ارائه می شود.

از میان سه معیار مطرح شده در بالا، معیار شباهت عناصر سرویس معیاری مناسب برای امتیازدهی به سرویس می باشد. علت مناسب بودن این معیار این است که در کشف سرویس آنچه که حائز اهمیت است میزان شباهت عناصر سرویس و درخواست به یکدیگر می باشد. معیارهایی از قبیل زمان پاسخدهی سرویس و کیفیت سرویس بیشتر برای پالایش سرویس مناسب هستند. بنابراین در الگوریتم ارائه شده برای امتیازدهی و در نتیجه آن رتبه بندی سرویس از معیار شباهت عناصر سرویس به یکدیگر استفاده شده است. مقصود از عناصر سرویس همان عناصر پروفایل سرویس یعنی ورودی و خروجی و پیش شرط و اثر می باشد. مشخص نمودن امتیاز سرویس بر اساس معیار شباهت عناصر سرویس به دو بخش تقسیم می شود. یک بخش میزان شباهت مفاهیمی است که عناصر پروفایل سرویس به آن ها اشاره می کنند و بخش دیگر میزان شباهت ویژگی های مفاهیم است. هر یک از این بخش ها سهمی در تعیین امتیاز سرویس دارند. این سهم همان طور که در ادامه خواهد آمد به میزان اهمیت هر یک بستگی دارد.

### 1-3-3 نحوه امتیازدهی به سرویس‌ها بر اساس شباهت میان مفاهیم

الگوریتم تطبیق شامل فراخوانی تابع اندازه‌گیری شباهت می‌باشد. تابع اندازه‌گیری بر اساس درخت سلسله مراتبی که نشان دهنده رابطه شمولی میان تمامی مفاهیم هستان شناسی است که شامل مفاهیم مورد مقایسه هستند، و عددی در بازه صفر و یک که نشان دهنده درجه شباهت می‌باشد را به‌عنوان نتیجه باز می‌گرداند. این تابع از تعدادی مقادیر طبقه‌بندی گره‌های درختی که مفاهیم در آن قرار دارند استفاده می‌کند. این مقادیر شامل کوتاهترین فاصله میان گره‌ها، عمق گره‌های مفاهیم، و عمق گره‌ها برای مفاهیم مورد مقایسه مشترک می‌شود.  $R$  را ریشه درخت سلسله مراتب مفاهیم در نظر می‌گیریم.  $a$  و  $b$  دو مفهوم هستند،  $d$  مینیمم فاصله میان مفاهیم  $L_a$  و  $L_b$  عمق‌های مرتبط با نقطه مشترک میان دو مفهوم می‌باشند. برای مثال اگر  $a$  مفهوم  $C_{10}$  درخت در شکل 2 و  $b$  مفهوم  $C_5$  باشند، بنابراین  $d=3$  و  $I_a=3$  و  $L_b=2$  و  $m_a=2$  و  $m_b=1$  می‌باشند.

در ادامه برخی ویژگی‌های محاسبه شباهت معنایی و مقادیر طبقه‌بندی را بیان می‌کنیم.

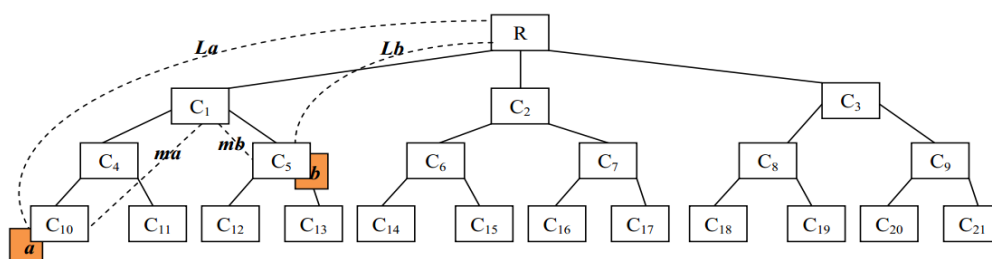
1. شباهت معنایی میان مفاهیم  $a$  و  $b$  با افزایش  $d$ ، به‌صورت خطی کاهش می‌یابد.
  2. شباهت معنایی میان مفاهیم  $a$  و  $b$  با افزایش  $(I_a+I_b)$ ، به‌صورت خطی افزایش می‌یابد.
  3. شباهت معنایی مفاهیم  $a$  و  $b$  با افزایش  $|I_a-I_b|$ ، به‌طور خطی کاهش می‌یابد.
  4. شباهت معنایی مفاهیم  $a$  و  $b$  با افزایش  $\max(m_a, m_b)$ ، به‌طور خطی کاهش می‌یابد.
  5. شباهت معنایی مفاهیم  $a$  و  $b$  با افزایش  $|m_a-m_b|$ ، به‌طور خطی کاهش می‌یابد.
- تابع  $S(a,b)$  باید ویژگی‌های زیر را داشته باشد.

$$0 \leq S(a,b) \leq 1$$

$$a: S(a,a)=1 \forall$$

$$a,b: S(a,b)=S(b,a) \forall$$

اولین ویژگی برد تابع را مشخص می‌کند، همان‌طور که شباهت معنایی دو مفهوم که کاملاً نامرتبط هستند حداقل صفر است و اگر دو مفهوم یکسان باشند ماکزیمم میزان شباهت آن‌ها یک یا همان 100٪ است. دومین ویژگی خاصیت انعکاسی تابع را بیان می‌دارد که شباهت میان هر مفهوم و خودش، یک می‌باشد، و این نشان دهنده این است که هر مفهوم همانند خودش است. آخرین ویژگی خاصیت تقارنی است.



شکل 2: مثال از درخت طبقه‌بندی هستان شناسی

پارامتر تنظیم‌پذیری است.  $\beta$  حال با ویژگی‌های مطرح شده در بالا فرمول 1 معرفی می‌شود. در این فرمول

$$\text{Concept-Sim} = \begin{cases} 1 & \max(m_a, m_b) = 0 \\ \frac{B \cdot (L_a + L_b)}{\max(m_a, m_b) + d + |m_a - m_b| + |L_a - L_b|} & \text{etc} \end{cases} \quad (\text{فرمول 1})$$

همان‌طور که ملاحظه می‌کنید این تابع پارامتری است.  $\text{Max}(m_a, m_b)$  اگر صفر باشد بدین معنی است که هر دو مفهوم یکسان هستند و بنابراین میزان شباهت آنان بیشترین مقدار یعنی یک می‌باشد. و اگر  $\text{Max}(m_a, m_b)$  صفر نباشد بدین معنی است که مفاهیم یکسان نیستند و میزان شباهت آنان باید به‌وسیله فرمول ارائه شده محاسبه شود.

### 2-3-3 نحوه تعیین میزان شباهت ویژگی مفاهیم

همان‌طور که در بخش راهکارهای پیشین مطرح شد روش‌های کمی به بررسی ویژگی میان مفاهیم پرداخته‌اند. در الگوریتم ارائه شده مقایسه میان ویژگی مفاهیم بیشتر در مورد انتساب امتیاز به سرویس به‌کار می‌رود. این‌که ویژگی‌های مفاهیم مورد اشاره عناصر درخواست / سرویس مورد مقایسه قرار گیرند، برای کمک به تشخیص بهتر رابطه میان مفاهیم است. بدیهی است که رابطه شمولی می‌تواند رابطه میان مفاهیم را کشف کند اما جواب استنتاج بر روی درخت سلسله مراتب هستان شناسی یک جواب درست و غلط است و بنابراین به فهم دقیق‌تر رابطه میان دو مفهوم کمکی نمی‌کند. بر همین اساس نیاز به مقایسه میان مفاهیم در کشف وب سرویس‌ها احساس می‌شود.

برای مقایسه میان مفاهیم ابتدا یک آرایه برای هر یک از مفاهیم در نظر گرفته شده که در آن‌ها ویژگی‌های هر یک از مفاهیم درج می‌شود. سپس این مفاهیم با یکدیگر مقایسه می‌شود. این مقایسه از نوع معنایی نیست بلکه از نوع لغوی می‌باشد. پس از انجام مقایسه میان آرایه ویژگی‌های مفهوم درخواست با سرویس نامزد تعداد ویژگی‌های مشترک به تعداد کل ویژگی‌های دو مفهوم تقسیم شده و در نتیجه امتیاز ویژگی‌های مفاهیم بدین صورت مشخص می‌شود. از این امتیاز برای رتبه‌بندی سرویس نامزد استفاده می‌شود. امتیاز شباهت ویژگی‌های مفاهیم مطابق فرمول 2 به‌دست می‌آید.

$$\text{att-Sim} = \frac{\text{Number of shared attributes}}{\text{Total number of attributes}}$$

(فرمول 2)

که در فرمول 2 att-Sim امتیاز شباهت ویژگی‌ها است که از تقسیم ویژگی‌های مشترک دو مفهوم بر تعداد کل ویژگی‌های هر دو مفهوم به‌دست می‌آید. امتیازدهی و رتبه‌بندی سرویس‌های تبلیغ در بخش بعد مورد بررسی قرار خواهد گرفت.

### 3-3-3 نحوه محاسبه امتیاز هر عنصر پروفایل سرویس

در این بخش به روش محاسبه امتیاز هر یک از ورودی‌ها، خروجی‌ها، پیش شرط‌ها و اثرها می‌پردازیم. در این روش از میزان شباهت مفاهیم<sup>16</sup> و میزان شباهت ویژگی‌های مفاهیم<sup>17</sup> برای محاسبه امتیاز هر زوجی از مفاهیم که با یکدیگر مقایسه می‌شوند استفاده شده است. در محاسبه امتیاز، امتیاز هر زوج از مفاهیم که با یکدیگر مقایسه می‌شوند محاسبه می‌شود. این زوج که شامل یک مفهوم از درخواست و یک مفهوم از سرویس نامزد می‌باشد از نظر روابط معنایی در سلسله مراتب هستان شناسی با یکدیگر مقایسه شده و نسبت به یکدیگر آبرمفهوم، زیرمفهوم یا یکسان هستند. لازم به ذکر است که هر عنصر از درخواست با تمامی عناصر نظیرش در سرویس نامزد مقایسه شده و زوجی که دارای امتیاز ماکزیمم باشد برای محاسبه امتیاز سرویس در نظر گرفته می‌شود. امتیاز هر عنصر از ورودی، خروجی، پیش شرط و اثر توسط فرمول 3 محاسبه می‌شود:

$$\text{Subscore} = w_1 * \text{Concept\_Sim} + w_2 * \text{att\_Sim} \quad (\text{فرمول 3})$$

که در آن  $w_1$  و  $w_2$  به ترتیب وزن‌هایی هستند که به شباهت مفاهیم و شباهت ویژگی‌های مفاهیم داده می‌شود. این وزن بسته به میزان اهمیت هر یک از شباهت مفاهیم و شباهت ویژگی‌ها برای کاربر دارد. باید به این نکته توجه شود که  $w_1 + w_2 = 1$ .

<sup>16</sup>Concept-Sim

<sup>17</sup>att-Sim

پس از این که امتیاز هر زوج<sup>۱۸</sup> محاسبه شد، امتیاز سرویس محاسبه می‌شود. در واقع امتیاز سرویس میانگین امتیاز تمام زوج‌هایی که با یکدیگر مقایسه شده و امتیاز آن‌ها محاسبه شده است، می‌باشد. علت در نظر گرفتن میانگین سخت‌گیری بیشتر در کشف سرویس بوده است. این بدین معنی است که اگر در میان زوج‌های مقایسه شده، زوج یا زوج‌هایی یافت شد که دارای امتیاز پایین‌تری نسبت به بقیه بود، امتیاز سرویس را تا حدودی کاهش دهد.

برای محاسبه امتیاز کلی سرویس تمامی امتیازات محاسبه شده از مقایسه مفاهیم درخواست و سرویس که از نظر سلسله مراتب هستان شناسی با یکدیگر مشابهت داشتند با یکدیگر جمع شده و بر تعداد زوج‌هایی که با یکدیگر مقایسه شده و مشابهت آن‌ها تعیین شده است تقسیم می‌شوند. فرمول محاسبه امتیاز کل سرویس به صورت زیر می‌باشد که در آن TotalScore امتیاز کل سرویس و SubScore<sub>i</sub> امتیاز مشابهت عنصر i ام و n تعداد زوج‌های مشابه می‌باشد.

$$\text{TotalScore} = \frac{\sum_{i=1}^n \text{subscore } i}{n} \quad (\text{فرمول 4})$$

پس از محاسبه امتیاز سرویس (TotalScore) هر درجه تطبیق را به صورت مجموعه‌ای در نظر گرفته و سرویس‌ها را بر اساس TotalScore در آن مرتب می‌نماییم.

#### 4- ارزیابی روش پیشنهادی

الگوریتم ارائه شده برای کشف وب سرویس معنایی را با استفاده از وب سرویس‌های معنایی مجموعه داده [37] OWLS-TC ارزیابی می‌نماییم. از 554 وب سرویس معنایی این مجموعه 100 وب سرویس را در دامنه‌های مختلفی انتخاب نمودیم. تعداد 20 پرس‌وجو از میان 28 پرس‌وجوی موجود در OWLS-TC انتخاب شده است. پرس‌وجوها به نحوی انتخاب شده‌اند که در رابطه با مجموعه سرویس‌های انتخاب شده باشند. پرس‌وجوها نیز مانند سرویس‌ها از دامنه‌های مختلفی انتخاب شده‌اند.

#### 4-1 معیارهای ارزیابی

برای ارزیابی میزان دقت الگوریتم FSM معیارهای Precision، Recall و F-score مورد استفاده قرار گرفته‌اند. Precision به صورت توانایی فراهم آوردن وب سرویس‌های مرتبط از مجموعه وب سرویس‌های بازگردانده شده به عنوان نتیجه، تعریف می‌شود. Recall نیز به عنوان توانایی فراهم آوردن تعداد ماکزیمم سرویس‌های مرتبط از مجموعه وب سرویس‌های مرتبط، تعریف می‌شود. F-score میانگین Precision و Recall می‌باشد. تعریف ریاضی این مفاهیم به صورت زیر است:

(فرمول 5)

$$\text{Precision} = \frac{\text{Number of relevant web Services retrieved}}{\text{Total Number OF retrieved Web Services from dataset}}$$

(فرمول 6)

$$\text{Recall} = \frac{\text{Number of relevant Web Services retrieved}}{\text{Total Numver of relevant Web Services in the dataset}}$$

(فرمول 7)

$$F - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

<sup>18</sup>SubScore

نتایج به دست آمده از این معیارها از 20 پرس و جو می‌باشد. برای ارزیابی از macro-average استفاده شده است که در آن میانگین مقادیر Precision مجموعه جواب‌های بازگردانده شده به وسیله تطبیق دهنده برای تمام درخواست‌ها (پرس‌وجوها) موجود در مجموعه آزمایش در سطوح Recall استاندارد ( $0 \leq i < \lambda$ ) محاسبه می‌شود. درون‌یابی برای تخمین مقادیر Precision که در مجموعه جواب برای برخی درخواست‌ها در سطوح Recall مورد نظر مشاهده نشده‌اند، استفاده شده است. تعداد سطوح Recall از 0 تا 1 می‌باشد. (با گام‌های،  $n=1..10$ ). بنابراین Precision که macro-average باشد به صورت زیر تعریف می‌شود:

(فرمول 8)

$$Precision_i = \frac{1}{|Q|} \times \sum_{q \in Q} \max\{P_0 | R_0 \geq Recall_i \wedge (R_0, P_0) \in O_q\}$$

که در آن  $O_q$  مجموعه مقادیر recall/precision مشاهده شده برای پرس‌وجوی  $q$  است.

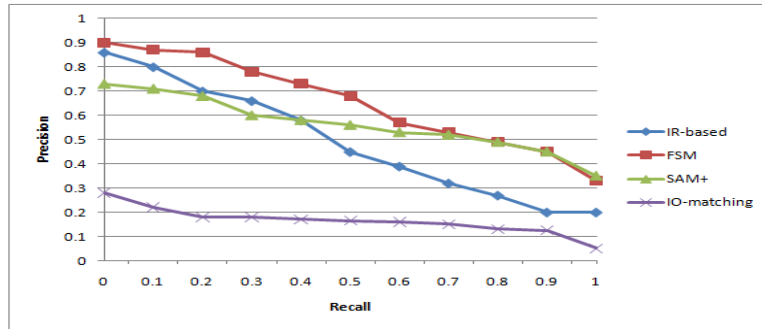
پس در ابتدا Precision و Recall اولیه برای هر پرس‌وجو محاسبه می‌شود. برای محاسبه Precision و Recall اولیه از تعداد سرویس‌های مختلف در هر مرحله استفاده شده است. برای مثال در محاسبه Precision و Recall برای پرس‌وجوی شماره 1 در مرحله اول 10 سرویس، در مرحله دوم 20 سرویس و همین‌طور تا مرحله دهم که 100 سرویس را به عنوان مجموعه داده در نظر گرفته‌ایم. از این رو مجموعه سرویس‌های مرتبط بسته به سرویس‌هایی که درون مجموعه مورد آزمایش قرار دارند در هر مرحله تغییر می‌کند.

پس از به دست آوردن Precision و Recall‌های مختلف به محاسبه Precision برای سطوح Recall مختلف می‌پردازیم. همان‌طور که گفته شد سطوح Recall به صورت  $\frac{n}{\lambda}$  محاسبه می‌شود که مقدار Recall در اینجا 10 در نظر گرفته شده و  $n=1..10$  می‌باشد. بنابراین سطوح Recall از 0.1 تا 1 است. محاسبه F-Score نیز بدین صورت است که Precision‌های به دست آمده را برای هر یک از سطوح Recall در فرمول 7 قرار می‌دهیم و مقدار آن محاسبه می‌شود. محاسبه میانگین زمان پاسخ نیز بدین صورت است که زمان اجرای الگوریتم به ازای 20 پرس‌وجو برای تعداد معینی سرویس در هر مرحله محاسبه می‌شود که تعداد سرویس‌ها در هر مرحله 100 و ... و 20 و 10 سرویس می‌باشد. سپس میانگین زمان اجرای به دست آمده از 20 پرس‌وجو در هر مرحله محاسبه می‌شود.

## 2-4 ارزیابی نتایج

برای مقایسه الگوریتم پیشنهادی با کارهای مشابه از سه الگوریتم دیگر نیز استفاده شده است. الگوریتم اول، تطبیق ورودی-خروجی می‌باشد و همان‌طور که از نام آن پیداست سرویس را فقط براساس مقایسه ورودی و خروجی کشف می‌کند. الگوریتم دوم از روش‌های بازیابی اطلاعات که بر مبنای بررسی متون هستند، برای کشف سرویس استفاده می‌نماید. الگوریتم سوم الگوریتم SAM+ [48] است. این الگوریتم مقایسه را بر مبنای IOPE انجام می‌دهد اما از ماژول شباهت WordNet برای تشخیص شباهت میان مفاهیم استفاده نموده است.

در ابتدا سه الگوریتم از نظر معیار Precision/Recall مورد مقایسه قرار گرفته‌اند که نتایج آن در نمودار شکل 3 آمده است.



شکل 3: نمودار Precision/Recall

همان‌طور که در نمودار شکل 3 مشاهده می‌شود، الگوریتم پیشنهادی نسبت به دیگر الگوریتم‌های مورد مقایسه دارای Precision/Recall بالاتری می‌باشد. تمامی الگوریتم‌ها در سطوح Recall پایین یعنی زمانی که نسبت سرویس‌های مرتبط کشف شده توسط الگوریتم به کلیه سرویس‌های مرتبط موجود در مجموعه سرویس‌های مرتبط کوچک می‌باشد، دارای Precision بالایی است. یعنی نسبت سرویس‌های مرتبطی که کشف کرده است به تمامی سرویس‌های کشف شده، نسبت بالایی است. اما به تدریج با افزایش Recall مشاهده می‌شود که مقدار Precision کاهش می‌یابد. این بدین معنی است که هر چه سرویس‌های مرتبط بیشتری توسط الگوریتم‌ها یافت می‌شوند تعداد سرویس‌های نامرتبط کشف شده بیشتر می‌شود. این روند در همه الگوریتم‌های مورد مقایسه دیده می‌شود اما الگوریتمی کارا تر و موثرتر است که در Recall‌های بالاتر دارای Precision بالاتری باشد. این بدین معنی است که هر چه تعداد سرویس‌های مرتبط بیشتری یافت می‌شود، تعداد سرویس‌های نامرتبط کمتری یافت شود. در واقع می‌توان گفت که این یک حالت آرمانی است که با افزایش تعداد سرویس‌های مرتبط کشف شده تعداد سرویس‌های نامرتبط کاهش قابل توجهی داشته باشد. اما الگوریتم‌هایی که برای کشف سرویس ارائه طراحی شده‌اند، سعی دارند به این حالت نزدیک شوند. از این رو الگوریتمی کارایی بهتری دارد که در سطوح Recall بالا Precision بالایی نیز داشته باشد.

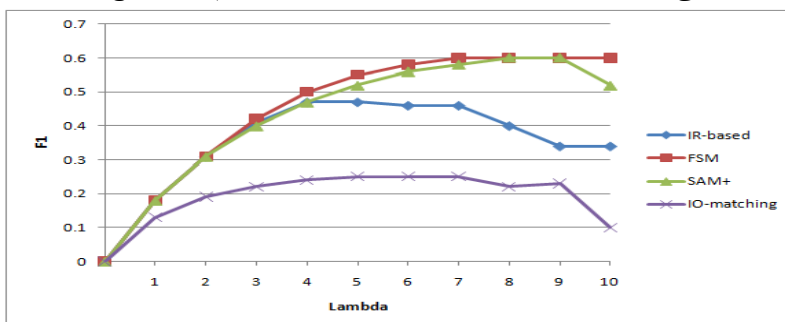
همان‌طور که در نمودار شکل 3 نشان داده شده است، الگوریتم پیشنهادی دارای ویژگی مذکور می‌باشد. یعنی الگوریتم پیشنهادی نسبت به دیگر الگوریتم‌های مورد مقایسه در سطوح Recall بالا دارای Precision بالاتری می‌باشد که این گویای کارا تر و مؤثرتر بودن الگوریتم پیشنهادی می‌باشد. این مطلب نیز قابل توجه است که در ابتدا الگوریتمی که از روش‌های بازیابی اطلاعات استفاده نموده دارای Precision بالایی است اما با افزایش Recall مقدار Precision به شدت کاهش پیدا می‌کند تا جایی که حتی از الگوریتم تطبیق IOPE نیز پایین‌تر می‌رود. این روند نشان دهنده کارایی پایین الگوریتم‌هایی است که از روش‌های سنتی بازیابی اطلاعات استفاده می‌کنند.

دیگر الگوریتم مورد بررسی از تطبیق ورودی و خروجی برای کشف سرویس استفاده می‌کند. همان‌طور که مشاهده می‌شود مقادیر Precision این الگوریتم در سطوح Recall مورد بررسی از الگوریتم پیشنهادی بسیار پایین‌تر می‌باشد. علت آن نیز به وضوح مشخص است زیرا که این الگوریتم تنها از تطبیق ورودی و خروجی برای کشف سرویس استفاده نموده و بنابراین دقت کمتری را نسبت به الگوریتم‌هایی که علاوه بر ورودی و خروجی از پیش شرط و اثر استفاده می‌کنند، دارد. حتی میزان دقت این الگوریتم از دیگر الگوریتم‌های مورد مقایسه نیز کمتر است و علت آن فقط مقایسه لغوی ورودی و خروجی به صورتی ابتدایی (مقایسه رشته‌ها و نویسه‌ها با یکدیگر) می‌باشد.

در میان الگوریتم‌های مورد بررسی الگوریتم SAM+ میزان Precision‌های نزدیکتری به الگوریتم پیشنهادی دارد ولی میزان دقت Precision آن همواره کمتر از الگوریتم پیشنهادی است. علت این است که این الگوریتم از ماژول شباهت WordNet برای تشخیص مشابهت میان مفاهیم استفاده نموده است و دیگر این که از مقایسه ویژگی‌های مفاهیم با یکدیگر استفاده نکرده است. به همین علت دارای میزان دقت پایین‌تری نسبت به الگوریتم پیشنهادی می‌باشد. نکته دیگری که در مقایسه نمودار این دو الگوریتم وجود دارد تغییرات مقدار Precision دو الگوریتم است. الگوریتم پیشنهادی با این که تغییرات بیشتری در مقدار Precision نسبت

به الگوریتم SAM+ دارد اما در نهایت مشاهده می‌شود که مقدار Precision آن در سطح Recall=1.0 به شدت کاهش پیدا می‌کند، در صورتی که در الگوریتم پیشنهادی این اتفاق مشاهده نمی‌شود و علت این امر آنست که با افزایش میزان سرویس‌ها تعداد سرویس‌های نامرتب در مجموعه سرویس‌های کشف شده توسط الگوریتم تطبیق IOPE به‌طور قابل ملاحظه‌ای افزایش می‌یابد اما در الگوریتم پیشنهادی اینطور نیست. علاوه بر این ملاحظه می‌شود که در نمودار الگوریتم پیشنهادی تغییرات ناگهانی در مقدار Precision با افزایش recall به تدریج کاهش می‌یابد اما در الگوریتم SAM+ کاهش ناگهانی مقدار Precision در سطوح Recall بالا اتفاق می‌افتد.

پس از بررسی Precision و Recall به بررسی F1-score می‌پردازیم. همان‌طور که از فرمول F1-score پیداست، مقدار آن به مقادیر Precision و Recall بستگی دارد. نمودار شکل 4 F1-score را برای الگوریتم‌های مطرح شده در بالا نشان داده است.



شکل 4: نمودار F1-score الگوریتم‌ها

محور افقی مقدار را نشان می‌دهد. همان‌طور که در بخش قبل گفته شد، حداکثر مقدار، 10 در نظر گرفته شده است. مطابق آنچه که در نمودار Precision/Recall دیده شد، در این نمودار نیز الگوریتم پیشنهادی دارای بالاترین میزان F1-score نسبت به سایر الگوریتم‌ها می‌باشد. ارزیابی در جدول 1 با توجه به نمودار شکل 4 مشاهده می‌شود. میزان دقت الگوریتم پیشنهادی نسبت به روش تطبیق ورودی- خروجی، در حدود 48.2 درصد، نسبت به روش IR-based در حدود 16.4 درصد و نسبت به روش SAM+ در حدود 7.6 درصد بهبود داشته است.

جدول 1: F1-score الگوریتم‌ها

|      |      |      |      |      |      |      |      |      |      |             |
|------|------|------|------|------|------|------|------|------|------|-------------|
| 10   | 9    | 8    | 7    | 6    | 5    | 4    | 3    | 2    | 1    | -----       |
| 0.1  | 0.23 | 0.22 | 0.25 | 0.25 | 0.25 | 0.24 | 0.22 | 0.19 | 0.13 | IO-matching |
| 0.34 | 0.34 | 0.4  | 0.46 | 0.46 | 0.47 | 0.47 | 0.41 | 0.31 | 0.18 | IR-based    |
| 0.52 | 0.6  | 0.6  | 0.58 | 0.56 | 0.52 | 0.47 | 0.4  | 0.31 | 0.18 | SAM+        |
| 0.6  | 0.6  | 0.6  | 0.6  | 0.58 | 0.55 | 0.5  | 0.42 | 0.32 | 0.19 | FSM         |

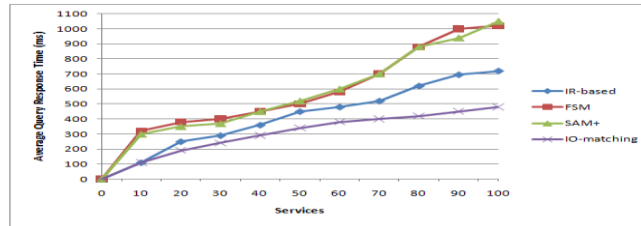
روش پیشنهادی دارای دقت بالاتری نسبت به روش‌های قبلی می‌باشد زیرا نسبت سرویس‌های مرتبطی که کشف کرده است به تمامی سرویس‌های کشف شده، بالا است و بر اساس معنا می‌باشد.

روش بر مبنای بازبازی اطلاعات یک روش سنتی است و دارای کارایی پایین در کشف سرویس می‌باشد. این روش بر روی واژگان و اهمیت آن‌ها در متون تاکید دارد و برای کشف بهترین سرویس‌ها کافی نیستند زیرا متون موجود در وب سرویس‌ها بسیار خلاصه و چکیده شده هستند و اطلاعات ساختاری که به اخذ معانی عملیات کمک می‌کند را نادیده می‌گیرد و در نتیجه دارای دقت پایینی در کشف سرویس می‌باشد.

روش تطبیق ورودی- خروجی تنها از تطبیق ورودی و خروجی برای کشف سرویس استفاده نموده و بنابراین دقت کمتری را نسبت به الگوریتم‌هایی که علاوه بر ورودی و خروجی از پیش شرط و اثر استفاده می‌کنند، دارد. این روش از مقایسه لغوی ورودی و خروجی به‌صورتی ابتدایی استفاده می‌کند.



روش SAM+ از ماژول شباهت WordNet برای تشخیص مشابهت میان مفاهیم استفاده نموده است و از مقایسه ویژگی‌های مفاهیم با یکدیگر استفاده نکرده است. به همین علت دارای میزان دقت پایین‌تری نسبت به الگوریتم پیشنهادی می‌باشد. اکنون به بررسی سومین معیار مقایسه یعنی میانگین زمان پاسخ می‌پردازیم. نمودار میانگین زمان پاسخ در شکل 5 آمده است.



شکل 5: نمودار میانگین زمان پاسخ الگوریتم‌ها

همان‌طور که مشاهده می‌شود، میانگین زمان پاسخ الگوریتم ارائه از الگوریتم‌های مقایسه ورودی و خروجی، الگوریتم بر مبنای روش‌های سنتی بازیابی اطلاعات بالاتر می‌باشد. دلیل این مطلب این است که الگوریتم‌های فوق فقط ورودی و خروجی را با یکدیگر مقایسه می‌نمایند و بنابراین در زمان کمتری به پاسخ خواهند رسید. اما الگوریتم SAM+ و الگوریتم پیشنهادی علاوه بر مقایسه ورودی و خروجی، پیش‌شرط و اثر را نیز مقایسه می‌نمایند که این خود بر زمان پاسخ‌دهی می‌افزاید. در مورد مقایسه الگوریتم پیشنهادی با الگوریتم مقایسه IOPE ملاحظه می‌شود که الگوریتم پیشنهادی در برخی از موارد میانگین زمان پاسخ بالاتری نسبت به الگوریتم مقایسه IOPE دارد. علت این امر این است که الگوریتم مقایسه IOPE با استفاده از ماژول WordNet به تعیین میزان مشابهت میان مفاهیم می‌پردازد.

## 5- آنالیز آماری

در این بخش نتایج از لحاظ آماری بررسی می‌شود. هدف از آزمون آماری که در این بخش بر روی داده‌ها انجام می‌شود، اثبات تصادفی نبودن داده‌ها و این که داده‌ها از روی تصادف به‌دست نیامده‌اند، می‌باشد. آزمون آماری که بر روی داده‌ها انجام شده است، آزمون مان-ویتنی می‌باشد که در ادامه توضیح داده خواهد شد.

### 5-1 آزمون مان-ویتنی

آزمون مان-ویتنی، آزمونی ناپارامتری است که برای آزمودن این فرض که دو نمونه از یک جامعه اخذ شده‌اند یا این که آیا مشاهدات در یک نمونه بزرگتر از مشاهدات در نمونه دیگر است، استفاده می‌شود. با این که این آزمون، یک آزمون ناپارامتری است که فرض می‌کند دو توزیع آماری دارای یک شکل هستند.

نحوه انجام این آزمون بدین صورت است که دو نمونه را به‌صورت نزولی در یک مجموعه مرتب کرده و آن‌ها را رتبه‌بندی می‌نماید. سپس با توجه به رتبه‌بندی‌های انجام شده و براساس فرمولی که در ادامه می‌آید، مقداری به نام  $U$  را به‌دست آورده و این مقدار را با جدول آزمون مان-ویتنی مقایسه نموده و اگر این مقدار کمتر از مقدار  $U$  جدول یا همان  $U_{critical}$  بود فرض را رد می‌کند و اگر این مقدار بزرگتر بود، فرض را می‌پذیرد.

فرمول به‌دست آوردن  $U$  به صورت زیر است:

$$u = n_a n_b + \frac{n_a(n_a+1)}{2} - T_A \quad \text{فرمول (9)}$$

که  $n_a$  و  $n_b$  ترتیب اندازه نمونه  $A$  و اندازه نمونه  $B$  هستند و  $T_A$  مجموع رتبه‌های نمونه  $A$  است.

### 5-2 نتایج آنالیز آماری

آزمون آماری بر روی داده‌های Precision الگوریتم‌های مختلف انجام گرفته است. دلیل این امر این است که داده‌های Precision از تمامی داده‌های دیگر برای مقایسه الگوریتم‌ها با اهمیت‌تر هستند زیرا میزان دقت الگوریتم کشف را می‌سنجند. با توجه به توضیحات بالا فرض‌های صفر و یک به صورت زیر می‌شوند:

فرض صفر ( $H_0$ ): Precision الگوریتم پیشنهادی از دیگر الگوریتم‌های مورد مقایسه بیشتر نیست.

فرض یک ( $H_1$ ): Precision الگوریتم پیشنهادی از دیگر الگوریتم‌ها بیشتر است.

با توجه به فرض‌های صفر و یک بالا آزمون‌های مان‌ویتنی برای هر یک از الگوریتم‌های مورد مقایسه در برابر الگوریتم ارائه شده انجام شد و نتایج آن در جدول زیر آمده است. (لازم به ذکر است که مقایسه‌ها در سطح  $\alpha=0.05$  صورت گرفته است و  $n_a=n_b=20$ ).

جدول 2: مقادیر  $U_A$  بدست آمده برای الگوریتم‌های مختلف

| نام الگوریتم | $U_A$ | $U_{critical}$ |
|--------------|-------|----------------|
| SAM+         | 19.5  | 127            |
| برمبنای IR   | 53.5  | 127            |
| تطبیق IO     | 14.5  | 127            |

همان‌طور که در جدول 2 مشاهده می‌شود، تمامی مقادیر  $U_A$  از  $U_{critical}$  کوچک‌تر است. بنابراین در تمامی موارد فرض صفر رد می‌شود و این بدین معنی است که Precision الگوریتم پیشنهادی از تمامی الگوریتم‌های مورد مقایسه بیشتر است.

## 6- نتیجه‌گیری

کارایی و میزان دقت الگوریتم ارائه شده به وسیله نسخه سوم مجموعه داده OWLS-TC ارزیابی و با سه الگوریتم کشف سرویس که یکی از روش‌های سنتی بازبازی اطلاعات برای کشف سرویس استفاده می‌کند و دیگری از تطبیق ورودی و خروجی و دیگری از تطبیق IOPE و ماژول WordNet برای تشخیص شباهت میان مفاهیم استفاده کرده است، مقایسه شده است. از نتایج آزمایش‌های انجام شده می‌توان موارد زیر را استنتاج نمود:

1. الگوریتم ارائه شده در این مقاله از نظر دقت از دیگر الگوریتم‌های مورد مقایسه بالاتر می‌باشد. نتایج نشان می‌دهد که در نظر گرفتن پیش‌شرط و اثر تا چه حد در میزان کارایی و دقت الگوریتم کشف تاثیرگذار است.
2. میانگین زمان پاسخ الگوریتم ارائه شده در این مقاله از دیگر الگوریتم‌های مورد مقایسه تا حدودی بالاتر است و علت آن در نظر گرفتن پیش‌شرط و اثر و تعیین مشابهت میان مفاهیم از طریق استدلال بر روی سلسله مراتب هستان شناسی است که این خود باعث بالا رفتن میانگین زمان پاسخ می‌شود.

### 1-6 کارهای آینده

کارهای آینده می‌توانند در زمینه‌های زیر انجام شوند:

1. بهبود یا گسترش مفهوم درجه تطبیق جایگزین
2. بهبود روش‌های تعیین مشابهت میان مفاهیم
3. ترکیب الگوریتم‌های تصمیم‌گیری با الگوریتم کشف وب سرویس معنایی

### 6-1-1 بهبود یا گسترش مفهوم درجه تطبیق جایگزین

درجه تطبیق جایگزین که در این مقاله ارائه شده است، به جای درجه تطبیق Fail در نظر گرفته شده است. این درجه تطبیق سرویس‌هایی را کشف می‌کند که گرچه نیاز کاربر را برآورده نمی‌نمایند، اما جایگزینی برای آن ارائه می‌دهند. با توجه به این که این درجه تطبیق به تازگی مطرح شده و در الگوریتم ارائه شده در این مقاله فقط با مقایسه ورودی و خروجی به سرویس منتسب شده لذا در آینده می‌توان روشی برای مقایسه پیش‌شرط و اثر ارائه نمود که به کشف سرویس جایگزین پردازد. این بدین معنی است که کشف سرویس جایگزین از طریق مقایسه IOPE صورت گیرد.

### 6-1-2 بهبود روش‌های تعیین مشابهت میان مفاهیم

روش تعیین مشابهت میان مفاهیم ارائه شده در این مقاله به صورتی است که زمان انجام آن بر روی زمان پاسخ الگوریتم تا حدودی تاثیر دارد. لذا بهبود روش حاضر با ارائه روش‌هایی که بتواند زمان پاسخ الگوریتم را به طور موثر کاهش دهد، از کارهایی است که می‌تواند در آینده انجام شود.

### 6-1-3 ترکیب الگوریتم تصمیم‌گیری با الگوریتم‌های کشف وب سرویس معنایی

الگوریتم‌هایی که تاکنون ارائه شده از کمبود یا نبود الگوریتم‌ها یا فرآیندهای تصمیم‌گیری رنج می‌برند. می‌توان با اضافه نمودن چنین الگوریتم‌هایی به فرآیند کشف وب سرویس معنایی سرویس‌های کشف شده را به درخواست کاربر نزدیک‌تر نمود و این به معنی افزایش دقت فرآیند کشف و در نتیجه رضایت بیشتر کاربر از سرویس‌های کشف شده می‌باشد.

## منابع

- [1] Friesen, A., Borger, E. (2006). A High-Level Specification for semantic web service Discovery services. in (ICWE'06), palo Alto, CA, 16-26.
- [2] Cardoso, J., Barros, A., May, N. and Kylau, U. (2010). Towards a Unified Service Description Language for the Internet of Services: Requirements and First Developments. in IEEE International Conference on Services Computing, Germany, 602-609.
- [3] Cardoso, J., Voigt, K. and Winkler, M. (2009). Service Engineering for the Internet of services. in Enterprise Information Systems. J.Filipe and J.Cordeiro, Eds, 1 ed Springer Berlin Heidelberg, 19, 15-27.
- [4] Schroth, C. and Janner, T. (2007). Web 2.0 and SOA: Converging Concepts Enabling the Internet of Services. IT Professional, 9, 36-41.
- [5] Clement, L., Hatley, A., von Riegen, C. and Rogers. (2004). T.(eds): UDDI Version 3.0.2, OASIS.
- [6] Pathak, J., Koul, N., Caragea, D., Honavar, V. G. (2005). A Framework for semantic web service Discovery. in proc, of the WIDM'05, Bermen, Germany, 45-50.
- [7] Aversano, L., Canfora, G., Ciampi, A. (2004). An algorithm for web service discovery through their composition. in proc. of the IEEE International Conference on web service (ICWS'04), 25-31.
- [8] Ahmadi, N., Binder, W. (2007). Flexible Matching and Ranking of Web service Advertisement. IN (MW4SOC'07), Newport Beach, CA, 30-35.
- [9] Sahai, A., Graupner, S. (2005). Overview of web services. in web services in the Enterprise Concepts, Standards, Solutions, and Management, 4<sup>th</sup> ed. New York: ch. 2, sec. 2, 17-25, Springer.
- [10] Kanneganti R., Chodavarapu, P. (2008). SOA Security. 1 ed: Manning.
- [11] Cecilia, P., Lee, L., Fleischer, R. (2007). Service Oriented Architecture (SOA)-Security Challenges and Mitigation Strategies. in Military Communications Conference IEEE, 7.
- [12] Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M. and Orchard, D. (2004). Web Service Architecture. W3C Working Group Note, February.
- [13] Daconta, M. C., Obrst, L. J. and Smith, K. T. (2003). The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management. Wiley Publishing, Indiana.

- [14] Studer, R., Grimm, S., Abecker, A. (2007). Architecture and Standardization of web services. in semantic web services concepts, Technologies, and Applications, 2<sup>nd</sup>ed . New York: Springer, ch.2,sec.1, 25-51.
- [15] Christensen, E., Curbera, F., Greg Meredith, and Weerawarana, S. (2001). Web Services Description Language (WSDL) 1.1. W3C.
- [16] Christensen E. and Curbera, F. (2007). Web Services Description Language (WSDL) Version 2.0. W3C.
- [17] Aguilera, U., Abaitua, J., Diaz, J. (2007). A Semantic Matching Algorithm for Discovery in UDDI. In International Conf. in semantic Computing, 751-758.
- [18] Jurie, M. B., Sasa, A., Brumen, B. and Rozman, I. (2009). WSDL and UDDI extensions for version support in web services. at Elsevier The Journal of Systems and Software, 82, 1326-1343.
- [19] Christensen, E., Curbera, F., Meredith, G. and Weerawarana, S. (2001). Web Services Description Language (WSDL) 1.1. W3C Note, 15 March.
- [20] Chinnici, R., Moreau, J. Ryman, A. and Weerawarana, S. (2007). Web Services Description Language (WSDL) Version 2.0. W3C Recommendation.
- [21] Kawamura, T., Hasegawa, T., Paolucci, M. and Sycara, K. (2005). web service Lockup: A Matchmaker Experiment. in IT professional, Vol.7, No.2, 36-41.
- [22] Cardoso, J., Sheth, A. (2003). Semantic E-Workflow Composition. in Intelligent information systems journal ,vol.21, No.3, 191-225.
- [23] Hebel, J., Fisher, M., Blace, R., Perez-Lopez, A. (2009). Preparing to program a semantic web of Data. in semantic web programming, 2<sup>nd</sup> ed. Indianapolis: Wiley, ch. 1, 1-34.
- [24] Studer, R., Grimm, S., Abecker, A. (2007). Architecture and standardization of web services “ in semantic web services concepts, Technologies, and Applications, 2<sup>nd</sup>ed New York: Springer, ch.3, sec.1, 51-105.
- [25] Studer, R., Grimm, S., Abecker, A. (2007). Ontology Development. in semantic web services concepts, Technologies, and Applications, 2<sup>nd</sup> ed. New York: Springer, ch.8, sec.1, 210-242.
- [26] Studer, R., Grimm, S., Abecker, A. (2007). Description. in semantic web services concepts, Technologies, and Applications, 2<sup>nd</sup> ed. New York: Springer, ch.7, sec.1, 161-179.
- [27] Ganjisaffar, Y., Ablhassani H. and Neshati, M. (2006). A Similarity Measure for OWL-S annotated web services. in proc. of the 2006 IEEE/WIC/ACM International Conference on web Intelligence, Hong kong, 102-109.
- [28] Studer, R., Grimm, S., Abecker, A. (2007). Description. in semantic web services concepts, Technologies, and Applications, 2<sup>nd</sup>ed New York: Springer, ch.6, sec.3, 161-179.
- [29] Talantikite, H. C., Aissani, D., Boudjlida, N. (2008). Semantic annotations for web service discovery and composition. in journal of computer standards interfaces, vol. 36, no.1, 472-482.
- [30] Li Kuang, Wu, J., Shuiguang Deng, Ying Li, Wei Shi, Zhaohui Wu, (2005). Exploring semantic Technologies in service Matchmaking. in proc. Of the 3th IEEE Europa Conference on web services (ECOWS 2005), Vaxji, Sweden, 62-67.
- [31] Studer, R., Grimm, S., Abecker, A. (2007). Ontology Development. in semantic web services concepts, Technologies, and Applications, 2<sup>nd</sup> ed. New York: Springer, ch.4, sec.1, 71-160.
- [32] Salhofer, P., Tretter, G. and Stadlhofer, B. (2008). Goal-Oriented Service Selection. in (ICEGOV'08), Cairo, Egypt, 60-66 .
- [33] Guo, R., Dong, D., Le, J. (2006). Discovery for web services Based on Relationship Model. in proc. of the 6<sup>th</sup> IEEE International Conference on Computer and Information Technology (CIT'06), 385-391.
- [34] Dong, X., Halevy, A. Y., Madhavan, J., Nemes, E., Zhang, J. (2004). similarity search for web services. in 'VLDB', 372-383.
- [35] Klusch, M., Kaufner, F. (2008). WSMO-MX: A Hybrid semantic web service Matchmaker. in web Intelligence and Agent systems: An Intelligence and Agent systems :An International journal, 5, 1-20.

- [36] Kuster, U., Konig-Ries, B., Stern, M., Klein, M. (2007). DIANE-An Integrated Approach to Automated Service Discovery, Matchmaking and composition. in proc. of WWW 2007, Banff, Alberta, Canada, 1033-1049.
- [37] OWLS Test collection, <http://projects.semwebcentral.org/projects/owlstc,September>, 2010.
- [38] <http://www.w3.org/submission/2004/SUBM-SWRL-20040521,september>, 2010.
- [39] Lonsdale, D., Embley, D. W., Piny, Y., Xu, L. and Hepp, M. () " Reusing ontologies and language components for ontology generation ", Data & knowledge Engineering , 2009.
- [40] Bianchini, D., De Antonellis, V., Pernici, pierluigiplebani, B. (2006). Ontology-base methodology for e-service discovery. Journal on information systems, special on semantic web and web services, Vol.31, NO.4-5, 361-380.
- [41] Klusch , M., Kaufer, F. (2009). S3 Contest, <http://www.ags.dfki.uni-sb.de/~klusch/s3>, December.
- [42] The OWL Services coalition, <http://www.darnl.org/services/owl-s/>
- [43] Guo, R., Le, J., Xia, X. (2005). Capability Matching of web services Based on OWL-S. in proceedings of the 16<sup>th</sup> International workshop on Database and Expert system Applications (DEXA), shanghai, 529-534.
- [44] Nayak, R., Lee, B. (2007). Web Service Discovery with additional Semantics and Clustering.in IEEE/WIC/ACM International Conference on Web Intelligence, 555-558.
- [45] Paolucci, M., Kawmura, T., Payne, T. and Sycara, K. (2002). semantic matching of web services capabilities. in proc of the int. semantic web conf, 103-111.
- [46] Chaiyakul, S., Limapichat, K., Dixit, A., Nantajeewarawat, E. (2006). A Framework for semantic web service Discovery and planning.in IEEE Conference on Cybernetics and Intelligent systems, 1-5
- [47] Sycara, K., Paolucci, M., Ankolekar, A. and Srinivasan, N. (2003). Automated Discovery , Interaction and composition of semantic web services. in Journal of web semantics, vol. 1, no.1, 27-46, December.
- [48] Bener, A. B., Ozadali, V., Ilhan, E. S. (2009). semantic matchmaker with precondition and effect matching using SWRL. in journal of Expert system Applications, vol.36, 9371-937, March.
- [49] Lu, H. (2005). semantic web services discovery and ranking. in WI'05: Proceedings of the 2005 IEEE/WIC/ACM International Conference on web Intelligence, Washington, DC, USA, 157-160.
- [50] Bin-hong, X., Yinh-jun, Z., Yong-yi, G. (2010). A Web service Matchmaker on Fuzzy Logic and OWL-S. in proc. of International conference on computational Aspects of social Networks, China, 591-599.
- [51] Gunay, A., Yolum, p. (2008). semantic Matchmaking of web service using Model Checking. inproc, of 7<sup>th</sup>Int I conf, on Autonomous Agents and Multiagent systems (AAMAS'08), Estroil, Portugal, 273-280.
- [52] Pan, S., Zhang, Y. (2009). Ranked web service matching for service description using OWL-S.in International conference on web Information systems and mining, 427-431.
- [53] Hao, Y., Zhang, Y., Cao, J. (2010). web services discovery and rank. An information retrieval approach.in journal of Future Generation Computer systems, 26, 1053-1062, june.
- [54] Bianchini, D., Antonellis, V. D. and Melchiori, M. (2006). Hybrid Ontology-Based Matchmaking for service Discovery.inproc, ACM symp. Applied Computing (SAC'06), Djon, France, 1707-1708.
- [55] Rong, W. and K. Liu. *A survey of context aware web service discovery: from user's perspective*. in *Service Oriented System Engineering (SOSE), 2010 Fifth IEEE International Symposium on*: IEEE.
- [56] Zhang, P. and W. Su. *Statistical inference on recall, precision and average precision under random selection*. in *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*. 2012: IEEE.
- [57] Dharanyadevi, P., et al. *Context aware SOA based discovery in mobile environment by bi-matchmaking technique*. in *Recent Trends in Information Technology (ICRTIT), 2011 International Conference on*. 2011: IEEE.
- [58] Maamar, Z., et al., *Using social networks for web services discovery*. Internet Computing, IEEE, 2011. 15(4): p. 48-54.

[59] Neepa Shah Assistant Professor and Sunita Mahajan PhD., "Document Clustering: A Detailed Review," *International Journal of Applied Information Systems (IJ AIS)* , vol. 4, p. 9, Octobr 2012.