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

République Algérienne Démocratique et Populaire
Université Abou Bekr 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

La protection de la vie privée des utilisateurs dans les services web

Réalisé par :

- **LAGHA Omar**
- **TEBBAL Asma**

Présenté le 24 Juin 2014 devant le jury composé de MM.

- *Mr HADJILA. F* *(Président)*
- *Mr* BELABED. *A* *(Encadreur)*
- *Mr BRIKCI. A* *(Examinateur)*
- *Mme KAZI. A* *(Examinateur)*

Année universitaire: 2013-2014

Remerciements

Nous tenons à remercier

Allah le tout puissant qui nous a donné santé, courage pour terminer ce travail.

Mr BELABED AMINE, notre encadreur, pour ses conseils, sa disponibilité et son encouragement qui nous ont permis de réaliser ce travail dans les meilleures conditions.

Les jurys pour leurs efforts et leur soin apporté à notre : travail :

- *Mr HADJILA. F*
- *Mr BRIKCI. A*
- *Mme KAZI. A*

Aux enseignants de notre université et du département d'informatique.

Dédicaces

Je dédie ce travail à

Mes très chers parents par les quelles j'aurais jamais atteindre la place où je suis sans leurs admirable rôles et si encouragent pour accomplir mes études.

Mon frère, mes sœurs et leurs familles.

Mes collègues et mes amies.

Omar

Dédicaces

Au nom Dieu le clément et le miséricordieux louange à Allah le tout puissant

Je remercie premièrement infiniment mon DIEU qui ma aidé dans mon travail.

A nos très chère parents, qui ont toujours été la pour nous, et nous ont donne un magnifique modèle de labeur et de persévérance.

A plus merveilleuse mère qui a sacrifie sa vie pour nous, sa présence et petits soins qu'elle ma donné m'ont procuré beaucoup de réconfort et maitrise pour attendre mon but, qu'elle trouve ici le témoignage de ma profond affectation.

A mes chères frères et sœurs pour leur soutien moral.

A mon très cher binôme Omar pour sa participation et son aide à accomplir ce projet.

A mes meilleurs amis de ma promo.

A mes tantes et à mes oncles.

A chaque cousin et cousines spécialement Isma, chahinez, et la petite zineb.

A toute la famille, grande et petite.

A tous ceux et toutes celle qui m'ont accompagné et soutenu de prés ou de loin durant cette année.

Asma

Résumé

Bien que les technologies de composition et de sélection des services web sont considérées comme les technologies les plus prometteuses pour l'intégration des sources de données hétérogènes et multiples, ainsi que pour la réalisation d'opérations complexes. La question de la protection de la vie privée des utilisateurs demeure l'une des préoccupations majeures liées à ces technologies. Notre objectif dans ce travail est le développement d'une application permettant d'améliorer la sélection des services web avec des mécanismes de protection des données personnelles des utilisateurs. Pour atteindre cet objectif, nous avons proposé une approche qui se base sur un algorithme de compatibilité qui permet de vérifier formellement la compatibilité entre les exigences et les politiques de confidentialité de tous les services lors d'un processus de sélection d'une composition.

Mots clés : services web, sélection des services web, la vie privée.

Abstract

The issue of the protection of users' privacy remains on of the most important preoccupation related to technologies composition and selection of web services, even though, these ones are considered as the promising technologies for integration of heterogeneous and multiple sources of data as well as performing complex operation. Our goal in this work is to develop an application in an effort to improve the selection of web services with mechanisms to protect users' personal data. In order to reach this goal, we have proposed an approach based on an algorithm of accounting system that permit the account examination between demand and privacy policies of all services during a process of selecting composition.

Keywords : web services, selection of web services, privacy.

ملخص

على الرغم من أن تكنولوجيا تركيب واختيار خدمات الويب تعتبر من أكثر التكنولوجيات الواعدة لدمج مصادر البيانات المتعددة والغير متجانسة، وكذلك لتحقيق عمليات معقدة. مسألة حماية الحياة الخاصة للمستعملين لا تزال شكل أحد العوائق الرئيسية المتعلقة بهذه التكنولوجيات. هدفنا في هذا العمل هو تحسين اختيار خدمات الويب مع آليات لحماية البيانات الشخصية للمستخدمين. لتحقيق هذا الهدف، اقترحنا نهج يقوم على خوارزمية التوافق التي تتيح التحقيق بين متطلبات وسياسات الخصوصية لجميع الخدمات أثناء عملية الاختيار.

الكلمات المفتاحية : خدمات الويب، اختيار الخدمات، الحياة الخاصة.

Table des matières

Introduction générale 1

Chapitre I : Les services web

I. Introduction 4

II. Définition des services web 5

II.1. IBM 5

II.2. W3C 5

III. Architecture des services web 5

IV. Cycle de vie d'un service web 7

V. Utilisation des services web 8

V.1. Description des services web 8

V.2. Publication et découvert des services web 8

V.3. Invocation des services web 8

VI. Les technologies des services web 9

VI.1. XML 9

VI.2. SOAP 10

VI.3. WSDL 11

VI.4. UDDI 12

VII. Composition des services web 13

VII.1. Définitions 13

VII.2. Types de composition de services web 14

VIII. Sélection des services web 14

IX. Quelques domaines d'application des services web 15

X. Avantages et inconvénients des services web 15

X.1. Avantages 15

X.2. Inconvénients 16

XI. Conclusion 17

Chapitre II : La vie privée dans les web services

I. Introduction 18

II. Définition 19

II.1. La vie privée (Sphère privée) 19

II.2. Droit à la vie privée 19

II.3. Protection de la vie privée (ou de la sphère) 19
II.4. La vie privée sur Internet.....19
II.5. Surveillance 20
II.6. Sécurité.....20
III. Les attaques de la vie privée 21
III.1. Vol d'identité.....21
III.2. L'hameçonnage (phishing).....21
IV. Technologies de protection de la vie privée 22
IV.1. Privacy by design 22
IV.2. Privacy Enhancing Technologies (PET).....24
IV.2.1. Les systèmes de communications et accès anonymes.....24
IV.2.2. Les systèmes de gestion d'identités 25
V. Les langages de protection de la vie privée.....25
V.1. Le Langage d'expression de politique de la vie privée.....26
V.1.1. Platform for Privacy Preferences (P3P) [LGMM ,2005] 26
V.2. Le Langage de préférence en termes de vie privée.....27
V.2.1. APPEL.....28
VI. Synthèse des travaux de recherche 29
VII. Conclusion.....30

Chapitre III : Conception et Implémentation du prototype
I. Introduction 31
II. Scénario de motivation 31
III. Approche proposée.....32
III.1. Les règles de confidentialité.....32
III.2. Matching entre services 33
III.2.1. Critères de matching.....33
III.2.2. Subsomption de confidentialité.....34
III.2.3. Exemple 35
III.3. Algorithme proposé 37
IV. Conception.....38
IV.1. Diagramme de cas d'utilisation.....38
IV.2. Diagramme de séquence 38

IV.3. Diagramme de classes....39
V. Présentation des outils technologiques utilisés....40
VI. Présentation de l'IHM....41
VII. Expérimentation....45
VII.1. Description de la base....45
VII.2. Expérimentation 1....45
VII.3. Expérimentation 2....46
VII.4. Expérimentation 3....48
VIII. Discussion....48
IX. Conclusion....49
Conclusion générale....50
Références Bibliographiques....51

Liste des figures

Figure I.1 : Architecture du service Web [HUBERT ET AL., 2003] 6
Figure I.2 : Cycle de vie d'un service web [DRISS, 2011] 7
Figure I.3 : Structure d'une enveloppe SOAP [TALEB, 2012 11
Figure I.4 : Structure d'un document WSDL [TAHIR, 2013] 11
Figure I.5 : Les trois facettes de l'annuaire UDDI [HUBERT ET AL., 2003] 13
Figure II.1 : Phishing visant les clients de la Société Générale, BNP Paribas, CIC et CCF [ACOCA, 2007] 22
Figure II.2 : Mixnets moyenne d'accès anonyme [SEBASTIEN, 2010] 24
Figure II.3 : P3P policy [LGMM05] 27
Figure II.4 : Exemple d'un fichier APPEL 28
Figure III.1 : Présentation du scénario de l'agence de voyage 31
Figure III.2 : Principe de la sélection 32
Figure III.3 : Matching entre deux services web 33
Figure III.4 : La notion de subsomption 34
Figure III.5 : Notre algorithme de sélection 37
Figure III.6 : Diagramme de cas d'utilisation 38
Figure III.7 : Diagramme de séquence du cas « Générer la sélection » 39
Figure III.8 : Diagramme de classes 40
Figure III.9 : Fenêtre d'authentification 42
Figure III.10 : Fenêtre principale 42
Figure III.11 : Exemple de base de données 43
Figure III.12 : Affichage de résultat d'une sélection 44
Figure III.13 : Message d'erreur 44
Figure III.14 : L'impact de la taille d'attributs sur le temps d'exécution 45
Figure III.15 : L'impact de la taille d'attributs sur le temps d'exécution (PCM) 46
Figure III.16 : L'impact de la taille de la composition sur le temps d'exécution 47
Figure III.17 : L'impact de la taille de la composition sur le temps d'exécution (PCM) 47
Figure III.18 : L'impact de la taille des services sur le temps d'exécution 48

Liste des Tableaux

Tableau III.1 : Domaine des valeurs prises par les critères de matching 34

Introduction générale

Contexte

L'internet du futur sera caractérisé par une nouvelle génération d'applications construites par la composition des services et des données provenant de différents fournisseurs et d'organisations, afin de fournir aux utilisateurs des services à valeur ajoutée adaptés à leurs besoins. Les services web jouent un rôle clé dans la réalisation de cette vision, car ils peuvent être publiés, découverts, et composés par Internet en utilisant des standards comme WSDL, UDDI, SOAP. La sélection d'une composition n'est pas simplement un regroupement quelconque de services web, mais un ensemble dont les tâches sont ordonnées en fonction des relations reliant ces services web.

Les services web sont généralement fournis par des organisations différentes et indépendamment de tout contexte d'exécution. Puisque chaque organisation possède ses propres règles de travail, les services web doivent être traités comme des unités strictement autonomes. En raison de l'augmentation du nombre de services disponibles et offrant des fonctionnalités similaires, il est difficile pour les utilisateurs de sélectionner une composition de service optimale parmi une liste de services candidats qui répondent à leurs besoins. Par conséquent, la sélection des services est un défi majeur dans l'Internet du futur.

La littérature offre une grande quantité de travaux sur la composition et la sélection des services web. Malgré ces efforts, le concept de la vie privée pose toujours des problèmes dans le composition et la sélection des services web. Généralement ces tâches impliquent la recueille d'une grande quantité de données personnelles sur les clients, et leurs partage avec les fournisseurs des services. Ceci peut entraîner des risques d'utilisation abusive des données, par exemple : un fournisseur de services peut utiliser les données des clients à des fins illicites. En conséquence, de plus en plus d'utilisateurs envisagent des pratiques de confidentialité adoptées par les fournisseurs des services web comme un facteur important pour le choix de service. Les utilisateurs des services web doivent exiger leurs propres préférences de confidentialité pour que les fournisseurs respectent les données confidentielles de ces clients.

Avec des utilisateurs qui exigent de plus en plus des mesures de protection de la vie privée, il est essentiel que les nouvelles technologies répondent à ce besoin de manière adéquate. L'une des questions de recherche les plus importantes est de savoir si les services web peuvent être exploités pour assurer une protection efficace de la vie privée.

Problématique

Au cours des dernières années, la vie privée dans les services web attire de plus en plus l'attention de la communauté de l'industrie et de la recherche. Un service web a généralement sa propre politique de confidentialité qui définit un ensemble de règles applicables à tous les utilisateurs. La vie privée des services spécifie généralement trois types de politique: la politique d'utilisation, la politique de stockage et la politique de sécurité de l'information.

Dans le cadre de ce mémoire, nous traitons le problème de la sélection des services web. Notre travail consiste à trouver une composition qui respecte les préférences des utilisateurs en terme de vie privée.

Contribution

Nous proposons dans de ce mémoire, une approche qui aide les utilisateurs et les fournisseurs des services web dans la sélection des services optimaux à l'égard de leurs préférences de confidentialité.

L'approche proposée se compose de deux étapes :

- La première étape, formalise les préférences des utilisateurs et les politiques des fournisseurs de service web.
- La deuxième étape, présente un algorithme qui détermine les compositions de services web compatibles aux préférences de confidentialité de l'utilisateur, c'est à dire les services composites dont la politique de confidentialité répond aux préférences de l'utilisateur par rapport à leurs niveaux de confidentialité.

Plan du mémoire

Ce manuscrit est composé de trois chapitres et une conclusion générale:

Chapitre I : Ce chapitre est consacré à l'étude des services web et leur composition. Tout d'abord, nous commençons par quelques définitions des services web, ainsi que leurs architectures. Ensuite, nous détaillons les différentes technologies des services web et nous abordons l'utilisation et la sélection de ces services web. Enfin, nous citons quelques domaines d'applications, les avantages et les inconvénients des services web.

Chapitre II : Ce chapitre est divisé en deux parties : dans la première partie nous présentons une vue générale sur la vie privée, pour cela nous commençons par la définition de certains aspects ayant une relation avec la vie privée, ainsi que quelques méthodes d'attaques et enfin les technologies et les langages permettant la protection de la vie privée. Dans la deuxième partie, nous présentons une synthèse des travaux de recherche proposant des approches dans le domaine de la protection des données dans les services web.

Chapitre III : Ce chapitre est consacré à la conception, l'implémentation et l'expérimentation de notre approche.

Conclusion générale : Résume les résultats de notre travail, et présente les perspectives que nous souhaitons réaliser dans le futur.

Chapitre I :

Les Services Web

I. Introduction

L'accès aux systèmes d'informations s'appuie de plus en plus sur des technologies Internet. Les efforts de standardisation de ces technologies ont accentué l'engouement des personnes et des organisations (académiques, industrielles, commerciales ou institutionnelles) à l'utilisation de l'Internet et ont permis l'émergence des développements d'applications accessibles par Internet, nommées par les services web. Ainsi, les technologies associées aux services web sont devenues incontournables pour le développement des applications interagissant les unes avec les autres par le biais de l'Internet.

Par ailleurs, le service web tend à donner plus d'interactions pour permettre à deux entités hétérogènes (entreprises, clients, applications, etc.) de dialoguer à travers le réseau Internet. Ainsi, les logiciels écrits par divers langages de programmation (C, Visual Basic, Java, etc.), sur différentes plateformes (Linux, Windows, etc.) et avec plusieurs architectures peuvent employer des services web pour échanger les données à travers des réseaux informatique. Chaque service web doit pouvoir être découvert et invoqué dynamiquement par les applications.

D'autre part, l'architecture des services web s'est imposée (tout comme le langage XML) grâce à sa simplicité, à sa lisibilité et à ses fondations normalisées. L'objectif principal des services web est de faciliter l'accès aux applications entre entités et ainsi de simplifier les échanges de données. Cette interopérabilité est due à l'utilisation de normes ouvertes. L'OSI et le W3C sont les comités de coordination responsables de l'architecture et de la standardisation des services web. Pour améliorer l'interopérabilité entre les réalisateurs des services web, l'organisation WS-I a développée une série de profils pour faire évoluer les futures normes impliquées. L'aspect le plus important dans les services web est le fait qu'ils reposent sur plusieurs standards permettant ainsi la structuration de leurs architectures. Cette collection de normes et de protocoles est appelée services web Protocol Stack. Elle contient entre autre XML et SOAP pour le formatage des données, WSDL pour la description des services web et UDDI pour la recherche des services web, nécessaires au bon fonctionnement des applications.

Par conséquent, nous présentons dans ce chapitre les aspects conceptuels de la modélisation des services web et les aspects liés à leur implémentation.

II. Définition des services web

Un service web est un composant logiciel indépendant, qui rassemble un ensemble de standards web et de langages dérivés du XML. Ces derniers permettent sa publication, sa découverte et son invocation à distance. Il expose des fonctionnalités via une interface publique et permet de communiquer avec d'autres applications et services web en utilisant des messages XML transportés par des protocoles internet comme le HTTP.

Plusieurs définitions des services web ont été mises en avant par différents auteurs. Citons :

II.1. IBM

« Les services web sont la nouvelle vague des applications web. Ce sont des applications modulaires, auto-contenues et auto-descriptives qui peuvent être publiées, localisées et invoquées depuis le web. Les services web effectuent des actions allant de simples requêtes à des processus métier complexes. Une fois qu'un service web est déployé, d'autres applications (y compris des services web) peuvent le découvrir et l'invoquer » [PONGE, 2004].

II.2. W3C

« Un service web est un système logiciel identifié par un URI dont les interfaces publiques et les incarnations sont définies et décrites en XML. Sa définition peut être découverte dynamiquement par d'autres systèmes logiciels. Ces derniers peuvent ensuite interagir avec le service web en utilisant des messages XML transportés par des protocoles Internet » [RICARDO, 2004].

III. Architecture des services web

L'exposition et l'utilisation des services se fait dans un contexte particulier qui définit clairement les interactions entre le service et ses utilisateurs. Les services web communiquent via un ensemble de technologies fondamentales qui partagent une architecture commune. Ils ont été conçus pour être réalisés sur de nombreux systèmes développés et déployés de façon indépendante.

Les technologies utilisées par les services web sont : HTTP, WSDL, XML-RPC, SOAP et UDDI. L'architecture standard d'un service web comme illustre la Figure I.1 est organisée

en plusieurs couches, chacune d'elles répond à des préoccupations fonctionnelles différentes telles que la publication, la description, la messagerie et le transport.

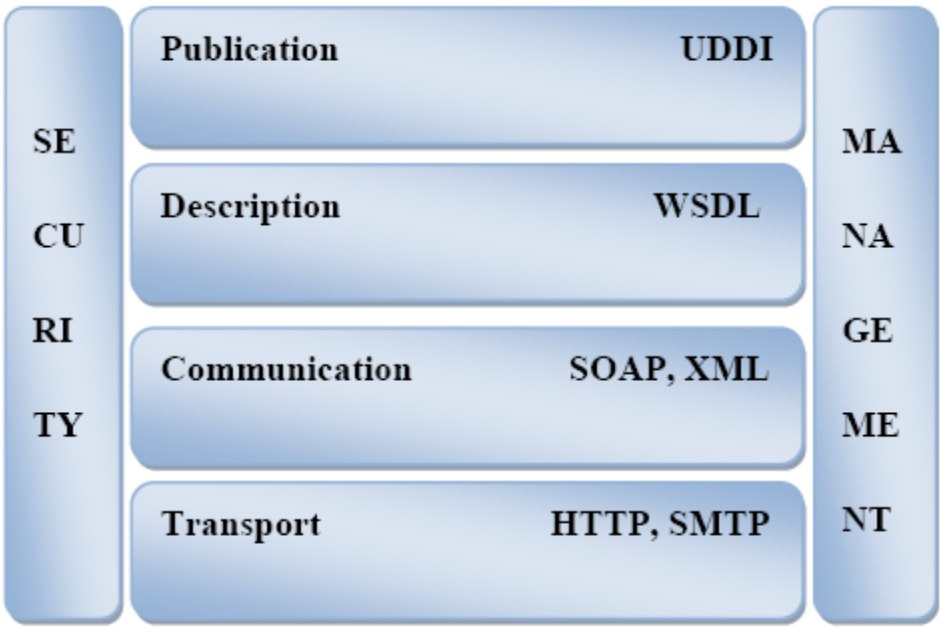


Figure I.1 : Architecture du service web [HUBERT ET AL., 2003]

Les différentes couches de l'architecture d'un service web s'interfacent avec des standards, comme suit [DEHANE, 2012]:

- **La couche de publication:** repose sur le protocole UDDI (Universal Description, Discovery and Integration) qui assure le regroupement, le stockage et la diffusion des descriptions des services web.
- **La couche description:** est prise en charge par le langage WSDL (Web Service Description Language) [CHRISTENSEN ET AL., 2001] qui décrit les fonctionnalités fournies par le service web, les messages reçus et envoyés pour chaque fonctionnalité, ainsi que le protocole utilisé pour la communication.
- **La couche communication:** la couche de communication des messages propose différents mécanismes liés à l'acheminement des messages (format de communication des messages, adressage, routage...etc.). Cette couche utilise des protocoles reposants sur le langage XML, qui résout les conflits syntaxiques lors de l'encodage des données et cela grâce à sa syntaxe unique. Actuellement SOAP (Simple Object Access Protocol) est le protocole le plus utilisé pour cette couche.

- **La couche transport:** le protocole le plus utilisé dans cette couche est l'HTTP (Hyper Text Transfert Protocol). Cependant, d'autres protocoles peuvent être utilisés, tels que le SMTP (Simple Mail Transfer Protocol) ou le FTP (File Transfer Protocol), permettant ainsi aux services web de rester indépendants du mode de transport utilisé.

IV. Cycle de vie d'un service web

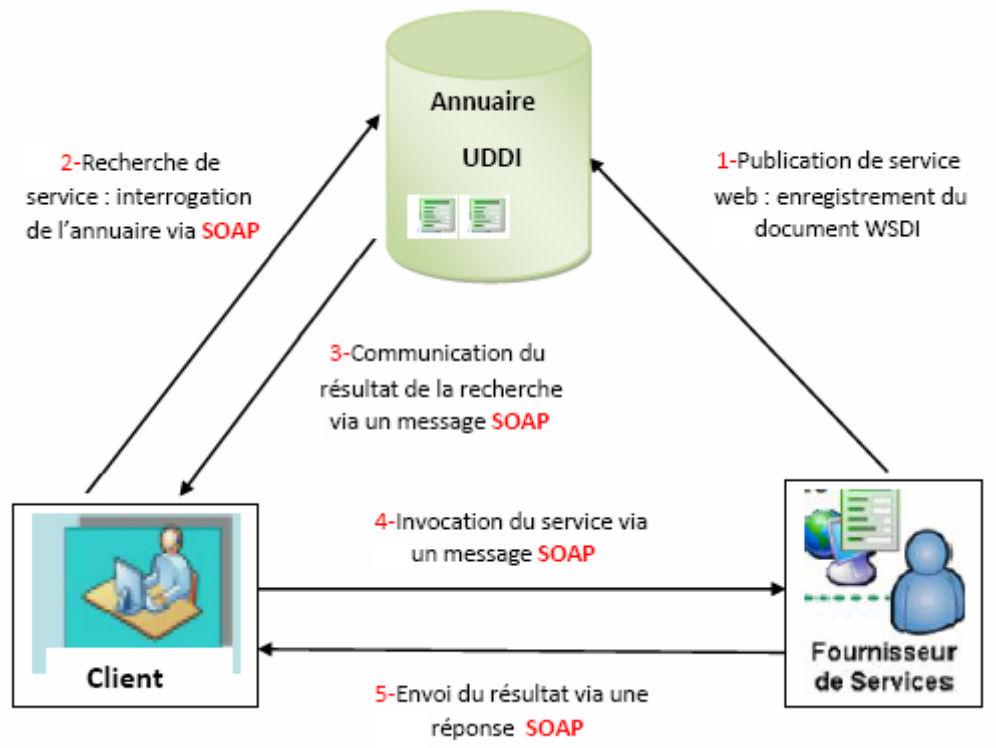


Figure I.2 : Cycle de vie d'un service web [DRISS, 2011]

Ce scénario se déroule en plusieurs étapes qui sont les suivantes [DRISS, 2011] :

1. Le fournisseur définit la description de son service dans un document WSDL et la publie dans l'annuaire UDDI.

2. Le client, désirant trouver un service interroge l'annuaire UDDI via un message SOAP.

3. L'annuaire retourne, via un message SOAP aussi une liste de services qui répondent à la requête du client. Le client n'a qu'à choisir un parmi la liste.

4. Le client récupère le document WSDL du service choisi. Ensuite, il examine ce document afin de récupérer les informations nécessaires lui permettant de se connecter au

fournisseur et d'interagir avec le service considéré. Enfin, il invoque l'opération désirée par le biais d'une requête SOAP renfermant les paramètres d'entrée de l'opération.

5. Le fournisseur de services reçoit la requête, la traite, formule la réponse SOAP et l'envoie au client.

V. Utilisation des services web

Les services web sont une nouvelle voie dans le développement des logiciels, en d'autres termes, se sont des composants logiciels qui peuvent cohabiter dans une application sur un réseau local ou Internet, et sont accessibles par des applications tiers. Les trois composants essentiels pour la fonction et l'interopérabilité de ces services web sont les suivants [CORTES, 2003].

V.1. Description des services web

Un service web a besoin d'être décrit pour autoriser des organisations de l'utiliser. WSDL est utilisé dans ce cas pour décrire un service web. La description du service web indique les fonctions de ce service, comme les paramètres entrée/sortie et le protocole de transport.

V.2. Publication et découvert des services web

Une organisation a besoin de publier les services web qu'elle possède pour les autres organisations, afin que ces dernières puissent les découvrir. Universal Description, Discovery and Intégration (UDDI) sont utilisés pour publier les services web sur un dépôt central UDDI. D'autres organisations peuvent exécuter les opérations UDDI pour accéder au dépôt UDDI et découvrir les services web qui les intéressent.

V.3. Invocation des services web

Une fois que l'organisation a découvert le service web via l'interface UDDI et qu'elle a prit la décision d'utiliser ce service dans son application, elle a besoin d'invoquer le service web. L'invocation d'un service web est faite via le protocole SOAP qui est implémenté au-dessus du protocole HTTP.

VI. Les technologies des services web

XML, SOAP, WSDL, et UDDI sont les technologies dominantes des services web. Une pléthore d'autres technologies viendront aux fils du temps pour enrichir l'architecture de ces derniers. Sans entrer dans les détails techniques, nous vous exposons dans le paragraphe suivant les grandes lignes de ces technologies.

VI.1. XML

XML (eXtensible Markup Language) est un format texte simple très flexible tiré du SGML (Standard Generalized Markup Language) [RAGGETT, 1999]. À l'origine conçu pour la publication électronique à grande échelle, XML joue aussi un rôle de plus en plus important dans l'échange d'une large variété de données sur le Web et ailleurs. W3C le recommande depuis 1998 comme standard de description de données.

XML est un méta langage qui permet d'identifier la structure d'un document. Un document est composé d'une définition de sa structure et d'un contenu. La structure d'un document XML est souvent représentée graphiquement comme un arbre. La racine du document constitue le sujet du document, et les feuilles sont les éléments de ce sujet. Grâce à sa flexibilité et extensibilité, XML est devenu rapidement le standard d'échange de données sur le web.

Les avantages du XML sont multiples. Citons :

- **Lisibilité :** il est facile pour un humain de lire un fichier XML car le code est structuré et simple à comprendre. Il est même possible de dire qu'aucune connaissance spécifique n'est nécessaire pour comprendre les données comprises à l'intérieur d'un document XML.
- **Disponibilité :** un fichier XML peut être créé à partir d'un simple logiciel de traitement de texte (exemple : bloc-note).
- **Interopérabilité :** ce langage est lu avec n'importe quel système d'exploitation. Les technologies utilisées sont transparentes lors de la lecture d'un fichier XML.
- **Extensibilité :** en fonction des besoins, des nouvelles balises peuvent être ajoutées.
- Divers parseurs XML doivent produire le même résultat, à condition qu'ils soient bien codés.

- Tous les récents navigateurs Internet intègrent un parseur XML pour lire les documents de ce langage informatique.

VI.2. SOAP

SOAP (Simple Object Access Protocol), ce qui signifie « Protocole Simple d'Accès aux Objets », consiste à faire circuler du XML via le HTTP sur le port 80. Cela induit à la facilité des communications, car ce langage est standard et l'utilisation du port 80 ne pose aucun problème pour les firewalls.

SOAP peut donc être utilisé dans tous les styles de communications : synchrone ou asynchrone, point à point ou multipoint, Intranet ou Internet [KULCHENKO, 2001].
Ce protocole est caractérisé par sa facilité d'implémentation dans les serveurs Web, destiné à l'échange d'informations dans un environnement distribué et décentralisé [GARDIEN, 2002].

SOAP utilise principalement les deux standards HTTP et XML :

- HTTP comme protocole de transport des messages SOAP. Il constitue un bon moyen de transport en raison de sa popularité sur le web.
- XML pour structurer les requêtes et les réponses, pour indiquer les paramètres des méthodes et les valeurs de retour, ainsi pour les éventuelles erreurs de traitements.

Les messages SOAP sont englobés dans une enveloppe constituée d'un entête et d'un corps [DRISS, 2011]:

- L'enveloppe (obligatoire), contient le nom du message et l'espace de nom (namespace).
- L'entête (facultatif), apporte des données supplémentaires au message SOAP, comme des informations concernant l'authentification et la gestion des transactions.
- Le corps (obligatoire) renferme :
    - du côté client, l'opération du service invoquée ainsi que les valeurs et les paramètres nécessaires à cette invocation.
    - du coté service, le résultat de l'exécution de l'opération invoquée.

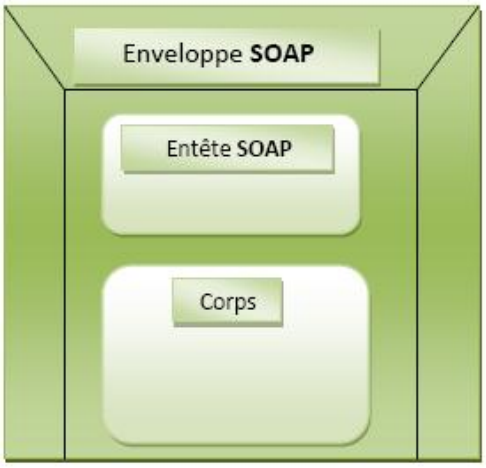


Figure I.3 : Structure d'une enveloppe SOAP [TALEB, 2012]

VI.3. WSDL

WSDL (Web Services Description Language) [W3C-WSD-GROUP, 2002] est un langage de la famille XML, il permet de décrire les types de données supportés et les fonctions offertes par un service web. L'objectif est de fournir la description en XML des services, indépendamment de la plate-forme et du langage utilisé, ainsi que sous une forme interprétable par des personnes ou des programmes.

Les interfaces IDL (Interface Definition Language) de CORBA, sont équivalentes aux descriptions WSDL. Le schéma suivant illustre la structure du langage WSDL sous la forme d'un document XML. Il décrit les relations entre les sections qui constituent un document WSDL.

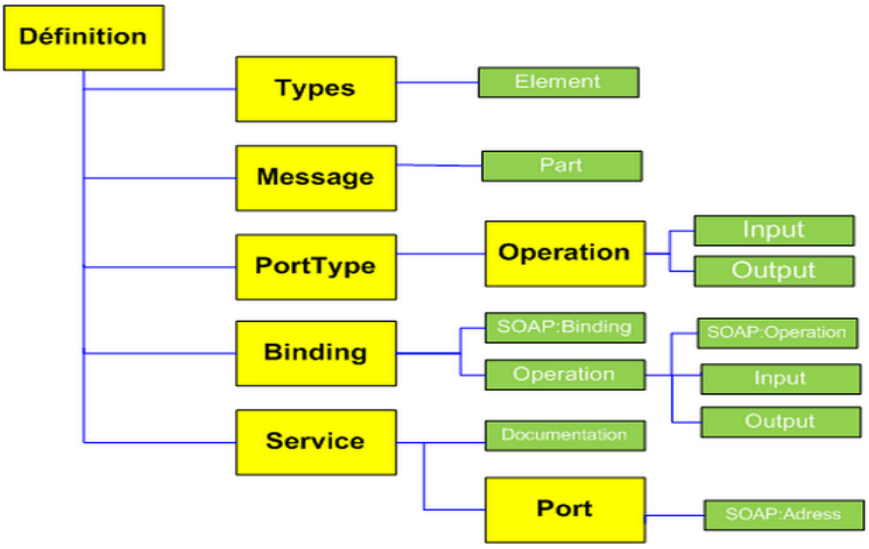


Figure I.4 : Structure d'un document WSDL [TAHIR, 2013]

Un fichier WSDL contient donc sept éléments:

- **Types** : fournissent la définition des types de données utilisés pour décrire les messages échangés.
- **Messages** : représentent une définition abstraire (noms et types) des données en cours de transmission.
- **PortTypes** : décrit un ensemble d'opérations. Chaque opération à zéro ou un message en entrée, zéro ou plusieurs messages de sortie ou d'erreurs.
- **Binding** : spécifie une liaison entre un <portType> et un protocole concret (SOAP, HTTP...etc).
- **Service** : indique les adresses des ports des différentes liaisons.
- **Port** : représente un point d'accès des services. Il est défini par une adresse réseau et une liaison.
- **Opération** : correspond à la description d'une action exposée dans le port.

VI.4. UDDI

UDDI (Universel Description, Discovery and Integration) est un standard édité par OASIS (*Organization for the Advancement of Structured Information Standards*). Il définit la structure d'un annuaire de services web, et celle de la gestion de services (publication, localisation, découverte) sous forme de répertoire. Il permet aussi de stocker les informations nécessaires pour retrouver et accéder à un service, telles que les informations techniques, l'adresse des services web, le nom de la personne/société qui gère un service donné et la description des fonctionnalités [CLEMENT ET AL., 2004].

Un service d'annuaire UDDI est un service web qui gère les méta-données des services, l'information sur les fournisseurs de services et les implémentations des services. Afin de trouver un service web, il est possible d'utiliser un annuaire UDDI en précisant des exigences concernant le service requis. On cherche le service par son nom et/ou par des mots clés.

Les informations sur un service publié dans un annuaire UDDI se présentent sous trois facettes comme illustre la figure I.5 [HUBERT ET AL., 2003] :

- Les pages blanches comprennent la liste des entreprises ainsi que des informations associées à ces dernières (coordonnées, description de l'entreprise, identifiants).

- Les pages jaunes recensent les services web de chacune des entreprises sous le standard WSDL.
- Les pages vertes fournissent des informations techniques précises sur les services fournis.

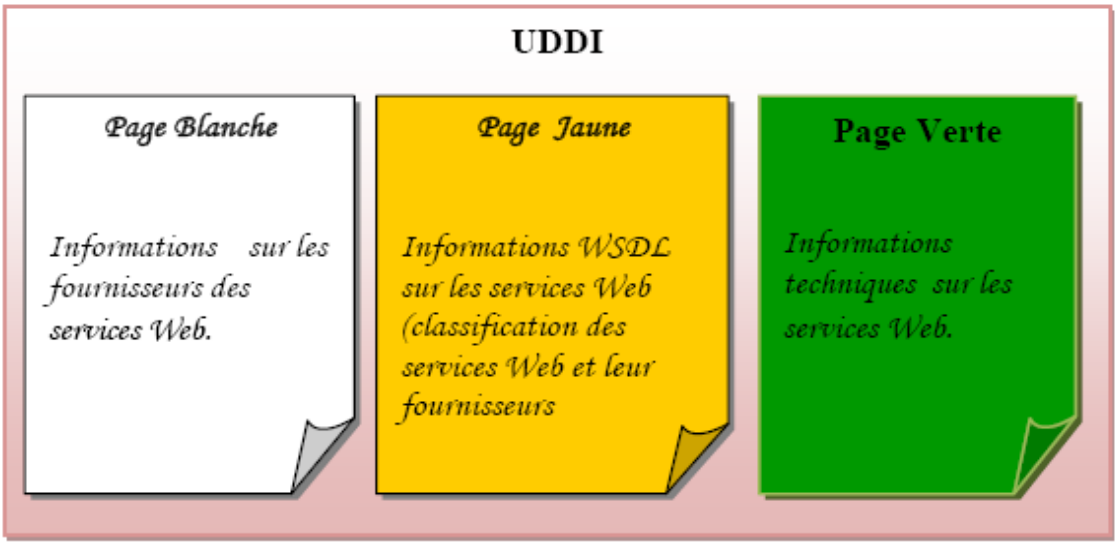


Figure I.5 : Les trois facettes de l'annuaire UDDI [HUBERT ET AL., 2003]

VII. Composition des services web

VII.1. Définitions

La composition des services web est le processus de construction de nouveaux services web à valeur ajoutée, à partir de deux ou plusieurs services web déjà présents et publiés sur le Web. En d'autres termes, la composition des services web est la combinaison des services web existants pour former de nouveaux services. L'objectif principal de la composition d'un nouveau service web est la possibilité de créer de nouvelles fonctionnalités grâce à la combinaison de plusieurs fonctionnalités offertes par d'autres services web existants. Elle implique la capacité de sélectionner, de coordonner, d'interagir et de faire inter-opérer des services web existants.

La composition de service web spécifie donc, quels services ont besoin d'être invoqués, dans quel ordre et comment gérer les conditions d'interaction [ARENAZA, 2006].

Un service web est dit composé ou composite lorsque son exécution implique des interactions avec d'autres services web, via des échanges de messages afin de faire appel à leurs fonctionnalités.

VII.2. Types de composition de services web

La composition des services web peut être soit une composition statique soit une composition dynamique :

A. **La composition statique :** dans cette approche, les services web à composer sont choisis pour concevoir l'architecture et le design. Les composants sont présélectionnés et reliés ensemble avant d'être compilés et déployés [DUSTDAR ET SCHREINER, 2005]. Ceci n'est possible que si l'environnement des services web, les partenaires d'affaires, ainsi que les composants ne changent pas ou peu. Microsoft Biztalk et Bea Weblogic sont deux exemples de moteurs de composition statique des services web [SUN ET AL., 2003]. Si les fournisseurs de services proposent d'autres services ou changent les anciens services, des incohérences peuvent avoir lieu, ce qui demanderait un changement de l'architecture du logiciel, voir même de la définition du processus et nécessiterait une nouvelle conception du système. Dans ce cas, la composition statique des services web est considérée trop restrictive, à cause de la limitation des composants à s'adapter d'une manière automatique aux changements imprévisibles [SUN ET AL., 2003].

B. **La composition dynamique :** dans cette approche, les services sont sélectionnés et composés à la volée en fonction des besoins formulés par l'utilisateur [OSMAN ET AL., 2005]. Cette approche offre le potentiel de réaliser des applications flexibles et adaptables, grâce à la sélection et à la combinaison des services de manière appropriée en se basant sur la requête et sur le contexte de l'utilisateur. Ce type de composition peut engendrer de nombreuses applications utiles qui n'ont pas été prévues à l'étape de conception. Par conséquent, la composition dynamique des services web est propice dans un environnement tel que le web et l'informatique pervasive, où les composants disponibles sont dynamiques et les attentes des utilisateurs sont variables et personnalisées.

VIII. Sélection des services web

Avec la sélection des services web, on cherche à choisir le meilleur fournisseur d'un service web, étant donné un ensemble de fournisseurs de ce service. La sélection consiste à choisir, parmi les services web découverts, ceux qui répondent au mieux aux exigences de l'utilisateur sur la base des besoins fonctionnels et/ou non fonctionnels.

Les besoins non fonctionnels des services web sont généralement exprimés à l'aide des critères de confidentialité des données. Dans une sélection de services web basée sur la confidentialité deux cas se présentent [JAEGER, 2006] :

Si la réponse à la requête d'un client exige la sélection d'un seul service web non composite, alors la sélection est très simple. Le candidat qui présente le meilleure critère de confidentialité de donné sera désigné pour répondre à la demande du client.

Si la réponse à la requête d'un client exige la combinaison de plusieurs services existants, alors la sélection dans ce cas sera plus complexe du fait qu'il faut choisir la combinaison des services composants qui répond mieux aux besoins des clients.

IX. Quelques domaines d'application des services web

Les services web ont été utilisés dans différents domaines, citons [DEHANE, 2012]:

- dans la plupart des scénarios applicatifs lorsque la communication peut être établit sur un modèle bidirectionnel (requête/réponse).
- dans les domaines du B2B, du B2C et de gestion de stock, etc.
    - B2B (Business to Business) : qualifie une application de type site internet destiné au commerce professionnel à professionnel.
    - B2C (Business to Consumer) : qualifie une application de type site internet destiné au grand public.

X. Avantages et inconvénients des services web

X.1. Avantages

La technologie des services web est populaire et couramment utilisée grâce aux avantages intéressants qu'elle offre aux utilisateurs des systèmes distribués [BEKKOUCHE, 2012] :

- Les services web réduisent le temps de mise en marché des services offerts par les diverses entreprises.
- Les services web permettent à des programmes écrits par différents langages et sur diverses plateformes, de communiquer entre eux par le biais de certaines normes.

En d'autres termes, les services web offrent une meilleure interopérabilité entre les logiciels.

- Les services web utilisent les normes et les protocoles ouverts.
- Grâce au protocole HTTP, les services web peuvent fonctionner malgré les pare-feu, sans pour autant avoir la nécessité de changer les critères de filtrage.
- Les protocoles et les modèles de données sont pratiquement sous format texte, afin de rendre plus intuitif le fonctionnement des échanges.
- Grâce aux services web, les coûts sont réduits par l'automatisation interne et externe des processus commerciaux.

X.2. Inconvénients

La technologie des services web comporte plusieurs inconvénients dont [BEKKOUCHE, 2012]:

- Problèmes de sécurité : Il est facile de contourner les mesures de sécurité mise en place par les pare-feu à cause de l'utilisation du protocole http, alors que d'autres approches sont plus mûres et plus sécurisées (exemple COBRA).
- Problèmes de performance : Les services web sont encore relativement faibles par rapport à d'autres approches de l'informatique réparties telles que COBRA ou RMI.
- Confiance : Les relations de confiance entre différentes composantes d'un service web sont difficiles à bâtir, car parfois ces mêmes composantes ne se connaissent même pas.
- Syntaxe et sémantique : On se concentre beaucoup sur comment invoquer les services (syntaxe) et pas assez sur ce qu'offrent les services web (sémantique).
- Fiabilité : Il est difficile de s'assurer de la fiabilité d'un service web, car on ne peut garantir que ses fournisseurs ainsi que les personnes qui l'invoquent travaillent d'une façon efficace.
- Disponibilité : Les services web peuvent bien satisfaire un ou plusieurs besoins du client. Seront-ils pour autant toujours disponibles et utilisables ? Ça reste un défi pour les services web.

XI. Conclusion

Nous avons vu dans ce chapitre la simplicité de la mise en œuvre des services web, qui permettent l'accès aux fonctionnalités d'une application existante via Internet, sans la modification du système d'information d'une entreprise, et cela grâce à l'ajout des briques supplémentaires.

Le protocole d'échange de messages SOAP et le langage WSDL pour la définition de l'interface, standardisent la couche de transport. Se sont des bases solides, qui ont prouvé leur maturité et efficacité.

Grâce à la grande interopérabilité de ces différents protocoles (SOAP, WSDL et UDDI), les services web avancent vers plus de normalisation. Cette standardisation assure une indépendance par rapport au système d'exploitation, à l'architecture ou au standard utilisé, pour l'implémentation des éléments essentiels d'un service web.

Après la présentation des différents protocoles des services web, il est important dans notre projet de consacrer un chapitre sur la vie privée dans ces services web. Dans la deuxième partie de notre mémoire, nous allons décrire la protection des données privées, ainsi que les moyens techniques mis en œuvres pour préserver les informations personnelles.

Chapitre II :

La vie privée dans les services web

I. Introduction

Les services web ont récemment émergés comme un moyen populaire pour la publication et le partage des données sur le web [MEDJAHED, 2003]. Les entreprises modernes à travers tous les spectres s'orientent vers une architecture basée sur la mise des bases de données derrière des services web, fournissant ainsi un bien documenté, une plate-forme indépendante et une méthode interopérable d'interaction avec leurs données.
Par ailleurs, dans notre civilisation, la protection de la vie privée est considérée comme l'une des libertés individuelles fondamentales. Dans le monde réel, cette protection repose sur des lois, sur des difficultés matérielles et sur le coût de collecte d'informations qui porteraient atteinte à la vie privée des individus. En revanche, dans le monde virtuel de l'Internet, une telle collecte est à la fois facile et peu coûteuse.

D'autre part, cette vie privée des services est considérée comme un élément à trois dimensions pour les infrastructures des services web. Ces dimensions sont : la vie privée de l'utilisateur (telle que perçue par l'utilisateur d'un service web), la confidentialité des données (soutenue au niveau des données), et la vie privée des services (comme exposée aux utilisateurs à travers un service web). L'exécution des conditions de ces trois niveaux d'intimité lance un certain nombre de défis, par exemple: l'exécution de l'intimité au taux de disponibilité, peut exiger une manipulation complexe et dynamique sur les conditions de confidentialité des données utilisées par les individus.

Par conséquent, nous présentons dans ce chapitre des notions de la vie privée et la protection des données personnelles, telles qu'elles sont considérées dans le domaine de l'informatique. Nous étudions ainsi, la nature de ces concepts et les moyens techniques qui leur sont rattachés. Nous finissons par une synthèse des travaux de recherche, qui proposent un certains nombres d'approches de la vie privée dans les services web.

II. Définition

Nous définissons d'une façon générale dans cette partie, les concepts de la vie privée, la notion de la vie privée sur Internet, ainsi que la surveillance et la sécurité.

II.1. La vie privée (Sphère privée)

La vie privée est l'ensemble des activités d'une personne qui relève de son intimité par opposition à la vie publique. Le droit au respect de la vie privée est proclamé par la loi [GANGER, 2010]. La sphère privée d'un individu est l'ensemble de ses informations qu'il les considère comme sensibles et donc dignes d'être protégées. Cette sphère est personnelle (l'individu est le propriétaire des informations qu'elle contient), personnalisable (l'individu décide des informations qu'elle contient), dynamique (les informations peuvent y être ajoutées ou en être retirées) et dépendante du contexte (les informations qu'elle contient peuvent en nature et en nombre, dépendre du temps, des activités de l'individu ou d'autres paramètres) [BEKARA, 2012].

II.2. Droit à la vie privée

Le droit à la vie privée d'un individu est sa prétention aux caractères personnel, personnalisable, dynamique et contextuel de sa sphère privée ainsi qu'au contrôle de la diffusion, de l'utilisation et de la conservation des informations contenues dans sa sphère privée quelles que soient leur représentation et la localisation de cette dernière. Nous incluons ainsi explicitement dans le droit à la vie privée la notion de propriété des données, fortement liée à celle du contrôle [BEKARA, 2012].

II.3. Protection de la vie privée (ou de la sphère)

La protection de la vie privée est l'ensemble des mesures techniques visant à assurer le respect du droit à la vie privée [BEKARA, 2012].

II.4. La vie privée sur Internet

La vie privée sur Internet est différente de la vie privée normale, du fait qu'il soit un réseau international et que n'importe quelle personne dans ce monde peut y accéder. Internet n'est pas confidentielle à 100%, les pirates informatique peuvent casser les barrières de sécurité et accéder aux informations privées (e-mail, compte bancaire…etc.),

ou nous pouvons même être espionné par l'état créateur d'Internet (USA), comme c'était révélé récemment (scandale des écoutes et d'espionnage sur l'Europe par les Etats unis).

Les communications électroniques constituent aussi des données à caractère personnel. Ces données peuvent correspondre à toutes informations relatives à une personne physique identifiée, ou qui peut être identifiée directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres [ARTICLE 1, 2004].

II.5. Surveillance

La loi luxembourgeoise sur la protection des personnes à l'égard du traitement des données à caractère personnel du 2 août 2002, définit la surveillance comme « toute activité opérée au moyen d'instruments techniques et consiste en: l'observation, la collecte ou l'enregistrement de manière non occasionnelle des données à caractère personnel d'une ou de plusieurs personnes, relatives à des comportements, des mouvements, des communications ou à l'utilisation d'appareils électroniques et informatisés » [LAURENT, 2009].

II.6. Sécurité

La sécurité est l'état de tranquillité et de confiance sentit par une personne. C'est le sentiment bien ou mal fondé, ainsi que d'être à l'abri de tout danger et risque. Ce sentiment associe calme, confiance, quiétude, sérénité, tranquillité, assurance et sûreté.
La sécurité informatique, d'une manière générale consiste à assurer que les ressources matérielles ou logicielles d'une organisation soient uniquement utilisées dans le cadre prévu [LAURENT, 2009].

La sécurité informatique vise généralement cinq principaux objectifs [LAURENT, 2009]:

- L'intégrité : consiste à garantir que les données soient bien celles que l'on pense.
- La confidentialité : consiste à assurer l'accès aux ressources échangées que par les personnes autorisées.
- La disponibilité : permet de maintenir le bon fonctionnement du système d'information.
- Le non répudiation : permet de garantir qu'aucune transaction ne soit niée.

- L'authentification : consiste à assurer l'accès aux ressources que par les personnes autorisées.

III. Les attaques de la vie privée

La vie privée sur Internet est devenue moins sécurisée à cause des nombreux hackers, ce qui expose les informations personnelles à des attaques récurrentes. Si les moyens de sécurisation sont insuffisants, l'identité d'un individu sera facilement volée et utilisée par une tierce pour des fins non légitimes.
Deux exemples d'attaques les plus couramment utilisées contre la vie privée des individus seront présentés les paragraphes suivants.

III.1. Vol d'identité

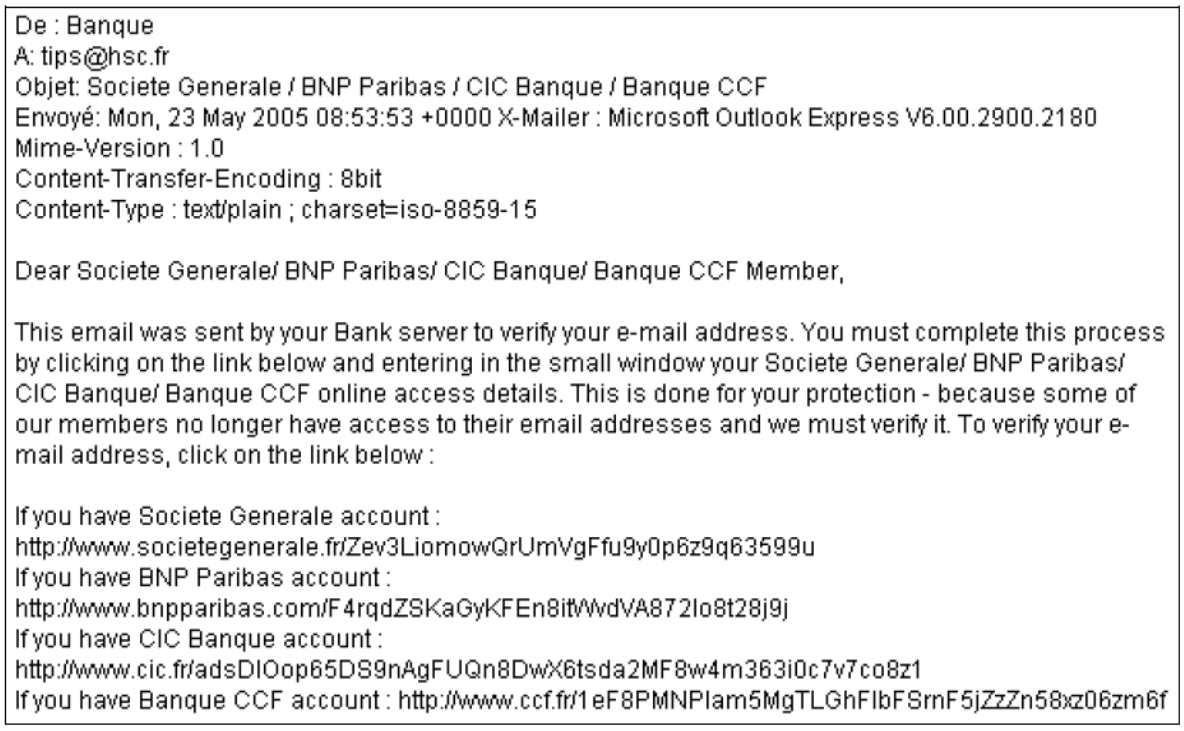
Le vol d'identité consiste en l'acquisition, la possession, le transfert ou l'utilisation non autorisées des informations personnelles d'une personne physique ou morale dans l'intention de commettre (ou en relation avec) des actes frauduleux ou autres délits [ACOCA, 2007]. Cette définition s'applique sur n'importe quel support utilisé pour que le vol d'identité soit perpétré.

L'usurpation d'identité débute toujours par la collecte des renseignements personnels sur l'individu à frauder. Ces renseignements peuvent être: le nom, le numéro de téléphone, la date de naissance, l'adresse, le numéro d'assurance sociale ou toutes autres informations permettant d'identifier la personne. La victime de l'usurpation d'identité possède la faculté de défendre ses droits. Les usurpateurs utilisent ensuite ces informations pour effectuer une ou des transactions en simulant l'identité de la personne fraudée. Par exemple, un fraudeur peut effectuer des appels téléphoniques ou faire des achats importants moyennant l'argent du compte bancaire de cet individu fraudé [BEAULIEU, 2003].

III.2. L'hameçonnage (phishing)

L'hameçonnage (phishing) est un moyen avec lequel les voleurs leurrent les internautes en leur envoyant des messages électroniques trompeurs souvent sous format de courriels diffusés en masse (spam), dans le but d'installer des logiciels malveillants dans les ordinateurs de leurs destinataires, ou via de faux sites Internet pour les amener par la ruse à révéler leurs informations personnelles [ACOCA, 2007].

L'hameçonnage ou phishing est décrit en général, comme une méthode de tromperie que les voleurs utilisent pour «pêcher» les informations sur les identités personnelles des utilisateurs peu méfiants d'Internet, moyennant des sites Internet miroirs et des messages électroniques revêtant l'apparence des messages venant d'entreprises légitimes; telles que les établissements financiers ou les administrations publiques. Comme le montre la FigureII.1, un exemple bien connu de phishing est le courriel prétendant d'être envoyé par une banque dont le destinataire est client pour vérifier ses identifiants. En France, par exemple en 2005, une attaque d'hameçonnage a visée en même temps les clients de quatre banques.



De : Banque
A: tips@hsc.fr
Objet: Societe Generale / BNP Paribas / CIC Banque / Banque CCF
Envoyé: Mon, 23 May 2005 08:53:53 +0000 X-Mailer : Microsoft Outlook Express V6.00.2900.2180
Mime-Version : 1.0
Content-Transfer-Encoding : 8bit
Content-Type : text/plain ; charset=iso-8859-15

Dear Societe Generale/ BNP Paribas/ CIC Banque/ Banque CCF Member,

This email was sent by your Bank server to verify your e-mail address. You must complete this process by clicking on the link below and entering in the small window your Societe Generale/ BNP Paribas/ CIC Banque/ Banque CCF online access details. This is done for your protection - because some of our members no longer have access to their email addresses and we must verify it. To verify your e-mail address, click on the link below :

If you have Societe Generale account :
http://www.societegenerale.fr/Zev3LiomowQrUmVgFfu9y0p6z9q63599u
If you have BNP Paribas account :
http://www.bnpparibas.com/F4rqdZSKaGyKFEn8itWvdVA872Io8t28j9j
If you have CIC Banque account :
http://www.cic.fr/adsDIOop65DS9nAgFUQn8DwX6tsda2MF8w4m363i0c7v7co8z1
If you have Banque CCF account : http://www.ccf.fr/1eF8PMNPIam5MgTLGhFIbFSrnF5jZzZn58xz06zm6f

Figure II.1 : Phishing visant les clients de la Société Générale, BNP Paribas, CIC et CCF [ACOCA, 2007]

IV. Technologies de protection de la vie privée

IV.1. Privacy by design

Le concept de « Privacy by Design »: est un remède à l'insuffisance des moyens actuels de protection de la vie privée. Il est aussi une démarche indispensable, car c'est très difficile d'améliorer la protection de la vie privée dans des systèmes qui n'ont pas été conçus avec cette exigence et cela suscite de plus en plus de travaux de recherche en informatique.

Cette protection intégrée de la vie privée repose sur sept principes fondamentaux [VINCENT, 2012] :

- ❖ Prendre des mesures proactives et non réactives, des mesures préventives et non correctives.
- ❖ Assurer la protection implicite de la vie privée.
- ❖ Intégrer la protection de la vie privée dans la conception des systèmes et des pratiques.
- ❖ Assurer une fonctionnalité intégrale selon un paradigme à somme positive et non à somme nulle.
- ❖ Assurer la sécurité de bout en bout, pendant toute la période de conservation des renseignements.
- ❖ Assurer la visibilité et la transparence.
- ❖ Respecter la vie privée des utilisateurs.

Parmi les exemples de «Privacy by Design», on peut citer [METAYER, 2010] :

- ✓ La conception des systèmes de péage routier reposant sur la géolocalisation des véhicules tout en préservant la vie privée des conducteurs, et cela grâce au calcul tarifaire effectué par le matériel de bord du véhicule.

- ✓ La conception des cartes d'identité blanches ou d'accréditation anonymes, permettant de prouver de nombreux statuts et attributs; comme le fait d'être membre d'une association, d'être titulaire d'un permis de conduire, d'avoir le droit de séjour, ou le droit de vote, etc., sans pour autant divulguer l'identité personnelle.

- ✓ La conception des systèmes permettant de mieux protéger le consentement des individus avant toute divulgation de données personnelles, et cela à travers l'assistance d'agents logiciels mettant en œuvre des politiques décidées au préalable.

- ✓ La conception d'architectures permettant d'éviter la «pollution de données» et de garder la maîtrise des informations personnelles des individus.

IV.2. Privacy Enhancing Technologies (PET)

PET est un ensemble de techniques et d'applications qui permettent à un individu de protéger ses informations personnelles pendant qu'il est en ligne. Ces technologies

regroupent un très grand nombre d'outils, mais ceux-ci demeurent complexes, peu standardisés et au final très peu utilisés. Pour qu'ils le deviennent, il leur faut répondre aux attentes de commodité qu'expriment les utilisateurs, se standardiser et se répandre très largement. [GAMBS, 2010], **[**GROUPE, 2009**].**

IV.2.1. Les systèmes de communications et accès anonymes

Les PETs permettent aux utilisateurs de communiquer de manière anonyme dans un réseau, c'est à dire ils protégent l'identité de l'envoyeur et/ou du receveur du message. Exemples : Mixnets, Crowds, Tor etc.

A. Mixnets : Concept introduit par Chaum en 1981 pour empêcher l'analyse de trafic. Le Mix est un routeur qui cache le lien entre les messages entrants et sortants, par un mécanisme de chiffrement et de permutation des messages, dans le but de faire face aux espions observant les communications échangées. Parmi ceux qui ont appliqué le Mixnets, le service de courriel anonyme (Mixmaster) [SEBASTIEN, 2010].

Fonctionnement d'un Mix simple :

1. Reçoit en entrée plusieurs paires du type (message; adresse du destinataire) qui ont été préalablement chiffrées.
2. Déchiffre les messages.
3. Envoie en sortie les messages à leurs destinataires correspondants (possiblement chiffrés).

La figure suivant montre le fonctionnement d'un Mixnets :

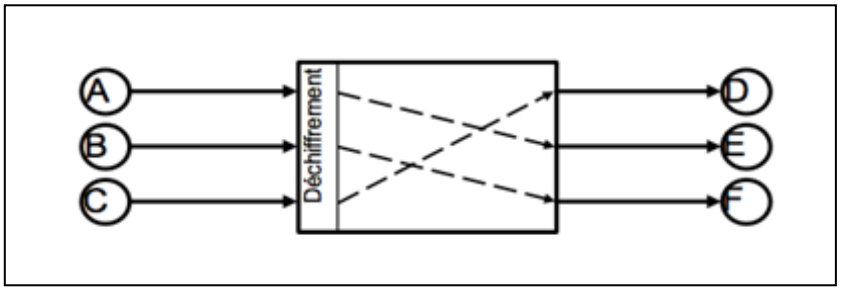


Figure II.2 : Mixnets moyenne d'accès anonyme [SEBASTIEN, 2010]

B. Crowd : Protocole de communication anonyme qui protège l'anonymat de l'envoyeur d'un message en le routant de manière aléatoire vers des groupes d'utilisateurs similaires. L'idée principale est de cacher l'origine d'un message en le dispersant. Son

fonctionnement est le suivant : chaque nouvel utilisateur s'enregistre en tant que membre d'un groupe (appelé « Crowd ») en contactant le responsable du groupe. Quand un utilisateur rejoint un groupe, tous les membres de ce dernier en sont notifiés. Le responsable du groupe est aussi chargé de la distribution des clés symétriques assurant la confidentialité entre paires de nœuds [SEBASTIEN, 2010].

C. Tor : The Onion Router ou Tor (le routage en oignon) est un réseau mondial décentralisé et organisé en couches, dont la tâche est de transmettre de manière anonyme les paquets TCP. Tout échange Internet basé sur le TCP peut être anonyme en utilisant le Tor [SEBASTIEN, 2010].

Tor fonctionne avec de nombreuses applications comme les navigateurs web, les clients de messagerie instantanée, les connexions à distance et tout un nombre d'applications basées sur le protocole TCP **[**GUINAULT, 2008].

IV.2.2. Les systèmes de gestion d'identités

Les divers usages du réseau Internet ont fait naître un peu partout dans le monde des comportements atypiques, tels que la multiplication des adresses électroniques, le recours aux pseudonymes dans les blogs et aux avatars dans les mondes virtuels, etc. Ces « identités multiples » sont plus difficiles à saisir qu'un numéro de passeport, de sécurité sociale ou de compte bancaire **[**BENGHOZI, 2008**]**. Parmi les systèmes de gestion d'identités, nous citons : Windows Live ID, Single Sign-On (SSO) et OpenID.

V. Les langages de protection de la vie privée

La protection de la vie privée doit offrir aux personnes la possibilité de contrôler la façon dont leurs données sont gérées et utilisées par une organisation particulière. Pour cela de nombreux organismes proposent des solutions. Nous présenterons dans les paragraphes suivants un langage très utilisé, il s'agit du P3P pour l'expression des la vie privée et APPEL pour la spécification des préférences de l'utilisateur.

V.1. Le Langage d'expression de politique de la vie privée

C'est un langage parmi les langages d'expression de politique de la vie privée, qui permet aux sites web d'informer les utilisateurs de leurs politiques vis-à-vis du respect de la vie privée.

V.1.1. Platform for Privacy Preferences (P3P) [LGMM ,2005]

La plateforme P3P est une combinaison de protocoles et d'architectures conçues pour informer les utilisateurs web des usages en matière de collecte de données pour les sites web qu'ils visitent. Le but est de donner aux utilisateurs un meilleur contrôle sur leurs informations personnelles. P3P a été développée par le W3C et elle a été adoptée officiellement le 16 Avril 2002. Pour établir leur propre politique de la vie privée, les utilisateurs doivent remplir un questionnaire standardisé à choix multiples dont les réponses sont compilées dans un document XML (les préférences de sa vie privée).

Les navigateurs peuvent ainsi facilement décrypter cette déclaration sur la vie privée, ce qui leur permet de bloquer automatiquement les actions pour lesquelles les internautes ont préalablement établi leur désaccord. Cette pratique a pour but d'établir une relation de confiance entre les usagers et Internet. P3P est déjà intégré dans plusieurs navigateurs (Internet Explorer, FireFox et Opera).

A. Objectifs du protocole P3P : Les deux principaux objectifs de la plateforme P3P sont les suivants:

- Permettre aux sites web de présenter leurs usages en matière de collecte de données d'une manière standardisée, lisible par l'ordinateur et facile à localiser.
- Permettre aux utilisateurs du web de savoir quelles données seront collectées par quels sites, comment ces données seront utilisées et quelles données et utilisations peuvent ils autoriser ou refuser.

B. La spécification P3P : La spécification P3P définit:

- Un fichier Référence écrit en XML qui fait correspondre à un ensemble de pages web leur politique de la vie privée.
- Un ou plusieurs fichiers politiques écrit en XML qui définissent la politique de la vie privée.
- Un fichier de préférence codé en langage APPEL (lui-même basé sur XML) pour exprimer les préférences de la vie privée de l'utilisateur.

C. Les Politiques de la vie privée : Les politiques P3P sont des documents XML qui définissent les pratiques de la vie privée d'un site web. Les informations de ce fichier sont

plus faciles à déchiffrer par la machine que par l'être humain. La figure ci-dessous montre un exemple d'un fichier politique.

```
<POLICIES xmlns="http://www.w3.org  /2002/01/P3Pv1">
<POLICY d1scur1="http://p3pbook.com/privacy.html"  name="policy">
   <ENTITY>
     <DATA-GROUP>
      <DATA ref="#business.contact-info.online.email"> privacy@p3pbook.com </DATA>
      <DATA ref="#business.contact-info.online.uri">http://p3pbook.com</DATA>
      <DATA ref="#business.name">Web Privacy With P3P<DATA>
     </DATA-GROUP>
  </ENTITY>
 <ACCESS><nonident/></ACCESS>
 <STATEMENT>
     <CONSEQUENCE> Our Web server collects access logs containing this information. </CONSEQUENCE>
     <PURPOSE><admin/><current/><develop/></PURPOSE>
     <RECIPIENT><ours/></RECIPIENT>
     <RETENTION><indefinitely/><RETENTION>
     <DATA-GROUP>
       <DATA ref="#dynamic.clickstrean"/>
       <DATA ref="#dynamic.http"/>
       </DATA-GROUP>
  </STATEMENT>
</POLICY>
</POLICIES>
```

Figure II.3 : P3P policy [LGMM,2005]

Ce fichier exemple présente bien quelques éléments et attributs utilisés dans un document XML d'une politique P3P.

V.2. Le Langage de préférence en termes de vie privée

Pour désigner les préférences de l'utilisateur, les concepteurs de P3P ont choisi d'utiliser une syntaxe différente de la spécification de la politique. Il s'agit du langage APPEL (A P3P Preference Exchange Language).

V.2.1. APPEL

C'est un langage proche de P3P, défini aussi par le w3c. Il vise à fournir un moyen aux internautes pour décrire leurs préférences personnelles en matière d'utilisation de leurs données. Ces préférences sont décrites en langage XML, qui ne peut être écrit par un

utilisateur basique, car ce fichier XML contient des connecteurs logiques, des expressions régulières, des balises et des attributs bien définis. Par contre, il est possible d'importer des préférences par défaut répondant aux attentes de la plupart des utilisateurs, ou bien c'est à eux mêmes de définir leurs propres fichiers APPEL **[**VINCENT, 2003].

Les préférences en termes de protection des données exprimées par APPEL, correspondent principalement à un ensemble de règles (chacune est exprimée par l'élément RULE) contenues dans un élément RULESET. Chaque élément RULE contient un attribut «body», un attribut «behaviour» et optionnellement un attribut «description» qui fournit une explication simple sur la règle [BEKARA, 2012].

La figure II.4 illustre la formulation des préférences utilisateur dans une politique APPEL :

```
<appel :RULE behavior="request">
      <p3p :POLICY>
            <p3p :STATEMENT>
                  <p3p :RECIPIENT appel :connective="and">
                        <p3p :ours/>
                  </p3p :RECIPIENT>
                  <p3p :PURPOSE appel :connective="non-and">
                        <p3p :tailoring/>
                  </p3p :PURPOSE>
                  <p3p :DATA-GROUP>
                        <p3p :DATA ref="#dynamic.cookies">
                              <p3p :CATEGORIES appel :connective="or">
                                    <state/>
                              </p3p :CATEGORIES>
                        </p3p :DATA>
                  </p3p :DATA-GROUP>
            </p3p :STATEMENT>
      </p3p :POLICY>
</appel :RULE>
```

Figure II.4 : Exemple d'un fichier APPEL

VI. Synthèse des travaux de recherche

Nombreux sont les travaux de recherche qui s'intéressent à la protection de la vie privée dans les services web. Dans cette partie nous présentons quelques propositions avec leurs approches et leurs principes, les plus couramment utilisés pour la protection de la vie privée.

Dans [WEIXU, 2006], les auteurs proposent une structure (qui s'inspire aussi de P3P) de la vie privée avec une composition de services web, qui repose sur le mode de fonctionnement suivant: Le client demande un modèle du service fournisseur. Ce dernier résume la manière dont il utilise les données du client (la politique). Le client vérifie par des techniques automatiques ses exigences (ses préférences). Et s'il y a une violation, il demande au fournisseur de s'adapter en lui envoyant en retour des obligations à respecter. Le fournisseur doit s'adapter dynamiquement aux nouvelles exigences.

Dans [SALAH, 2013], les auteurs proposent un modèle formel de la vie privée (étendu par des capacités de confidentialité pour les services web), qui permet à un service de définir une politique et un ensemble d'exigences (des préférences) de confidentialité. Ensuite, ils proposent un algorithme appelé PCM (Privacy Compatibility Matching) pour vérifier la compatibilité de la vie privée dans une composition, entre les différents services. La compatibilité correspondante est basée sur la notion de sous consommation de la vie privée et sur un modèle de coût. Un seuil est mis en place par les services pour répondre à la compatibilité de la vie privée partielle et totale. Ainsi, Ils proposent une approche de négociation pour aborder les incompatibilités entre les politiques de la vie privée et les exigences.

Dans [Elisa, 2013], les auteurs proposent une approche basée sur les buts pour aider les utilisateurs et les fournisseurs de services web dans la composition. Cette approche permet de vérifier la conformité entre les besoins de la vie privée des utilisateurs et les politiques de confidentialité des fournisseurs de services, tout en sélectionnant les services optimaux d'un point de vue préférences de confidentialité. Son principe est basé, d'un coté sur un modèle qui exprime les politiques de confidentialité des fournisseurs de services web et les préférences de confidentialité des utilisateurs à plusieurs dimensions, citons : sensibilité, but, période de rétention et visibilité. Et d'un autre coté sur un algorithme -*Service Composition*- de composition de services web fusionnés en une seule étape. La plupart des approches existantes exécutent le choix des services satisfaisant les besoins fonctionnels des utilisateurs et la sélection des services conformes à leurs besoins de la vie privée en deux étapes séparées.

Dans [ABDELMOUNAAM, 2002], L'objectif des auteurs est d'établir la faisabilité et la fiabilité prouvable, basées sur une technologie qui préserve les solutions des infrastructures de service web. Pour cela, ils présentent une architecture nommée : Digital

government (DG), qui fournit un ensemble de mécanismes coopérant pour préserver la vie privée des citoyens. La conception est basée sur les informations d'identification de la vie privée numérique, sur les agents de conservation de la vie privée mobile et sur les filtres des données. Ces derniers sont utilisés pour contrôler l'accès aux informations stockées dans les bases des données de différents organismes gouvernementaux. Une information privée locale est livrée à une entité distante par un agent mobile, cet agent fonctionne comme un hôte d'interrogation et applique les politiques de confidentialités locaux et globaux lorsque l'information qui lui est associée est accessible par l'entité distante.

Dans [GERMOUCHE, 2007] les auteurs fournissent un formalisme de P3P transporté pour exposer des politiques et des préférences dans les services web. Ils débutent par compléter la notion des protocoles business temporisés, c'est-à-dire par spécifier la séquence des messages échangés, en lui associant la notion de la vie privée. Ensuite, ils proposent des algorithmes pour décider si un service web peut remplacer un autre en respectant les contraintes de la vie privée (politiques et préférences).

VII. Conclusion

Au cours de ces dernières années, les problèmes liés à la vie privée dans les services web sont devenus très importants aux yeux des utilisateurs. Les usagers devraient savoir que tous les outils n'offrent pas des moyens efficaces pour protéger la vie privée, et cela est du généralement à l'incapacité d'aborder cette protection une fois que les données sont collectées.

Dans le chapitre suivant, nous mettons en pratique les notions de la protection de la vie privée des utilisateurs de services web, présentés dans ce chapitre, ainsi que dans le premier. A partir d'un exemple de composition de services web, nous allons choisir et sélectionner que ceux qui répondent aux préférences des utilisateurs.

Chapitre III :

Conception et Implémentation du prototype

I. Introduction

Notre travail consiste à développer un système de sélection des services web en respectant la vie privée des utilisateurs.

Nous allons commencer par présenter notre approche, le prototype de l'application, ensuite les différents outils techniques liés à l'implémentation, et enfin nous présentons quelques expérimentations.

II. Scénario de motivation

Dans cette partie, nous présentons un scénario dont l'objectif est d'illustrer l'intérêt de notre approche pour la sélection des services web, en prenant comme système d'information celui d'une agence de voyages. Le processus de réservation fait appel à plusieurs services selon le besoin, citons : un service pour la réservation du vol, un service pour la réservation d'hôtel, un service pour la location de voiture (ou de taxi) et un service pour le paiement. Le scénario est illustré par la Figure III.1.

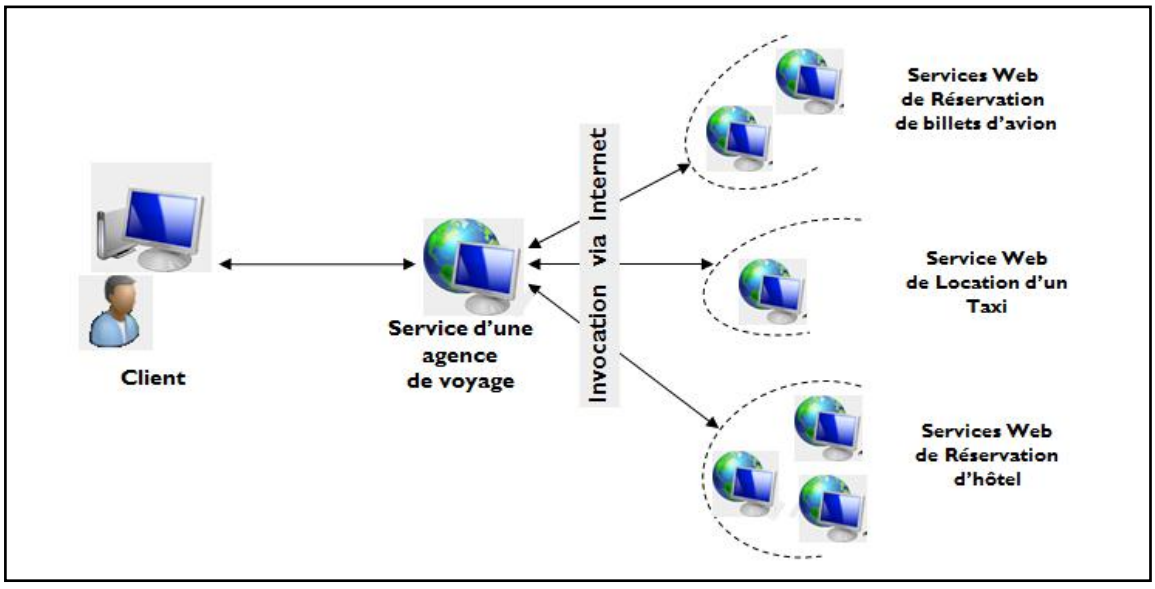


Figure III.1 : Présentation du scénario de l'agence de voyage

Le service « Agence de voyages » reçoit les informations nécessaires pour l'organisation et pour l'élaboration du voyage (Name, Numéro de Sécurité Sociale NSS, E-mail, CreditCard, …etc.). L'agence de voyage collabore dans le respect des préférences des clients, avec des compagnies aériennes, des hôtels, des agences de location de voitures et des banques, pour effectuer les réservations voulues par les utilisateurs. À la fin de la

réservation, les clients reçoivent des confirmations de la part de cette agence. La question qui se pose maintenant, comment l'agence de voyage peut trouver les sélections optimales à partir d'une composition de services web qui respectent les préférences des utilisateurs d'un coté et celles de chaque fournisseur de service (réservation de billets d'avion, location d'un taxi, …etc.) d'un autre coté.

III. Approche proposée

Le mode de sélection sur lequel est basé notre approche est représenté dans l'architecture de la figure III.2. Cette architecture est organisée en plusieurs ensembles de compositions, nommés classes, chaque composition contient un ensemble de services web dans lesquels une politique et une préférence sont définies.

L'exemple ci-dessous illustre notre approche pour la sélection des services web des différentes classes (C2, C3), qui respectent les préférences en termes de vie privée du service web1 (WS1) de la classe 1 (C1).

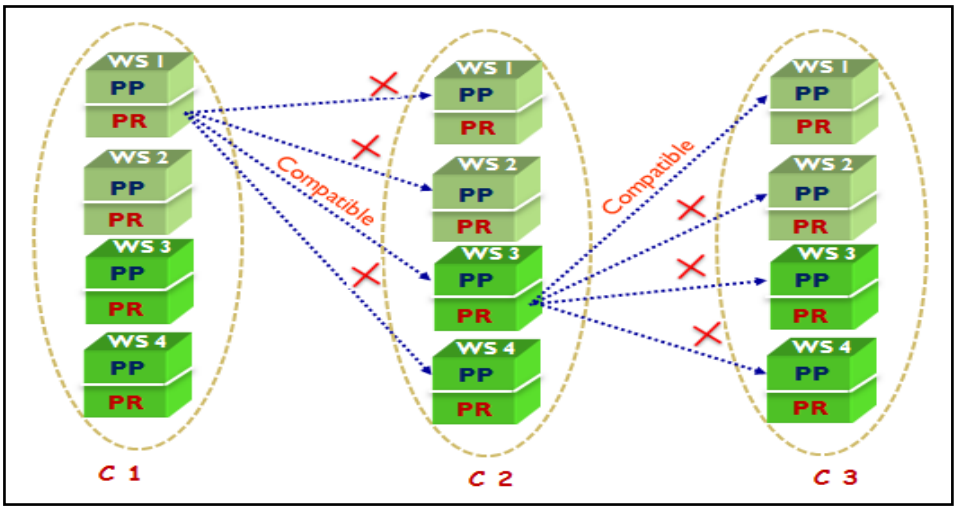


Figure III.2 : Principe de la sélection

Les règles de confidentialité, le processus de matching, ainsi que l'algorithme proposé pour la sélection seront décrits dans les paragraphes suivants.

III.1. Les règles de confidentialité

Chaque service web a une politique et une préférence de confidentialité.

- **Politiques de confidentialité:** Les politiques privées notées PP (en anglais Privacy Policy) sont l'ensemble des règles spécifiant comment un fournisseur utilise les données privées de ses clients.

- **Préférences de confidentialité:** Les préférences privées notées PR (en anglais Privacy Preference) sont l'ensemble des règles indiquant comment un client souhaite que le fournisseur de service utilise ses données privées.

Remarque : Le fournisseur de service peut prendre le rôle d'un service client, et donc envoyer les données privées collectées à un service tiers.

III.2. Matching entre services

Le matching (appariement) est un ensemble de mécanismes permettant d'identifier des correspondances entre certaines propriétés de services web, tout en utilisant un ensemble de critères. Ceci peut être réalisé grâce à une ou plusieurs méthodes de comparaison. Dans notre travail nous avons opté pour la subsomption privée (voir section III.4).

Comme le représente la figure III.3, un mécanisme de matching entre deux services web S_1 et S_2 repose sur la compatibilité entre la politique de confidentialité de S_2 et la préférence de confidentialité de S_1.

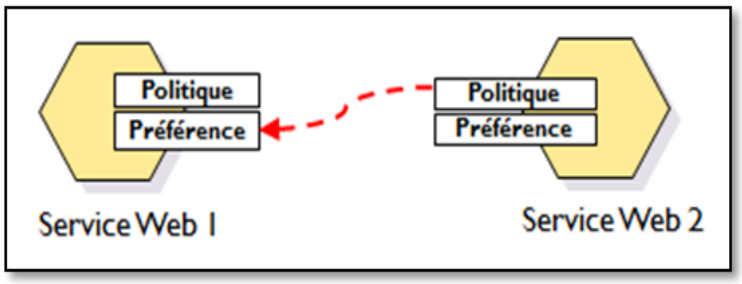


Figure III.3 : Matching entre deux services web

Pour que la politique de S_2 soit compatible avec la préférence de S_1, on applique un test de subsomption entre les deux services (Voir la figure III.4).
En détaillant les critères de matching et grâce à un exemple de subsomption privée, nous allons présenter donc les notions utilisées pour l'élaboration de notre projet.

III.2.1. Critères de matching

Les critères de confidentialité d'un service composite peuvent être définis comme suit :

- **La Durée:** est la période du temps pour laquelle le fournisseur de services garde les données, qui peuvent être consultées ou manipulées par les clients utilisant ces services. L'unité principale de la durée est le «**Jour**».

- **Le Partage:** est la mise à disposition des données aux organismes référencés.
- **La Profondeur:** est le nombre d'intermédiaire accepté pour le partage des données.
- **Le But:** est la raison pour laquelle les données sont sauvegardées ou stockées.
- **La Sensibilité** : est le degré ou le niveau de la confidentialité des informations dont l'accès est limité qu'aux organismes référencés.

Les données du fichier sont générées aléatoirement. Le tableau III.1 représente un exemple sur les domaines des valeurs prises par les cinq paramètres précédents :

Critères	Domaines
Durée	0, 1, …, indéfinie (jour)
Partage	Public, Gouvernement, Université, Laboratoire, Hôpital
Profondeur	0, 1, …, indéfinie
But	Requête, Statistique, Indéfinie
Sensibilité	Obligatoire, Optionnel

Tableau III.1 : Domaine des valeurs prises par les critères de confidentialité

III.2.2. Subsomption de confidentialité

Les valeurs de chaque critère sont liées par une relation de généralisation. Par exemple, considérons le domaine de partage {public, gouvernement, université, laboratoire, hôpital}, la figure III.4 illustre la notion de subsomption entre les entités de ce domaine.

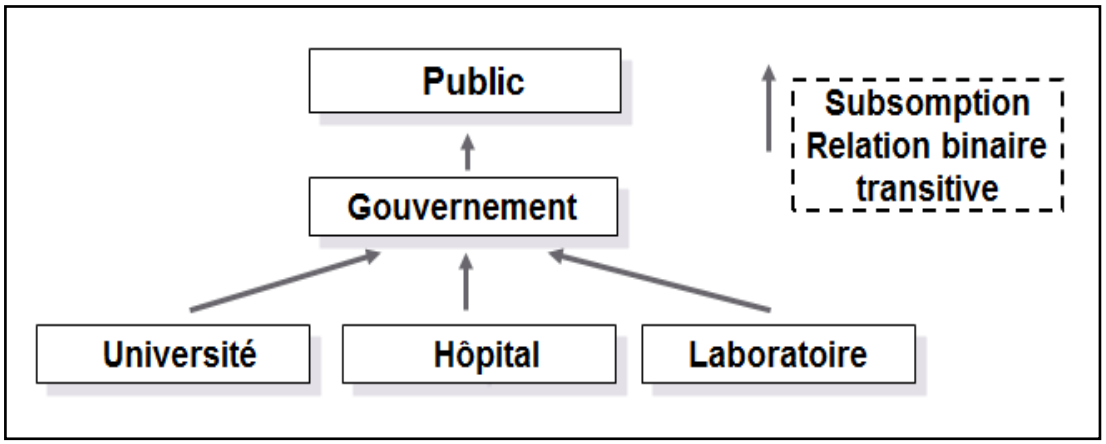


Figure III.4 : La notion de subsomption

La valeur «public» est plus générale que les autres valeurs de partage. En effet, si l'utilisateur a déclaré comme valeur de partage l'attribut public (partager avec toutes entités), alors le partage avec les entités : gouvernement, université, laboratoire, hôpital reste toujours valable. De même, la valeur «gouvernement» est plus générale que université, laboratoire et hôpital, puisque les trois sont des organismes gouvernementaux.

Pour le domaine de la durée {0, 1, …, indéfinie}, la valeur indéfinie est plus générale que les autres valeurs. Déclarée à ce montant, le destinataire peut avoir toutes les autres valeurs. Le même principe est pour les valeurs des deux domaines : profondeur et but.

Pour saisir la relation sémantique entre les valeurs de domaine, nous introduisons la notion de vie privée, nommée : subsomption (noté $\subseteq$).

Par exemple, les subsomptions suivantes peuvent êtres établies: gouvernement $\subseteq$ public, université $\subseteq$ gouvernement, laboratoire $\subseteq$ gouvernement, hôpital $\subseteq$ gouvernement.

III.2.3. Exemple

Par exemple, supposons que nous avons deux services S_1 et S_2 tel que :

- La préférence de S_1et la politique de S_2 contient un attribut « Name » qui est caractérisé par ses propres critères de matching (Voir le tableau III.1).

Préférence de S_1
Attribut = « Name »
Durée = 100
Partage = Public
Profondeur = 3
But = indéfinie
Sensibilité = Obligatoire

Politique de S_2
Attribut = « Name »
Durée = 60
Partage = Laboratoire
Profondeur = 1
But = indéfinie
Sensibilité = Optionnel

Si on applique un test de subsomption entre les deux services, on obtient les résultats suivants :

$(Durée_{(S2)} = 60 \leq Durée_{(S1)} = 100) \Rightarrow (Durée_{(S2)} \subseteq Durée_{(S1)})$

$(Partage_{(S2)} = Laboratoire \leq Partage_{(S1)} = Public) \Rightarrow (Partage_{(S2)} \subseteq Partage_{(S1)})$

$(Profondeur_{(S2)} = 1 \leq Profondeur_{(S1)} = 3) \Rightarrow (Profondeur_{(S2)} \subseteq Profondeur_{(S1)})$

$(But_{(S2)} = indefinite \leq But_{(S1)} = indefinite) \Rightarrow (But_{(S2)} \subseteq But_{(S1)})$

S_1 match S_2

S_1 et S_2 sont compatibles

Le résultat de la fin d'exécution du test de subsomption entre ces deux services révèle que *S_1 match* S_2.

S_1 ***match S_2*** signifie que S_1 et S_2 sont compatible, c.-à-d. la politique de confidentialité de S_2 est compatible avec la préférence de confidentialité de S_1.

Remarque : Si la valeur de sensibilité d'un attribut est égale à «Obligatoire», cela signifie que cet attribut est très sensible, et exige une utilisation sécurisée (c.-à-d. il faut que tous les critères de matching de cet attribut soient respectés). Par contre s'il est «Optionnel», l'utilisation sécurisée de ce dernier n'est pas exigée.

III.3. Algorithme proposé

La figure III.5 décrit notre algorithme de sélection.

Entre : Base(C, S, PPij, PRij)

$C=\{C_0, C_1, \ldots, C_i\} \mid 0=< i =< Nbr\_Classe$

$S=\{ S_{j,i}\} \mid 0=<j<Nbr\_service$ // Service j de la classe i

$PP(S_{j,i})$ // Politique Privée de $S_{j,i}$

$PR(S_{j,i})$ // Préference Privée de $S_{j,i}$

Sortie : COM //COM : une liste contient les compositions privées

Début

COM ⟵ Fonction Dominat (C_i, $S_{k,i}$)

Fin

Fonction Dominat (C_i, $S_{k,i}$)

Début

Si $|C_i|< |C|$ **Alors**

Pour j=0 jusqu'à |S| faire

Si $PP(S_{j,i+1})$ compatible avec $PR(S_{k,i})$ **Alors**

$PP(S_{j,i+1}) \subseteq PR(S_{k,i})$

Ajouter $S_{j,i+1}$ dans L // L : une liste contient les services compatible

Sinon $PP(S_{j,i+1}) \not\subseteq PR(S_{k,i})$

Fin Si

Fin Pour

Pour k'= 0 jusqu'à |L| faire

Dominat (C_{i+1}, $L_{k'}$) //fait un appel à la fonction Dominant

Fin Pour

Sinon Ajouter le groupe de la composition dans la liste COM

Fin Si

Return COM

Fin

Figure III.5 : Notre algorithme de sélection

IV. Conception

IV.1. Diagramme de cas d'utilisation

Afin de faciliter la compréhension du projet, une représentation de notre application a été modélisée sous forme de diagramme de cas d'utilisation, illustré dans la figure III.6.

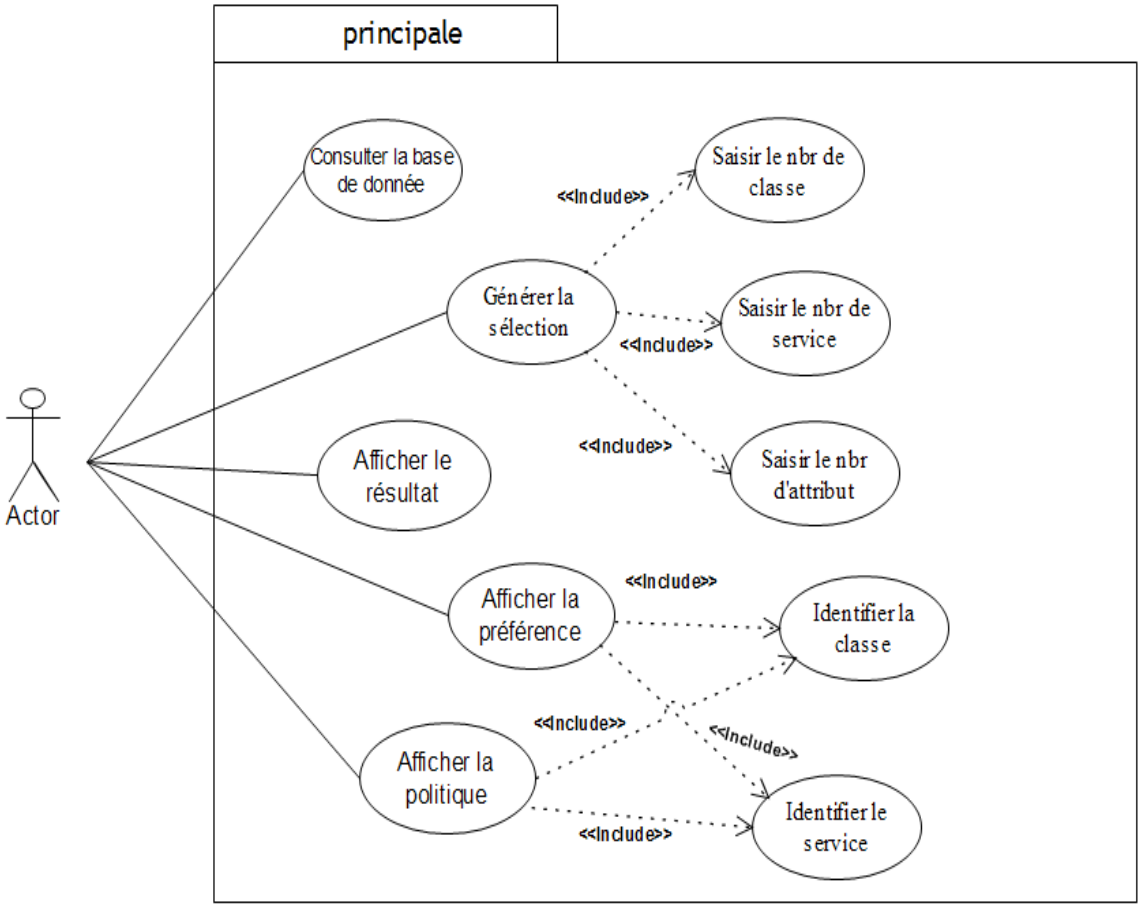


Figure III.6: Diagramme de cas d'utilisation

IV.2. Diagramme de séquence

Sachant que chaque tâche a son propre diagramme de séquence, nous avons choisi de détailler la partie la plus pertinente de notre projet qui est celle du digramme de séquence du cas «Générer la sélection», représentée dans la figure III.7.

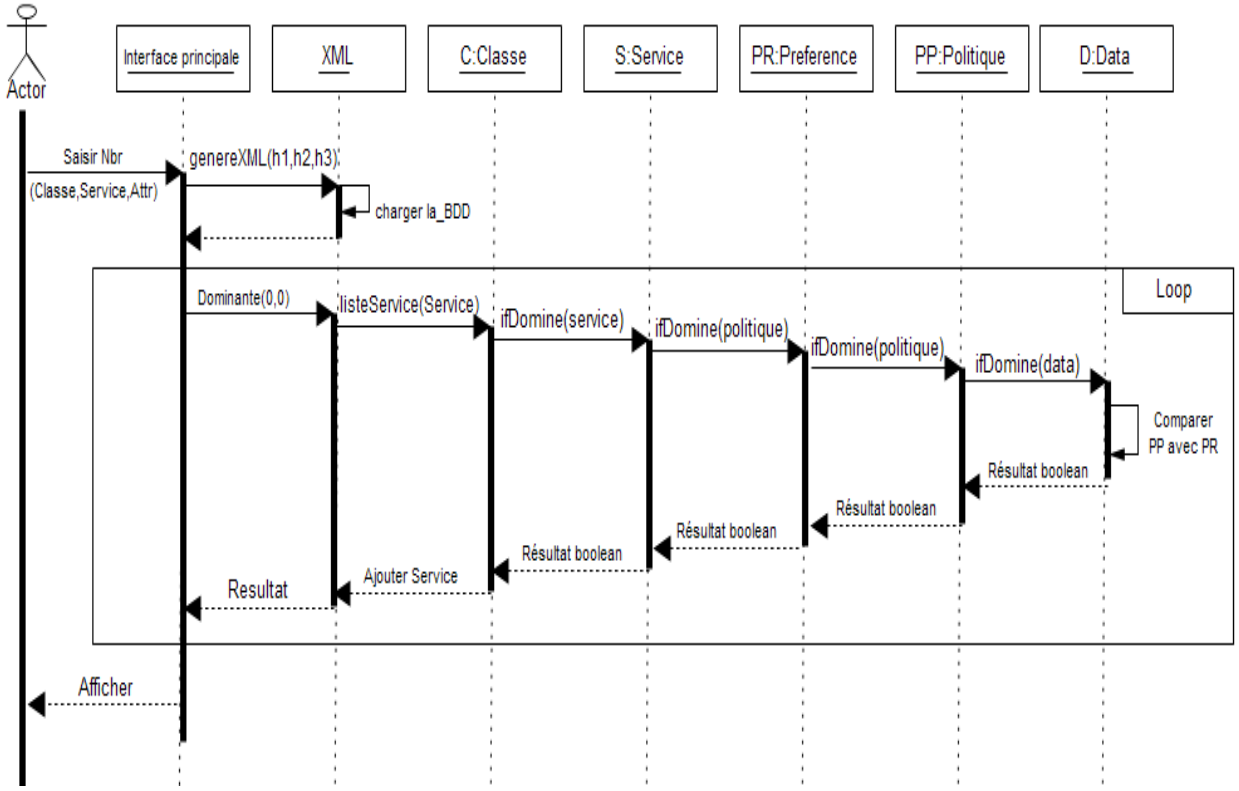


Figure III.7 : Diagramme de séquence du cas « Générer la sélection »

IV.4. Diagramme de classes

Le principe d'utilisation de notre outil de protection de la vie privée dans les services web est basé sur la sélection qui répond aux besoins en terme de préférences de la vie privée des utilisateurs. Ce principe est résumé dans le diagramme de classe de la figure III.8, qui modélise la fenêtre principale de notre application :

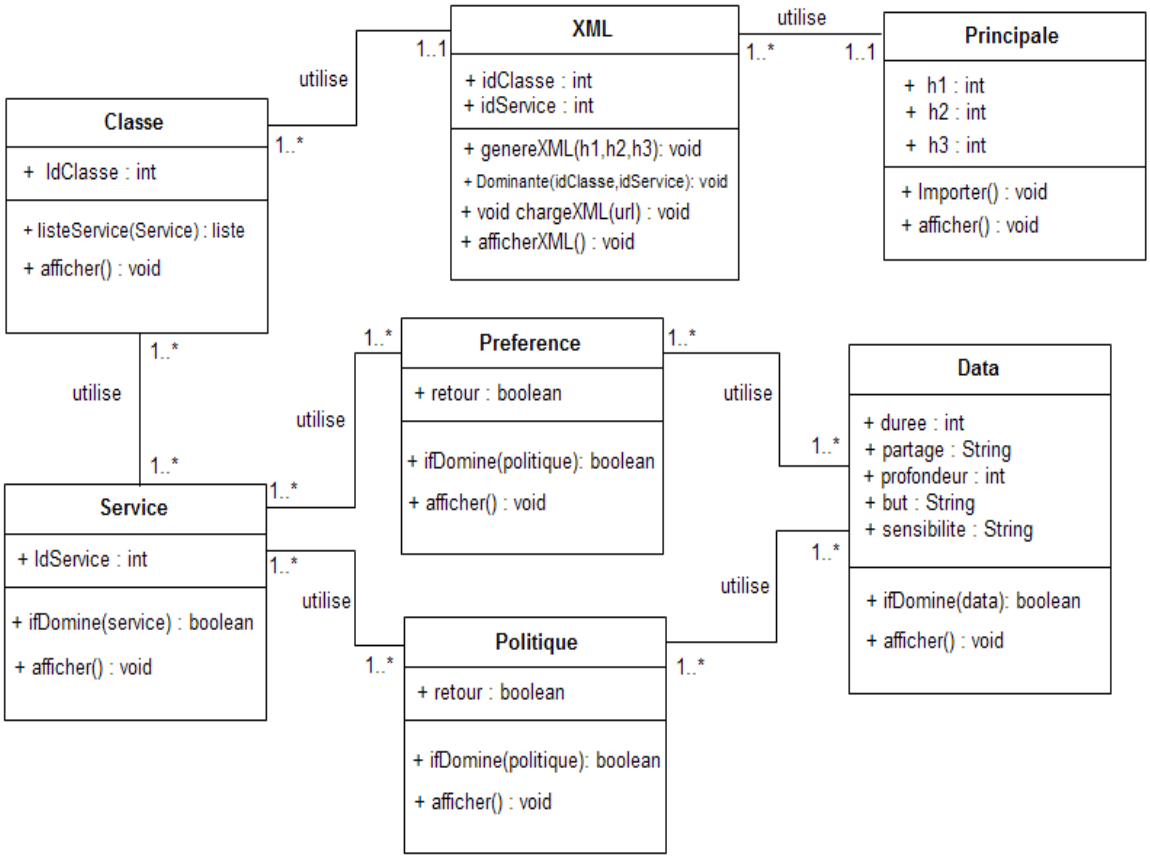


Figure III.8 : Diagramme de classes

V. Présentation des outils technologiques utilisés

Le choix des outils technologiques pour l'implémentation de notre application a été fait selon les avantages qu'offrent ces derniers. Son développement a nécessité l'utilisation de :

- **JAVA**

Comme langage de programmation, notre choix s'est porté sur le langage JAVA. C'est un simple langage orienté objet qui réduit les risques d'incohérences. Autres avantages de son utilisation est sa portabilité, ainsi que la possibilité de son utilisation sur différents systèmes (Windows, Linux, Macintosh, …etc), sans avoir le besoin de le modifier. Enfin il possède une riche bibliothèque de classes comprenant des fonctions diverses, telles que : les fonctions standards, le système de gestion de fichiers, les fonctions multimédia et beaucoup d'autres fonctionnalités.

- **NetBeans**

Pour l'environnement de développement, nous avons opté pour NetBeans version 6.9.1, car il possède de nombreux points forts qui sont à l'origine de son énorme succès, dont les principaux sont :

- Il fournit un environnement standard de développement pour la création rapide des interfaces.
- Support plusieurs plates-formes pour son exécution : Windows, Linux, Mac OS.

- **L'API JDOM**

Afin de représenter et de manipuler les documents XML d'une manière intuitive, nous avons utilisé JDOM (Java-based Document Object Model), qui est une API open source Java, simple à manipuler et ce qui rend son utilisation assez répandue. Vue comme un modèle de documents objets, il ne nécessite pas une connaissance pointue de XML pour pouvoir le maîtriser par un développeur JAVA. Il permet de vérifier que les données contenues dans les éléments, respectent cette norme XML.

VI. Présentation de l'IHM

L'interface homme/machine représente l'élément clé dans l'utilisation de tout système informatique, c'est pour cette raison que nous avons créé notre propre prototype qui sera représenté avec ses différentes interfaces graphiques.

- **Fenêtre d'authentification**

Dans n'importe quel système informatique sécurisé, ce niveau de protection est indispensable. Pour accéder donc à notre application, l'utilisateur doit s'authentifier avec le nom et le mot de passe qui lui sont attribués, cela lui permet par la suite de bénéficier des fonctionnalités qu'offre notre système. (Voir la figure III.9).

Figure III.9 : Fenêtre d'authentification

- **Fenêtre principale**

Une fois l'utilisateur authentifié, il accédera automatiquement à notre fenêtre principale. Cette dernière contient toutes les fonctionnalités de notre application (voir la figure III.10).

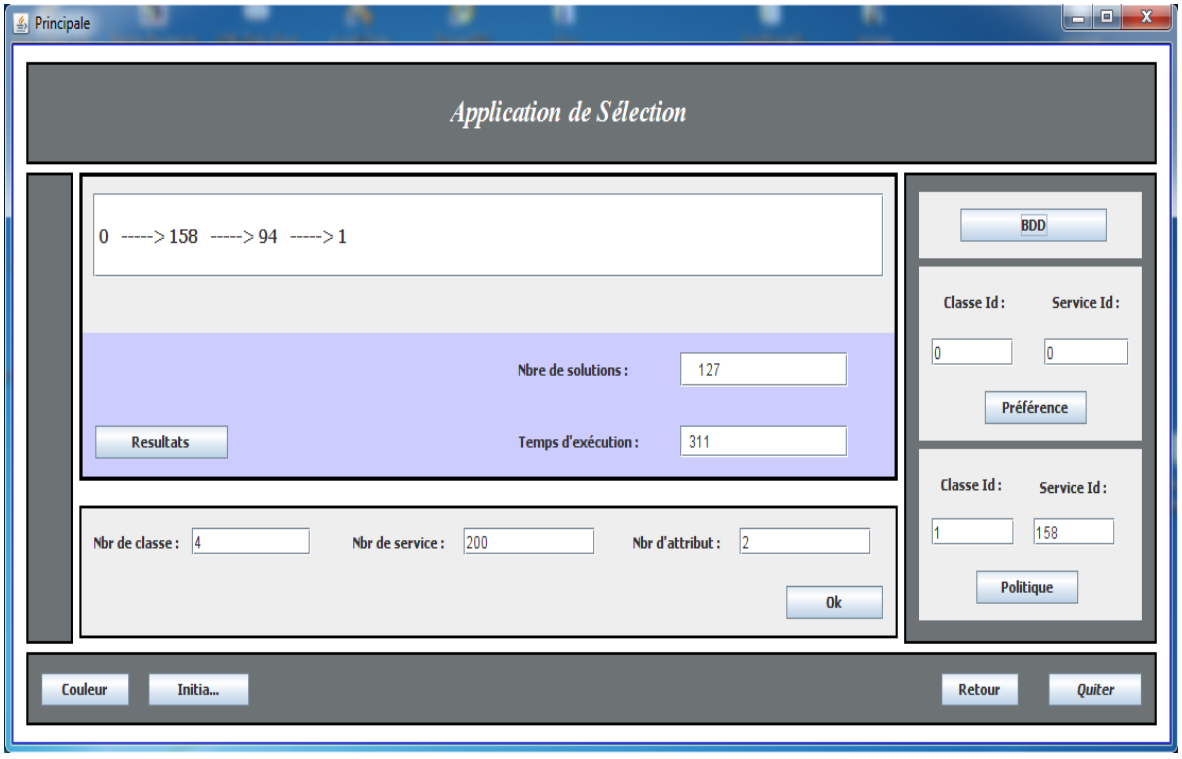


Figure III.10 : Fenêtre principale

- **Mode d'emploi**

  - L'utilisateur peut saisir dans les trois champs textes qui lui sont réservés, le nombre des classes qui correspond à la taille de la composition, le nombre de services qui correspond à la taille des services pour chaque composition, ainsi que le nombre d'attributs qui correspond à la taille de chaque politique et préférence. Le bouton OK de la figure III.10, permettra de charger et de sauvegarder la base dans un fichier (.xml) et de lancer la sélection avec les paramètres choisis initialement.
  - Le bouton BDD, permet la représentation de la base de données sous forme d'arbre **(**Figure III.11), simple pour la consultation et pour la compréhension.

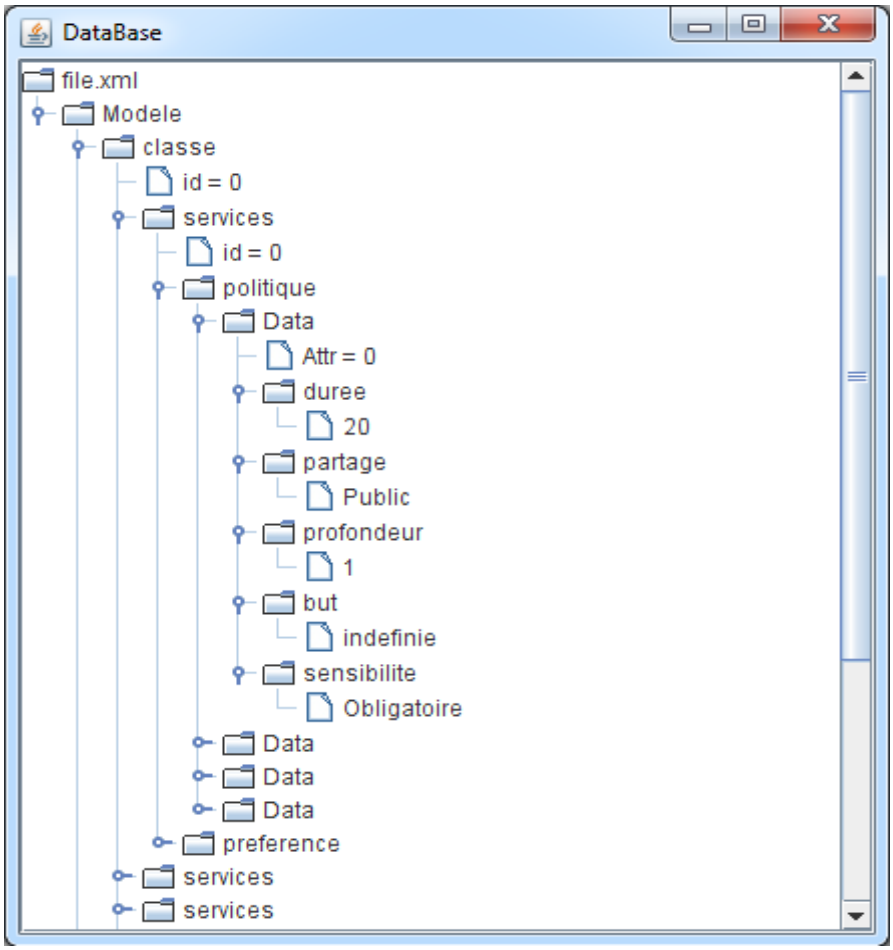


Figure III.11 : Exemple de base de données

  - L'affichage de la préférence du service nécessite la saisie de l'identifiant de service, ainsi que la composition qui contient ce dernier, dans les champs

textes positionnés au dessus du bouton «Préférence» (Voir figure III.10). Il est de même, pour le bouton «Politique».

- Le résultat d'exécution est consultable à partir d'une fenêtre d'affichage lancée grâce au bouton «Résultat». Cette fenêtre affiche tous les groupes de sélection, comme le représente la Figure III.12.

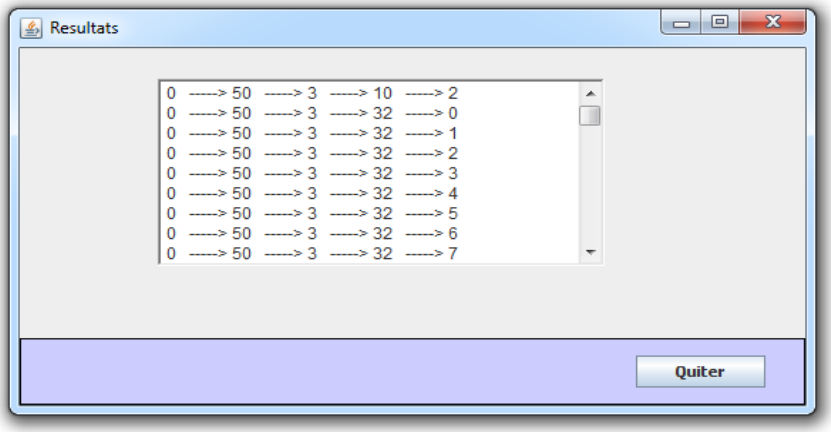


Figure III.12 : Affichage de résultat d'une sélection

- Enfin, le temps d'exécution nécessaire pour trouver les meilleurs groupes de sélection pour l'utilisateur, sera affiché dans le champ de texte qui est en bas du panel de l'affichage. Ce champ nous indiquera la nécessité ou non d'utiliser une machine plus puissante pour améliorer la rapidité d'exécution, qui dépend de l'ensemble des paramètres et des bases nécessaires pour obtenir nos sélections.

- **Exceptions**

Si la préférence de l'utilisateur et la politique des fournisseurs de services ne sont pas compatibles, un message d'erreur s'affichera pour indiquer le non-matching, comme la montre la Figure III.13.

Figure III.13 : Message d'erreur

VII. Expérimentation

Les expériences menées pour analyser les performances de notre approche seront décrites dans différentes expérimentations.

Nous avons implémenté notre algorithme de sélection en NetBeans IDE 6.9.1 et son exécution a été faite sur un ordinateur portable avec un processeur de 2,53 GHz Intel Core I3, de 4 Go de RAM et sous le système d'exploitation Windows 7 (64 bits).

VII.1. Description de la base

Notre base d'exemples comporte un fichier XML contenant une liste de «n» classes, chacune d'elles comprend une liste de «m» services. Ces services possèdent une politique et une préférence avec «k» d'attributs (Name, Numéro de Sécurité Sociale NSS, E-mail, CreditCard, …etc.). Tous les attributs sont caractérisés par les cinq critères de confidentialité de données présentés précédemment (Voir le paragraphe III.2.1).

Dans chacune des expérimentations qui suivent, nous avons fixé deux des paramètres et nous avons calculé par la suite le temps d'exécution en fonction des variations du troisième paramètre. Deux des résultats ont été comparés avec celles de l'algorithme PCM [Salah, 2013].

VII.2. Expérimentation 1 :

Dans la première expérimentation, nous avons mis l'accent principalement sur la phase de vérification de la compatibilité. Nous considérons les paramètres suivants :

- La taille de la composition (|classe|) = 2.
- La taille de service = 100.

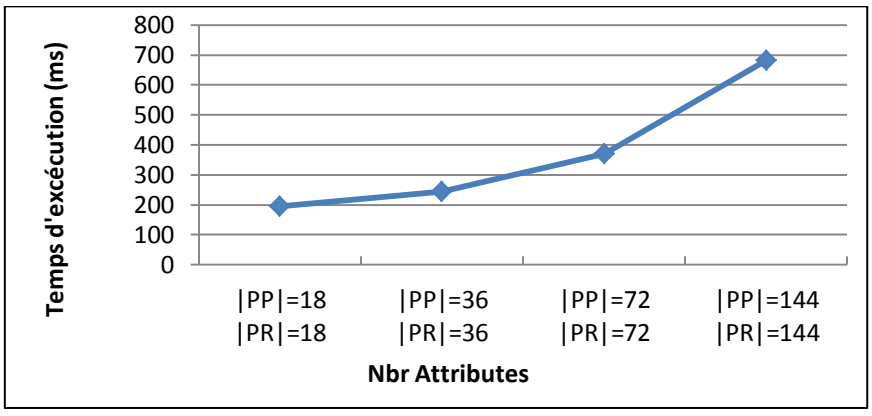


Figure III.14 : L'impact de la taille d'attributs sur le temps d'exécution

Nous remarquons dans les résultats de cette expérimentation que le temps d'exécution est proportionnel à la taille des PP et PR (|PP|, |PR|). Donc si cette dernière augmente, le temps d'exécution sera élevé.

Si nous comparons cette expérimentation avec celle de l'algorithme PCM [Salah, 2013] (La figure III.15), nous remarquons que pour | PP | = 18 et | PR | = 18, la performance du PCM est meilleur que celle de notre algorithme (60ms pour PCM, 195ms pour notre algorithme), mais quand |PP| = 36 et |PR| = 36, les deux algorithmes sont de même performance (250ms). Au delà de ces valeurs (|PP|= 36 et |PR| = 36), notre algorithme prend le dessus et devient plus performant que celui de PCM.

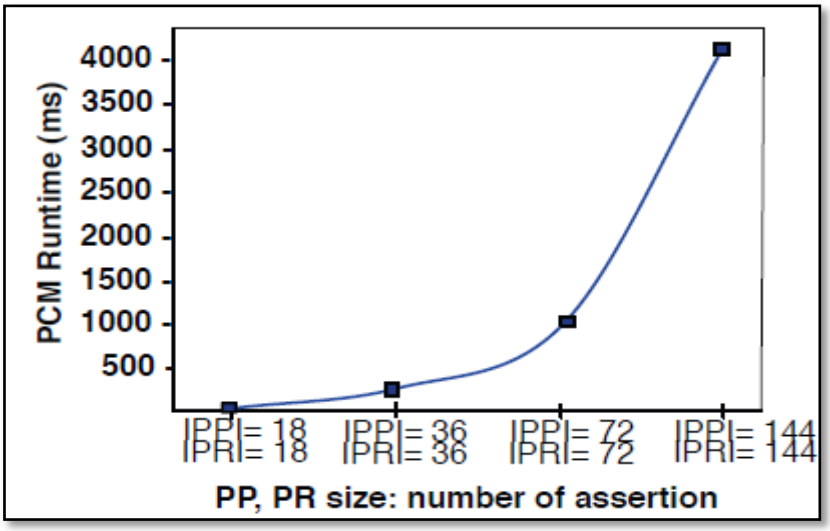


Figure III.15 : L'impact de la taille d'attributs sur le temps d'exécution (PCM)

VII.3. Expérimentation 2 :

Dans cette deuxième expérimentation, nous avons étudié l'impact de la taille de la composition sur le temps d'exécution. Pour cela nous considérons les paramètres suivants :

- La taille de service = 100
- Les tailles des PP et PR sont fixées à la fois à 10, puis à 20.

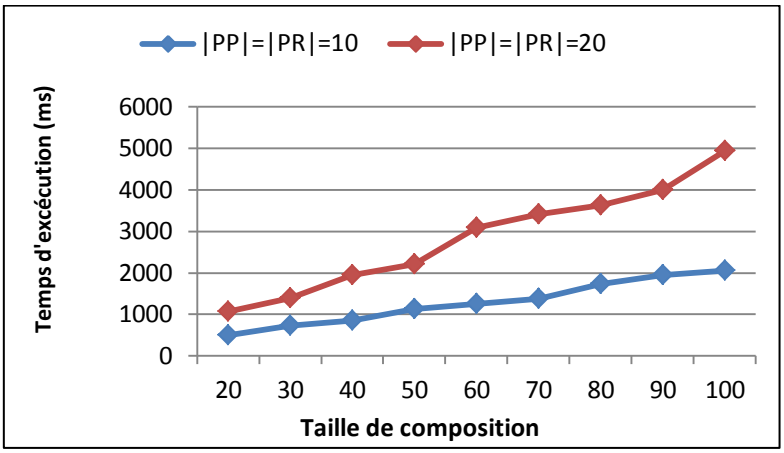


Figure III.16 : L'impact de la taille de la composition sur le temps d'exécution

Les résultats de cette expérimentation sont représentés dans la Figure III.16. Nous remarquons que le temps d'exécution est très important et proportionnel à la taille de la composition.

De la même manière, si nous comparons cette expérimentation avec celle de l'algorithme PCM [Salah, 2013] (La figure III.17), nous constatons que pour |PP| = 10 et |PR| = 10, PCM est plus performant que notre algorithme, mais à partir des valeurs |PP| = 20 et |PR| = 20, la performance de notre algorithme est plus meilleure.

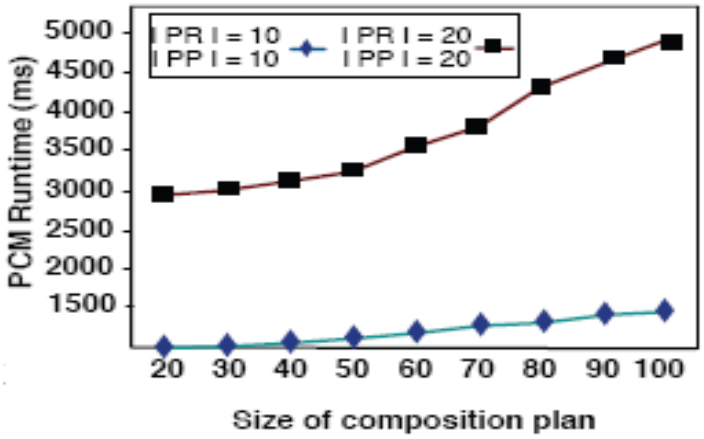


Figure III.17 : L'impact de la taille de la composition sur le temps d'exécution (PCM)

Remarque : L'environnement (la machine) d'expérimentation des deux évaluations possède les mêmes caractéristiques, seulement au niveau du CPU l'algorithme PCM a été exécuté par un Intel Core 2 Duo.

VII.3. Expérimentation 3 :

Dans la dernière expérimentation, nous avons voulu étudier l'impact de la taille des services sur le temps d'exécution de l'algorithme. Dans cette manipulation, nous considérons les paramètres suivants :

- La taille de la composition (|classe|) sera fixée à 9, à 12 et enfin à 15.
- Le nombre d'attribut = 7.

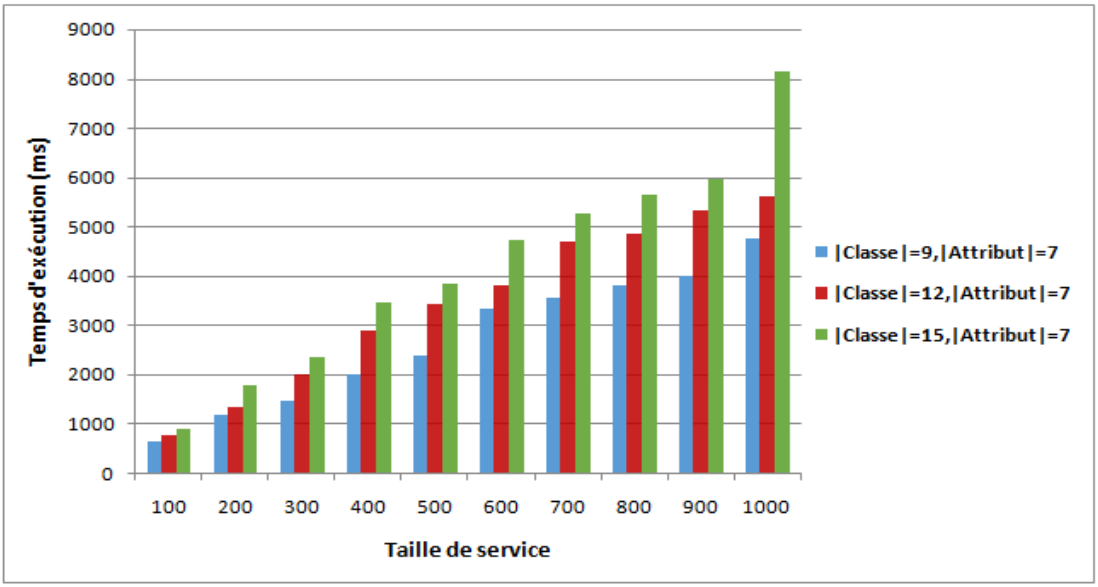


Figure III.18 : L'impact de la taille des services sur le temps d'exécution

Les résultats de cette expérimentation, comme l'indique la Figure III.18 sont aussi proportionnels à la taille de service. Si la taille de ce dernier augmente, le temps d'exécution sera trop élevé.

VIII. Discussion

- Globalement, l'impact de la taille de la composition et des services sur le temps de traitement de l'algorithme est plus important que celui de la taille d'attribut.
- Nous pouvons confirmer que le temps d'exécution de l'algorithme est proportionnel à la taille de la base utilisée.
- L'approche proposée offre des solutions qui respectent d'une manière précise les contraintes prédéfinies selon le besoin en terme de vie privée.

- Notre algorithme possède aussi un temps d'exécution acceptable pour les compositions ayant une petite taille (≤ 5), mais pour des compositions de grande échelle, l'algorithme présente un temps de réponse très élevé.

IX. Conclusion

Dans ce chapitre nous avons présenté la partie pratique de notre travail, qui regroupe l'ensemble des notions évoquées au début de ce chapitre, ainsi que dans les deux premiers.

Le prototype que nous avons développé a pour but la protection de la vie privée des utilisateurs dans les services web et l'amélioration du temps de traitement, par rapport à d'autres travaux existants (algorithme PCM), et cela grâce à son algorithme de sélection des services web qui respectent les préférences des utilisateurs.

En plus de notre algorithme de sélection, la performance de notre approche dépend aussi de celle de la machine utilisée.

Conclusion générale

Les services web sont une technologie très répondue pour l'intégration et l'interopérabilité des systèmes répartis. Ils sont caractérisés par leurs indépendances aux plateformes et aux systèmes d'exploitation, ce qui a impliqué leur adoption par les différentes organisations commerciales et industrielles offrant leurs services à travers le web, et par conséquence l'augmentation du nombre de services offerts. Lorsqu'une application est de type web, ses tâches composantes peuvent être exécutées à l'aide des services web. En particulier, pour une même tâche, on peut découvrir plusieurs services web aptes à l'exécuter, on serait alors face à un problème de sélection de services web afin de choisir la combinaison de services qui fournit le meilleurs offre, cette combinaison nécessite un partage des données personnelles entre les différentes entités participantes au service, cela peut entraîner des risques d'utilisation abusive des données qui peut atteinte à la vie privée des utilisateurs.

Dans ce travail de mémoire, nous avons proposé une approche de sélection pour les services web en considérant la protection de la vie privée comme critère principal. Pour se faire nous avons utilisé un algorithme qui permet de vérifier la compatibilité entre les exigences (préférences) et les politiques de confidentialité de tous les services lors d'un processus de sélection d'une composition.

En comparent l'évolution de notre algorithme avec l'évolution de de l'algorithme proposé dans le travail de recherche [Salah, 2013]. Les résultats obtenus sont très encourageants et montrent l'efficacité de notre approche.

Tout travail est amené à être amélioré, en ce sens, notre système peut encore évoluer et se voir améliorer. Parmi les perspectives à prendre en compte nous citons notamment :

- Proposer un mécanisme de négociation pour aborder les incompatibilités entre les politiques et les exigences de confidentialité.
- Implémenter une métrique (score) pour classer les compositions sélectionnées.
- Etudier la possibilité d'utilisation des Algorithmes à base des méta-heuristiques dans le cas où le nombre de services est très important.

Références Bibliographiques

[ABDELMOUNAAM, 2002] Abdelmounaam Rezgui, Mourad Ouzzani, Athman Bouguettaya, Brahim Medjahed, Preserving Privacy in Web Services, Proceedings of the 4th international workshop on Web information and data management, Pages 56 – 62, New York, 2002, ISBN:1-58113-593-9.

[ACOCA, 2007] Brigitte Acoca, Document Exploratoire sur le Vol d'identité en Ligne, Comité à l'égard de la Politique des Consommateurs (« CPC »), DSTI/CP(2007)3/FINAL.

[ARENAZA, 2006] Nerea Arenaza, Composition semi-automatique des Web services, Projet de Master, Ecole Polytechnique Fédérale de Lausane, Février 2006.

[ARTICLE 1, 2004] Loi n° 2004-801, Article 1, « La définition des données personnelles », relative à la protection des personnes physique à l'égard des traitements de données à caractère personnel, 06 août 2004.

[BEAULIEU, 2003] Marie-Hélène Beaulieu France Gaignard et Stéphanie Poulin, « Usurpation d'identité », Revue, Association des consommateurs de québec, Novembre 2003.

[BEKARA, 2012] Kheira Dari Bekara, protection des donnees personnelles côte utilisateur dans le E-COMMERCE, thèse de doctorat, l'universite Piere et Marie Curie, 18 Décembre 2012.

[BEKKOUCHE, 2012] Amina Bekkouche, Composition des Services Web Sémantiques À base d'Algorithmes Génétiques, thèse de magistère, Université Abou-bekr Belkaid Tlemcen, 2011-2012.

[BENGHOZI, 2008] P. Benghozi, S. Bureau, F. Massit-Folléa, « Vox Internet - Internet governance the democratic construction of standards», http://www.voxinternet.org/IMG/pdf/IdO.pdf, Octobre 2008.

[CHRISTENSEN ET AL., 2001] E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana, Web Services Description Language (WSDL) 1.1, rapport téchnique,W3C, http://www.w3.org/TR/wsdl, 2001.

[CLEMENT ET AL, 2004] Clement L., Hately A., von Riegen C. et Rogers T., UDDI v.3.0.2, OASIS Specification, http://uddi.org/pubs/uddi_v3, 2004.

[CORTES, 2003] Damien Cortès, Sid Ali Guebli, les services web et l'impact sur le eBusiness, Université Claude Bernard Lyon1, Décembre 2003.

[DEHANE, 2012] A. Dehane et Z. Kebir, Evaluation des techniques de codage d'ontologies sur les performances de la composition de services Web, Université de Tlemcen,2011/2012,http://bibfac.univ-tlemcen.dz/bibfs/opac_css/doc_num.php?explnum_id=268.

[DRISS, 2011] Maha Driss, Approche multi-perspective centrée exigences de composition de services Web, thèse de doctorat, Université de Rennes, 2011.

[DUSTDAR ET SCHREINER, 2005] Dustdar, S., Schreiner, W. 2005. « A Survey on Web Services Composition », InJ.. Web and Grid Services, Vol.l, No.l.

[ELISA, 2013] Elisa Costante, Federica Paciy, Nicola Zannone, Privacy-Aware Web Service Composition and Ranking, Proceedings of the IEEE 20th International Conference on Web Services, IEEE Computer Society Washington, 2013, ISBN: 978-0-7695-5025-1.

[GAMBS, 2010] Sebastien Gambs, « Introduction à la protection de la vie privée », Cours, .IRISA. Institut de Recherche en Informatique et Systèmes Aléatoires, 2010.

[GANGER, 2010] François GRANGER, « SansFiltre », Mars 2010, www.fgranger.com/ dotclear/index.php/post/2010/02/13/Vie-priv%C3%A9ed % C3%A9finition.

[GARDIEN, 2002] Georges Gardien, XML des bases de données aux Services Web, édition Dunod, 2002.

[GERMOUCHE, 2007] N. Guermouche, S. Benbernou, Coquery, C E; hacid, M, Privacy-Aware Web Service Protocol Replaceability. In: IEEE International Conference on Web Services ICWS'07. Salt Lake City, Utah, USA, IEEE Computer Society (2007).

[GROUPE, 2009] GROUPE « INFORMATIQUE & LIBERTES 2.0 ?», « Le nouveau paysage des données personnelles: Quelles conséquences sur les droits des individus ? », Janvier 2009.

[GUINAULT, 2008] Adrien GUINAULT, « Retour sur le Hack Of The Year : TOR, votre meilleur ennemi », Janvier 2008, « XMCO - Consultants en sécurité informatique (PCI QSA et ISO 27001 Lead Auditor)».

[HUBERT ET AL., 2003] K. Hubert, M. Valérie, LES WEB SERVICES, Edition DUNOD, 2003.

[JAEGER, 2006] Jaeger, M.C. and Muhl, G.. Soft real-time aspects for service-oriented architectures. In proceedings of the 8th IEEE International Conference on E Commerce Technology and the 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, San Francisco, CA, 5-5. (2006).

[KULCHENKO, 2001] Pavel Kulchenko, James Snell, et Doug Tidwell, Programming Web Services with SOAP, édition O'Reilly, Décembre 2001.

[LAURENT, 2009] Laurent COLLÉE, « Sécurité et vie privée sur les réseaux sociaux », Mémoire de master, Université du Luxembourg, 2009.

[LGMM, 2005] L. Cranor, G. Hogben, M. Langheinrich; m.marchiori; m.presler-marshall; j.reagle; m.schunter: the platform for privacy preference 1.1({p3p} 1.1) specification. in: technical report, w3c working draft (2005).

[MEDJAHED, 2003] B. Medjahed, B. Benatallah, A. Bouguettaya, A. H. H. Ngu, and A. K. Elmagarmid. Business-to-business interactions: issues and enabling technologies. The VLDB Journal, 12:59–85, May 2003.

[METAYER, 2010] Daniel Le Métayer et Guillaume Piolle, « Droits et obligations à l'ère numérique : protection de la vie privée », 30 septembre 2010.

[OSMAN ET AL., 2005] T.Osman, D. Thakker, and D. Al-Dabass. Bridging the Grap between Work-flow and Semantic-based Web Services Composition. In Proc. Of the Web Service Composition Workshop WSCOMPS05, 2005.

[PONGE, 2004] Julien Ponge ; Comptibilité et substitution dynamique des web Services ; Mémoire de fin d'étude, Université Blaise Pascal Clermont II ; Juillet 2004.

[RAGGETT, 1999] D. Raggett, A. Le Hors et I. Jacobs, HTML 4.01 Specification, W3C Recommandation 24 Décembre 1999, http://www.w3.org/TR/1999/REC-html401-19991224/.

[RICARDO, 2004] Ricardo DE LA ROSA-ROSERO ; Découverte et Sélection de Service Web pour une application Mélusine ; l'Institut d'Informatique et de Mathématique Appliquées de Grenoble, le 15 septembre 2004.

[SALAH, 2013] Salah-Eddine Tbahriti, Chirine Ghedira, Brahim Medjahed and Michael Mrissa, Privacy-Enhanced Web Service Composition, Services Computing, IEEE Transactions on, vol.PP, no.99, pp.1,1, 0, 08 mars 2013, ISSN :1939-1374.

[SEBASTIEN, 2010] Sebastien Gambs, «Réseaux de communication anonyme», «IRISA». Institut de Recherche en Informatique et Systèmes Aléatoires, 19 novembre 2010.

[SUN ET AL., 2003] H. Sun, Wang, X., Zhou, B. and Zou, P. 2003. « Research and Implementation of Dynamic Web Services Composition », APPT 2003,LNCS 2834,Springer-Verlag Berlin Hbidelberg, pp.457 -- 466.

[TAHIR, 2013] F. Tahir, A. Ghaffour, SELECTION DES SERVICES WEB : Une approche à base de Skylines et Clustering Hiérarchique Ascendant, Université Abou-bekr Belkaid Tlemcen, 2012-2013.

[TALEB, 2012] A. Taleb, Conception et Réalisation pour la médecine de travail basé sur : Les Services Web Sémantique, 2011/2012.

[VINCENT, 2003] Vincent CRIDLIG-Olivier FESTOR-Jacques GUYARD-Pierre Etienne MOREAU, «Formalisation et évaluation de politiques», LORIA - Campus Scientifique -, CFIP'02, 5 mai 2003.

[VINCENT, 2012] Vincent Regnault, « Protection de la vie privée des patients par la traçabilité des accès aux applications médicales », Fécamp : Centre Hospitalier Intercommunal, 2012.

[WEIXU, 2006] WEI XU V.N. Venkatakrishnan r. Sekar, I V Ramakrishnan: A Framework for Building Privacy-Conscious Composite Web Services. In: IEEE International Confernce on Web Serviced (ICWS'06) (2006).

[W3C-WSD-GROUP, 2002] W3C-WSD-Group. Web Services Description Working Group. http://www.w3.org/2002/ws/desc.

[VINCENT, 2012] Vincent Regnault, « Protection de la vie privée des patients par la traçabilité des accès aux applications médicales », Fécamp : Centre Hospitalier Intercommunal, 2012.

[WEIXU, 2006] WEI XU V.N. Venkatakrishnan r. Sekar, I V Ramakrishnan: A Framework for Building Privacy-Conscious Composite Web Services. In: IEEE International Confernce on Web Serviced (ICWS'06) (2006).

[W3C-WSD-GROUP, 2002] W3C-WSD-Group. Web Services Description Working Group. http://www.w3.org/2002/ws/desc.

Résumé

Bien que les technologies de composition et de sélection des services web sont considérées comme les technologies les plus prometteuses pour l'intégration des sources de données hétérogènes et multiples, ainsi que pour la réalisation d'opérations complexes. La question de la protection de la vie privée des utilisateurs demeure l'une des préoccupations majeures liées à ces technologies. Notre objectif dans ce travail est le développement d'une application permettant d'améliorer la sélection des services web avec des mécanismes de protection des données personnelles des utilisateurs. Pour atteindre cet objectif, nous avons proposé une approche qui se base sur un algorithme de compatibilité qui permet de vérifier formellement la compatibilité entre les exigences et les politiques de confidentialité de tous les services lors d'un processus de sélection d'une composition.

Mots clés : services web, sélection des services web, la vie privée.

Abstract

The issue of the protection of users' privacy remains on of the most important preoccupation related to technologies composition and selection of web services, even though, these ones are considered as the promising technologies for integration of heterogeneous and multiple sources of data as well as performing complex operation. Our goal in this work is to develop an application in an effort to improve the selection of web services with mechanisms to protect users' personal data. In order to reach this goal, we have proposed an approach based on an algorithm of accounting system that permit the account examination between demand and privacy policies of all services during a process of selecting composition.

Keywords : web services, selection of web services, privacy.

ملخص

على الرغم من أن تكنولوجيا تركيب واختيار خدمات الويب تعتبر من أكثر التكنولوجيات الواعدة لدمج مصادر البيانات المتعددة والغير متجانسة، وكذلك لتحقيق عمليات معقدة. مسألة حماية الحياة الخاصة للمستعملين لا تزال شكل أحد العوائق الرئيسية المتعلقة بهذه التكنولوجيات. هدفنا في هذا العمل هو تحسين اختيار خدمات الويب مع آليات لحماية البيانات الشخصية للمستخدمين. لتحقيق هذا الهدف، اقترحنا نهج يقوم على خوارزمية التوافق التي تتيح التحقيق بين متطلبات وسياسات الخصوصية لجميع الخدمات أثناء عملية الاختيار.

الكلمات المفتاحية : خدمات الويب، اختيار الخدمات، الحياة الخاصة.