Université Abou Bekr Belkaid
Tlemcen Algérie
جامعة أبي بكر بلقايد
تلمسان الجزائر

République Algérienne Démocratique et Populaire
Université Abou Bakr Belkaid– Tlemcen
Faculté des Sciences
Département d'Informatique

Mémoire de fin d'études

pour l'obtention du diplôme de Master en Informatique

Option : *Système d'Information et de Connaissances (S.I.C)*

Thème

Une approche de sélection des services Web composites à base de critères de vie Privée

Réalisé par :

- **ELOUCHDI Faysal**
- **KHITRI Mohammed El Mahdi**

Présenté le 6 juillet 2019 devant le jury composé de :

- *Mr. BENAZZOUZ Mortada* *(Président)*
- *Mr. BELABED Amine* *(Encadreur)*
- *Mr. HADJILA Fethallah* *(Examinateur)*

Année universitaire: 2018-2019

Remerciements

Nous tenons tout d'abord à exprimer notre gratitude à Monsieur A. BELABED enseignant à l'université Abou Bekr Belkaid pour nous avoir encadrés, pour son aide, son soutien et ses conseils. Sa disponibilité et ouverture d'esprit nous ont permis de travailler dans un cadre agréable et d'acquérir de nouvelles connaissances.

Nous souhaitons remercier les membres du jury d'avoir accepté d'évaluer ce modeste travail, Monsieur M. BENAZZOUZ et Monsieur F. HADJILA.

Nous sommes honorés de les avoir tous eus comme enseignants durant notre cursus.

Sommaire

Sommaire

Liste de figures

Liste de figures

Liste de Table

Introduction générale

❖ **Contexte :**

Nous vivons et partageons une grande partie de notre vie en ligne. Les entreprises et les gouvernements mettent en œuvre des services et des technologies dans de vastes réseaux qui accumulent nos données sans prendre dûment compte des risques de sécurité ou de vie privée. L'évolution constante des technologies du Web utilise de plus en plus nos données personnelles, à savoir, les identités biométriques ainsi que nos traces et dossiers numériques. Les services Web proposés, ainsi que le nombre des entreprises qui migrent vers cette technologie sont en forte croissance. Ce déploiement massif des services Web, augmente la préoccupation des clients concernant la confidentialité de leurs données privées. L'implication de ces services dans différentes compositions d'application menace la vie privée des utilisateurs à cause des transitions des données confidentielles entre services Web.

La sélection des services Web, qui consiste à choisir parmi un ensemble de services candidats, ceux qui répondent le mieux aux besoins des utilisateurs, est majoritairement basée sur des critères de qualité de service (QoS). Les critères de confidentialité et de vie privée sont souvent négligés au détriment des besoins non fonctionnelles [1] [2]. Cette négligence augmente le risque de violation de la vie privée des utilisateurs des services Web résultant de ce type de sélection.

❖ **Contribution :**

Dans le présent travail nous traitons la problématique de protection de la vie privée dans le contexte de sélection de services Web. Nous proposons comme solution à cette dernière, un Framework de sélection à base de critères de vie privée. Ce Framework se base sur un système multi-agents [3], son objectif est de trouver une composition qui respecte les contraintes de confidentialité des utilisateurs ainsi que les fournisseurs de services. Nos contributions se focalisent principalement sur l'amélioration du Framework proposé par [4]. À cette fin, nous avons proposé un modèle de négociation qui intervient dans le cas où il n'existe pas une composition qui préserve les contraintes de vie privée. Dans un tel cas, le Framework de sélection interagit avec les fournisseurs de services afin de relaxer leurs contraintes de vie privée. Notre deuxième contribution consiste à proposer une reformulation de l'encodage ASP (Answer Set Programminng) de l'algorithme de sélection proposé par [4]. Cette reformulation a induit un gain considérable dans le temps de réponse de l'algorithme de sélection.

❖ **Organisation du document :**

- Dans un premier chapitre, nous donnons un aperçu général sur la vie privée sur Internet avec quelques définitions de celle-ci et des différents dangers qui la menace, ainsi que quelques technologies de protection.
- Dans un second chapitre, nous définissons les services Web avec leurs différentes technologies permettant leurs implémentations. Nous exposons aussi les méthodes de sélection des services suivies d'un état de l'art sur la composition et la sélection à base de vie privée.
- Dans un troisième chapitre, nous présentons les détails de notre contribution. Et enfin, une conclusion générale où nous résumons l'essentiel de notre travail et donnons un aperçu sur les perspectives envisagées.

Chapitre 1 : La vie privée sur Internet

1. Introduction

La vie privée comme toute propriété de la personne doit être évaluée dans le nouveau monde numérique à cause de l'apparition des médias numériques, les réseaux sociaux et la vulgarisation de l'utilisation des outils informatiques dans la société.

Dans ce chapitre, nous allons revenir sur la définition et les principes de la vie privée, puis nous exposerons les risques sur celle-ci dans le Web, ainsi que les moyens mis en œuvre pour essayer de la protéger.

2. Définitions

La vie privée est définie juridiquement comme un droit civil s'appliquant à la vie sentimentale, familiale, à la santé, au droit à l'image et au secret de la résidence. [5]

3. Principes de la protection de la vie privée

La protection de la vie privée est régie par des règles qui sont résumées selon [6] comme suit : «

- **Autorité de collecter :** limiter la collecte des renseignements personnels par les institutions à la réalisation des activités autorisées.
- **Mode de collecte :** assurer la collecte directe des renseignements personnels auprès d'une personne, sauf dans quelques rares cas.
- **Exigences en matière d'avis :** informer une personne de la collecte de ses renseignements personnels.
- **Utilisation et divulgation nécessaires :** limiter l'utilisation et le partage du contrôle ou la diffusion des renseignements personnels à la réalisation des activités autorisées.
- **Exactitude :** assurer la mise en place de processus grâce auxquels les renseignements personnels sont toujours exacts.
- **Conservation** : assurer l'accès d'une personne à ses propres renseignements personnels pendant une certaine période.
- **Sécurité :** assurer la sécurité et la confidentialité des renseignements personnels.
- **Transfert à des archives et destruction :** assurer l'autorisation et la sécurité du transfert aux archives et de la destruction des renseignements personnels. »

4. Les Menaces sur la Vie privée

L'existence des données sensibles à la vie privée sur des machines distantes non contrôlées directement par les entités propriétaires des données (personne physique/morale, application, etc.) présente plusieurs risques de confidentialité parmi eux :

4.1. Divulgation des données personnelles : l'Atteinte à la réputation et à l'intimité

La propagation des nouvelles dans le Web est une affaire de seconde. Une atteinte à la réputation d'une personne dans cette circonstance peut être dramatique et irréparable. Ce type de violation de vie privée est défini par une diffusion de déclaration diffamatoire visant la réputation d'une personne ou pour le rejet de ce dernier [7]. Les données relevant de l'intimité des personnes transitant sur les réseaux ou stockées dans les différents appareils doivent être exploitées de façon à ne pas violer l'intimité des propriétaires, comme le souligne la déclaration des Nations Unis de 1948 dans l'article 12 : « Nul ne peut être soumis à une ingérence arbitraire dans sa vie privée, sa famille, son domicile ou sa correspondance, ni aux attaques contre son honneur et sa réputation. Tout le monde a droit à la protection de la loi contre de telles interférences ou attaques » [8].

L'utilisation et la divulgation des renseignements personnels sont régulées par la loi. Cette exploitation de donnée n'est possible que dans les cas suivants. Le premier, c'est la compatibilité du but de l'utilisation avec celui indiqué lors de la collecte de la donnée ou bien par consentement. Le second cas se présente dans la conformité à d'autres lois ou pour l'exécution de la loi, comme le code de la route. Le troisième se présente dans l'utilisation ou la divulgation pour des fins d'exercice de fonction d'une institution, exemple l'utilisation de données pour la rédaction d'ordre de mission. Quatrièmement, les organisations peuvent échanger des données privées, comme exemple lors de transfert d'employé. Un cas spécial permet dans un état d'urgence où la santé ou la sécurité d'une personne sont en jeu des institutions peuvent divulguer des informations dans ce cadre avec un avis de divulgation à la personne concernée. [6].

4.2. Vol et usurpation d'identité

Les failles de sécurité dans la protection des données peuvent être exploitées par des personnes mal intentionnées. L'exploitation de ces brèches induit des alternances des données (suppression, modification ou ajout). Elle peut aussi avoir comme but un vol de données. Nous citons comme exemple la violation en 2016 des données du système électoral du Philippine, contenant les données personnel, les détails du passeport et les empreintes digitales de 70 millions électeurs inscrits [9].

Les vols de donnée combinés à de l'ingénierie sociale engendrent une usurpation d'identité, c'est l'utilisation d'information privée propre à la personne, comme le nom d'une personne, le numéro d'un compte bancaire, d'adresse, la date de naissance et le numéro de sécurité sociale à l'insu de son propriétaire ainsi que tout autre identifiant électronique. Nous prenons comme exemple le recensement par Perceval[1] de 17 831 signalements d'escroquerie aux cartes bancaires en France, d'un préjudice total de plus de 5,5 millions d'euros [10].

4.3. Le Profilage

Le profilage est une technique pour traiter automatiquement les données personnelles et non-personnelles dans le but de développer des connaissances prédictives à partir des données dans la forme de profil en construction pouvant être utilisées par la suite comme une base de prise de décision.

Un profil est un ensemble de données corrélées qui représente un sujet humain ou non, individuel ou en groupe. Les profils constructifs sont le processus de découverte de patterns (modèles) inattendus entre des données dans de grands ensembles de données qui peuvent être utilisés pour la création de profil.

L'application de profilage est le processus d'identification et de représentation d'un sujet spécifique ou d'identifier un sujet comme étant un membre d'un groupe spécifique ou d'une catégorie et à prendre une forme de décision en se basant sur cette identification et représentation [11].

[1] Plateforme française de lutte contre les fraudes à la carte bancaire sur internet.

5. Les techniques d'attaques

L'exploitation des brèches par les pirates est réalisée par des techniques différentes. Chaque méthode d'attaque utilise un concept et un modèle applicable bien donné. Nous citions parmi ces méthodes :

5.1. Les Malwares

L'exploitation mal intentionnée de l'outil informatique est réalisée principalement avec les Malwares. Dont la définition est la suivante : « Les Malwares (contraction des mots « malicious » et « software » en anglais) sont des logiciels malveillants, dont le but est d'accéder à l'appareil d'un utilisateur à son insu. Ces types de logiciels incluent les logiciels espions, les logiciels publicitaires, les virus, les chevaux de Troie, les vers informatiques, les rootkits, les logiciels de rançon et les détourneurs de navigateur. » [12].

5.2. Les Cookies

C'est un fichier enregistré dans la partie client a la demande du serveur. Ce fichier contient des informations du client aidant à retenir ces préférences et détails de connexion. Un cookie utilisé pour permettre à un utilisateur de s'authentifier à un serveur distant peut être volé, cette action ce nom Hijacking. Par exemple : les cookies HTTP utilisés pour maintenir une session sur plusieurs sites web peuvent être volés à l'aide d'un ordinateur intermédiaire ou en accédant aux cookies enregistrés sur l'ordinateur de la victime. Si un pirate peut voler les cookies de l'authentification, il peut lancer une requête comme s'il est l'utilisateur réel, il peut accéder ou modifier des informations confidentielles. Si ces cookies sont persistants, l'usurpation d'identité de l'utilisateur victime peut continuer pour une durée de temps considérable. N'importe quel protocole dans lequel l'état est maintenu à l'aide d'une clé passée entre deux parties est vulnérable, en particulier si elle n'est pas chiffrée [13].

5.3. L'hameçonnage (Phishing)

L'hameçonnage repose sur l'exploitation du manque de vigilance et d'information des personnes visée. Nous citons [14] pour une définition formelle : « L'hameçonnage (appelés également « phishing ») est une approche détournée qu'utilisent les cyber-escrocs pour vous pousser à révéler des informations personnelles, comme des mots de passe ou des numéros de carte de crédit, de sécurité sociale ou de compte bancaire. Ils le font en vous envoyant des e-mails contrefaits ou en vous dirigeant sur un site web contrefait. ».

5.4. Attaques par Inférence

Une attaque d'inférence est une technique d'exploration de données utilisée pour accéder illégalement à des informations sur un sujet ou une base de données. Une telle attaque se produit lorsqu'un utilisateur est en mesure de déduire des informations clé ou critiques d'une base de données à partir d'une analyse d'informations triviales, sans y accéder directement [15].

L'attaque par inférence peut être utilisée dans le profilage et l'analyse de vie privée de grands ensembles de données géolocalisées (attaque basée sur le paradigme de MapReduce[2]) [16, 17].

6. Les Technologies de Protection de la Vie privée

Le danger qui plane sur les données de vie privée sur internet, force à mettre en place des technologies de protection. La diversité des techniques d'attaque diversifie les méthodes de protection. Nous citons parmi ces dernières :

6.1. Les systèmes de gestion d'identité

Les concepteurs des nouveaux systèmes offrant davantage de services Web sont dans l'obligation de garantir l'authentification et la vie privée des utilisateurs. L'authentification permet de reconnaître formellement le demandeur du service par son prestataire, alors que la vie privée veille à la non-divulgation des informations pour éviter le profilage et le suivi des utilisateurs. Pour la résolution de ce conflit, des systèmes de gestion d'identité sont mise en place. Deux principaux systèmes sont adoptés, U-Prove[3] proposé par Microsoft ainsi

[2] https://fr.talend.com/resources/what-is-mapreduce/
[3] https://www.microsoft.com/en-us/research/wp-ontent/uploads/2016/02/Privacy20and20accountability20-20best20of20both20worlds.pdf

qu'Idemix [4] proposé par IBM [18] afin de garantir l'authentification et la confidentialité.

6.2. Accréditations anonymes

L'accréditation est émise par une autorité. Elle prouve une qualification ou autorisation à une personne. Elle contient des données à caractère personnelle telles que le nom, l'adresse, la date de naissance ou encore le profil biométrique d'un individu [17]. Nous pouvons citer comme exemple d'accréditation le permis de conduire d'une personne.

L'accréditation anonyme est une version anonyme des accréditations qui permet de prouver une propriété ou un droit lié à son possesseur, mais sans révéler l'identité de celui-ci [17]. L'accréditation anonyme permet essentiellement à l'utilisateur de transformer le certificat en une nouvelle, contenant seulement un sous-ensemble des attributs du certificat original (cela permet de prouver seulement un sous-ensemble des attributs au vérificateur - propriété de divulgation sélective.). Au lieu de révéler la valeur exacte de l'attribut, les systèmes d'accréditation anonyme permettent à l'utilisateur de la transformation d'appliquer n'importe quelle fonction mathématique à la valeur de l'attribut originale, lui permettant de prouver seulement les attributs de propriétés sans révéler l'attribut lui-même.

Comme exemple nous citons Idmix, il compte sur des procédés de certification avec une suite de protocoles cryptographiques pour garantir l'anonymat et l'inapplicabilité. Supposons que Alice utilise Idemix pour générer une accréditation anonyme pour Bob, qui révèle uniquement qu'elle possède un permis de conduire valide et rien d'autre. La preuve d'accréditation serra :

- Bob n'apprend aucune information supplémentaire sur Alice à part le fait qu'elle possède une licence valide (anonymat).
- Si Alice se rend plusieurs fois au magasin et génère une épreuve à chaque fois pour Bob, celui-ci ne pourrait pas en déduire qu'il s'agissait de la même personne (inapplicabilité). [19]

[4] https://www.zurich.ibm.com/pdf/csc/Identity_Mixer_Nov_2015.pdf

6.3. Réseaux de communication anonyme

La préservation de la confidentialité sur les réseaux de communication est primordiale lors de la recherche de l'anonymat dans les transactions par réseaux. Le premier protocole appelé Mix network a été décrit en 1981 par David Chaum [20]. Les réseaux de communication impliquant ce genre de protocole de routage créent des communications difficiles à tracer en utilisant une chaîne de serveurs proxy connus sous le nom de mixes. Ces serveurs prennent des messages provenant de plusieurs expéditeurs, les mélangent et les envoient de nouveau dans un ordre aléatoire à la prochaine destination (éventuellement un autre nœud de mixage). Cela brise le lien entre la source de la demande et la destination, brouillant toute possibilité de traçabilité.

La non-divulgation de la source originaire et la destination final étant inconnu au proxy (nœuds de mixage) les rendent moins susceptibles d'être malveillant. Parmi les réseaux de communication anonyme, il existe : DC Network [21], Mix nets [20], Tor (Onion Routing) [22].

6.4. Protection à base de langages de spécification des exigences de vie privée

Afin d'appuyer les technologies de protection de la vie privée des langages sont mis en place pour exprimer les exigences de vie privée des utilisateurs et les spécifications des fournisseurs. Un modèle de correspondance sert à vérifier la conformité entre exigences et spécifications exprimées. Le principe de fonctionnement de ce modèle est illustré dans la Figure 1-1.

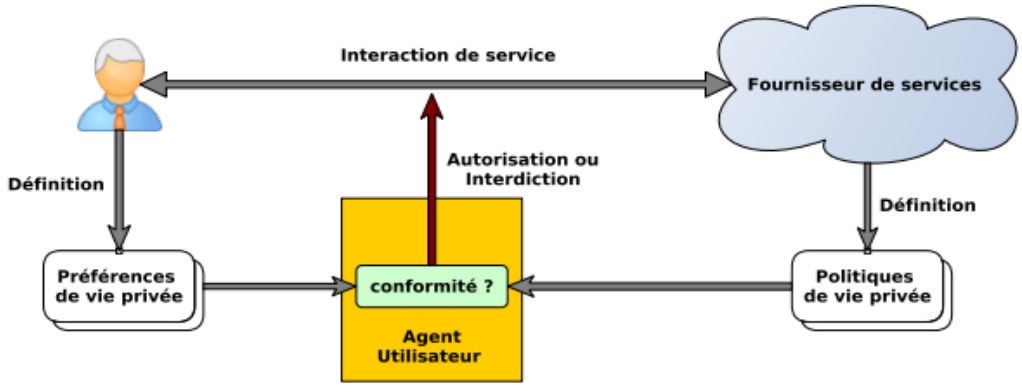

FIGURE 1-1 MODELE DE BASE POUR LA MISE EN CORRESPONDANCE AUTOMATIQUE DES PREFERENCES ET LA POLITIQUE DE VIE PRIVEE D'UN UTILISATEUR FINAL ET UN FOURNISSEUR DE SERVICES [23]

Un agents utilisateur est implémenté pour vérifier la conformité des politiques avec la préférence. Dans le cas de compatibilité, l'utilisateur peut utiliser le service avec une garantie de protection [23].

Des langages standardisés sont proposés, parmi eux on cite :

- **Platform for Privacy Preferences (P3P) :** Projet a l'initiative du consortium W3C, qui le définit comme suit : Le projet P3P (Platform for Privacy Preferences) permet aux sites Web d'exprimer leurs pratiques en matière de confidentialité dans un format standard. Ce format peut être récupéré automatiquement et interprété facilement par les agents utilisateurs. Les agents utilisateurs P3P permettront aux utilisateurs d'être informés des pratiques du site (dans des formats lisibles par machine et par l'homme) et d'automatiser la prise de décision en fonction de ces pratiques. Ainsi, les utilisateurs n'ont pas besoin de lire les politiques de confidentialité de chaque site visité [24].
- **A P3P Preference Exchange Language (APPEL) :** APPEL complète la spécification P3P en spécifiant un langage permettant de décrire les ensembles de préférences en matière de règles P3P entre agents P3P. Ce langage permet aux utilisateurs d'exprimer un ensemble de règles de préférence (appelé RULESET), qui peuvent ensuite être utilisées par leurs agents pour prendre des décisions automatisées ou semi-automatisées concernant l'acceptabilité de politiques de vie privé lisibles par machine à partir du P3P du Site Web consulté [25].
- **eXtensible Access Control Language (XACML) :** OASIS[5] propose le standard XACML basé sur le langage XML pour exprimer et échanger les politiques de contrôle d'accès. Il n'est pas spécifiquement conçu pour la gestion de la vie privée, mais il représente une innovation pertinente dans le domaine des politiques de contrôle

[5] https://www.oasis-open.org/org

d’accès et a été utilisé comme base pour suivre les langages d’autorisation relatif à la vie privée [26].

- **Enterprise Privacy Authorization Language (EPAL) :** le langage d'autorisation de confidentialité d'entreprise (EPAL) [27] est un langage basé sur XML permettant de spécifier et d'appliquer des règles de confidentialité d'entreprise. EPAL est spécialement conçu pour permettre aux organisations de traduire leurs politiques de confidentialité en déclarations de contrôle informatique et d'appliquer des politiques qui peuvent être déclarées et communiquées conformément aux spécifications P3P [26].

7. Conclusion :

Nous avons exposé dans ce chapitre une vue générale sur la vie privée sur Internet. Nous avons ainsi présenté la définition de la vie privée, les menaces sur celle-ci, ainsi que les différentes méthodes d’attaque. Comme tout problème qui demande des solutions, des technologies de protection de la vie privée, ont été présentées.

Chapitre 2 : Web Services

1. Introduction

Dans ce chapitre nous allons définir les services Web et relater leurs différents types et propriété. Nous exposons aussi le problème de sélection des services Web ainsi que les travaux qui ont été faite dans ce domaine.

2. Définitions & caractéristiques des services Web

Le consortium W3C définit un service Web comme un système logiciel conçu pour prendre en charge une interaction interopérable de machine à machine sur un réseau. Il possède une interface décrite dans un format pouvant être traité par une machine (en particulier WSDL). D'autres systèmes interagissent avec le service Web de la manière spécifiée par sa description à l'aide de messages SOAP, généralement acheminés via HTTP avec une sérialisation XML conjointement avec d'autres normes relatives au Web [28] .

L'intégration du service web dans un environnement distribué est réalisé grâce à Universal Discovery Description and Integration (UDDI) et l'exploitation de ses interfaces le protocole est fourni avec un descriptif en Web Service Description Language (WSDL), généralement XML, précisant les méthodes qui peuvent être réalisé par le service avec leurs signatures réciproques ainsi que les points d'accès (Url, port) du service web [28]. La Figure 2-1 schématise la structure d'un service Web.

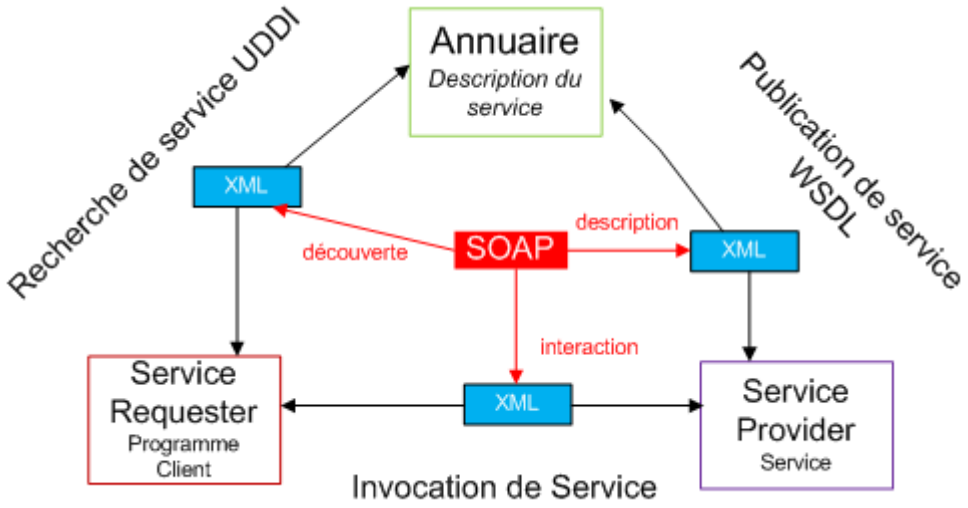

FIGURE 2-1 STRUCTURE D'UN SERVICE WEB [29]

3. Les types des services Web

3.1. Services Web REST

Ce type de service utilise l'architecture REST dans la communication entre services Web qui sont légers, maintenables et évolutifs. Les services Web utilisant l'architecture REST sont nommé comme service RESTful. Le protocole HTTP vient accompagner le protocole REST dans la transmission. RESTful est implémenté comme suit sur six clés principales selon [30] :

- Les ressources : elles représentent la ressource elle-même.
- Les verbes de requêtes : décrivent l'action voulu sur les ressources, parmi eux on trouve GET, POST, PUT et DELET.
- L'en-têtes de demande : ce sont des instructions supplémentaires envoyées avec la demande. Ceux-ci peuvent définir le type de réponse requise ou les détails d'autorisation.
- Le corps de la demande : les données sont envoyées normalement lors qu'une demande POST est faite au service Web.
- Le corps de la réponse : il s'agit du corps principal de la réponse, renvoyé par un document XML.
- Les codes d'état de réponse : ces codes sont les codes généraux qui sont retournés avec la réponse du serveur Web.

3.2. Services Web SOAP

Ce type de service utilise le protocole SOAP (voir section 4.2), un protocole souple et standardisé pour les échanges de messages en ajoutant le protocole HTTP (voir section 4.1). L'objectif d'utiliser ce protocole est de définir la structure générale des messages échangés entre les composants, sans pour autant en définir la structure du contenu. En outre, il laisse aux services de définir le format du contenu du message. Les messages SOAP échangés n'ont pas de relation entre eux à la base (des messages « one-way »), mais en pratique dans les services web on leurs affecte un mécanisme de requête/réponse. [31](voir la Figure 2-1).

4. Architecture & protocoles

4.1. HTTP

Le protocole HTTP (Hypertext Transfer Protocol) est un protocole de couche d'application pour la transmission de documents hypermédia, tels que HTML. Il a été conçu pour la communication entre les navigateurs Web et les serveurs Web, mais il peut également être utilisé à d'autres fins. HTTP suit un modèle client-serveur classique, avec un client ouvrant une connexion pour faire une demande, puis en attente jusqu'à ce qu'il reçoive une réponse. HTTP est un protocole apatride, ce qui signifie que le serveur ne conserve aucune donnée (État) entre deux demandes. Bien que souvent basé sur une couche TCP/IP, il peut être utilisé sur n'importe quelle couche de transport fiable [32].

4.2. SOAP

SOAP (Simple Object Access Protocol) est un protocole, sa place dans la pile de technologie des services Web est l'emballage normalisé pour les messages partagés par les applications. La spécification ne définit rien de plus qu'une simple enveloppe basée sur XML pour les informations transférées, et un ensemble de règles pour traduire des types de données spécifiques à une application et à une plateforme dans des représentations XML. La conception de SOAP le rend convenable à une grande variété de modèles de messagerie et d'intégration d'applications. En d'autres termes, SOAP est une application de la spécification XML. Il s'appuie fortement sur des normes XML telles que XML Schema et XML Namespaces pour sa définition et son fonctionnement. [33]

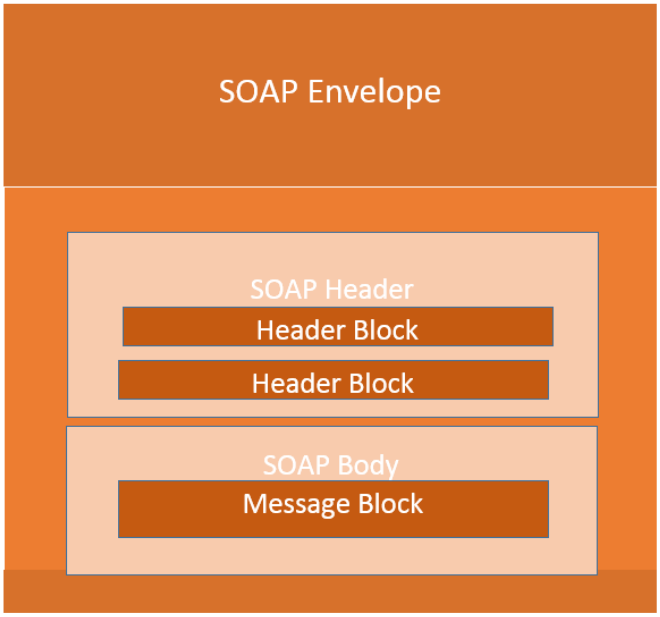

FIGURE 2-2 STRUCTURE DU MESSAGE SOAP [34]

4.3. WSDL (Web Services Description Language)

WSDL signifie Web Services Description Language, qui est un format du langage XML. L'utilité du WSDL est de décrire les services réseau sous la forme d'un ensemble de nœuds de terminaison fonctionnant sur des messages contenant des informations orientées document ou orientées procédure, ainsi que les services Web. Les opérations et les messages sont décrits de manière abstraite, puis liés à un protocole de réseau et à un format de message concrets pour définir un nœud de terminaison. Les nœuds de terminaison concrets associés sont combinés en nœuds de terminaison abstraits (services). WSDL est une recommandation du W3C depuis le 26 juin 2007 [35] [36].

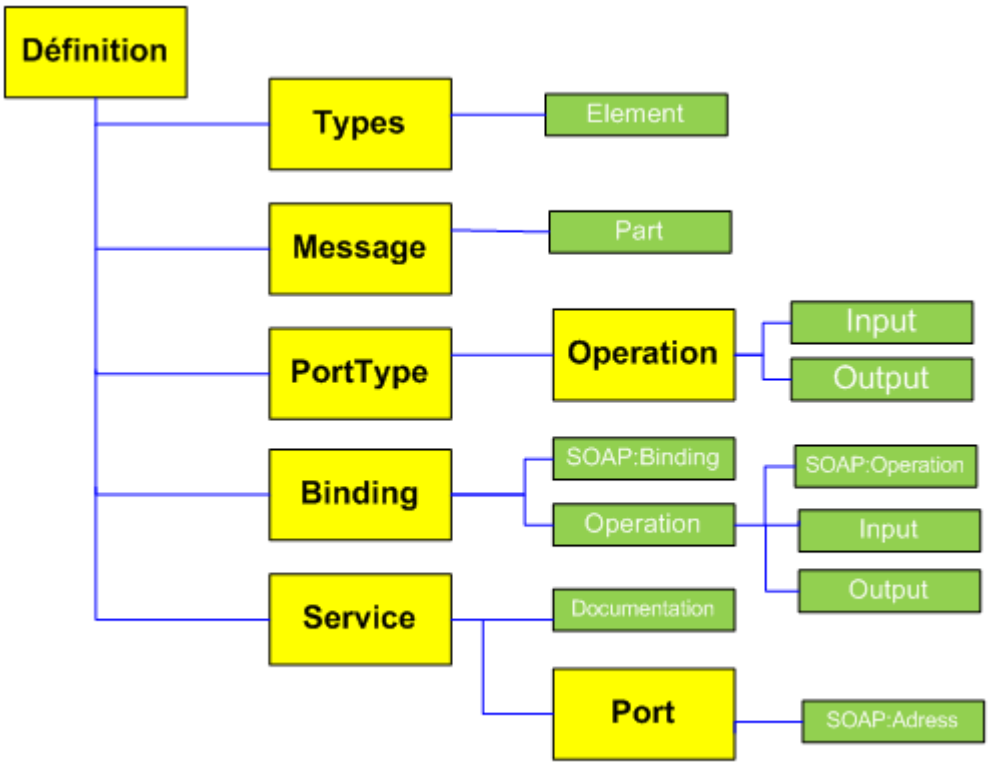

FIGURE 2-3 STRUCTURE D'UN DOCUMENT WSDL [29]

4.4. UDDI (Universal Description Discovery and Integration)

Comme décrit part [33], Le projet UDDI est un l'effort de l'industrie pour définir un registre de services interrogeables et leurs descriptions où les consommateurs peuvent automatiquement découvrir les services dont ils ont besoin.

UDDI se compose de deux parties : un registre de toutes les métadonnées d'un service Web (y compris un pointeur vers la description WSDL d'un service), et un ensemble de définitions de type de port WSDL pour manipuler et rechercher ce registre. La dernière spécification UDDI est la version 2,0.

Grâce à un document de bonne pratique lors de l'utilisation de UDDI et WSDL ensemble, on peut générer un enregistrement UDDI d'un service à partir de sa description WSDL. Le document UDDI est obtenu en divisant la description WSDL en deux parties (deux fichiers WSDL distincts). Le premier fichier devient la description de l'interface. Il inclut les types de données, les messages, les types de ports et liaisons. Le deuxième fichier est connu comme la description de l'implémentation. Il ne comprend que les définitions de service.

UDDI n'est pas la seule option pour la découverte de service. IBM et Microsoft ont récemment annoncé le Web services inspection Language (WS-inspection), un langage basé sur XML qui fournit un index de tous les services Web à un emplacement Web donné [37].

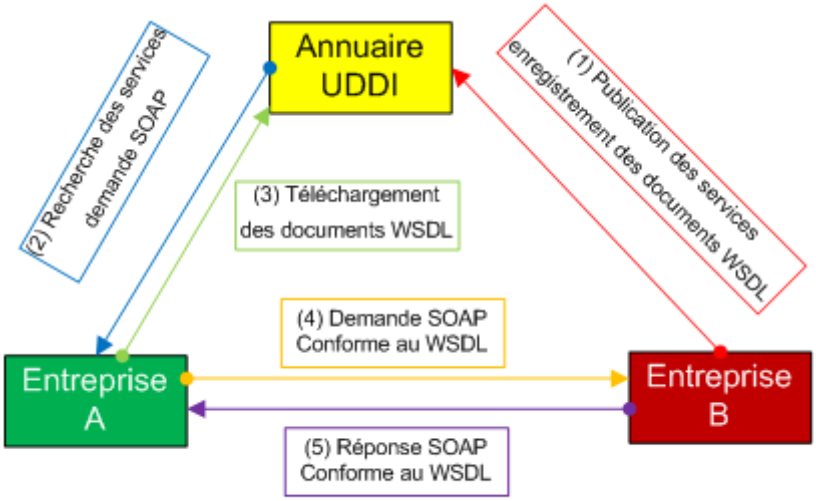

FIGURE 2-4 SCHEMA GENERALE DE L'ANNUAIRE UDDI [29]

5. Composition des services Web

La composition de service Web est une mise en relation collaborative de plusieurs services nécessaire dans la formation d'un service composite demandé par l'utilisateur. Les technologies de base, décrite précédemment (section 4), utilisé dans la conception des services Web ne permettent pas d'avoir une collaboration entre services. Un processus métier avec les détails comportementaux sur le rôle de chaque service pour une collaboration doit être mis en place. Ce dernier peut être construit de deux manière : l'orchestration de services ou la chorégraphie de services [38].

5.1. Modèle d'orchestration

Le modèle d'orchestration utilise une approche centralisée dans la composition des services où un service de composition central est chargé de gérer la collaboration entre les services et des échanges entre eux [38]. Les interactions dans ce modèle suivent le graphe présenté dans la Figure 2-5

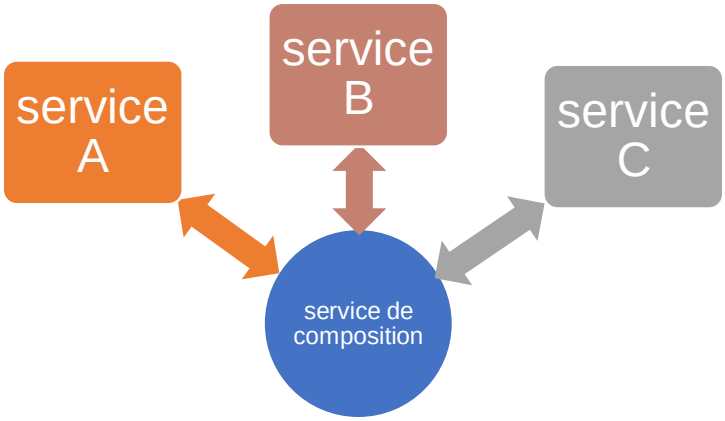

FIGURE 2-5 GRAPHE D'ECHANGE DE MESSAGES DANS LE MODELE D'ORCHESTRATION

5.2. Modèle de chorégraphie

Le modèle de chorégraphie utilise une approche décentralisée pour la composition des services. La chorégraphie de service se fait par des règles d'interaction et accord entre deux ou plusieurs points d'extrémité des services [38]. Les interactions dans ce modèle suivent le graphe présenté dans la Figure 2-6

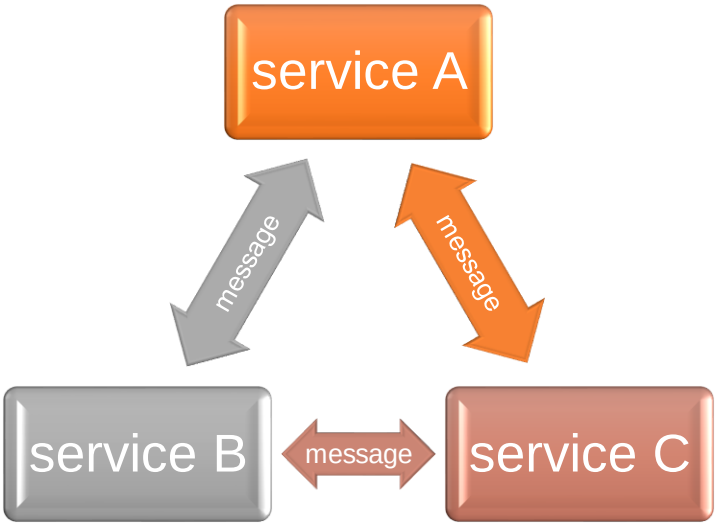

FIGURE 2-6 GRAPHE D'ECHANGE DE MESSAGE DANS LE MODELE CHOREGRAPHIE

6. Sélection des services Web

La migration que suivent les entreprises vers les services Web et le développement de ce mode de service a induit à une croissance très prononcée dans le nombre de services proposé. Dans le dilemme du choix qui se pose aux utilisateurs des moyens de comparaison et de sélection doivent les guider. La sélection des services doit répondre aux besoins des utilisateurs fonctionnels (à base de qualité) ou non-fonctionnels soit elle (à base de critère de vie privée). Nous citons donc ces deux bases de sélection :

6.1. Sélection à base de qualité de service QoS-Aware

La sélection à base de qualité de service dans les services Web œuvre pour satisfaire les différents critères non-fonctionnels des utilisateurs (exemple : temps, prix, marque, origine…). Les indicateurs de qualité de service sont affectés aux services Web composants pour optimiser la qualité de service du service composite. La valeur de ces indicateurs est dynamique. Elles varient pour des causes qui peuvent être internes ou externes, les charges du système d'environnement où se trouve le service changent, il y'a ceux qui modifient leurs QoS et ceux qui disparaissent, et aussi des nouveaux services qui apparaissent. Etant donné les préférences et exigences variées des utilisateurs, ces indicateurs sont à la base de la sélection des services composant du service web composite proposé à l'utilisateur.

6.2. Sélection à base de vie privée

La sélection de service Web à base de vie privée touche aussi le côté des besoins non fonctionnels des utilisateurs. Les exploitations de données de vie privée sans avoir de permission des utilisateurs des services doit être résolu pour fournir un niveau acceptable de confidentialité. La sélection à base de vie privée consiste à prendre en charge les exigences et politiques de vie privée lors de la composition des services. Chaque utilisateur fournit les exigences que le service composite doit respecter. Les services formant le service composite exposent eux aussi des politiques et exigences qui doivent être respecter mutuellement dans la chaîne de composition. Trouver le service Web composite optimal vis-à-vis aux exigences de la vie privée se voit reflété un problème NP-difficile [4]. Si la sélection d'un seul service semble évidente la composition de service l'est

beaucoup moins dû à la variance des exigences et politiques des services Web, et à leurs niveaux de tolérance pour l'adaptation aux critères de vie privée de l'utilisateur.

7. État de l'art sur la composition et Sélection à base de vie privée

Le nombre des services Web de différents types est en croissance permanente. La facilité d'intégration et la mobilité des services Web les mettent dans les technologies les plus utilisées dans le Web, que ça soit pour des communications entre application d'un réseau local ou sur Internet. Les services Web ont comme utilisateurs des entités variées qui peuvent demander certains besoins spécifiques. Les besoins peuvent avoir un caractère fonctionnel ou non fonctionnel.

Pour la satisfaction des utilisateurs, des modes de sélection des services Web supervisent la composition du service composite. Nous distinguons deux types de choix (voir section 6). Le premier, étant la sélection à base de qualité de service QoS, qui se focalise sur les besoins non fonctionnels de l'utilisateur. La négligence des paramètres de vie privée, induit des brèches d'exploitation de données privées. Pour remédier à ce problème, un deuxième type de sélection à base de vie privée est formulé.

Un modèle de sélection capable de préserver les exigences de vie privée des utilisateurs est formulé moyennant trois principaux axes. La représentation des politiques de confidentialité qui se partagent entre les utilisateurs d'une part et les fournisseurs d'autre part constitue le premier axe. Une gestion efficace de ces politiques de spécification pour trouver la combinaison de services Web adéquate est l'objet du deuxième axe. Enfin, une proposition d'un service web composite avec un niveau acceptable de protection de la vie privé est exprimée.

Un bon nombre de travaux ont été consacrées à la protection de la vie privée des services Web. Dans la suite de cette section nous citons brièvement les approches les plus proche à notre travail.

Dans le travail de [39], les auteurs utilisent un modèle de chorégraphie (voir section 5.2) avec un mécanisme à base de cryptographie pour la composition de service. Dans ce modèle, les besoins de vie privée sont communiqués à travers un

protocole de chiffrement dans un composant appelé « ElitePicker application ». Cette approche reste limitée à un modèle de chorégraphie.

Dans une autre approche proposée par [40], et afin d'améliorer la confidentialité dans des services Web DaaS (données en tant que service) les auteurs proposant un modèle de confidentialité dynamique pour les services Web. C'est un modèle formel permettant aux utilisateurs et aux fournisseurs de services de définir un ensemble de règles de confidentialité afin d'exprimer leurs politiques et leurs exigences en matière de confidentialité. Cette approche décrit un algorithme appelé PCM (Privacy Compatibility Matching) pour vérifier la compatibilité de la confidentialité entre les stratégies et les exigences d'une composition DaaS avec un mécanisme de négociation qui établit une réconciliation dynamique entre les services en cas d'incompatibilité. L'algorithme proposé ne contenant pas d'heuristique peut causer des problèmes dans la mise à l'échelle.

Dans le travail [41], les auteurs exposent un Framework de sélection de services Web dont l'objectif est de sécuriser les besoins des utilisateurs et du fournisseur de services liés à la confidentialité. Cette approche protégée les règles de fourniture de services contre les divulgations non désirées grâce à des méthodes cryptographiques. Le composant principal de la sélection de services proposé s'appelle le négociateur privé. Ce dernier, estime avec une analyse de sécurité formelle la quantité d'informations éventuellement déduite par les fournisseurs de services et les utilisateurs au cours des protocoles d'échanges. Cette sélection ne se base donc pas sur la correspondance de règles de confidentialité.

Dans [42], les auteurs présentent une approche utilisant la sensibilité des attributs de données conjointement avec l'objectif, la visibilité et la durée de rétention pour définir les besoins de vie privée (besoin non-fonctionnel) des utilisateurs et des fournisseurs de services Web, mais toute en prenant en compte des paramètres de qualité de service dans l'algorithme de composition et ainsi sélectionner le service Web composite qui a le plus haut niveau de confidentialité. Étant donné que l'implémentation effectuée dans ce travail utilise seulement un service Web d'agence de voyage spécifique ne fournit pas une véritable preuve d'efficacité de l'approche proposée.

Dans [43], l'approche présentée repose sur un mécanisme de négociation de la confidentialité qui repose sur la théorie de Galois. L'objectif et les résultats de ce mécanisme sont de parvenir à un accord entre l'utilisateur et le fournisseur de service, qui répond aux préférences de chacun en ce qui concerne le volume optimal d'informations de vie privée transmise lors de l'utilisation du service. La négociation est possible en variant un paramètre θ dans la plage de 1 à 5 pour changer le réseau correspondant, et ainsi déterminer quels éléments de données doivent ou non être requis par les utilisateurs. Ce faisant, il serait possible de dériver une politique de confidentialité optimale. Les auteurs ont aussi présenté un système prototype qui suit l'algorithme de construction de réseau conceptuel afin de générer des règles de négociation et à montrer comment réduire l'écart entre les informations requises par les fournisseurs de services après la négociation d'une politique de confidentialité. Cette approche met un grand désagrément à l'utilisateur qui se voit négocier les termes avec les fournisseurs à chaque demande.

Dans cette approche [4], les auteurs proposent l'intégration d'un Framework de sélection de services web à base de critères de vie privée. Le Framework utilise un modèle d'orchestration pour la composition de service. Son objectif est de trouver une composition de services qui répond aux contraintes de confidentialité des utilisateurs ainsi que celles des fournisseurs de services. Le Framework a été modélisé et testé sur une base de données générée aléatoirement. [4] Propose plusieurs types d'algorithmes pour résoudre le problème de sélection, en commençant par un algorithme Best First Search, qui consiste à trouver le chemin le plus court entre le nœud « requête » et le nœud « puits », le chemin doit respecter la contrainte de ne contenir que des services valides. Puis, il introduit deux approches déclaratives, l'approche MaxSAT (partial weighted Max-SAT) et l'approche ASP (Answer Set Programming) [44].

8. Conclusion

Nous avons défini dans ce chapitre les services Web et relater leurs différents types et propriétés. Le développement des approches de sélection des services Web ouvre de nombreuses voies dans ce domaine, mais il n'existe pas de protocole normalisé pour la protection de la vie privée. À cet effet, nous avons exposé un état de l'art sur la composition et Sélection à base de vie privée. Ce dernier, nous a permis d'opter pour une amélioration du Framework de sélection proposé par [4], détaillée dans le chapitre suivant. Cette amélioration rend le Framework capable de prendre en charge une négociation avec les services Web. En l'accompagnant d'une optimisation de l'approche ASP, puisque cette dernière a prouvé son efficacité dans la mise à l'échelle.

Chapitre 3 : Une approche de sélection à base de critères de vie privée

1. Introduction

Dans ce chapitre, nous abordons le Problème de Sélection des services Web avec Préservation de la vie Privée (PSWPP). Notre principale contribution consiste à étendre le modèle de composition présenté par [4], par un mécanisme de négociation qui se base sur un système multi-agent. Nous proposons aussi un nouvel encodage ASP du problème de sélection PSWPP, cet encodage améliore le temps de réponse du système de sélection.

Premièrement, nous mettons en avant le Framework de sélection et son architecture principale. Ce Framework est utilisé pour la résolution du problème de sélection (PSWPP). Par la suite, nous proposons le modèle de négociation, ce dernier permet de trouver des solutions alternatives en cas d'échec de l'approche proposé par [4] au moyen d'un système multi-agent. Ce système peut négocier les termes de vie privée. Puis nous donnons une reformulation de l'encodage ASP (Answer Set Programming) [44] avec une évaluation ainsi qu'une comparaison avec les performances de l'encodage ASP proposé par [4].

2. Le Framework de sélection

Dans cette partie, nous présentons un Framework de sélection, son objectif est de résoudre le problème de sélection de service en proposant une composition respectant les contraintes de confidentialité.

L'architecture du Framework repose sur un gestionnaire de vie privée générant un service Web composite (une composition de service) grâce à un système multi-agents [3] en utilisant un algorithme de sélection (voir Figure 3-1 architecture du Framework). Il prend en entrée :

• Une composition abstraite qui spécifie le flux de données et le flux de contrôle des services composants (tâches abstraites).

• Un ensemble de services concrets qui implémente les tâches abstraites précédentes, en plus de la spécification de leurs contraintes de confidentialité (stockées dans le registre des services).

• Une requête de l'utilisateur qui spécifie ses exigences de confidentialité.

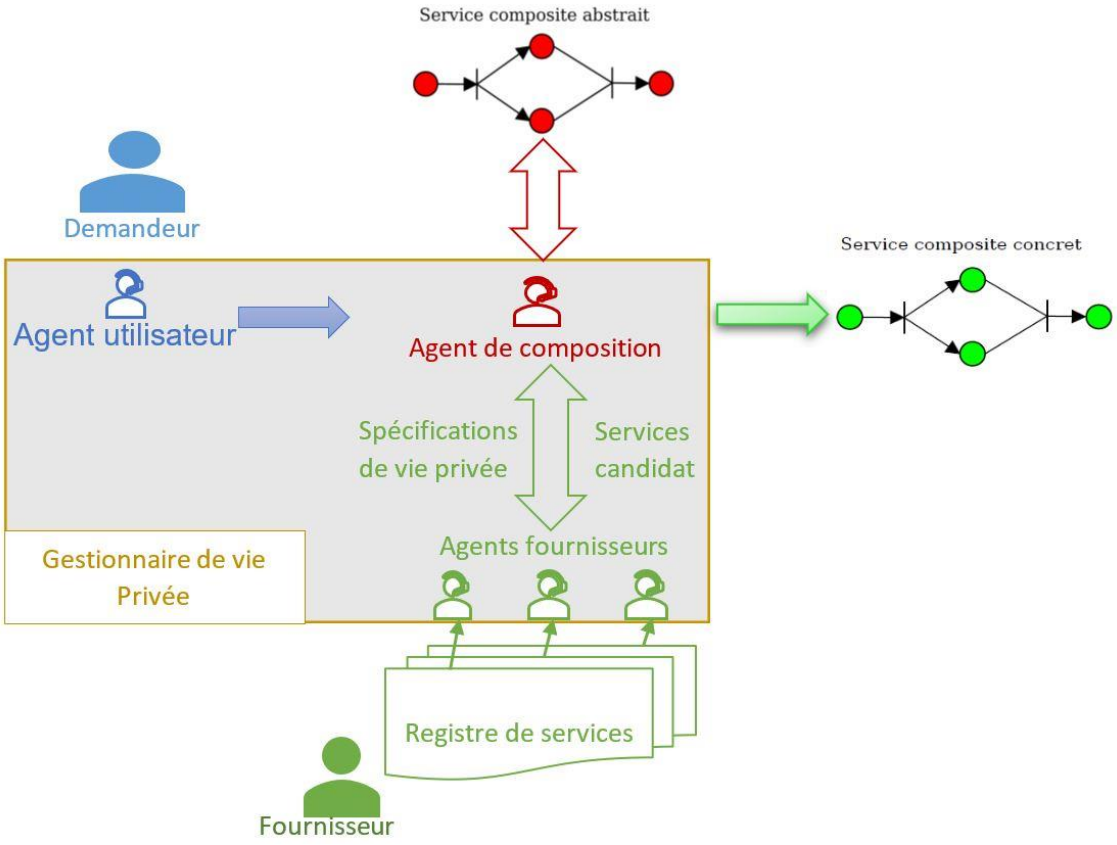

FIGURE 3-1 ARCHITECTURE DU FRAMEWORK

Le gestionnaire de vie privée est un système multi-agents composé de trois types d'agents :

a. L'agent utilisateur : il représente l'entité de l'utilisateur dans le système multi-agents. Ce dernier envoi les exigences (contraintes) de confidentialité avec la requête de composition à l'agent de composition.
b. Les agents fournisseur : chaque agent fournisseur envoi la spécification des contraintes de confidentialité à l'agent de composition pour qu'elles soient prises en compte lors de la combinaison des services.
c. L'agent de composition : c'est l'agent principal chargé de chercher la composition et de retourner le service composite concret après la réception de la requête de l'agent utilisateur.

Le scénario suivant concrétise l'échange de message qui s'opère entre les différents agents du système multi-agents :

Séquence	Détails
1	Requête de prise en charge (enregistrement dans le système)
1.1	Réponse à la requête (acceptation ou rejet)
1.1.1	Envoi des spécifications des contraintes de confidentialité
2	Envoi de la requête de composition
2.1	Enregistrement des spécifications (préférences de l'utilisateur)
2.2	Recherche de la composition
2.3	Envoi du service composite concret

TABLE 3-1 DESCRIPTION DES SEQUENCES DE LA SELECTION D'UN SERVICE COMPOSITE CONCRET

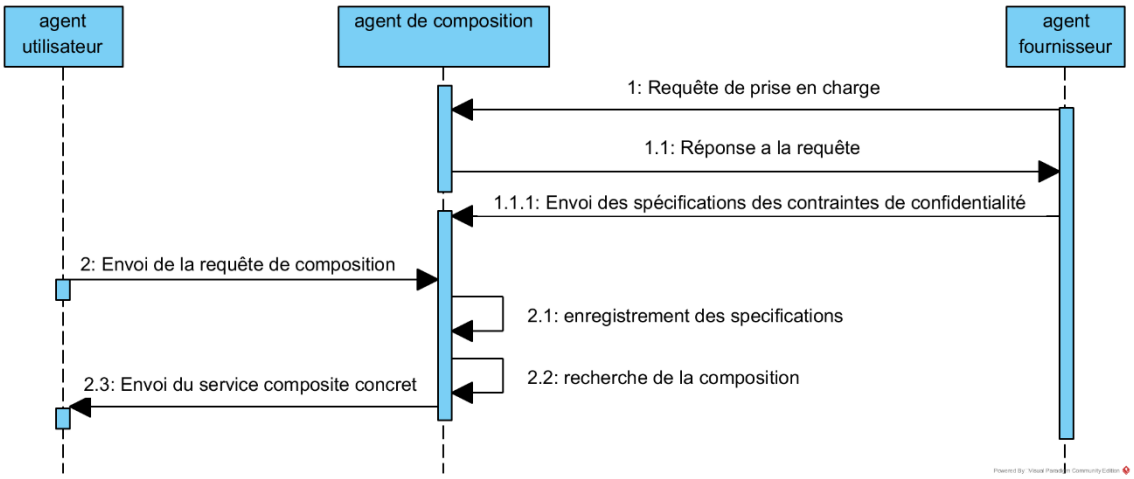

FIGURE 3-2 DIAGRAMME DE SEQUENCE DE LA SELECTION DU SERVICE COMPOSITE CONCRET

3. Les modèles du Framework

Le gestionnaire de vie privée repose sur trois modèles. Le modèle de composition, le modèle de vie privée [4] et le modèle de négociation, le rôle de ce dernier modèle est de définir un processus de négociation qui essaie de trouver un arrangement en cas d'absence d'une composition qui préserve toutes les contraintes de vie privée.

3.1. Le modèle de composition

Ce modèle est dédié à la représentation du flux de données.

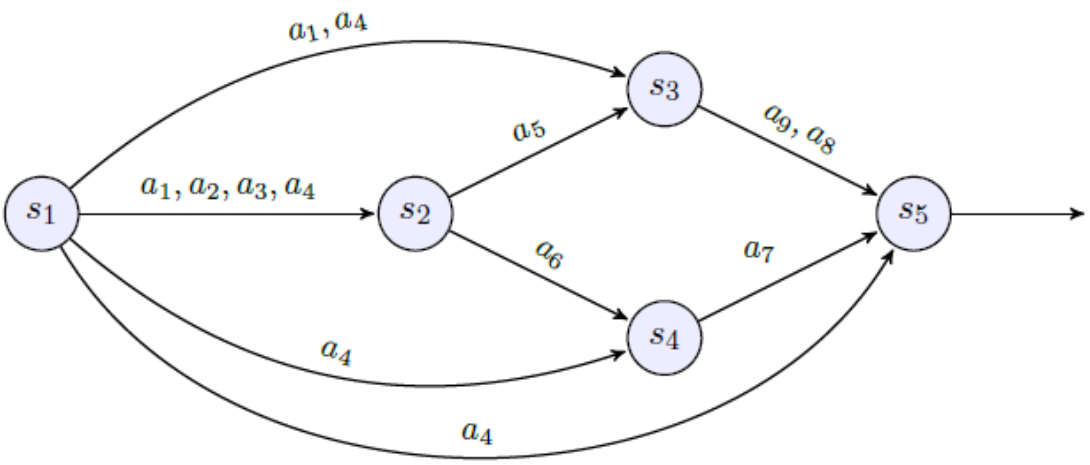

FIGURE 3-3 GRAPHE DE VIE PRIVEE CORRESPONDANT AU GRAPHE DE FLUX DE DONNEES [4]

La Figure 3-3 représente un exemple. Pour l'interprétation du graphe on peut prendre comme exemple l'arc entre s_1 et s_2, qui signifie que les politiques de vie privée du service s_2 relatives aux attributs a_1, a_2, a_3 et a_4 doivent être conformes aux exigences de vie privée du service s_1, sinon le service s_2 n'est pas éligible de fournir son service.

3.2. Le modèle de politique de vie privée

Ce modèle est dédié à la représentation des spécifications pour représenter les besoins de vie privée. Il se compose de :

3.2.1. Les Prédicats de vie privée

- L'objectif (Obj) : définit comment les données peuvent être utilisées une fois collectées.
- La visibilité (V): définit qui sont autorisés à voir les données fournies.
- La granularité (G) : définit le degré de précision des données fournies.
- Le temps de rétention (T) : définit la durée de conservation des données par le collecteur de ces dernières.

Afin de normaliser les prédicats de visibilité V et de granularité G des valeurs numériques sont attribuées à leurs valeurs. Elles varient selon des niveaux. Plus la valeur numérique de ces derniers est grande plus le niveau de visibilité et de divulgation des données de vie privée est élevé. Les figures suivantes illustrent des exemples de normalisation.

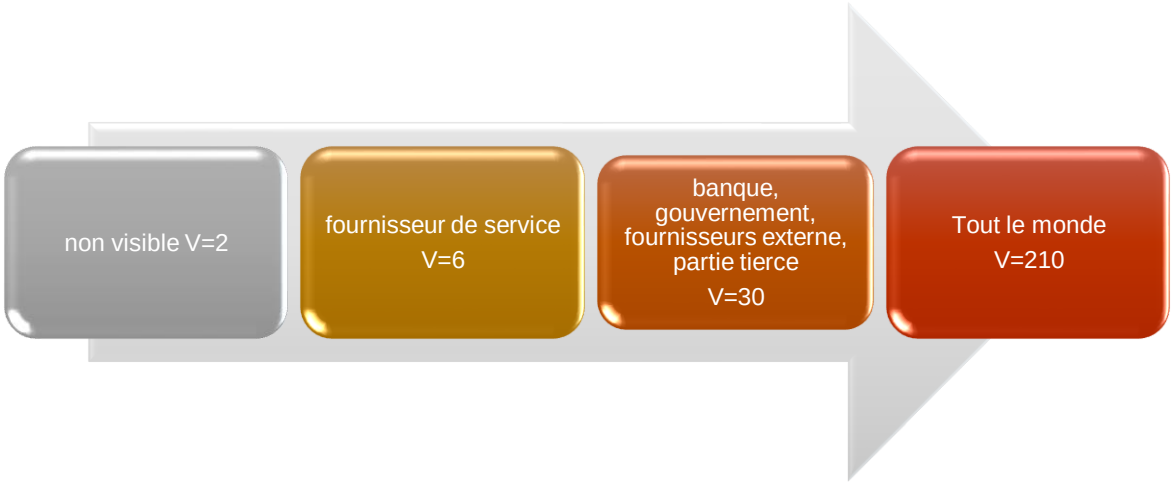

FIGURE 3-4 EXEMPLE DE NORMALISATION DU PREDICAT DE VISIBILITE V

Remarque : Nous avons mis pour l'exemple la divulgation à la banque, au gouvernement, au fournisseur externe et aux parties tierces un même niveau (V=30).

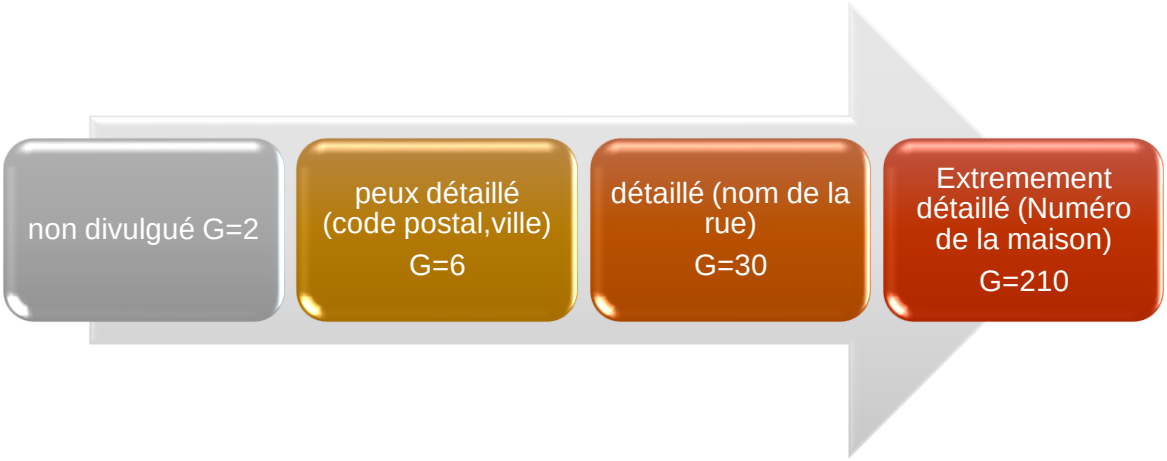

FIGURE 3-5 EXEMPLE DE NORMALISATION DU PREDICAT DE GRANULARITE *G*

3.2.2. Un ensemble de règles de vie privée

Ces règles sont formulées de sorte à garantir une représentation formelle des besoins en termes de vie privée en reprenant les prédicats précédents à savoir l'Objectif (Obj), la Visibilité (V), la Granularité (G) et le Temps de rétention (T). Chaque règle est sous le format suivant :

$$R_i = \{a_i, Obj_i, V_i, G_i, T_i\}$$

Tel que $a_i \in I$, est un attribut de données privées, (Ex : Nom, SSN, âge, adresse, etc.).

3.2.3. Politiques et Exigences de vie privée formulées grâce aux règles précédentes

- **Les politiques de vie privée (PP) :** sont un ensemble de règles de vie privée spécifiant l'ensemble des pratiques de vie privée applicables à toute donnée collectée par chaque fournisseur de services.

  Formellement défini comme suit : $PP^{s_x} = \{R_i \in P | R_i[I] \in I^{s_x}\}$. Où $I^{s_x} \subseteq \{Input\ de\ s_x\}$ et P représente l'ensemble des règles. [4]

Exemple : Soit un service de vente dont l'ensemble des fournisseurs est $Vendeur$, tel que pour chaque $vendeur_i \in Vendeur$ correspond l'ensemble de règles $\{PP^{vendeur_i}\}$, qui définissent comment $vendeur_i$ va utiliser les attributs $I^{vendeur} \subseteq I$.

Nous posons $PP_1^{vendeur_1} \in \{PP^{vendeur_1}\}$ et $PP_1^{vendeur_2} \in \{PP^{vendeur_2}\}$ avec les formalisations suivantes :

$$PP_1^{vendeur_1} = \{nom, "\ but\ d'achat", "\ partie\ tierce", détaillé\ , 120jrs\}$$

La Normalisation effectué grasse aux Figure 3-4 et Figure 3-5 donne : $PP_1^{vendeur_1} = \{nom, "\ but\ d'achat", 30, 30, 120jrs\}$

Telle que la règle $PP_1^{vendeur_1} \in PP^{vendeur_1}$ a été définie pour exprimer que $vendeur_1$ partage le nom du client avec une partie tierce, pour une durée de 120 jours dans le but d'effectuer l'achat où le nom du client complet va apparaître.

$$PP_1^{vendeur_2} = \{adresse, statistique, "\ partie\ tierce", "peux\ détaillé", 120jrs\}$$

La Normalisation effectué grasse aux Figure 3-4 et Figure 3-5 donne : $PP_1^{vendeur_2} = \{adresse, statistique, 30{,}6, 120\}$

Telle que la règle $PP_1^{vendeur_2} \in PP^{vendeur_2}$ a été définie pour exprimer que $vendeur_2$ partage l'adresse du client avec une partie tierce, dans une durée de 120 jours dans le but d'effectuer des statistiques où le code postal d'adresse du client va apparaître.

- **Les exigences de vie privée (PR) :** sont un ensemble de règles de vie privée spécifiant l'ensemble des conditions de vie privée qu'un fournisseur de service doit respecter en utilisant les données collectées. Cet ensemble est formalisé comme suit :
  $PR^{s_x} = \{R_i \in P | R_i[Att] \in O^{s_x}\}$. Où $O^{s_x} \subseteq \{Output\ de\ s_x\}$. [4]

  Exemple : Soit un utilisateur avec l'ensemble de règles $\{PR^{utilisateur}\}$. L' utilisateur défini une règle
  $$PR_1^{utilisateur} \in \{PR^{utilisateur}\}$$
  exprimant qu'il ne veut pas révéler les info-carte-crédit
  $$O^{utilisateur} \subseteq \{info - carte - crédit\}$$
  extrêmement détaillé qu'au service de payement uniquement pour effectuer le payement avec une durée de détention maximale d'une journée.

$$PR_1^{utilisateur} = \{info - carte - credit, payement, \\ " service\ fournis", " extrêmement\ détaillé", 1jrs\ \}$$

La Normalisation effectué grasse aux Figure 3-4 et Figure 3-5 donne : $PR_1^{utilisateur} = \{info - carte - credit, payement, 6, 210, 1\ \}$

3.2.4. Une fonction de risque

Quantifie le degré de divulgation des informations privées dans les compositions qui préservent les contraintes de vie privée. Cette fonction permet de sélectionner une composition avec un risque minimum (section 3.8).

3.3. Règles comparables

Deux règles de vie privée sont dites comparables, si elles sont associées au même attribut et si les valeurs de leurs prédicats « objectif » appartiennent au même ensemble. Formellement [4], les deux règles R_i et R_j sont comparable si :

$$\left(R_i[I] = R_j[I]\right) \wedge \left(\left(R_i[Obj], R_j[Obj]\right) \in \mathrm{Gl} \times \mathrm{Gl}\right) \qquad 3\text{-}1$$

Où, $\mathrm{Gl} \subseteq Obj$, représente l'ensemble des objectifs d'une composition donnée.

Nous définissons la fonction $comp(R_i, j)$ qui renvoie 1 si les deux règles (R_i, j) sont comparables.

$$comp(R_i, j) = \begin{cases} 1\ Si\ \left(R_i[I] = R_j[I]\right) \wedge \left(\left(R_i[Obj], R_j[Obj]\right) \in \mathrm{Gl} \times \mathrm{Gl}\right) \\ 0\ Sinon \end{cases} \qquad 3\text{-}2$$

3.4. Règles conformes

Une règle de vie privée R_i est dite conforme avec une règle de vie privée R_j $\left(R_i \sim R_j\right)$ si :

1. Les deux règles sont comparables ;
2. Les valeurs : visibilité, granularité et temps de rétention de Ri sont supérieurs ou égales à celles de R_j.

Formellement [4] :

$$R_i \sim R_j \Leftrightarrow \begin{cases} comp(R_i, j) = 1 \wedge \\ R_i[V] \geq R_j[V] \wedge \\ R_i[G] \geq R_j[G] \wedge \\ R_i[T] \geq R_j[T] \end{cases} \quad 3\text{-}3$$

3.5. Services conformes

Un service s_x avec une exigence de vie privée $PR^{s_{xy}}$ est conforme à un service s_y qui définit une politique de vie privée $PP^{s_{xy}}$, si la fonction $Nconf(s_x, s_y)$ renvoie zéro. Cette fonction représente le nombre de règles de vie privée non conformes entre les deux services s_x et s_y. Cette fonction est définie comme suit [4] :

$$Nconf(s_x, s_y) = \begin{cases} 0\ Si\ \forall\ R_i \in\ PR^{s_{xy}}, \exists R_j \in PP^{s_{xy}} : R_i \sim R_j \\ k\ sinon, (k = |NC|) \end{cases} \quad 3\text{-}4$$

Où, $NC = \{R_i \in\ PR^{s_{xy}} | \nexists R_j \in PP^{s_{xy}} : R_i \sim R_j\}$. $PR^{s_{xy}} \in PR^{s_x}$, $PP^{s_{xy}} \in PP^{s_x}$ est définies par : $PR^{s_{xy}} = \{R_i \in\ PR^{s_{xy}} | R_i[I] \in DEP^{s_x, s_y}\}$, $PP^{s_{xy}} = \{R_i \in\ PP^{s_{xy}} | R_i[I] \in DEP^{s_x, s_y}\}$, ou, PR^{s_x} et PP^{s_x} représentent respectivement les ensembles des exigences et des politiques de vie privée définies par les services s_x et s_y, alors que DEP^{s_x, s_y} représente l'ensemble des dépendances entre les services s_x et s_y.

3.6. Service valide

Un service s_x est dit valide, s'il est conforme à tous les services dont il dépend. Formellement [4]:

$$vld(s_x) = \begin{cases} 1\ Si\ \forall s_x \in PRD^{s_x} : Nconf(s_x, s_y) = 0 \\ 0\ Sinon \end{cases} \quad 3\text{-}5$$

Où, PRD^{s_x} représente l'ensemble de précédences du service s_x.

3.7. Composition valide

Une composition $C = \{s_1, s_2, \ldots, s_n\}$ est dite valide si tous ses services composants sont valides [4]. Nous définissons la fonction $vld(C)$, qui retourne 1 si la composition C est valide :

$$vld(C) = \begin{cases} 1\ Si\ \forall s_x\ \in C : vld(s_x) = 1 \\ 0\ Sinon \end{cases} \quad 3\text{-}6$$

3.8. Fonction de Risque

Afin de retourner la meilleure composition à l'utilisateur, le gestionnaire de vie privée inclut une fonction de minimisation du risque. Le risque d'une composition est défini par l'agrégation des risques des services qui la compose. Le risque de chaque service s_x est calculé grâce à la fonction $risk(s_x)$ pour l'ensemble des politiques de vie privée PP.

$$risk(s_x) = \frac{1}{\sum_{i=1}^{i=|PP|} \lambda_i} \sum_{R_i=1}^{R_{i=|PP|}} \lambda_i \times Agr(R_i[V], R_i[G], R_i[T]) \qquad 3\text{-}7$$

Où, λ_i représente le degré de sensibilité de l'attribut $R_i[I] = a_i$ de la règle R_i. Notons que le degré de sensibilité est une valeur entre 0 et 1, définie par le propriétaire de l'attribut de données (utilisateur ou fournisseur de services), et Agr représente l'intégrale de Choquet [45] [46], employée comme fonction d'agrégation [4].

Exemple d'application de la formule :

Soit deux services s_1 et s_2 avec les politiques de vie privée PP^{s_1} et PP^{s_2} pour un même attribut $O^s \subseteq \{adresse\}$, dans le but de faire des statistiques commerciales. La différence entre les deux fournisseurs de service est que l'un fait les statistiques localement $PP_1^{s_2}$, tant dis que le deuxième les confie à une entreprise d'analyse de données $PP_1^{s_1}$. Les règles sont définies après normalisation selon la Figure 3-4 et la Figure 3-5 comme suit :

$PP_1^{s_1} = \{adresse, statistique, 30{,}6, 120\}$

$PP_1^{s_2} = \{adresse, statistique, 6{,}6, 120\}$

A partir de ces règles, et en appliquant la fonction de calcul du risque avec $\lambda_i = 1$ (sensibilité égale), et une simple somme des prédicats $(R_i[V], R_i[G], R_i[T])$ comme fonction d'agrégation nous obtenons $risk(s_1) = 156$ et $risk(s_2) = 132$, Les résultats démontrent que s_1 présente un risque plus élevé de divulgation des données par rapport à s_2 dans la composition des services.

3.9. Modèle de négociation :

La modélisation de base représentée par le modèle de composition et le modèle de vie privée sert principalement à trouver une combinaison qui respecte les contraintes de vie privée fixées par les fournisseurs de services et l'utilisateur, or dans le cas où le gestionnaire de vie privée ne parvient pas à trouver une solution, l'utilisateur n'aura aucune composition en réponse. Pour remédier à ce problème nous proposons une extension des modèles précédents par un mécanisme qui a pour but de trouver une composition alternative en utilisant un processus de négociation. Ce dernier comprend un ensemble de traitements qui s'ajoutent au Framework de base. Le déroulement du processus de négociation est représenté par la Figure 3-6, et se déroule comme suit :

a. L'agent de composition : il dirige la négociation. Son rôle est d'envoyer, après que le processus de recherche de composition échoue des requêtes de négociation aux agents fournisseurs (Figure 3-6 5.1). Les réponses reçues seront utilisées pour la recherche d'une composition secondaire (Figure 3-6 5.3).
b. Les agents fournisseurs : ils sont en charge de fournir de nouvelles règles de vie privée si la négociation est acceptée. Dans le cas où l'agent a des règles fixes pour son service il renvoi un refus à l'agent de négociation (Figure 3-6 5.2).

Cette modélisation permet d'avoir une composition secondaire tout en gardant la transparence du processus de négociation vis-à-vis de l'utilisateur.

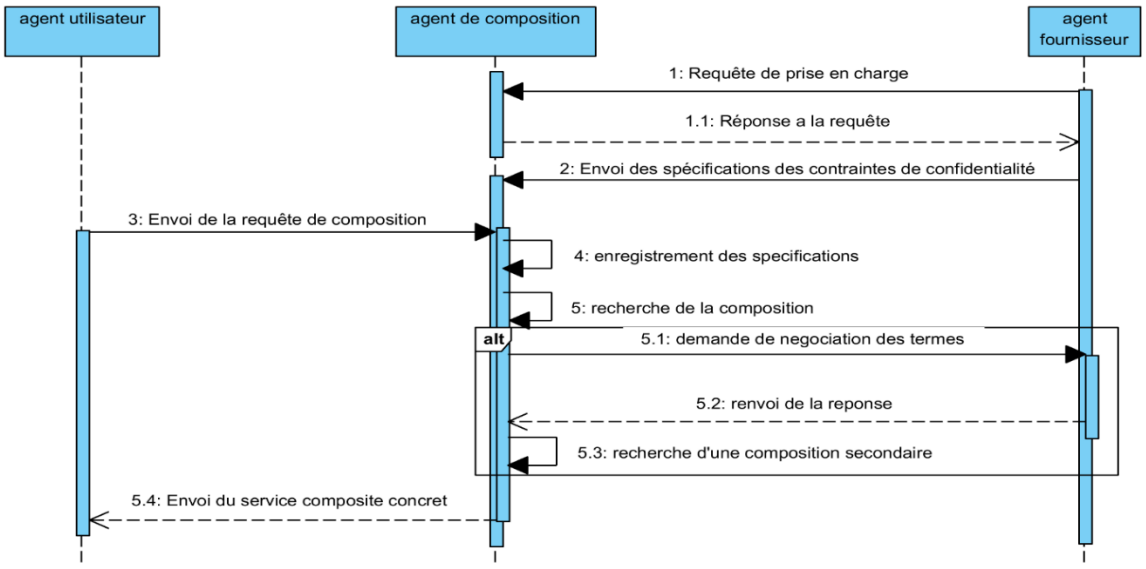

FIGURE 3-6 DIAGRAMME DE SEQUENCE DU PROCESSUS DE NEGOCIATION ET DE LA SELECTION DU SERVICE COMPOSITE CONCRET

4. Problème de sélection des services Web avec Préservation de la vie Privée (PSWPP)

4.1. Présentation du problème

La sélection des services Web préservant la vie privée demeure l'axe de notre travail. Les modèles et concepts définis dans le Framework de sélection (voir section 2) permettent de tracer un mode de sélection (voir section 4.2). L'objectif de ce processus est de trouver des compositions de services concrètes à partir des compositions abstraites. La sélection doit respecter les exigences et politiques de vie privée et passer l'évaluation du risque du service. Les compositions obtenues sont évaluées par la suite avec la fonction de risque de composition. La composition retournée est définie comme la composition concrète.

4.2. Mode de sélection

La sélection de service dans notre approche ce base sur le procédé suivant :

En premier lieu, l'algorithme cherche des services avec des règles conformes avec les règles de vie privée de l'utilisateur. Ces services sont définis comme des services candidats.

Dans un second lieu, à partir des services candidats, l'algorithme dégage un ensemble de services conformes, puis, cet ensemble est réduit à un ensemble de services valides.

En dernier lieu, notre algorithme classe les services qui constituent une composition valide selon la valeur de la fonction de risque.

En résumé, la sélection de services dans notre approche se base sur la recherche d'une composition valide (voir section 3.7) avec un risque minimal.

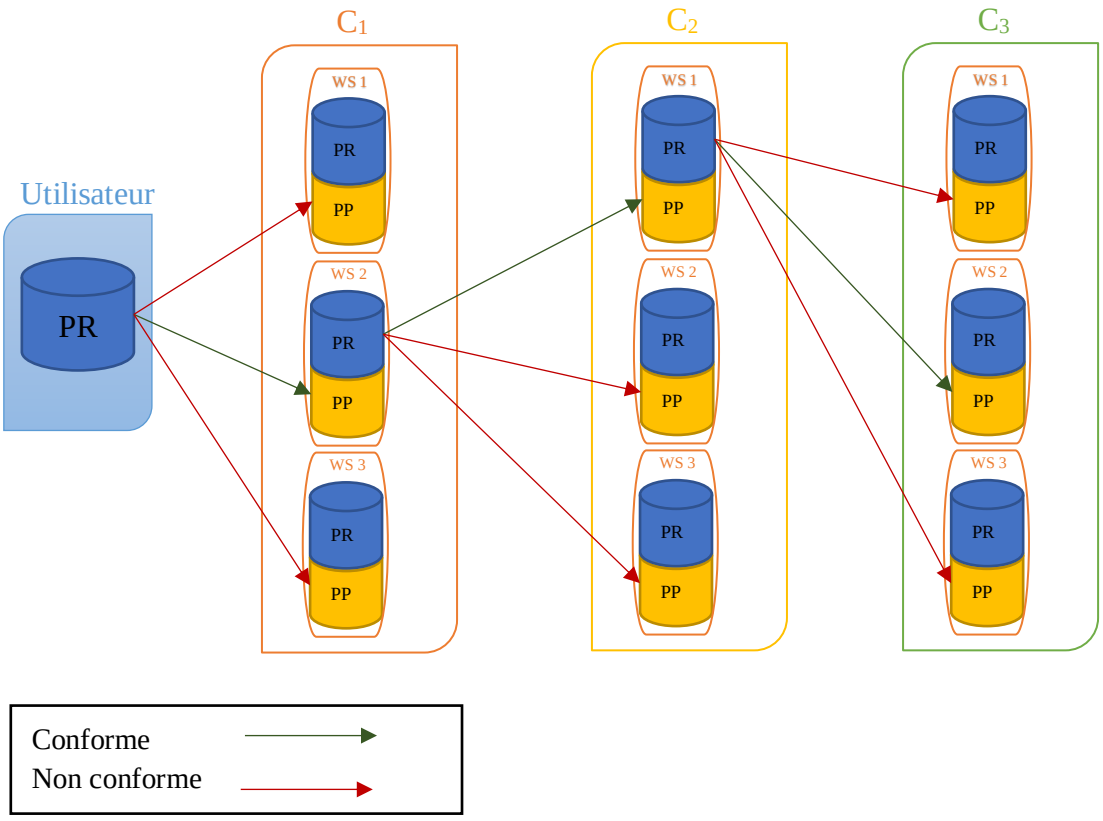

FIGURE 3-7 ILLUSTRATION DU PRINCIPE DE SELECTION DE SERVICE DANS PSWPP

Nous illustrons dans la Figure 3-7, l'utilisateur avec ces exigences de confidentialité PR, ainsi que trois classes de service C_1, C_2 et C_3, chacune d'elle propose trois services avec différentes politiques et préférences PP et PR. Chaque composition entre services conformes (voir équation3-4) et valides (voir équation 3-5) est liée avec des flèches vertes pour finalement avoir un service composite valide (voir équation 3-6). Durant les procédés de vérification de conformité et de validité, la fonction de risque évalue à chaque fois les risques des services candidats (conforme et valide) pour la sélection selon l'équation 3-7

5. Exemple illustratif du processus de négociation :

Nous simulons dans la Figure 3-8 un scénario d'achat en ligne qui nécessite un processus de négociation. L'agent utilisateur lance une demande d'achat R (requête) avec l'ensemble des attributs de vie privée

$$I = \{nom, adresse\}$$

. La transaction mobilise trois autres agents en plus de l'agent utilisateur : l'agent de composition, l'agent vendeur 1 et l'agent vendeur 2. La simulation se déroule selon les séquences suivantes :

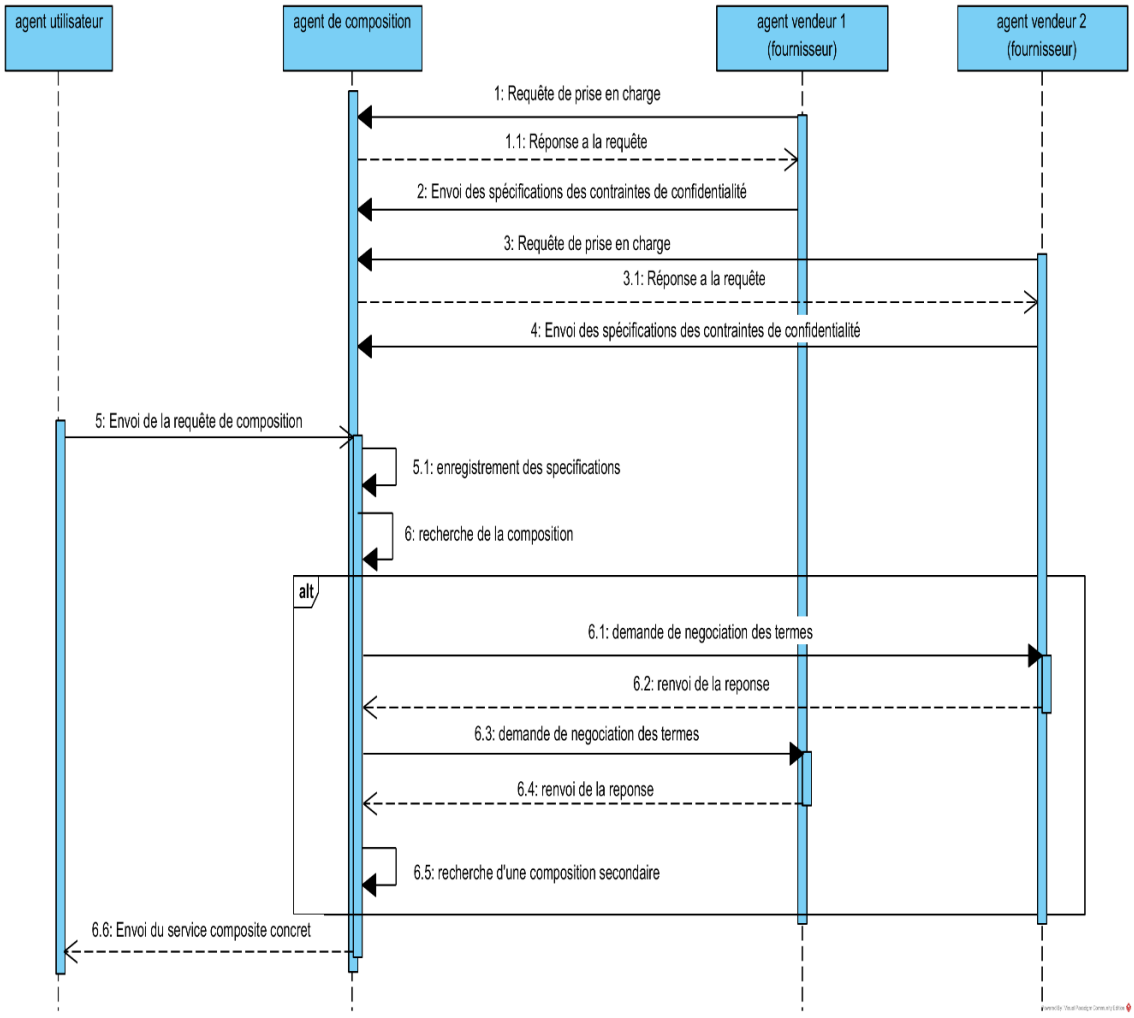

FIGURE 3-8 DIAGRAMME DE SEQUENCE DE LA SIMULATION

Séquence	Détails
1, 3	Envoi de requête de prise en charge de la part des agents fournisseurs pour qu'elle soit prise en charge dans le traitement des demandes de composition des agents utilisateurs.
1 :1,3 :1	L'agent de composition envoi des réponses affirmant la prise en charge
2	L'agent vendeur 1, envoi après la réception de 1 :1, les spécifications de contraintes de confidentialité : $PP_1^{vendeur_1}$= {nom, but d'achat, service-fourni, précis, 120jrs} $PP_2^{vendeur_1}$= {adresse, but d'achat, service-tiers, précis, 40jrs}
4	L'agent vendeur 2 envoi, après la réception de 3 :1, les spécifications de contrainte de confidentialité : $PP_1^{vendeur_2}$= {nom, but d'achat, service-fourni, précis, 20jrs} $PP_2^{vendeur_2}$= {adresse, but statistique, service-fourni, précis, 120jrs}
5	L'agent utilisateur envoi une demande de composition avec les spécifications de contraintes de confidentialité : $PR_1^{utilisateur}$ = {nom, but d'achat, service-fourni, précis, 20jrs} $PR_2^{utilisateur}$ = {adresse, but d'achat, service-fourni, précis, 40jrs}
5 :1	L'agent de composition enregistre les spécifications de l'utilisateur pour lancer la composition.
6	L'agent de composition lance le processus de résolution du problème PWPP en minimisant le risque. Le résultat de ce procédé est : ECHEC, causes : $PP_1^{vendeur_1}$ incompatible avec $PR_1^{utilisateur}$ et $PP_2^{vendeur_2}$ incompatible avec $PR_2^{utilisateur}$. Par conséquent le processus de négociation est enclenché.
6 :1, 6 :3	L'agent de composition envoie des demandes de négociation aux agents fournisseurs.
6 :2	L'agent vendeur 2 (fournisseur) répond par un refus de négociation
6 :4	L'agent vendeur 1 (fournisseur) répond par de nouvelles règles : $PP_1^{vendeur_1}$= {nom, but d'achat, service-fourni, précis, 10jrs} $PP_2^{vendeur_1}$= {adresse, but d'achat, service-fourni, précis, 40jrs}
6 :5	L'agent de composition met à jours les règles introduites, par les fournisseurs qui ont accepté la négociation et relance le processus de résolution du problème PSWPP qui va proposer une composition = agent vendeur 1, en conséquence de la compatibilité des contraintes de vie privée.
6 :6	L'agent de composition envoi le service composite concret qui répond à la requête de l'agent utilisateur (5)

TABLE 3-2 DESCRIPTION DES SEQUENCES DE LA SIMULATION

6. L'Algorithme de sélection

L'algorithme du procédé de sélection du service Web fait appel à un solveur ASP pour traiter le problème avec un encodage logique. Dans le but d'améliorer l'expérience de l'utilisateur, nous proposons une reformulation de l'encodage ASP proposé dans l'article [4], afin d'optimiser le temps de réponse du solveur et ainsi pouvoir assurer une gestion de composition des services Web plus fluide et acceptable par l'utilisateur.

6.1. Encodage ASP

6.1.1. La partie génération

La partie génération gère tous les services possibles de l'instance du problème. L'ensemble des prédicats utilisés pour cette tache sont :

$classe(1..n).$
$servicenb\ (0..m,n).$
$\{serviceId(I,J): servicenb(I,J)\} :- classe(J).$

Le prédicat $classe$ définie le nombre de classes dans l'instance du problème, tandis que le prédicat $servicenb$ précise le nombre de service de chaque classe. La règle avec l'entête $\{serviceId(I,J):\}$ génère l'ensemble de tous les services $J \in 0..m$ de la classe $I \in 1..n$.

6.1.2. La partie définition

Cette partie comporte tous les prédicats auxiliaires définissant l'instance du problème.

$$prNumber(I,J,N) :- depend(I,J), N = \#count\{A: pr(\_,I,A,\_,\_,\_), pp(\_,J,A,\_,\_,\_)\}.$$

Calcule le nombre des règles $(pp; pr)$ dépendant entre un service d'une classe I et un service d'une classe J . Avec, les prédicats $(pp/6; pr/6)$ représentent la spécification des règles PP , PR associées à un attribut A

$$\begin{aligned} conforme(S1,I,S2,J,A) :- & serviceId(S1,I), serviceId(S2,J), depend(I,J), \\ & pr(S1,I,A,V1,G1,T), pp(S2,J,A,V11,G11,T1), \\ & (V11 \setminus\setminus V1) == 0, (G11 \setminus\setminus G1) == 0, T \geq T1. \end{aligned}$$

Teste pour un attribut A, si une règle pr d'un service $S1$ de la classe I est conforme avec une règle pp d'un service $S2$ de la classe J. Répandant à la règle de conformité (de la section 3.4).

6.1.3. La partie test

$$:-serviceId(S1,I), serviceId(S2,J), depend(I,J), prNumber(I,J,N), N > 0,$$
$$N > \#count\{A: conforme(S1,I,S2,J,A)\}.$$

Décrit la contrainte qui exige que le nombre de règles conformes entre deux services dépendants ne doit pas être inférieur au nombre total des règles dépendantes de ces services.

6.1.4. La partie optimisation

$$\#minimize\{R,I,J : serviceId(I,J), risk(I,J,R)\}.$$

Cette partie indique au solveur de calculer le modèle stable optimal. Nous utilisons l'instruction suivante pour minimiser la valeur de risque de vie privée associée à une composition valide.

7. Evaluation

7.1. Evaluation de l'approche ASP proposée

Afin d'évaluer le gain obtenu par la reformulation nous avons utilisé deux types de base données : une base de données small-world et une base scale-free, ces deux bases ont étaient générées dans le travail de [4] en se basant sur les affirmations de plusieurs travaux [47], [48], [49], [50] assurant que la majorité des réseaux de service Web ont des caractéristiques des réseaux small-world [51] ou des réseaux scale-free [52]. La Table 3-3 la description des données.

Taille de la composition	La base small-world			La base scale-free		
	PP	PR	TOTAL	PP	PR	TOTAL
3	48	36	**84**	24	24	**48**
4	72	48	**120**	48	36	**84**
5	48	41	**89**	72	36	**108**
6	72	49	**121**	120	48	**168**
7	96	64	**160**	144	72	**216**
8	72	64	**136**	120	100	**220**
9	144	107	**251**	117	93	**210**
10	144	98	**242**	144	91	**235**

TABLE 3-3 DESCRIPTION DES ENSEMBLES DE DONNEES GENERES

Les expérimentations sont réalisées sur une machine dotée d'un processeur Intel(R) Core (TM) i7-5500U CPU @ 2.40GHz (4 CPUs), ~2.4GHz, avec une mémoire de 8192MB de RAM, et un processeur graphique AMD Radeon HD 8500M d'une mémoire dédiée de 2039MB, opérant sous OS Microsoft Windows 10 Professionnel 64-bit (10.0, Build 17134).

L'algorithme est implémenté dans Apache NetBeans IDE 10.0 (Build incubator-netbeans-release-380-on-20181217), Java : 1.8.0_112 ; Java Hotspot (TM) 64-Bit Server VM 25.112-b15, Runtime : Java (TM) SE Runtime Environment 1.8.0_112-b15.

Pour implémenter l'approche ASP nous avons utilisé le solveur Clingo [53]. Dans nos tests nous étudions l'influence du nombre de services candidats par classe ainsi que l'influence de la taille de composition sur l'efficacité de l'approche proposée.

7.1.1. L'influence du nombre de services candidats par classe sur l'efficacité de l'encodage proposé :

Dans cette expérimentation, nous fixons la taille de composition à 4 (voir Table 3-3) et nous varions le nombre de services par classe de 50 à 1000. Les résultats obtenus sont notés dans la Table 3-4 et représentés dans : la Figure 3-9 et la Figure 3-10.

Nombre de services	DATA SMALL WORLD	DATA SCALE FREE
50	159.482	113.744
100	377.222	230.246
150	719.226	405.86
200	1264.34	601.564
250	1932.994	816.416
300	2312.57	1133.772
350	3112.724	1496.934
400	4518.366	1762.474
450	5682.694	2198.75
500	6857.952	2553.568
600	9786.008	4117.898
700	14250.922	5617.394
800	17084.544	7152.348
900	21871.962	9574.542
1000	27224.708	11303.804

TABLE 3-4 L'INFLUENCE DE LA VARIANCE DU NOMBRE DE SERVICES SUR L'EFFICACITE DE L'ENCODAGE PROPOSE AVEC UNE COMPOSITION FIXE A 4

A. LA BASE SMALL WORLD

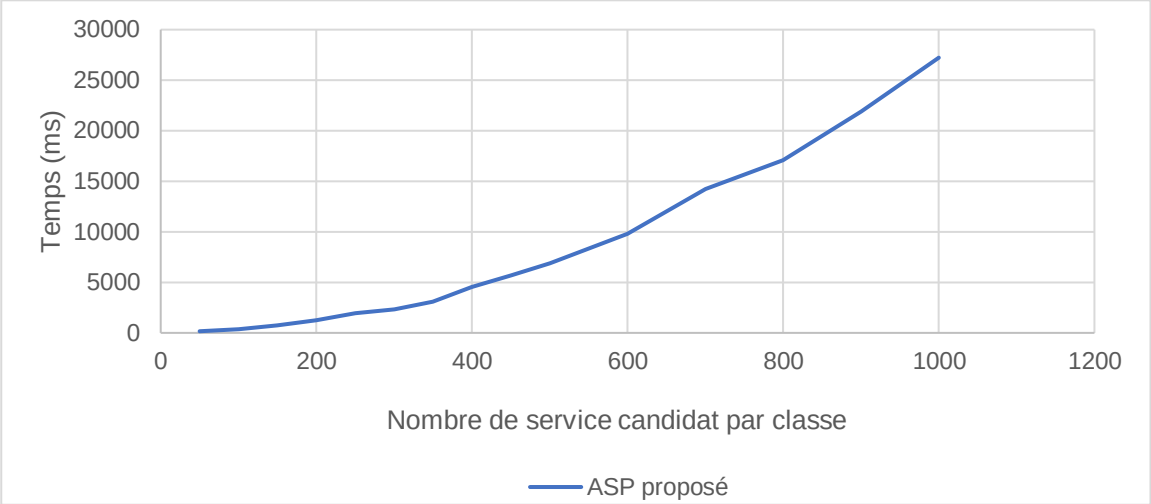

FIGURE 3-9 L'INFLUENCE DU NOMBRE DE SERVICES CANDIDATS PAR CLASSE DANS LA BASE SMALL WORLD

Dans la base small world, les résultats de cette expérimentation montrent clairement que le temps de calcul augmente proportionnellement avec l'augmentation du nombre de services par classe. Cette augmentation du temps de recherche prend une courbe qui suit une tendance polynomiale d'ordre 2 qui varie de 159.482 à 27 224.708 (ms).

B. LA BASE SCALE FREE

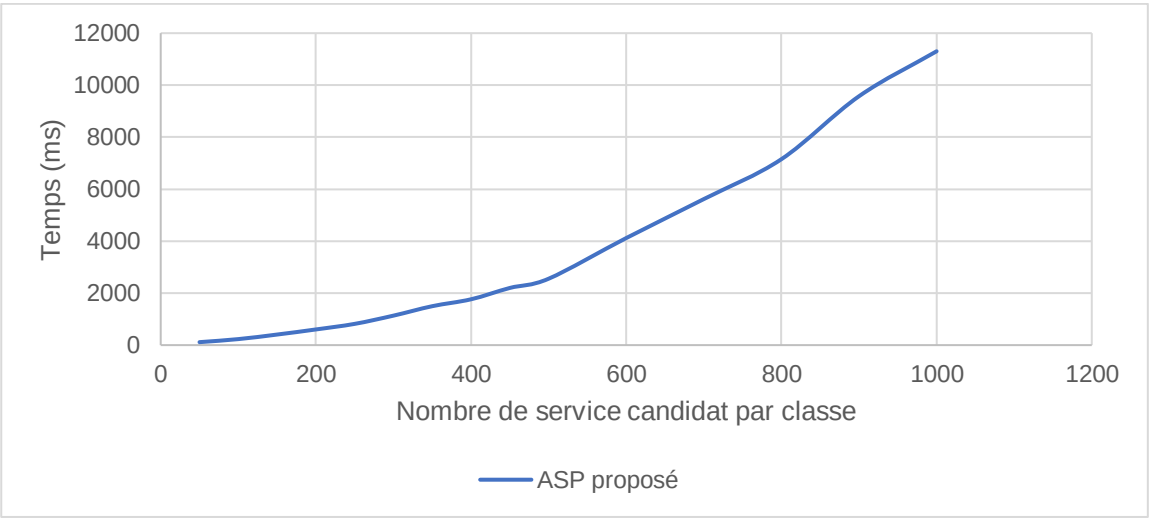

FIGURE 3-10 L'INFLUENCE DU NOMBRE DE SERVICES CANDIDATS PAR CLASSE DANS LA BASE SCALE FREE

Nous avons la même remarque que celle de la base small world. L'augmentation du temps d'exécution pour la base scall free prend aussi une courbe qui suit une tendance polynomiale d'ordre 2, variant de 113.744 à 11303.804 (ms).

7.1.2. L'influence de la taille de la composition sur l'efficacité de l'approche proposée :

Dans cette expérimentation nous évaluons les performances de l'algorithme de sélection en fixant le nombre de services candidats de chaque classe à 50 services en variant la taille de la composition de 3 à 10 (voir Table 3-3). Les résultats obtenus sont notés dans la Table 3-5 et représentés dans : la Figure 3-11 et la Figure 3-12.

Taille de la composition	DATA SMALL WORLD	DATA SCALE FREE
3	107,228	86,244
4	159,482	113,744
5	178,754	202,116
6	163,596	267,952
7	248,346	400,63
8	240,502	392,422
9	421,788	434,104
10	442,224	405,346

TABLE 3-5 L'INFLUENCE DE LA TAILLE DE COMPOSITION SUR L'EFFICACITE DE L'ENCODAGE PROPOSE AVEC UN NOMBRE DE CLASSE FIXE A 50

A. LA BASE SMALL WORLD

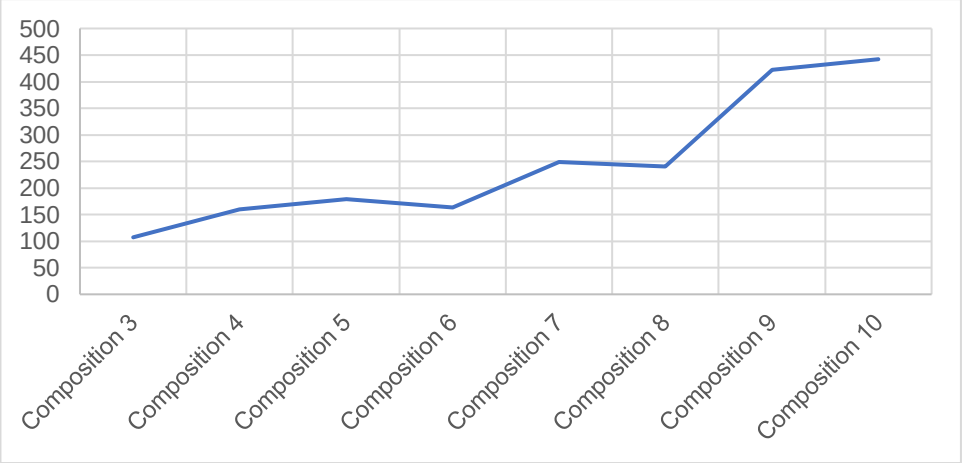

FIGURE 3-11 L'INFLUENCE DE LA TAILLE DE COMPOSITION SUR L'EFFICACITE DE L'ENCODAGE PROPOSE AVEC UN NOMBRE DE CLASSE FIXE A 50 DANS LA BASE SMALL WORLD

Ce graphe montre que les types de composition des graphes voir (Table 3-3) induisent à des augmentations saccadées du temps de composition dans les données small world. Comme nous pouvons remarquer un temps croissant entre la composition 3, 4 et 5 qui se casse dans la composition 6, puis il reprend la croissance, qui se casse a la composition 8 et reprend la croissance.

B. LA BASE SCALE FREE

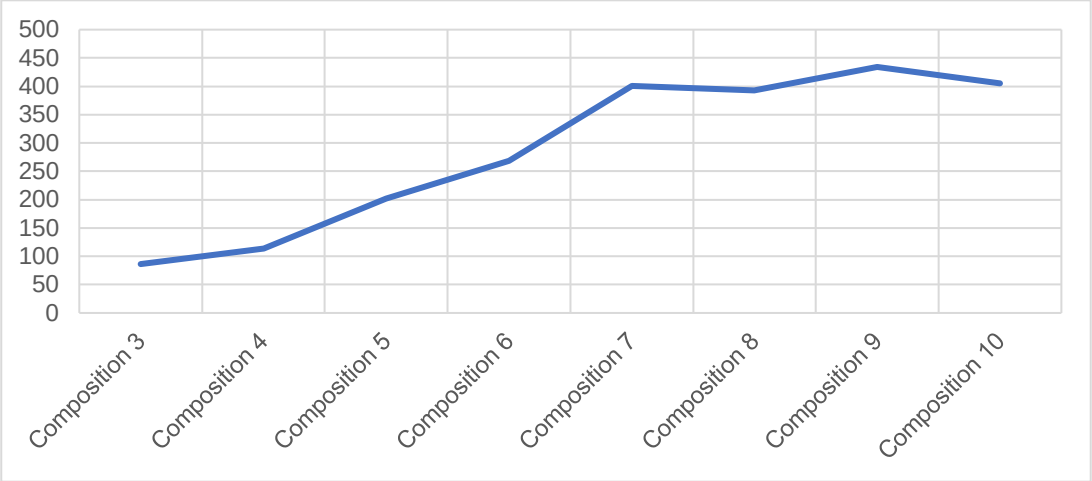

FIGURE 3-12 L'INFLUENCE DE LA TAILLE DE COMPOSITION SUR L'EFFICACITE DE L'ENCODAGE PROPOSE AVEC UN NOMBRE DE CLASSE FIXE A 50 DANS LA BASE SCALE FREE

Dans ce graphe on remarque une augmentation conséquente à partir de la composition 3 jusqu'à la composition 7, puis le rythme de croissance devient saccadé.

7.2. Comparaison

Dans cette partie nous présentons les résultats de la comparaison entre les performances de l'encodage proposé par [4] et notre encodage (7). Les comparaisons sont faites en se basant sur les mêmes expérimentations précédentes.

7.2.1. Comparaison du nombre de services candidats par classe sur l'efficacité de l'encodage proposé et l'encodage [4]

Cette expérimentation utilise les mêmes conditions que celle de la section 7.2.17.1.1, c'est à dire, la taille de la composition est fixée à 4, et le nombre de services par classe varie de 50 à 1000.

	DATA SMALL WORLD			DATA SCALE FREE		
Nombre de services	ASP [4]	Encodage proposé	Gain %	ASP [4]	Encodage proposé	Gain %
50	186.792	159.482	15%	166.946	113.744	32%
100	502.2	377.222	25%	479.812	230.246	52%
150	1024.314	719.226	30%	1005.046	405.86	60%
200	1793.212	1264.34	29%	1755	601.564	66%
250	2915.542	1932.994	34%	2779.87	816.416	71%
300	3968.096	2312.57	42%	3999.044	1133.772	72%
350	5562.176	3112.724	44%	5325.252	1496.934	72%
400	7030.232	4518.366	36%	7674.476	1762.474	77%
450	9207.17	5682.694	38%	9346.69	2198.75	76%
500	11620.816	6857.952	41%	11373.656	2553.568	78%
600	16409.832	9786.008	40%	17735.542	4117.898	77%
700	22533.132	14250.922	37%	23124.212	5617.394	76%
800	30555.038	17084.544	44%	31710.722	7152.348	77%
900	41444.274	21871.962	47%	41011.366	9574.542	77%
1000	50229.512	27224.708	46%	50132.5	11303.804	77%

TABLE 3-6 COMPARAISON DE L'INFLUENCE DE LA VARIANCE DU NOMBRE DE SERVICES SUR L'EFFICACITE DES APPROCHES PROPOSEES AVEC UNE COMPOSITION FIXE A 4

A. LA BASE SMALL WORLD

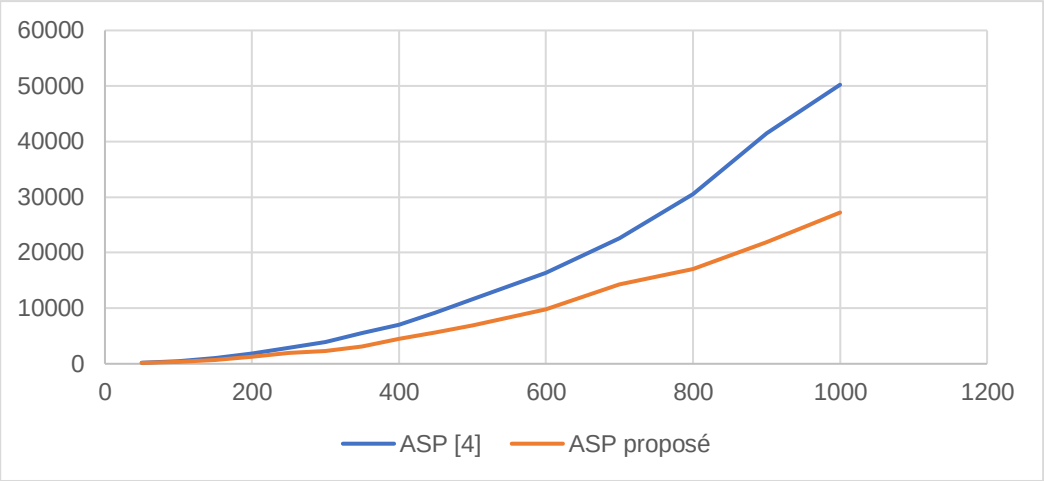

FIGURE 3-13 COMPARAISON DE L'EFFET DU NOMBRE DE SERVICES CANDIDATS PAR CLASSE SUR L'EFFICACITE DANS LA BASE SMALL WORLD DE L'ENCODAGE PROPOSE ET L'ENCODAGE [4]

Dans cette base, le temps de recherche du nouvel encodage est clairement diminué de sorte à avoir un gain de temps variant de 15 à 47%. Ce qui encourage l'adoption du nouvel en encodage dans la base small world.

B. LA BASE SCALE FREE

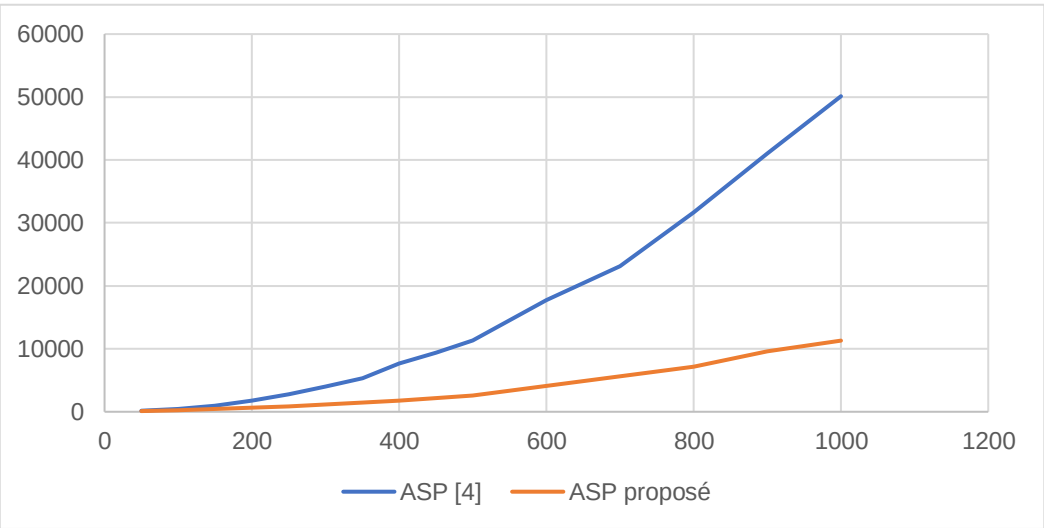

FIGURE 3-14 COMPARAISON DE L'EFFET DU NOMBRE DE SERVICES CANDIDATS PAR CLASSE SUR L'EFFICACITE DANS LA BASE SCALE FREE DE L'ENCODAGE PROPOSE ET L'ENCODAGE [4]

Dans cette base, nous remarquons aussi que le temps de réponse du nouvel encodage est clairement diminué de sorte à avoir un gain de temps variant de 32 à 77%. Le nouvel encodage s'adapte bien à la base scale free.

7.2.2. Comparaison l'influence de la taille de la composition sur l'efficacité de l'approche proposé et l'approche ASP [4]

Cette expérimentation utilise les mêmes conditions que celle de la section 7.1.2, c'est à dire, le nombre de services candidats de chaque classe est fixé à 50, et la taille de la composition varie de 3 à 10.

	DATA SMALL WORLD			DATA SCALE FREE		
Taille de la composition	ASP [4]	Notre ASP	Gain	ASP [4]	Notre ASP	Gain
3	113,336	107,228	5%	102,46	86,244	16%
4	186,792	159,482	15%	166,946	113,744	32%
5	264,2	178,754	32%	308,486	202,116	34%
6	332,55	163,596	51%	441,396	267,952	39%
7	568,25	248,346	56%	621,038	400,63	35%
8	685,05	240,502	65%	762,902	392,422	49%
9	1073,478	421,788	61%	999,366	434,104	57%
10	1687,048	442,224	74%	1194,99	405,346	66%

TABLE 3-7 COMPARAISON DE L'INFLUENCE DE LA TAILLE DE LA COMPOSITION SUR L'EFFICACITE DE L'APPROCHE PROPOSE ET L'APPROCHE [4]

A. LA BASE SMALL WORLD

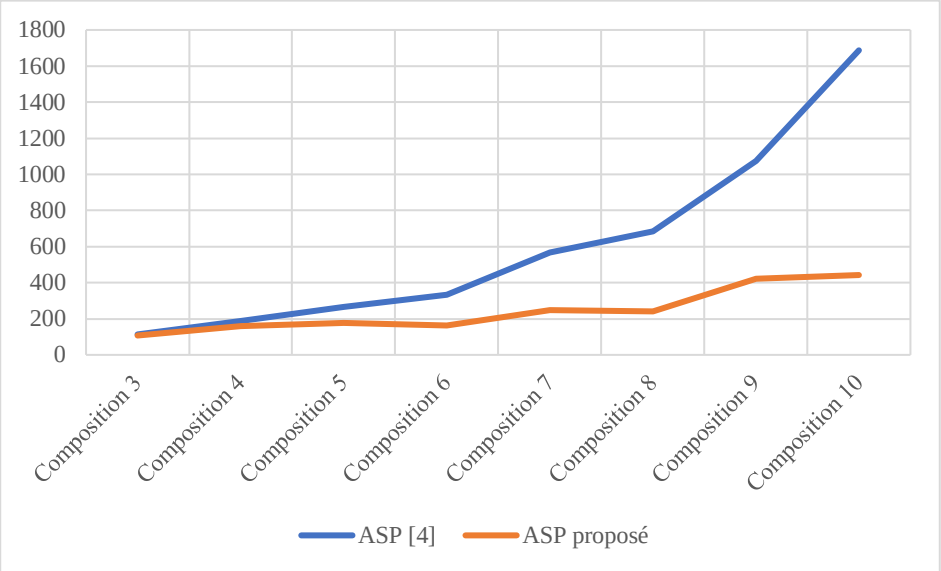

FIGURE 3-15 L'INFLUENCE DE LA TAILLE DE COMPOSITION SUR L'EFFICACITE DES DEUX ENCODAGES AVEC UN NOMBRE DE CLASSE FIXE A 50 DANS LA BASE SMALL WORLD

Pour base small world, les résultats montrent un net gain de temps d'exécution grâce à la reformulation de l'ASP. Un gain qui varie entre 5% pour la composition de taille 3 allant jusqu'à 74% de gain pour la composition de taille 10.

B. LA BASE SCALE FREE

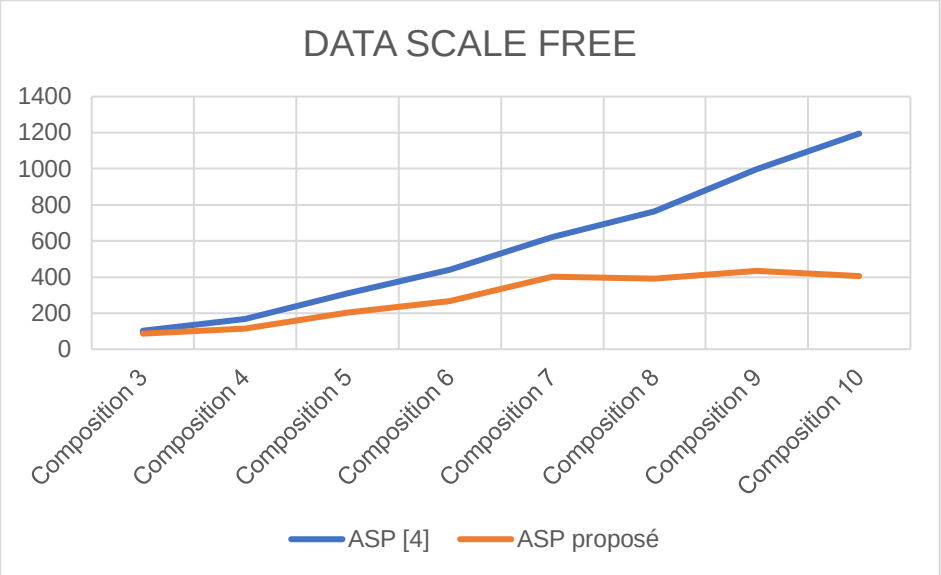

FIGURE 3-16 L'INFLUENCE DE LA TAILLE DE COMPOSITION SUR L'EFFICACITE DES DEUX ENCODAGES AVEC UN NOMBRE DE CLASSE FIXE A 50 DANS LA BASE SCALE FREE

Sur les données scale free, nous remarquons aussi un net gain de temps d'exécution grâce à la reformulation de l'ASP. Un pourcentage de gain allant de 16% pour la composition de taille 3 jusqu'à 66% de gain pour la composition de taille 10.

8. Conclusion

Nous avons proposé dans ce chapitre, un modèle de négociation complémentaire pour le Framework [4] ainsi qu'une reformulation de l'encodage ASP [4]. Ces contributions ont permis aux Framework de prendre en charge la négociation des contraintes de vie privée avec un meilleur temps de réponse.

Conclusion et perspectives

Dans ce travail, nous nous sommes intéressés à un problème épineux et d'actualité, celui de la protection de la vie privée dans les services Web. Nous avons proposé un Framework de sélection des services Web. Ce Framework, permet de sélectionner une composition avec un minimum de risque de menace sur la vie privée. Le Framework est doté d'un processus de négociation, qui se déclenche si la composition qui vérifie toutes les contraintes de confidentialité n'existe pas. Nous avons également proposé un encodage ASP de l'algorithme de sélection. Les tests effectués sur les performances de cet encodage, ont été très encourageants. Ces contributions ont permis aux Framework de sélection d'assurer plus de protection et plus de souplesse aux utilisateurs du Framework.

Comme tout travail, le présent travail est loin d'être complet, l'approche proposée nécessite plus d'améliorations et d'études sur l'efficacité de la protection des données. De ce fait, Plusieurs améliorations et extensions peuvent être envisagées pour enrichir l'approche proposée.

Du fait que le temps de réponse représente toujours un défi dans ce type de problème, chercher à le diminuer le met parmi les perspectives les plus importantes. L'amélioration peut être conduite en optimisant davantage l'encodage ASP, ou bien, en explorant d'autres approches que ce soit déclaratives comme le SMT (Satisfiability Modulo Theories), ou à base de méta-heuristiques.

L'un des axes qui nécessite plus d'attention, est la possibilité de joindre les deux aspects de sélection, à savoir, celui à base de qualité de service QoS et le modèle à base de vie privée, dans un seul Framework capable de satisfaire tous les besoins de l'utilisateur.

Aussi, implémenter le Framework proposé dans un environnement réel demeure une perspective importante, le but est d'évaluer l'efficacité et le comportement du Framework ainsi que l'algorithme de sélection dans un tel environnement.

Bibliographie

[1] F. HADJILA, Composition et interopération des services web sémantiques, Tlemcen: thèse de doctorat, 2014.

[2] A. HALFAOUI épouse GHERNAOUT, La sélection des services web dans une composition à base de critères non fonctionnels, Tlemcen: These de doctorat, 2017.

[3] P. G. a. D. S. Balaji, «An introduction to multi-agent systems,» chez *Innovations in multi-agent systems and applications-1*, Berlin, 2010.

[4] A. e. a. Belabed, «A Privacy-Preserving Approach for Composite Web Service Selection,» *Transactions on Data Privacy,* vol. 10, n° %12, pp. 83-115, 2017.

[5] linternaute, «Https://www.linternaute.fr/dictionnaire/fr/definition/vie-privee/,» 2019. [En ligne]. Available: Https://www.linternaute.fr/dictionnaire/fr/definition/vie-privee/. [Accès le 06 2019].

[6] I. d. l. R. p. l'Ontario, «https://www.ontario.ca/fr/document/manuel-sur-lacces-linformation-et-la-protection-de-la-vie-privee/chapitre-7-principes-fondamentaux-de-la-protection-de-la-vie-privee,» 2019. [En ligne]. [Accès le 06 2019].

[7] Google, «https://support.google.com/youtube/answer/6154230?hl=fr,» 06 2019. [En ligne]. [Accès le 06 06 2019].

[8] Assembly, UN General, «Universal declaration of human rights.,» 1948.

[9] privacyinternational, «Understanding Identity Systems Part 3: The Risks of ID,» 2019. [En ligne]. Available: https://www.privacyinternational.org/explainer/2672/understanding-identity-systems-part-3-risks-id. [Accès le 06 2019].

[10] E. NGUYEN, «Fraude à la carte bancaire : 18.000 signalements depuis juin, Perceval démarre en trombe,» *LA TRIBUNE,* 29 08 2018.

[11] V. F. B. a. E. D. Ferraris, The impact of profiling on fundamental rights, SSRN , 2013.

[12] Avast, «https://www.avast.com/fr-fr/c-malware,» 2019. [En ligne]. [Accès le 06 2019].

[13] M. Rhodes-Ousley, Information security: the complete reference, McGraw Hill Education, 2013.

[14] Avast, «https://www.avast.com/fr-fr/c-phishing,» 2019. [En ligne]. [Accès le 06 2019].

[15] Cybrary.it, «https://www.cybrary.it/glossary/i-the-glossary/inference-attack/,» 2019. [En ligne]. [Accès le 06 2019].

[16] S. Gambs, «https://www.irisa.fr/prive/sgambs/presentations/Sebastien_Gambs_LYRICS.pdf ,» 28 05 2015. [En ligne]. [Accès le 06 2019].

[17] S. Gambs, «https://www.irisa.fr/prive/sgambs/cours6_pvp.pdf,» 25 11 2015. [En ligne]. [Accès le 06 2019].

[18] J. R. Vacca, Computer and information security handbook, 3rd éd., Newnes, 2012, p. 770.

[19] «MSP Implementation with Identity Mixer,» 2019.

[20] D. L. Chaum, Untraceable electronic mail, return addresses, and digital pseudonyms, vol. 24, Communications of the ACM, 1981, pp. 84-90.

[21] D. Chaum, «The dining cryptographers problem: Unconditional sender and recipient untraceability,» *Journal of cryptology*, vol. 1, n° %11, pp. 65-75, 1988.

[22] D. M. M. G. R. a. P. F. S. Goldschlag, «Hiding routing information,» chez *International workshop on information hiding*, 1996.

[23] A. Belabed, La protection de la vie privée sur Internet, Tlemcen: thèse de doctorat, 2018.

[24] W3C, «https://www.w3.org/P3P/,» 2018. [En ligne]. Available: https://www.w3.org/P3P/. [Accès le 06 2019].

[25] W3C, «A P3P Preference Exchange Language 1.0 (APPEL1.0),» 2002. [En ligne]. Available: https://www.w3.org/TR/P3P-preferences/. [Accès le 06 2019].

[26] J. R. Vacca, Computer and information security handbook, 1st éd., Newnes, 2012.

[27] P. e. a. Ashley, «E-P3P privacy policies and privacy authorization,» chez *Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society*, 2002.

[28] W3C, «https://www.w3.org/TR/ws-arch/,» 2004. [En ligne]. [Accès le 06 2019].

[29] C. Refes, «Les services Web,» Openclassrooms, 22 02 2017. [En ligne]. Available: https://openclassrooms.com/fr/courses/219329-les-services-web. [Accès le 06 2019].

[30] Guru99, «https://www.guru99.com/restful-web-services.html#3,» 2019. [En ligne]. [Accès le 06 2019].

[31] Université Paris-Est Marne-la-Vallée , «Les spécifications WebService,» [En ligne]. Available: http://igm.univ-mlv.fr/~dr/XPOSE2005/rouvio_WebServices/soap.html. [Accès le 2019].

[32] mozilla, «https://developer.mozilla.org/en-US/docs/Web/HTTP,» 2019. [En ligne]. [Accès le 06 2019].

[33] J. D. T. a. P. K. Snell, Programming Web Services with SOAP: building distributed applications, O'Reilly Media, Inc, 2001.

[34] Guru99, «SOAP Web Services Tutorial: Simple Object Access Protocol EXAMPLE,» 2019. [En ligne]. Available: https://www.guru99.com/soap-simple-object-access-protocol.html. [Accès le 06 2019].

[35] W3C, «Web Services Description Language (WSDL) 1.1,» 15 03 2001. [En ligne]. Available: https://www.w3.org/TR/2001/NOTE-wsdl-20010315. [Accès le 2019].

[36] w3schools, «XML WSDL,» 2019. [En ligne]. [Accès le 06 2019].

[37] International Business Machines Corporation, Microsoft, «Web Services Inspection Language (WS-Inspection) 1.0,» 2002. [En ligne]. Available: https://svn.apache.org/repos/asf/webservices/archive/wsil4j/trunk/java/docs/wsinspection.html. [Accès le 06 2019].

[38] SYS-CON Media, «Web Services Orchestration and Choreography,» 2004. [En ligne]. Available: http://www2.sys-con.com/ITSG/virtualcd/WebServices/archives/0307/peltz/index.html. [Accès le 06 2019].

[39] E. F. a. N. H. T. Barbara Carminati, «A privacy-preserving framework for constrained choreographed service composition,» chez *IEEE International Conference onWeb Services*, 2015.

[40] C. G. B. M. a. M. M. Salah-Eddine Tbahriti, Privacy-enhanced web service composition, vol. 7, IEEE Transactions on Services Computing, 2013, pp. 210-222.

[41] A. C. B. C. a. S. K. Squicciarini, «Privacy aware service selection of composite web services invited paper,» chez *9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, 2013.

[42] E. F. P. a. N. Z. Costante, «Privacy-aware web service composition and ranking,» chez *2013 IEEE 20th International Conference on Web Services*, 2013.

[43] O. Y. L. a. D. S. Kwon, A Galois lattice approach to a context-aware privacy negotiation service, vol. 38, Expert Systems with Applications, 2011, pp. 12619-12629.

[44] T. a. I. N. Janhunen, «The answer set programming paradigm,» *AI Magazine*, vol. 37, n° %13, pp. 13-24, 2016.

[45] M. Grabisch, «The application of fuzzy integrals in multicriteria decision making,» *European*, vol. 89, n° %13, pp. 445-456, 1996.

[46] J.-L. Marichal, «An axiomatic approach of the discrete choquet integral as a tool to aggregate,» *IEEE Transactions on Fuzzy Systems*, vol. 8, n° %16, pp. 800-807, 2000.

[47] B. L. Z. Z. a. S. C. Zhiyong Feng, «A study of semantic web services,» *The Computer Journal*, vol. 58, n° %16, pp. 1293-1305, 2015.

[48] K. Y. F. a. W. T. Huang, «An empirical study of programmable web: A network analysis on a service-mashup system,» chez *2012 IEEE 19th International Conference on Web Services*, 2012.

[49] V. L. a. J.-F. S. Chantal Cherifi, «Benefits of semantics on web service,» chez *In International Conference on Networked*, 2010.

[50] S.-C. Oh, «Effective web-service composition in diverse and large-scale service networks,» *IEEE Transactions on Services Computing*, vol. 1, n° %11, pp. 15-32, 2008.

[51] D. J. a. S. H. S. Watts, «Collective dynamics of 'small-world'networks,» *nature*, vol. 393, n° %16684, pp. 440-442, 1998.

[52] B. e. a. Bollobás, «Directed scale-free graphs,» chez *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, 2003.

[53] M. e. a. Gebser, «Clingo= ASP+ control: Preliminary report,» *arXiv preprint arXiv*, vol. 1405, n° %13694, 2014.

[54] H. e. a. Kil, «Graph theoretic topological analysis of web service networks,» *World Wide Web*, vol. 12, n° %13, pp. 321-343, 2009.

[55] C. Hadnagy, Social engineering: The art of human hacking, John Wiley & Sons, 2010.

Résumé

Les applications Web composées de services Web atomiques sont utilisées dans de nombreux domaines. Afin de satisfaire les besoins des utilisateurs, des modèles de sélection des services Web ont été mis en place. Les principaux critères de sélection pris en charge par ces derniers sont les critères de qualité de services (QoS), en négligeant les paramètres de confidentialité. Dans ce travail, afin de pallier au problème de la protection de la vie privée dans les services Web, nous proposons un Framework de sélection des services Web à base de critères de vie privée. Ce Framework est constitué d'un système multi-agents et doté d'un processus de négociation. Le Framework proposé, permet de chercher une composition avec un minimum de risque de menace sur la vie privée. Nous avons également proposé une approche déclarative à base de l'Answer Set Programming (ASP) pour implémenter l'algorithme de sélection.

Mots-clés : Vie privée sur Internet, service Web, sélection des Services Web, négociation, composition.

Abstract

Web applications composed of atomic Web services are used in many fields. In order to meet user needs, Web services selection models have been implemented. The main selection criteria supported by the latter are the quality of service (QoS) criteria, neglecting confidentiality settings. In this work, in order to address the issue of privacy protection in Web services, we propose a Framework for selecting Web services based on privacy criteria. This Framework consists of a multi-agent system with a negotiation process. The proposed Framework is able to find a composition with a minimum risk of threat to privacy. We also proposed a declarative approach based on Answer Set Programming (ASP) to implement the selection algorithm.

Keywords : Internet privacy, web service, web services selection, negotiation, composition.

ملخص

تستخدم تطبيقات الويب (التي تتكون بدورها من عدة خدمات اخرى) في مجالات عديدة لتلبية احتياجات المستخدمين. تم وضع نماذج لاختيارات خدمات الويب من أجل تلبية تفضيلات المستخدمين. معايير الاختيار الرئيسية المعتمدة في أغلب الاحيان هي معايير جودة الخدمات (QoS) و التي غالبا ما تهمل إعدادات الحياة الخاصة للمستخدم. في هذا العمل، للتخفيف من مشكلة حماية الحياة الخاصة في خدمات الويب، نقترح اطار (Framework) اختيار خدمات الويب بمعايير الحياة الخاصة. يتكون هذا الإطار من نظام متعدد الوكلاء و يتضمن عملية تفاوض. يتيح الإطار المقترح البحث عن تركيب خدماتي مع حد أدنى من مخاطر التهديد على الحياة الخاصة. اقترحنا أيضا ترميز ASP لتجسيد خوارزمية الاختيار.

الكلمات الرئيسية: الحياة الخاصة على الإنترنت، خدمة الويب، اختيار خدمات الويب (الأقسام)، التفاوض، التكوين.