Solutions opérationnelles d'une transaction électronique sécurisée et respectueuse de la vie privée

Aude Plateaux

Université de Caen Basse-Normandie

École doctorale SIMEM

Thèse de doctorat

par

Aude Plateaux

pour obtenir le

Doctorat de l'Université de Caen Basse-Normandie
Spécialité : Informatique et Applications

Solutions opérationnelles d'une transaction électronique sécurisée et respectueuse de la vie privée

Directeur de thèse : Christophe Rosenberger
Co-directeur de thèse : Kumar Murty

Jury

Maryline LAURENT	Professeur, Telecom SudParis	(Rapporteur)
Audun JØSANG	Professeur, Université d'Oslo, Norvège	(Rapporteur)
Christophe BIDAN	Professeur, Supelec	(Examinateur)
Laurent VERCOUTER	Professeur des Universités, INSA de Rouen	(Examinateur)
Patrick LACHARME	Maître de conférences, ENSICAEN	(Encadrant université)
Vincent COQUET	Chef de projet, BULL SAS	(Encadrant entreprise)
Christophe ROSENBERGER	Professeur des Universités, ENSICAEN	(Directeur de thèse)
Kumar MURTY	Professeur, Université de Toronto, Canada	(co-Directeur de thèse)

À mes parents, Charline et Dominique,
qui ont toujours été dans ma petite poche...

Résumé

Les échanges à distance par mobile, le paiement par carte à puce ou encore l'emploi de badges sans contact dans les entreprises sont des transactions électroniques courantes qui entraînent de multiples échanges d'informations entre les différentes entités. Un grand nombre d'informations personnelles, souvent confidentielles, transitent durant ces transactions à travers les divers acteurs et doivent être protégées. Les Transactions Electroniques Sécurisées (ou TES) peuvent alors jouer ce rôle. Néanmoins, la sécurisation de ces informations personnelles doit prendre en compte la notion de protection de la vie privée des utilisateurs. Il convient de ne pas divulguer n'importe quelle information auprès de tous les acteurs de ces échanges, ni de la faire sans le consentement des utilisateurs. De même, il est nécessaire d'utiliser des éléments sécurisés afin de stocker ces informations de manière sûre. Cette protection doit respecter des principes de minimisation des données et de souveraineté des données, clairement encouragées par la Loi Informatique et Libertés de la *CNIL*.

Le sujet de cette thèse se situe dans le champ de la protection de la vie privée des individus. Dans un premier temps, nous définissons les notions importantes en termes de protection de la vie privée et de transactions électroniques sécurisées. Nous présentons ensuite différents outils et méthodes permettant de protéger les informations personnelles d'un individu. Une fois les principes définis et les outils explicités, trois axes de recherche sont explorés : la gestion des données personnelles des internautes, la e-santé, ainsi que le e-commerce. Pour chaque axes, des exigences de sécurité et de protection de la vie privée sont décrites et les solutions existantes dans la littérature sont énoncées. Nous proposons ensuite de nouvelles architectures, que nous détaillons, analysons et comparons à celles existantes. Les différentes implémentations de ces concepts sont également proposées à la fin de chaque chapitre. Dans un premier temps, nous nous intéressons à un système de gestion de la sécurité et de la protection des données personnelles sur Internet. Ce système offre différentes fonctionnalités qui permettent de protéger les données d'un internaute lors

d'un enregistrement en ligne par exemple. Le second aspect traité dans cette thèse est la protection des dossiers médicaux des patients au sein d'un ou plusieurs établissements de santé. Un système de gestion des dossiers médicaux des patients est proposé en prenant en compte une politique d'accès adaptée au contexte de la e-santé. Finalement, au vu de l'augmentation du nombre de transactions financières réalisées sur Internet, le commerce électronique est un domaine qui nécessite une attention particulière en termes de protection de la vie privée. En effet, lors d'un paiement en ligne, beaucoup de données sensibles transitent, comme le numéro de carte ou le cryptogramme visuel du client. Nous présentons une architecture de paiement en ligne ne divulguant aucune information bancaire du client et garantissant les différentes exigences en termes de sécurité et de protection des données personnelles.

Table des matières

INTRODUCTION

« La liberté n'est rien quand tout le monde est libre. »

Pierre Corneille

Contexte

Comme le définit l'article 2 de la Déclaration des droits de l'homme et du citoyen (1789), la liberté, la propriété, la sûreté et la résistance à l'oppression sont les "droits naturels et imprescriptibles de l'homme". Plus récemment, la Déclaration Universelle énonce les droits de l'individu dont celui de la protection de la vie privée. Malheureusement, les réseaux sociaux, où apparaissent noms, prénoms, origines, loisirs ou encore religions, deviennent les premiers services sur Internet en termes de nombre de personnes inscrites, sans que ces personnes n'aient un réel contrôle sur le devenir de ces données personnelles. D'autres sites prônent l'anonymat grâce aux pseudonymes, même si votre date de naissance et votre lieu d'habitation apparaissent très clairement.

Neuf milliards de cartes à puce et plus de cinq milliard de cartes SIM circulent actuellement dans le Monde. De même, la carte Vitale2 compte des dizaines de millions d'exemplaires en France et plus de 400 millions de passeports électroniques dans le Monde. Alors que les Transactions Électroniques Sécurisées (ou TES) ont été initialement mises en place à des fins financières afin de protéger les banques et les marchands, les utilisateurs souhaitent désormais que ces transactions servent davantage à protéger leurs données personnelles. En effet, les échanges à distance par mobile, le paiement par carte à puce, le contrôle d'accès par badge sans contact dans les entreprises, l'enregistrement auprès de sites marchands... sont des transactions électroniques courantes qui entraînent des échanges d'informations personnelles, souvent confidentielles, entre différentes entités. L'ensemble de ces données doit donc être protégé.

Cette informatisation force donc à revoir la notion de protection de la vie privée. En effet, bien qu'il s'agisse d'un progrès d'un point de vue technique, de nombreuses menaces apparaissent, notamment en ce qui concerne les données personnelles des individus. Qui plus est, même si des barrières légales existent, comme par exemple la loi informatique et libertés, celles-ci n'ont pas donné lieu à la création de techniques permettant d'assurer complètement la protection des informations sensibles. Il est nécessaire de faire appel à d'autres techniques, notamment cryptographiques, permettant entre autre de garantir la confidentialité des données, et d'utiliser des éléments sécurisés afin de stocker des informations de manière sûre. Cependant, ces méthodes ne sont pas suffisantes. Il faut prendre en compte leurs utilisations lors de transactions électroniques qu'il faudra nécessairement sécuriser. Ces échanges requièrent des protocoles particuliers basés sur les propriétés usuelles de sécurité : authentification, confidentialité, intégrité et non-répudiation. Il est donc important de se familiariser et de s'intéresser de près à ces concepts.

Objectifs

Le sujet de cette thèse se situe dans le champ de la protection de la vie privée des individus et s'intéresse plus particulièrement à trois applications des TES : la protection des données personnelles des internautes, la e-santé et le paiement en ligne. Cette problématique de protection de la vie privée est de plus en plus importante aujourd'hui. Effectivement, le développement du numérique engendre de gros problèmes d'anonymat, de traçabilité et plus largement de divulgation de données personnelles. Ainsi, comme le confirment les nouvelles recommandations de l'Union Européenne, il est urgent de proposer de nouvelles techniques permettant de garantir la confidentialité des données échangées et de les adapter aux différents contextes des transactions électroniques. Cette thèse propose ainsi différentes architectures répondant à des exigences en termes de sécurité et de protection des données personnelles. Ces exigences sont spécifiques aux contextes traités et donc à la nature des informations échangées. Ces architectures s'appuient aussi bien sur de nouveaux protocoles que sur l'utilisation d'éléments sécurisés.

Cette thèse étant cofinancée par la société informatique BULL, nous accordons une attention particulière aux outils et solutions qu'elle propose. Cette entreprise est spécialisée dans les infrastructures à clés publiques pour la production de certificats électroniques, dans les solutions de signature électronique, de logiciels et matériels cryptographiques, ainsi que dans la gestion d'identités et le contrôle d'accès. Elle se retrouve également dans le domaine de la sécurité monétique. BULL travaille sur la conception de Hardware Security Module (ou HSM) qui génèrent, stockent et protègent des clés cryptographiques, ainsi que sur les Field-Programmable Gate Array (ou $FPGA$) utilisés pour de nombreuses

applications comme les télécommunications, les transports et la monétique. Le *HSM* programmable de BULL, CRYPT2PAY, est ainsi appelé à jouer un rôle dans les différentes architectures proposées dans cette thèse.

Contributions

Dans un premier temps, nous commençons par positionner la problématique et présenter différents outils et méthodes permettant de protéger les informations personnelles d'un individu. Sur la base des outils énoncés, plusieurs axes sont explorés. En particulier, un système de gestion de la sécurité et de la protection des données personnelles sur Internet est proposé. Différentes fonctionnalités sont disponibles au sein de ce système permettant de protéger les données de l'internaute : vérification de signature, établissement d'une preuve de connaissance, authentification par biométrie révocable, vérification et analyse de conditions d'utilisation et d'accès ou encore utilisation d'un coffre fort électronique. Une preuve de concept y est également décrite.

Dans un second temps, une préoccupation d'actualité est étudiée : la protection des dossiers médicaux des patients. Plus particulièrement, nous nous intéressons à la gestion et à l'accès aux données personnelles d'un patient au sein d'un hôpital, ainsi qu'à leurs transferts entre institutions. Un protocole pour la gestion du chiffrement des données adapté au contexte médical est notamment proposé à partir d'un système de partage de secret. Ce protocole permet entre autre de garantir la minimisation des données personnelles.

La dernière partie se concentre sur un autre domaine en plein essor : le domaine du commerce électronique. Celui-ci nécessite des échanges de données particulièrement sensibles, notamment lors d'un paiement en ligne. Tout d'abord, nous analysons et améliorons le système 3D-Secure, ainsi qu'une architecture proposée par Ashrafi et Ng. Dans un second temps, nous proposons une architecture de paiement en ligne sécurisée minimisant la divulgation des informations du client et du commerçant et ne divulguant aucune information bancaire du client. Pour chacune de ces trois architectures, des exigences de sécurité et de protection de la vie privée sont décrites et les solutions existantes sont énoncées. Les différentes implémentations de ces trois applications y sont également commentées.

Première partie

Positionnement du problème et Travaux existants

Chapitre 1

Positionnement du problème

Cette thèse adresse la problématique générale de la protection de la vie privée d'un utilisateur lors d'une transaction électronique sécurisée. Ce chapitre définit les différents termes utilisés dans la suite de ce manuscrit, les propriétés attendues d'une transaction électronique sécurisée respectueuse de la vie privée, ainsi que les menaces pour les individus.

Sommaire

1.1 Introduction

Compte tenu de l'essor croissant des transactions électroniques, la vie privée des individus est fortement menacée. Le premier enjeu des Transactions Électroniques Sécurisées (TES) a cependant été financier. Les échanges d'argent devaient en effet être sécurisés au vu de l'augmentation de ces transactions financières. Ainsi, une fois l'enjeu financier pris en considération, les utilisateurs ont souhaité de plus en plus que ces transactions s'intéressent à la protection de leurs données personnelles.

En fonction de la population prise en compte lors de la transmission d'une donnée, la portée de la protection, et ainsi les informations à protéger, tout comme les propriétés attendues lors d'une telle transaction, sont différentes. En effet, il est possible de considérer un unique individu au sein d'une population qui sera le seul à posséder les informations et acceptera ou non de les partager ou il peut s'agir d'un groupe de personnes, par exemple l'ensemble des dirigeants d'une entreprise. Ils ont alors la possibilité de s'identifier ou de signer un document au nom de tous les dirigeants. Dans ce cas, les données à protéger sont les données de la société entière, et particulièrement celles qui concernent leur projet et leur ressource. Les dirigeants devront alors avoir défini la politique d'envoi en répondant aux questions suivantes :

- un seul membre du groupe pourra-t-il envoyer un document ?
- un individu devra-t-il se concerter avec ses collègues avant chaque envoi ?
- un membre devra-t-il avoir l'empreinte d'autres membres du groupe pour envoyer un dossier ?

De plus, si une option a été préalablement consentie par le groupe, le responsable du groupe peut être capable d'identifier précisément le signataire. Nous pouvons également considérer un ensemble de personnes d'environnements différents, par exemple des amis d'enfance, peuvent décider de créer un domaine (ou cercle) de confiance, cela revient alors à se poser les mêmes questions que précédemment. Finalement, il est également nécessaire d'envisager la population entière d'un pays, en opposition avec les citoyens des autres territoires. Les problèmes de transfert des données, de sécurité intérieure, de compréhension de la langue se posent alors en supplément. Ainsi, les menaces potentielles sont différentes en fonction du contexte de l'application, et par conséquent, les propriétés attendues pour protéger la vie privée des utilisateurs et la sécurité du système varient. Il est donc nécessaire dans un premier temps de définir clairement les termes étudiés dans cette thèse avant de proposer une quelconque solution assurant les exigences souhaitées.

La première section de ce chapitre définit ainsi explicitement ce qu'est une transaction électronique sécurisée (TES) ainsi que leurs enjeux. La suivante définit les notions de vie privée et de sécurité, ainsi que les attaques auxquelles les données sont confrontées. Ces menaces ainsi exposées, la section suivante se concentre sur les propriétés et les principes importants pour qu'une transaction électronique soit sécurisée et respectueuse de la vie privée. Pour finir, les sections 1.5 et 1.6 concluent ce chapitre et explicitent les objectifs de cette thèse.

1.2 Transaction Électronique Sécurisée

Une transaction électronique sécurisée (TES) est un protocole permettant l'échange d'informations dématérialisées entre entités (individus ou organisations) au travers de systèmes informatiques. L'objectif d'une telle transaction est de gagner en sécurité, en prenant en compte des propriétés telles que la confidentialité des données et l'authentification des acteurs, ainsi qu'en efficacité, en réduisant par exemple le coût d'une transaction ou en augmentant sa rapidité.

Les TES se retrouvent dans notre quotidien dans de nombreux domaines d'applications et via l'usage de différents outils, comme les cartes bancaires, de santé, d'entreprise, de fidélité ; ou encore durant un accès aux comptes bancaires à distance, mais également lors de micro-paiement via un mobile. Elles sont également nécessaires lors d'une commande de billet électronique ou d'une authentification par empreinte digitale pour l'accès à un ordinateur. Les TES se retrouvent ainsi dans trois grands domaines d'activités :

- Les transactions financières : la carte bancaire et le porte-monnaie électronique, les paiements en ligne, la carte de fidélité, la billettique...
- L'échange sécurisé de données : le domaine de la e-santé, la carte Sesam-Vitale, la carte SIM (Subscriber Identity Module) ...
- Le contrôle d'identité numérique : la carte d'identité, le passeport électronique, la signature électronique...

Comme l'illustre la Figure 1.1, une TES fait intervenir différentes technologies complexes. Afin de gérer ces transactions, l'évolution des technologies nous amène à utiliser aussi bien des mobiles, des réseaux sans fil ou encore de la cryptographie. Il en est de même pour les services proposés par ces TES. D'un point de vue technique, les identifications peuvent se faire à l'aide de puces, de signatures (section 2.2.3) et de certificats (section 2.2.3) ou encore grâce à la biométrie. L'accès peut être lancé via mobile NFC (Near Field Communication), Bluetooth ou Wifi... Une multitude de possibilités est envisageable.

Les TES peuvent être représentées par des échanges sur des réseaux de télécommunications comme le e-commerce, les transferts de documents électroniques, ... ; ou encore par des transactions initialisées mettant en œuvre une carte à puce (authentification, paiement). Ces applications touchent aussi bien les particuliers que les entreprises. Ainsi, plus de 9 milliards de cartes bancaires sont actuellement en circulation dans le monde. Étant donné l'omniprésence des TES, on peut imaginer les conséquences désastreuses qu'auraient la propagation des données personnelles et sensibles qu'elles exploitent.

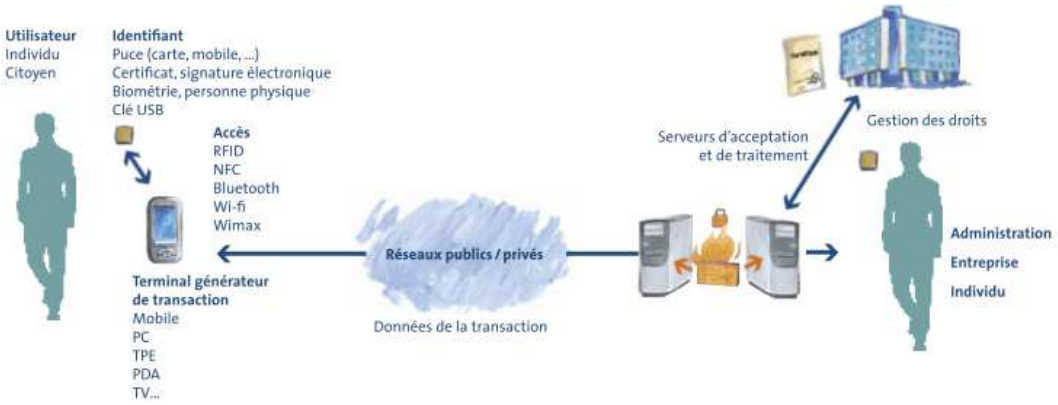


FIGURE 1.1 – Schéma général de la chaîne d'une transaction électronique sécurisée. Source : [1]

Un des enjeux majeurs des TES est la gestion d'identité numérique. Comme le montre la Figure 1.2, chaque individu manipule différents statuts et identités en fonction de son environnement professionnel et privé. Une identité correspond à un ensemble suffisant d'attributs permettant d'identifier un individu au sein d'une population donnée dont les caractéristiques globales sont connues. L'**identification** de l'entité est ainsi une technique permettant de reconnaître, de distinguer, ou de localiser une personne ou encore créer un contact dans un domaine précis à l'aide d'un ensemble d'attributs. La gestion des identités consiste donc en des *systèmes et processus qui gèrent et contrôlent ceux qui ont accès aux ressources, et ce que chaque utilisateur a le droit de faire avec ces ressources, en conformité avec les politiques de l'organisation* [2]. Actuellement, et à titre indicatif, les acteurs développant ce genre de systèmes sont Microsoft et Kontara Initiative. La personne liée à cette identité numérique est responsable de ses actes. Par conséquent, le vol d'identité constitue une forte menace pour les utilisateurs.

Ainsi, dans le monde réel, les attributs de la personnalité sont souvent liés à la filiation paternelle ou définit par les autorités publiques selon des critères légaux. Personne ne peut (ne devrait pouvoir) usurper une identité complète reconnue des autorités publiques. Dans le monde virtuel, l'adresse IP (Internet Protocol) de chaque machine connectée au réseau est une donnée possible d'identification. Cependant, bien qu'en 2004, la CNIL ait accepté de considérer cette adresse comme donnée personnelle, ce n'est qu'en mars 2010 que l'adresse IP est considérée comme telle. En novembre 2009, le Sénat adopte la proposition de loi présentée par Yves Détraigne et Anne-Marie Escoffier *"visant à mieux garantir le droit à la vie privée à l'heure du numérique"*. Ainsi, l'*article 2*, défini dans la section précédente, se voit ajouter la phrase : *"Tout numéro identifiant le titulaire d'un accès à des services de communication au public en ligne est visé par le présent alinéa"*. La question

se pose de la correspondance entre une identité physique et une identité virtuelle. En effet, alors qu'une authentification permet d'authentifier une personne physique ou virtuelle, la non-répudiation doit permettre à une telle personne de contredire son authentification en cas d'usurpation d'identité. Une dualité existe donc entre ces deux notions.

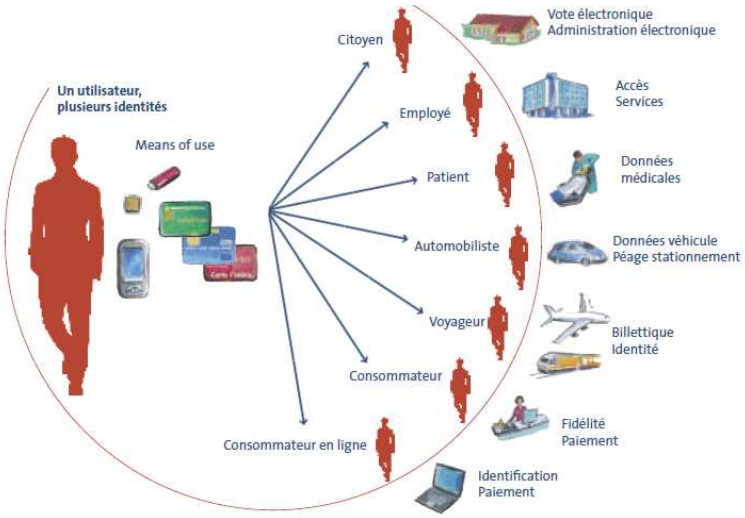


FIGURE 1.2 – Identités possibles pour un utilisateur. Source : [1]

L'**authentification** permet de garantir l'identité d'une entité à l'aide de : ce que l'on sait, ce que l'on possède et ce que l'on est. L'authentification peut être utilisée pour le **contrôle d'accès** des données pouvant se baser sur l'identité de l'entité, une liste de droits d'accès et la sensibilité des informations qu'il protège. Dans le but d'une telle authentification, toute identité est associée à un identifiant, à un secret et est liée à un support (mobile, carte, clé, ...). Il est alors difficile de gérer tous ces identifiants et mots de passe. Il ne faut cependant pas confondre identification et authentification. L'identification permet de déterminer l'identité d'une entité sans pour autant utiliser une information secrète, alors que l'authentification consiste à apporter la preuve de cette identité et est régie par la norme ISO/IEC 9798 ([3]), ainsi que par le $NIST$ dans [4]. Une preuve d'authentification peut être distribuée de différentes façons, via un ordinateur, un téléphone portable, un papier ou encore une clé USB (Universal Serial Bus). Lorsque la preuve d'identité est un élément unique (par exemple mot de passe, tag $RFID$ Radio Frequency IDentification, empreinte), on parle d'**authentification simple**. Cependant, cette forme d'authentification, notamment par mot de passe, n'offre pas suffisamment de sécurité pour assurer la protection de données sensibles et est vulnérable à de nombreuses attaques : force brute, par dictionnaire, "Man-in-the Middle, écoute,... La principale attaque concernant l'authentification est l'usurpation d'identité. Il est heureusement possible de limiter les risques d'usurpation en combinant différentes preuves d'identité. Ce type d'**authentification** est dite **forte**. De nombreuses formes de preuves existent :

- Mémorielle : ce que l'on sait. Ex. : mot de passe, code ;
- Matérielle : ce que l'on possède. Ex. : carte à puce, certificat électronique, téléphone, PDA (Personal Digital Assistan) ;
- Réactionnelle : ce que l'on sait faire. Ex. : dynamique, geste, signature ;
- Corporelle : ce que l'on est. Ex. : empreinte digitale, iris.

Dans le cas d'une authentification forte, un facteur d'authentification supplémentaire est utilisé : les OTP (One Time Password), les tags $RFID$, les certificats, ainsi que la biométrie. Par exemple, Krawczyk et Jain présentent dans [5] une méthode d'authentification basée sur la vérification vocale, utilisée pour les dossiers médicaux électroniques.

De plus, étant donnés les nombreux usages et types de supports, la mobilité est désormais un point important et nécessaire des TES. La miniaturisation permet l'intégration de nombreux composants électroniques dans des objets de la vie quotidienne. De plus, les réseaux sans fil et sans contact, de type NFC, offrent la mobilité des échanges de données et une grande facilité d'usage lors des transactions.

Il est donc nécessaire d'utiliser les moyens suffisants pour l'identification et l'authentification des acteurs d'une transaction et d'obtenir une preuve de transaction. La sécurisation de ces échanges est donc obligatoire afin d'assurer les différentes propriétés de sécurisation et de protection de la vie privée des différents acteurs. Heureusement, il existe de nombreuses technologies nous permettant de garantir certains principes et de définir différents niveaux de politique. Afin de garantir la sécurité et la protection de la vie privée lors d'une transaction, il est avant tout nécessaire de comprendre les termes intervenant durant ces échanges et principalement la notion de vie privée.

1.3 Vie privée et sécurité

L'*article 9* du Nouveau Code Civil ([6]) stipule que : "Chacun a droit au respect de sa vie privée". Cependant, cette notion est complexe et varie en fonction du contexte social et culturel, des sujets d'étude, des intérêts des actionnaires et des domaines d'application.

1.3.1 Aspects juridiques

En France, l'expression : "Droit de l'informatique" regroupe les droits relatifs aux nouvelles technologies de l'information et de la communication. Plus précisément, les "lois informatiques et libertés" du 6 janvier 1978 ([7]) définissent les lois destinées à garantir la protection de la vie privée des citoyens face aux moyens de traitements automatisés de données numériques. Afin de contrôler la bonne application de cette loi, il a fallu créer une

autorité de contrôle : la Commission Nationale Informatique et Libertés (*CNIL*, [8]). Cette dernière a six missions principales : **informer** les personnes sur leurs droits et obligations ; **réguler** le traitement des fichiers ; **sanctionner** les responsables de traitements ne respectant pas la loi ; **protéger** les citoyens ; **contrôler** les fichiers et leurs responsables, et enfin **anticiper** les développements technologiques.

Les données et traitements de données à caractère personnel sont à la base des "lois informatiques et libertés" (ou *LIL*). Ces lois relatives à l'informatique, aux fichiers et aux libertés sont décrites dans la loi numéro 78-17. Elle est composée de 13 chapitres, soit un total de 72 articles. Les trois premiers articles du chapitre 1 ("Principes et définitions") exposent les principes de cette loi. Plus précisément, l'*article 2* ([9]) définit certains termes importants pour la suite :

- **Donnée à caractère personnel** : "toute information relative à une personne physique identifiée ou qui peut être identifiée [...]" ;
- **Traitement de données à caractère personnel** : "toute opération ou tout ensemble d'opérations portant sur de telles données, quelque soit le procédé utilisé". Par exemple : "la collecte, l'enregistrement, l'organisation, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, [...], l'effacement ou la destruction" ;
- **Fichier de données à caractère personnel** : "tout ensemble structuré et stable de données à caractère personnel accessibles selon des critères déterminés." ;
- **Personne concernée par un traitement de données à caractère personnel** : "celle à laquelle se rapportent les données qui font l'objet du traitement".

L'*article 3* ([10]) quant à lui spécifie les obligations du responsable de traitement et les destinataires autorisés pour celui-ci. Le destinataire est une "personne habilitée à recevoir la communication de ces données autre que la personne concernée [...]".

Le second acteur à entrer en jeu lorsqu'il s'agit de la protection de la vie privée est évidemment l'utilisateur. L'utilisateur est une personne quelconque qui utilise un système informatique. Différents types d'utilisateurs existent. Les utilisateurs n'ayant aucune compétence particulière en informatique utilisent un système pour leurs loisirs ou leurs achats sur un site web notamment. C'est cet acteur que les publicitaires essaient de toucher. Il peut adopter des comportements d'utilisateur au sens commercial. Les utilisateurs professionnels sont ceux qui exploitent l'outil informatique en fonction de leur activité technique. L'utilisateur avancé quant à lui connaît le fonctionnement de son système, ses réactions, ainsi que plusieurs de ses limites. Il s'agit en général des personnes travaillant quotidiennement avec les nouvelles technologies. De plus, alors que l'utilisateur système

humain prend en compte le développeur de système ou son administrateur, l'utilisateur système machine est beaucoup plus spécialisé et permet généralement de décharger les utilisateurs professionnels de contraintes, voire de les remplacer dans certains cas. Enfin, afin de permettre l'utilisation d'une ressource, il est nécessaire de recenser des utilisateurs objets (ou fonctions).

Les droits des utilisateurs sont concentrés dans le chapitre 5 Section 1 de la loi du 6 janvier 1978 modifiée. L'*article 32* ([11]) traite du droit d'information, c'est-à-dire le droit de connaître les informations concernant la personne. Malheureusement, dans la pratique, les personnes sont informées lors de la collecte des données et non avant celle-ci. L'*article 38* ([12]) établit le droit d'opposition ou le droit d'être maître de ses propres données personnelles. Cependant, ce droit n'existe pas pour les fichiers du secteur public comme les services fiscaux, la police, la justice, la sécurité sociale ... Le droit d'accès est régi par l'*article 39* ([13]). Il s'agit du droit de consulter ses données personnelles et permet donc à une personne de vérifier l'exactitude des données la concernant. En général, ce droit s'exerce directement auprès de l'organisme gérant les informations. Cependant, comme le précisent les *articles 41* ([14]) et *42* ([15]), le droit d'accès est indirect pour des fichiers de police ou gendarmerie ou encore en matière d'infractions et d'impositions. Le dernier droit des particuliers est le droit de rectification décrit dans l'*article 40* ([16]) et est complémentaire au droit d'accès. Il permet à la personne de rectifier, compléter, actualiser, verrouiller ou faire effacer des données inexactes la concernant.

Lors d'une transaction, d'autres acteurs peuvent entrer en jeu. Les fournisseurs de services par exemple ont un rôle important lors de ces échanges. Effectivement, au sens de la Directive $2004/18/CE$ du Parlement Européen ([17]), *les termes* [...] *"fournisseur" et "prestataire de services" désignent toute* [...] *entité publique ou groupement de ces personnes* [...] *qui offre, la réalisation de travaux et/ou d'ouvrages, des produits ou des services sur le marché*. Il s'agit donc d'une entité physique ou morale capable de fournir des services. Ainsi, les SSII (Société de Services en Ingénierie Informatique), les opérateurs de télécommunications, les entreprises d'assistance et les écoles privées dans l'Union Européenne ([18]) sont considérées comme des prestataires de services. De plus, un individu peut avoir plusieurs identités au cours du temps et en fonction des transactions qu'il réalise. Ainsi, un fournisseur d'identité, comme Microsoft, Google, Verisign, ou tout autre Autorité de Certification (CA) est nécessaire. Effectivement, bien qu'ils fournissent et attestent principalement les identités, ils permettent également le dépôt de données personnelles, parfois sensibles, l'archivage de données de connexion et, par conséquent, l'accès à ces informations. Cependant, certains fournisseurs peuvent identifier une entité

sans pour autant l'authentifier [19]. L'État est un cas particulier de ces fournisseurs d'identité. Bien que la traçabilité des opérations ne doive être possible que par le serveur auquel l'internaute a accepté de se connecter, il peut être demandé à un prestataire de services, dans le cadre d'enquêtes judiciaires, de retrouver l'identité de l'accédant.

Pour finir, un autre acteur, bien qu'indésirable, peut également être présent, l'attaquant. Il s'agit d'un utilisateur malintentionné qui souhaite porter atteinte à une entité. Il entame le bon fonctionnement du système et met en danger les données qu'il contient. La version consolidée du décret 2007 – 451 du 25 mars 2007 ([20]) pour l'application de la loi du 6 janvier 1978 permet de punir les attaquants. La partie législative du Code Pénal s'intéresse également à ces sanctions. Dans le Livre III (*"Des crimes et délits contre les biens"*), Titre II (*"Des autres atteintes aux biens"*), Chapitre III, sept articles concernant les *"atteintes aux systèmes de traitement automatisé de données"* sont édités ([21]). Parmi eux, l'*article* 323 – 1 stipule que : *"Le fait d'accéder ou de se maintenir, frauduleusement, dans tout ou partie d'un système de traitement automatisé de données est puni de deux ans d'emprisonnement et de 30000 euros d'amende"*. De plus, s'il en résulte la *"suppression ou la modification de données"*, la peine s'élève à trois ans d'emprisonnement et 45000 euros d'amende. L'*article* 323 – 2 poursuit sur le fait qu'entraver ou fausser *"le fonctionnement d'un système de traitement automatisé de données* entraine un emprisonnement de cinq ans et une amende de 75000 euros. Enfin, l'*article* 323 – 3 – 1 ajoute que le *"fait, sans motif légitime, d'importer, de détenir, d'offrir, de céder ou de mettre à disposition* [...] *toute donnée* [...] *pour commettre une ou plusieurs des infractions prévues par les articles 323-1 à 323-3 est puni des peines prévues respectivement pour l'infraction elle-même ou pour l'infraction la plus sévèrement réprimée"*. Cependant, ces sanctions ne suffisent pas à empêcher les attaquants de nuire aux utilisateurs et plus précisément à leur vie privée.

Les attaquants ont en effet un large choix de techniques d'attaques et de nombreuses menaces mettent alors en péril la vie privée des individus. Nous abordons ce sujet dans la section suivante.

1.3.2 Menaces contre la sécurité et la vie privée

De nombreuses menaces en termes de vie privée et de sécurité existent et sont répertoriées dans la littérature. L'outil **STRIDE** (Spoofing, Tampering, Repudiation, Information disclosure, Denial of Service (DoS) et Elevation of privilege, [22]) développé par Microsoft permet par exemple de classifier six de ces dangers en termes de sécurité, alors qu'en termes de protection de la vie privée, Mina Deng propose dans sa thèse [23] une modélisation de ces dangers. Elle appuie son approche par un catalogue d'arbres les explicitant

et une cartographie des technologies améliorant la protection de la vie privée (Privacy Enhancing Technologies, ou *PET*). Sa modélisation systématique, nommée **LINDDUN** (Linkability, Identifiability, Non-repudiation, Detectability, Divulgation, Unawareness, Non-conformity) permet de représenter les exigences de la vie privée et de sélectionner les technologies améliorant cette dernière. Onze menaces principales en termes de protection de la vie privée et de sécurité peuvent ainsi être recensées :

1. L'**Usurpation d'identité** est un des risques majeur concernant l'authentification d'un individu. Il s'agit de la possibilité qu'une personne essaie de s'approprier le compte virtuel d'un individu. Cette usurpation d'identité est clairement définie par Lynn M. LoPucki dans son papier [24]. De façon similaire, l'usurpation d'adresse *IP* consiste à envoyer des paquets *IP* avec une adresse non attribuée à l'émetteur. Une solution à ce type de menace est proposée dans le $RFC - 1948$ (pour Request For Comments [25]) ;
2. La **Falsification** des données entraine une modification malveillante des données ;
3. La **Non-répudiation** est le fait qu'une des entités ne puisse pas remettre en cause une action. Notons que la non-répudiation est une menace envers la vie privée, alors qu'il s'agit d'une propriété importante en termes de sécurité comme le montre la Section 1.4 ;
4. La **Divulgation d'information** génère la propagation d'informations à des personnes non autorisées à y avoir accès ;
5. Le **Déni de service** entraine le refus de l'exécution d'un service permettant de valider l'utilisateur ;
6. L'**Élévation de privilèges** permet à un utilisateur non privilégié de se voir attribuer des accès privilégiés ;
7. L'**Associabilité** de deux ou plusieurs points d'intérêts permet à un attaquant de distinguer si ces points sont ou non liés au système ;
8. L'**Identifiabilité d'un sujet** permet à un attaquant de plus ou moins identifier un point du message, comme par exemple, l'expéditeur, le titre, . . . ;
9. La **Détectabilité d'un point d'intérêt** permet à un attaquant de suffisamment bien distinguer si un élément est présent ou non dans le message.
10. L'**Inconscience du contenu** implique que l'utilisateur ne connait pas l'information divulguée au système. S'il fournit trop de renseignements, un attaquant peut récupérer facilement son identité ou bien entraîner une mauvaise compréhension du message et ainsi causer de dangereuses décisions ;
11. La **Non-conformité du consentement et de la politique** qui entraine que, même si le système donne sa politique sur la protection de la vie privée aux utilisateurs, la

conformité de celle-ci n'est pas garantie. Les données de l'utilisateur peuvent donc être révélées sans le savoir.

Grâce aux méthodes d'authentification, d'identification et de contrôle d'accès, certaines de ces menaces peuvent être contrées. Effectivement, l'*usurpation d'identité* se voit contrer par l'**authentification** et le **contrôle d'accès** empêche une personne non autorisée de prendre connaissance de certaines données et rend impossible l'*élévation des privilèges*.

Ces menaces sont rendues possibles grâce aux grands nombres de techniques d'attaques existantes envers les données personnelles des individus. Ainsi, de nombreuses méthodes permettent d'usurper l'identité d'un individu, le **phishing** [26] ou hameçonnage consiste par exemple à aller à la "pêche" aux informations confidentielles. Une personne malveillante se fait passer pour un tiers de confiance afin d'obtenir des informations personnelles et confidentielles. Cette attaque est très répandue chez les particuliers, les attaquants peuvent ainsi voler une identité et notamment les coordonnées bancaires associées. Bien souvent, cette attaque prend deux formes : La première se traduit par la réception d'un e-mail dont la chartre graphique est rigoureusement celle de la banque de l'utilisateur. Cette astuce peut-être évitée en ouvrant uniquement les e-mails en mode "caractère" ; La seconde possibilité est la réception d'un e-mail contenant un lien "piégé" vers le site d'une des banques de l'internaute. Il s'agit d'une *usurpation d'interface* qui peut être évitée en allant directement sur la page de la banque, sans utiliser le lien présent dans l'e-mail.

Les **attaques par rejeu** ([27, 28]) sont quant à elles de type *"homme du milieu"* ("man in the middle" en anglais). Elle consiste à intercepter des communications entre deux entités puis à les rejouer. Ainsi, en retransmettant les paquets au serveur destinataire sans les déchiffrer, l'attaquant peut bénéficier des mêmes droits que l'utilisateur et ainsi avoir accès à ses données. De plus, si le système possède une fonction permettant de modifier le mot de passe, le pirate peut devenir le seul utilisateur du compte.

L'**espionnage** est une autre technique d'attaque, il en existe plusieurs types : le traçage d'individus, le profilage, le ciblage comportemental [29]. L'espionnage regroupe l'ensemble des moyens permettant de suivre et d'analyser le parcours des internautes afin de comprendre leur comportement et d'établir leur profil.Chaque page Internet comporte des marqueurs (ou tags) utilisés par des centaines de sociétés pour faire des statistiques, apprendre la localisation des visiteurs, leurs nombres et leurs actions sur le site, ainsi qu'à des fins publicitaires. De cette façon, la vie privée des utilisateurs devient publique et ceci sans leur consentement et sans qu'ils puissent facilement s'y opposer. Les cookies, les barres d'outils, les traces sont un exemple de moyens permettant cette atteinte à la vie

privée. De plus, les données, bien que personnelles, peuvent être utilisées par des banques lors d'une demande de prêt, par les ressources humaines lors d'un entretien ou encore par la police. Effectivement, toutes sortes de profils peuvent être établis via ces données : des profils psychologiques, de consommateur, socioprofessionnels, culturels, politiques, religieux, etc.... La plate-forme ATOM de Srivastava et Eustace est un exemple marquant de profilage étant donné qu'il permet de convertir un programme en un programme étant son propre espion [30]. Un autre type d'exemple est présenté dans l'article [31] qui propose un profilage à l'aide de l'analyse par dynamique de frappe au clavier.

L'**ingénierie sociale** [32] est une autre technique d'attaque courante sur Internet. Cette manipulation permet de contourner un dispositif de sécurité et se fait par téléphone, mail, courrier. Elle se compose de trois phases : Approche, Mise en alerte, Diversion. Elle entraîne la divulgation en toute conscience des noms, prénoms, adresses, comptes, numéros de cartes bancaires et codes confidentiels, logins/mots de passe, ... de la personne manipulée. Mitnick explicite plusieurs de ces attaques dans [33], alors qu'Hasle, Kristiansen, Kintel et Snekkenes proposent des contre-mesures à celles-ci dans leur livre *Measuring resistance to social engineering* [34].

Une technique d'attaque très connue est l'**injection de codes malicieux** dans un ordinateur : spyware (ou logiciel espion [35]) ou keylogger. Un spyware est un programme introduit dans l'ordinateur afin de collecter des données de l'utilisateur sans son consentement. Comme pour les exemples précédents, il peut permettre de déterminer les centres d'intérêts de l'usager. Cependant, par divers recoupements, il est désormais possible de dresser, à partir de données non confidentielles, des profils complets. Par exemple, une étude, décrite dans [36], a été menée durant une semaine auprès d'étudiants de l'Université de Washington. Celle-ci consistait à analyser leurs comportements afin d'obtenir des signatures qui permettaient ensuite de détecter leur présence sur des ordinateurs distants à l'aide d'une surveillance passive du réseau.

Les **keyloggers** sont des programmes, commerciaux ou non, installés silencieusement sur un poste de travail afin d'effectuer une surveillance invisible et totale. Ainsi, il note dans des fichiers cachés (et compressés) l'ensemble des activités de l'utilisateur et notamment toutes les touches frappées au clavier, et peut également noter la façon de taper au clavier (keystrocke dynamics, [37]). De plus, ces keyloggers peuvent travailler en temps réel ou différé, sur place ou à distance. Ce qui pose de nombreux problèmes en termes de protection de la vie privée et de sécurité. Des couples login/mot de passe peuvent en effet être clairement révélés.

L'étude de la **corrélation entre mots de passe** consiste notamment à examiner l'importance de la relation entre ceux-ci. Étant donné l'habitude des individus à choisir des mots de passe en relation avec leur vie personnelle, il est possible de lier des mots de passe distincts et de différentes bases de données pour ensuite définir s'il s'agit d'un même utilisateur. Le *NIST* (National Institute of Standards and Technology) fournit ainsi ses recommandations au sujet des mots de passe dans le tutorial [38].

Bien qu'un grand nombre de menaces soient connu et compte des contre-mesures, leur prise en compte n'est pas suffisante. Il est nécessaire de définir convenablement la portée de la protection d'une transaction avant de spécifier les propriétés permettant de contrer ces menaces.

1.4 Vers une TES respectueuse de la vie privée

En fonction de la portée de protection et donc du contexte, la notion de vie privée peut varier [39, 23] et comme le rappelle Bruce Schneier [40], cette notion est complémentaire avec la sécurité. Ainsi, la liberté requiert ces deux aspects. Anita Allen [41] déclare quant à elle qu'il existe trois sortes de vie privée : physique (isolement du territoire, solitude), informationnelle (confidentialité, protection de données, contrôle des informations personnelles) et décisionnelle (famille, sexe, religion, santé,...).

La protection de la vie privée et ainsi des données personnelles est particulièrement exposée au développement de nouvelles technologies comme Internet, les réseaux sociaux et plus généralement l'augmentation des bases de données sensibles. Dans ce contexte, de nombreuses règlementations tentent d'assurer la sécurité des systèmes d'information et la protection de la vie privée de l'utilisateur. La structure *International Organization for Standardization* (*ISO*) et le comité *International Electrotechnical Commission* (*IEC*) sont les deux organismes mondiaux spécialisés dans la normalisation. On rappelle qu'on entend par **information personnelle**, toute donnée pouvant identifier la personne à laquelle se rapporte l'information et par **identité**, un ensemble suffisant d'attributs pour identifier l'individu au sein d'une population donnée dont on connait les caractères globaux, la question se pose de la définition d'une **identification**. La résolution 45/95, adoptée par l'Assemblée générale des Nations Unies [42] en est un exemple. Elle présente plusieurs principes relatifs aux fichiers informatisés de données personnelles, dont les quatre suivants :

1. **Principe d'égalité et d'équité** : *Les informations concernant des personnes ne devront pas être collectées ou traitées de manière déloyale ou illicite [...].*

2. **Principe d'exactitude** : *Les personnes responsables de la compilation de fichiers ou de leur tenue ont l'obligation de procéder à des contrôles réguliers sur l'exactitude et la pertinence des données enregistrées [...].*

3. **Principe de définition des objectifs** : *Toutes les données personnelles collectées et enregistrées restent pertinentes par rapport à la finalité spécifique. Aucune de ces données à caractère personnel n'est utilisée ou communiquée [...]. La période de conservation de ces données personnelles ne dépasse pas ce qui permettrait la réalisation de l'objectif.*

4. **Principe de sécurité** : *Des mesures appropriées devraient être prises pour protéger les fichiers contre les risques naturels et humains, tels que l'accès non autorisé, l'utilisation frauduleuse des données ou la contamination par des virus informatiques.*

Avec ce dernier principe, d'autres critères, régis par la norme *ISO* 17799 ([43]) qui définit les principaux critères de sécurité en termes de sécurité, sont mis en place pour contourner les menaces :

1. **Confidentialité** : Garantissant que seules les personnes autorisées ont accès aux données ;

2. **Intégrité** : Assurant l'exactitude du contenu et donc de la non-altération des données durant leur transmission ou leur stockage ;

3. **Disponibilité** : Assurant la possibilité d'accéder aux informations.

Alors que la première propriété permet de contrer le problème de *divulgation d'information*, la seconde résout la menace de *falsification*. Le *déni de service* est quant à lui relié au critère de **disponibilité**.

De plus, les exigences de protection de la vie privée doivent être développées pour des renseignements personnels. Plusieurs règlements *concernant la protection des données personnelles et la vie privée de l'utilisateur* ont été créés, comme par exemple les directives de l'Union européenne (1995, 2002), ou l'article 8 de la Convention européenne de sauvegarde des droits de l'homme et des libertés fondamentales [44, 45, 46]. Cependant, les protections standards assurant la sécurité des systèmes d'information ne suffisent pas, il faut développer des exigences pour la vie privée afin de protéger les renseignements personnels. Ainsi, trois principes relatifs à la vie privée sont développés :

1. Le **principe de minimisation** des données (également appelé *contrôle utilisateur et consentement*) indique que la *divulgation de données personnelles doit être limitée à des données adéquates, pertinentes et non excessives.*

2. Le **principe de souveraineté** des données stipule que les données personnelles appartiennent à un individu, avec un contrôle et un consentement sur la façon dont ces données sont utilisées et à quelles fins.
3. Le **principe de sensibilité** des données stipule que les données personnelles traitées sont considérées comme sensibles et nécessitent une structure décentralisée pour leur stockage.

Ces deux premiers principes sont détaillés par Deswarte et Gambs pour la carte d'identité nationale dans [47], ainsi que dans le projet PRIME (*PRivacy and Identity Management for Europe*) dans [2] dont le but était d'élaborer un cadre pour une gestion d'identité protégeant la vie privée. En novembre 2010, la Commission européenne *a examiné les moyens de renforcer le principe de minimisation des données* [48]. Nous remarquons que le principes d'égalité et d'équité, et le principe d'exactitude définis par les Nations Unies sont pris en compte par les principes de minimisation des données et de sensibilité des données. Le principe de définition des objectifs est quant à lui similaire aux principes de souveraineté des données.

La classe FPR (protection des informations personnelles) dans les exigences fonctionnelles des Critères Communs établit des Politiques de Sécurité des Systèmes d'Information ($PSSI$) [49]. Cette classe décrit quatre contraintes : **anonymat**, **pseudonymat**, **intraçabilité** et **non-observabilité**. Ces deux premiers principes, ainsi que le principe de minimisation des données, sont également discutés dans le rapport technique de Pfitzmann et Hansen [50] et par Cameron dans [51] :

1. L'**intraçabilité** d'une ou plusieurs valeurs signifie qu'il n'est pas possible de distinguer de manière suffisante ces données. Cette notion est définie par la norme ISO 15408 [49].
2. L'**anonymat** d'une entité signifie que celle-ci n'est pas identifiable au sein d'un ensemble d'entités. Ce critère est également défini par la norme ISO 15408 [49].
3. Le **pseudonymat** est synonyme d'intraçabilité et nécessite l'utilisation d'un pseudonyme.
4. La **non-observabilité** garantit à un individu l'utilisation d'une ressource (ou d'un service) sans que des tiers soient en mesure d'observer la ressource utilisée. Il n'est donc pas possible de déterminer si une opération spécifique a été effectuée.

Ces quatre principes entrainent une limitation des attaques envers les données personnelles d'une entité et notamment celles citées dans [23]. En effet, la menace d'*identifiabilité* est contrée par l'**anonymat** et le **pseudonymat**. Ils sont notamment utilisés pour les architectures de gestion d'identité préservant la vie privée. Plus précisément, l'anonymat

assure que l'utilisateur peut accéder à une ressource, sans révéler son identité, alors que le pseudonymat exige que cet utilisateur soit responsable de cette utilisation. Cela signifie que, dans des cas exceptionnels, l'identité doit être récupérée par des personnes autorisées. Dans un tel contexte, une fonction de chiffrement réversible (voir la Section 2.2.1) est utilisée uniquement par des personnes de confiance et connaissant la clé secrète. Un système de chiffrement symétrique de type AES ([52]) est un exemple adapté pour le pseudonymat des données. Des exemples de systèmes de gestion d'identité centrés sur l'utilisateur sont Liberty Alliance [53] ou Idemix [54], développé par Camenisch et Van Herreweghen et basé sur l'identification anonyme [55, 56]. La *détectabilité* des données est quant à elle évitée grâce aux propriétés d'**intraçabilité** et de **non-observabilité**. La notion d'intraçabilité garantit que les données personnelles sont protégées contre la procédure d'agrégation. Cette notion est liée à l'anonymat étant donné que la traçabilité des données pourrait permettre de récupérer l'identité d'un individu. En outre, le principe d'intraçabilité doit prendre en compte la possibilité qu'un attaquant connaisse certaines informations. Par exemple, la date de naissance du patient peut être récupérée par des moyens tels que les réseaux sociaux. De plus, dans un cas particulier où les données de différentes organisations doivent être regroupées, le principe d'intraçabilité ne serait pas élémentaire.

Rappelons que Mina Deng observe également deux autres problèmes en termes de protection de la vie privée : l'*inconscience du contenu* et la *non-conformité du consentement et de la politique*. Ces menaces sont logiquement gérées par la conscience du contenu, par une politique et par un consentement conformes, et par conséquent par le principe de souveraineté des données. De plus, comme constaté précédemment, l'auteur considère la non-répudiation comme un problème et contre ainsi la possibilité d'avoir un déni.

Le Tableau 1.1 décrit les relations mises en évidence dans cette partie entre les propriétés et leurs menaces liées à la vie privée. Le Tableau 1.2 quant à lui fait le lien entre les critères de sécurité et leurs dangers associés.

D'autres propriétés peuvent également être attendues en fonction du contexte. Entre autre, un médecin doit pouvoir **accéder en temps réel** à la fiche d'un de ses patients. Dans un autre contexte, supposons qu'un notaire utilise des signatures électroniques pour faire signer le testament de ses clients, notamment les très jeunes clients, il est impératif que cette signature soit **valable sur du long terme**.

Propriété liée à la vie privée	**Menace correspondante**
Déni possible (Répudiation)	Non-répudiation
Souveraineté des données	Non-conformité du consentement
	Non-conformité de la politique
	Inconscience du contenu
Minimisation des données	Associabilité
	Détectabilité
Anonymat et Pseudonymat	Identifiabilité

TABLE 1.1: Relation entre propriétés et menaces relatives à la vie privée

Propriété liée à la sécurité	**Menace correspondante**
Authentification	Usurpation d'identité
Intégrité	Falsification
Confidentialité	Divulgation d'information
Contrôle d'accès	Élévation des privilèges
Non-répudiation	Déni possible (Répudiation)
Disponibilité	Déni de service

TABLE 1.2: Relation entre propriétés et menaces de sécurité

En ce qui concerne les informations personnelles en provenance d'organismes divers, leur **recoupement doit être impossible**. Effectivement, la traçabilité des opérations ne doit être possible que pour le site ou serveur ayant le consentement de l'internaute. Sans cette propriété, une relation entre un pseudonyme et des transactions effectuées auprès de différents prestataires de services peut être élaborée. De plus, dans le cadre d'enquêtes judiciaires ou médicales, un prestataire de services peut se voir demander l'identité d'un antécédent enregistré. La **réversibilité d'une information** doit donc être réalisable dans des cas de force majeure. Afin de permettre des travaux statistiques sur des bases de données personnelles, par exemple médicales, des **considérations d'anonymisation** peuvent être prises en compte.

Cette partie décrit ainsi les différents types de propriétés permettant d'inquiéter les menaces et les contraintes liées aux applications.

1.5 Objectifs de la thèse

L'objectif principal de cette thèse est de proposer de nouvelles architectures centrées sur l'utilisateur veillant à la protection de ses données personnelles. Diverses architectures de domaines variés ont été étudiées.

Dans un premier temps, lorsqu'on parle de protection des données sur Internet, les nombreuses informations à révéler aux fournisseurs de service pour s'enregistrer en ligne nous viennent à l'esprit. Nous avons donc choisi de commencer par étudier la gestion des données d'un utilisateur sur Internet afin de le guider lors de telles manipulations. Nous proposons ainsi un logiciel facile d'utilisation pour l'internaute et respectant les exigences attendues en termes de vie privée et de sécurité pour de tels systèmes.

Après avoir appris qu'"*un médecin du CHU de Caen s'est fait hospitaliser et que tous ses collègues ont eu accès à son dossier sans autorisation du patient*", nous avons ensuite décidé de nous concentrer sur le domaine hospitalier. De plus, alors que le *DMP* (Dossier Médical Patient) est en train d'être mis en place, nous devons nous interroger sur les faiblesses et sur les améliorations possibles d'un tel système. Nous proposons ainsi une nouvelle architecture de e-santé permettant de gérer les dossiers des patients aussi bien à l'intérieur d'un établissement de santé avec une gestion des accès différente en fonction des employées, qu'entre établissements de santé lors d'un transfert de dossier ou en cas d'urgence.

Pour finir, nous proposons une troisième architecture en relation avec le commerce électronique en plein essor depuis quelques années. Au vu des nombreux inconvénients et failles des systèmes de paiement en ligne actuelle, nous avons décidé de proposer une architecture de paiement en ligne résolvant l'ensemble de ces problèmes. Cette dernière architecture permet de faire un achat sur Internet de façon sécurisée et parfaitement respectueuse des données personnelles du client, notamment des données bancaires. Pour chacune de ces architectures, nous avons définis un ensemble d'exigences de sécurité et de protection des données personnelles spécifiques au contexte et donc à la nature des informations échangées. Nous analysons à l'aide de ces propriétés certaines solutions existantes dans la littérature, ainsi que notre architecture, afin de procéder à une comparaison de celles-ci. Nos différentes architectures sont enrichies avec des technologies et protocoles cryptographiques, ainsi que par l'usage d'éléments sécurisés.

1.6 Conclusion

Ce chapitre expose la problématique et formalise les notions de vie privée et de transactions électroniques sécurisées. De nombreux problèmes menacent ces propriétés et par conséquent la vie privée des individus. L'usurpation d'identité, l'espionnage informatique, l'ingénierie sociale et l'insertion de codes malicieux ne sont qu'une courte liste d'attaques possibles. Il est donc impératif de les contrer et de définir des exigences essentielles à la sécurité et à la protection des données personnelles. Différentes propriétés relatives aux TES ont ainsi été énoncées dans ce chapitre. Cependant, toutes n'ont pas la même importance et certaines dépendent du contexte. Dans cette thèse, certaines propriétés de sécurité et de protection des données personnelles sont toujours considérées. Concernant la sécurité, il s'agit des propriétés suivantes : confidentialité des données, contrôle d'accès aux informations, anonymat ou pseudonymat en fonction des cas d'usage, ainsi que l'authentification des différents acteurs. Concernant la protection des données personnelles, les principes essentiels sont les principes de minimisation des données, de souveraineté des données et de sensibilité des données. D'autres critères peuvent également être pris en compte lors d'une TES : l'intégrité des données, la non-observabilité, l'intraçabilité, ainsi que la non-répudiation de l'expéditeur. Sans oublier que toutes les données doivent pouvoir être contrôlées par leur utilisateur. Finalement, cinq propriétés potentiellement importantes en fonction des applications traitées sont rapportées dans ce chapitre : l'accès en temps réel à des données, la vérifiabilité au long terme d'une information, l'impossibilité de recoupements entre fichiers, la réversibilité d'une information, ainsi que l'anonymisation des bases de données.

Le chapitre suivant s'intéresse aux technologies et protocoles existants dans l'état de l'art pouvant contribuer à la protection des données personnelles. Dans la suite de cette thèse, les éléments sécurisés et protocoles cryptographiques déterminés sont utilisés pour mettre en place trois architectures respectueuses de la vie privée dans trois domaines d'applications différents : l'enregistrement en ligne d'un client, la gestion des dossiers médicaux au sein des hôpitaux et entre hôpital, ainsi que le paiement d'un achat sur Internet. Parmi les services mis en place, certains pourront être réalisés à l'aide d'un $HSM\ BULL$: le stockage de secrets ou encore certains calculs cryptographiques comme ceux nécessaires à une authentification zero-knowledge ou à l'utilisation de biométrie révocable.

Chapitre 2

Technologies disponibles pour la protection des données personnelles

Ce chapitre se concentre sur les solutions existantes dans la littérature en termes de protection de la vie privée lors d'une transaction électronique sécurisée. Des protocoles cryptographiques assurant la confidentialité des informations et/ou garantissant une divulgation minimale des données y sont présentés, ainsi que les propriétés qu'ils assurent.

Sommaire

2.1 Introduction

Le chapitre 1 a pu présenter les nombreux enjeux et propriétés liés à la protection des données personnelles. Nous nous concentrons dans ce deuxième chapitre sur les mécanismes permettant de protéger ces données à caractère personnel. Comme l'illustre la Figure 2.1, nous détaillons différents groupes de protocoles. Le premier groupe est explicité dans la première partie de ce chapitre. Il concerne les protocoles cryptographiques classiques : les chiffrements asymétriques, symétriques, les protocoles de signatures numériques, les fonctions de hachage, ainsi que les infrastructures à clés publiques et les

certificats. La deuxième section de ce chapitre présente les mécanismes dédiés à la protection de la vie privée. Il s'agit par exemple des preuves de connaissance sans divulgation de connaissance, ainsi que des protocoles spécifiques de signatures, comme la signature aveugle ou des systèmes d'accréditations anonymes. La troisième partie décrit les outils permettant de stocker ces données sensibles, c'est-à-dire des éléments sécurisés tels que les cartes à puce ou les *HSM*.

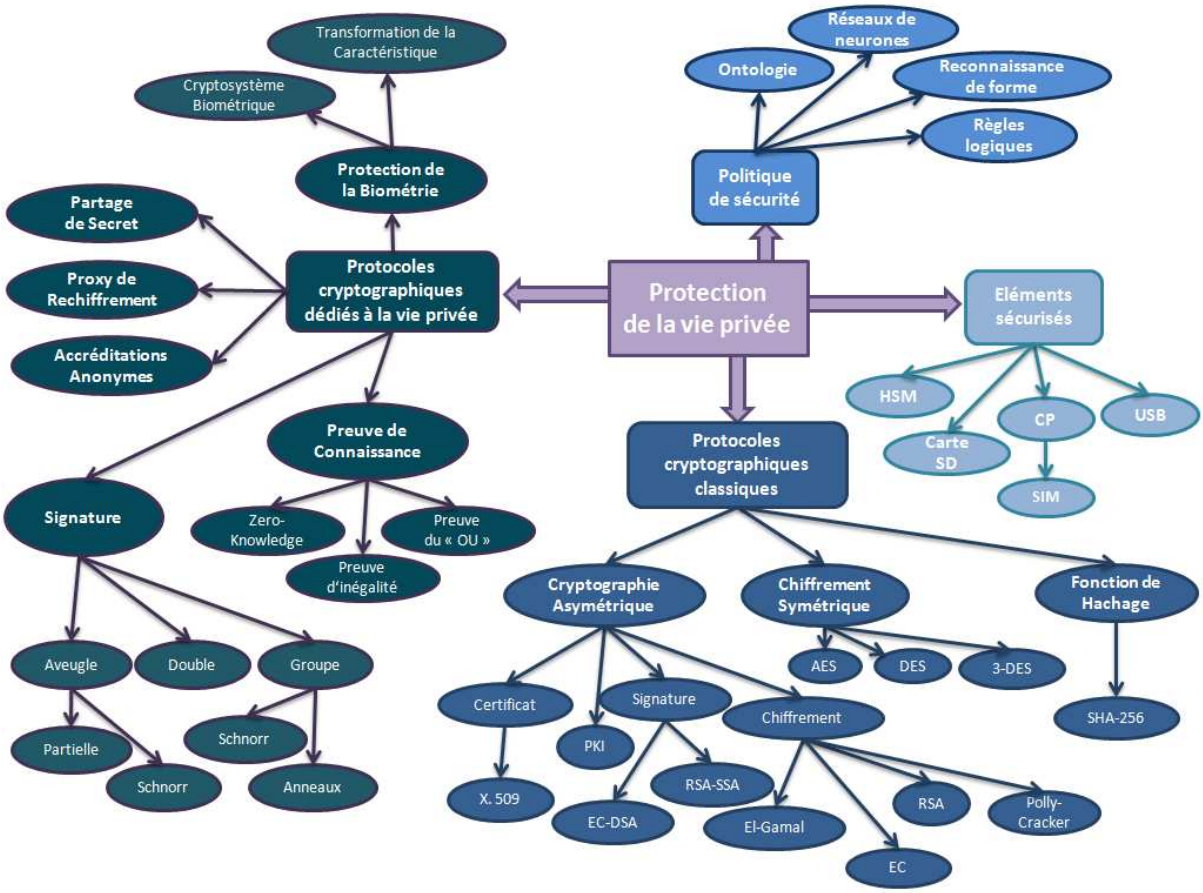


FIGURE 2.1 – Schéma des différents mécanismes et outils pour la protection de la vie privée

2.2 Protocoles cryptographiques classiques

Le chiffrement est le procédé cryptographique permettant de transformer un message afin de le rendre incompréhensible par un tiers. Le processus inverse est appelé le déchiffrement. Ces protocoles assurent la confidentialité des données. En France, l'usage du chiffrement a longtemps été considéré comme arme de deuxième catégorie, c'est-à-dire comme un matériel destiné à porter ou utiliser les armes de guerre, au même titre que les chars, les avions ou les navires. La loi pour la confiance dans l'économie numérique du 21

juin 2004 et consolidée le 11 juillet 2010 a enfin permis de libéraliser l'utilisation de la cryptologie [57]. Nous nous intéressons principalement dans cette section aux cryptosystèmes symétriques et asymétriques, ainsi qu'aux normes françaises et internationales les concernant.

2.2.1 Cryptographie symétrique

Généralités

Les systèmes symétriques sont également appelés systèmes à clés secrètes. Ils utilisent la même clé pour chiffrer et déchiffrer un message. Ainsi, cette clé est confidentielle et uniquement connue de l'expéditeur et du destinataire. Trois algorithmes symétriques sont principalement utilisés : le DES, le $Triple-DES$ et l'AES.

Le DES (Data Encryption Standard, [58]) repose sur le schéma de Feistel, une succession de permutations et de substitutions. La longueur de sa clé (56 bits) est désormais trop courte pour assurer un très bon niveau de sécurité. L'$ANSSI$ (Agence Nationale de la Sécurité des Systèmes d'Information) recommande en effet et ceci jusqu'en 2020, des clés d'au moins 100 bits ([59]) pour un chiffrement symétrique. Le DES n'est donc plus conforme aux recommandations de l'$ANSSI$. Afin de pallier aux faiblesses de longueur des clés du DES, le $Triple-DES$ ([58]) a été inventé. Il est composé de deux chiffrements DES de même clé, séparés par un déchiffrement DES utilisant une autre clé.

La clé $Triple-DES$ est donc composée de deux clés DES de 56 bits, c'est-à-dire de 112 bits. Il est hors de portée d'une attaque exhaustive. Cependant, bien qu'il soit encore utilisé dans le domaine commercial et bancaire, et qu'il ait une taille de clé supérieure à 100 bits, il ne suit pas les recommandations de l'$ANSSI$.

Ainsi, afin de trouver un successeur au $Triple-DES$ qui se révéla lent et moins sûr, le $NIST$ (National Institute of Standarts and Technology, [60]) lance un appel à projet. L'AES (Advanced Encryption Standard, [52, 61]) voit le jour en 2000. Ce nouvel algorithme peut être utilisé avec des tailles de clés différentes : 128, 192 et 256 bits. De plus, il est capable de chiffrer des blocs de 128 bits. Ce chiffrement est une succession de tours semblables. Le nombre de tours est calculé en fonction de la taille de la clé. De plus, l'AES décrit dans le $FIPS-197$ ([52]) est un mécanisme de chiffrement par bloc conforme pour l'$ANSSI$.

Propriétés des chiffrements symétriques

Afin de garantir la *confidentialité* des informations lors d'un chiffrement par bloc, il est nécessaire de choisir un mode opératoire non déterministe. Cela permettra d'éviter que plusieurs chiffrements d'un même message ne donnent le même chiffré. Il faut donc

introduire une valeur aléatoire, c'est-à-dire "randomiser" l'algorithme. Le mode Cipher-Block Chaining (CBC, [62]) est le plus connu. De plus, aucune attaque n'étant connue pour l'AES et le $Triple-DES$, ils vérifient tous deux une des propriétés fondamentales pour le chiffrement : la *non-inversibilité*, bien qu'elle ne soit pas suffisante pour des protocoles déterministes.

Les propriétés d'*authentification* et d'*intégrité* des messages sont quant à elles assurées par le calcul d'un Code d'Authentification de Message (MAC pour Message Authentification Code) et plus précisément grâce au mode d'authentification $CMAC$ décrit par le NIST dans [63].

2.2.2 Les fonctions de hachage

Ces fonctions produisent un condensat appelé empreinte (ou haché) de taille fixe et réduite à partir d'un message arbitrairement long. Elles peuvent donc jouer un rôle très important en authentification et en signature. Une fonction de hachage h doit respecter plusieurs propriétés. Premièrement, elle doit être à **sens unique** (ou **non-inversible**), c'est-à-dire que, pour toute empreinte y, il doit être difficile de trouver un message x tel que : $h(x) = y$. Deuxièmement, elle doit être **(fortement) résistante aux collisions** ou **sans collision**, il est alors difficile de trouver deux messages x et x' ayant un condensat identique : $h(x) = h(x')$. Il arrive cependant que certaines fonctions soient uniquement **faiblement résistantes aux collisions**, dans ce cas, pour (essentiellement) chaque message x, il est difficile de trouver un message $x' \neq x$ ayant une même empreinte. Ainsi, cette dernière propriété permet d'éviter l'attaque suivante : *Une personne malveillante en possession d'un message x, de son empreinte et de sa signature s peut calculer un message x' de même condensat que x. Ainsi, elle est capable de faire croire que x' est authentique en l'accompagnant de s*. Cependant, cette propriété ne protège pas des attaques génériques (ou **paradoxe des anniversaires**), d'où l'intérêt de la propriété *sans collision*.

L'$ANSSI$ recommande l'utilisation du mécanisme de hachage $SHA-256$ qui est défini dans $FIPS\ PUB\ 180-2$ [64] et déconseille fortement le précédent algorithme $SHA-1$. Effectivement, il est impératif qu'une empreinte générée par une fonction de hachage soit d'au moins 200 bits. De plus, il faut que la meilleure attaque permettant de trouver des collisions nécessite au minimum $2^{n/2}$ calculs d'empreintes. De plus, depuis octobre 2012, $SHA-3$ est la nouvelle fonction de hachage destinée à remplacer $SHA-2$ dans le cas où cette dernière serait compromise par une attaque significative [65]. Bien que le problème de la taille du message soit résolu par les fonctions de hachage, il reste celui de la vérification d'une signature. La création d'une autorité de confiance est une solution.

2.2.3 Cryptographie asymétrique

Généralités

Les systèmes asymétriques sont des algorithmes à clé publique. Ils nécessitent une clé privée, possédée uniquement par le receveur et une clé publique. Ainsi, une clé différente est utilisée pour chiffrer et déchiffrer le message. Les principaux algorithmes asymétriques connus sont RSA, *ElGamal* et les courbes elliptiques. Cependant, il en existe d'autres, comme ceux basés sur les codes correcteurs d'erreurs ou sur des problèmes NP-Complet.

RSA (Rivest Shamir Adleman) est l'algorithme le plus utilisé de nos jours ([66]). Il est présent dans les réseaux télécommunications, les réseaux Ethernet, des logiciels, cartes à puce mais également dans les industries. Dans le cas de RSA, les clés sont généralement de 1024 bits, sachant que nous trouvons de plus en plus des tailles de 2048 bits. En effet, bien qu'aucun module de 1024 bits n'est encore été factorisé en 2007, l'article [67] détaille la factorisation d'un module de 1039 bits. Ainsi, l'$ANSSI$ ([59]) recommande la prudence en considérant des clés de 2048 ou 4096 bits. Cependant, cet algorithme est basé sur la difficulté à factoriser des grands entiers premiers et la puissance croissante des supercalculateurs le rende vulnérable. Cependant, déchiffrer des messages chiffrés par RSA revient à extraire des racines nièmes modulo un grand nombre (produit de deux grands nombres premiers). Afin de rendre difficile la factorisation du grand produit ($N = p.q$, appelé module), l'$ANSSI$ recommande d'utiliser des exposants publiques strictement supérieur à $2^{16} = 65536$ et de respecter un certain nombre de conditions disponibles dans [59].

Un autre chiffrement asymétrique est le chiffrement Polly Cracker ([68]) basé sur un système NP-Complet : le problème d'appartenance à un idéal.

Afin d'éviter la détermination de la base de Gröbner de l'idéal choisi, il est important de choisir un idéal dont la base soit très difficile à trouver [69]. D'autre part, la cryptanalyse différentielle a donné d'excellents résultats et a ainsi rendu l'utilisation de Polly Cracker inefficace dans la plupart des cas. L'algorithme est alors considéré comme sûr avec une taille de clé privée de $2^{40} \sim 10^{12}$ bits et de clé publique de $2^{80} \sim 10^{24}$. Ainsi, l'idée de Fellows et Koblitz de se baser sur un problème NP-complet est excellente, mais dans ce cas, elle ne fournit pas un algorithme fonctionnel pour des tailles de clés raisonnables (ne dépassant pas 300 000 bits). A contrario, l'utilisation de Polly Cracker peut être pertinente si la durée de protection des données est courte.

Afin de rendre le problème du logarithme discret, réputé difficile, suffisamment sûr, il est nécessaire de faire attention aux choix du module. L'$ANSSI$ ([59]) recommande des sous-groupes d'ordre premier et conseille *"de privilégier des primitives à clé publique ne se fondant pas sur le problème du logarithme discret dans* $GF(2^n)$*"*. Cependant, pour des raisons de sécurité (et non de temps de calcul), il est préférable de travailler avec des corps

finis de type $GF(p)$. ElGamal ([70]) est un algorithme de chiffrement à clé publique dont la sécurité est basée sur le problème du logarithme discret.Ce système nécessite le choix d'un grand nombre premier p afin que la sécurité de cet algorithme soit assurée. En ce qui concerne la taille des clés, elle équivaut à celle du système RSA. De plus, afin de contrer la méthode de Pohlig-Hellman ([71]) de calcul du logarithme discret, $p-1$ doit avoir un grand facteur premier.

Les courbes elliptiques sont quant à elles des objets mathématiques permettant de faire également des opérations cryptographiques, comme des échanges de clés sur canal non-sécurisé ou du chiffrement asymétrique. Ce dernier est généralement appelé ECC pour Elliptic Curve Cryptography et existe depuis 1985 ([72, 73]). La sécurité de ce cryptosystème est également assurée par le problème du logarithme discret. Les corps de base pour cette méthode sont les mêmes que pour ElGamal, soit $GF(p)$ et $GF(2^n)$ où p et n doivent être premiers. De même, il est nécessaire d'employer des sous-groupes d'ordre multiple d'un nombre premier d'au moins 200 bits, voir d'ordre premier directement.

Les avantages de ce cryptosystème sont en premier lieu sa rapidité de chiffrement d'un message et le recours à des clés de tailles nettement plus petites (220 ou 240 bits, contre 1024 ou 2048 pour RSA). De plus, alors que la taille des clés RSA double, celle du chiffrement par les courbes elliptiques augmente de 20 bits uniquement. Enfin, les calculs de ce cryptosystème sont plus simples et plus rapides que ceux des autres algorithmes asymétriques.

Jacques Stern oppose ces différents cryptosystèmes dans [74] et arrive à différentes conclusions. Par rapport à RSA, les courbes elliptiques sont plus simples car elles ne font pas appel au groupe multiplicatif. De plus, elles sont plus sûres étant donné qu'aucun algorithme connu ne permet de résoudre le logarithme discret alors que ce n'est pas prouvé pour la sécurité de RSA. Ainsi, RSA résiste à de nombreuses attaques variées sous l'hypothèse qu'on ne puisse pas l'inverser. Stern traite également de l'Optimal Asymmetric Encryption Padding ($OAEP$) de Bellare et Rogaway ([75]). Celui-ci est un schéma de remplissage souvent utilisé par RSA et inspiré des réseaux de Feistel ([76]) qui utilise une source d'aléa et deux fonctions de hachage (Section 2.2.2).

Propriétés des chiffrements asymétriques

Contrairement aux algorithmes symétriques, la sécurité des chiffrements asymétriques est basée sur des problèmes difficiles. De plus, de la même manière que pour les algorithmes symétriques, la propriété de *confidentialité* est assurée par les cryptosystèmes à clé publique. Cependant, il est encore important de les *randomiser*. L'*authentification des entités* est possible en utilisant un protocole de signature ou des protocoles de "challenge-réponse" : une des entités demande à l'autre de signer un message avec sa clé secrète. La

deuxième peut vérifier qu'elle la possède bien en utilisant la clé publique de la première. Si elle retrouve le message, il s'agit bien de la bonne personne. Cette dernière peut par exemple se faire par des mécanismes "à divulgation nulle de connaissance" décrits dans la Section 2.3.1. Enfin, les *échanges de clés* peuvent être gérés par l'**échange de Diffie-Helmann** ([77]). L'*intégrité* des messages et la *non-répudiation* peuvent être garanties par des signatures cryptographiques basées sur ces mécanismes asymétriques (voir la Section 2.2.3). Leur authenticité est ainsi validée par la clé publique du signataire.

Le Tableau 2.1 récapitule les caractéristiques et propriétés assurées par les algorithmes de chiffrements symétriques et asymétriques.

Propriétés	Symétriques	Asymétriques
Confidentialité	✓	✓
Non-inversibilité	✓	✓
Authentification (messages)	✓	✓
Authentification (entités)	✓	✓
Intégrité	✓	✓
Non-répudiation		✓

TABLE 2.1: Tableau récapitulatif des caractéristiques et propriétés des algorithmes symétriques et asymétriques

La signature numérique

La cryptographie ne se limite plus à l'art de chiffrer, elle effectue désormais de nouvelles tâches. Cette science permet entre autres la signature numérique d'un document. En France, depuis mars 2000, elle est considérée comme ayant la même valeur légale que la signature manuscrite et est régie par la loi n°2000 – 230 du 13 mars 2000 sur l'*adaptation du droit de la preuve aux technologies de l'information et relative à la signature électronique* [78] et par son décret d'application n°2001 – 272 du 30 mars 2001 [79].

Deux types de schémas de signature existent dans la pratique. Les schémas avec restauration du message offrent une signature permettant de retrouver le document en garantissant son authenticité. Ainsi, bien que la signature soit nécessairement plus longue, il est inutile d'ajouter le message lors de l'envoi. Il est aussi possible d'avoir des schémas en appendice. Dans ce cas, afin de vérifier la signature, il faut être en possession de celle-ci et du message signé. Généralement, la signature est une donnée beaucoup plus courte que la valeur qu'elle authentifie. La signature numérique doit remplir certaines conditions :

1. **Authenticité** : on doit pouvoir retrouver de manière certaine l'identité du signataire ;
2. **Infalsifiabilité** : il est impossible de se faire passer pour un autre ;
3. **Non-réutilisabilité** : la signature fait partie du document signé, elle ne peut être réutilisée pour un autre message ;
4. **Inaltérabilité** : une fois le document signé, on ne peut plus le modifier ;
5. **Irrévocabilité** : la personne ayant signé le document ne peut le nier.

De nombreux protocoles de signature existent. La signature RSA est basée sur la factorisation des nombres premiers. Étant attaquable par contrefaçon existentielle, celle-ci s'est vu ajouter une gestion des clés publiques par certificats (voir 2.2.3) et des schémas de redondance. Ces schémas protègent également cette signature contre des inconvénients dus à la propriété multiplicative de RSA. Ils définissent la façon de formater les messages à signer avant d'appliquer une exponentiation RSA. Un des plus connus est celui défini par la norme ISO 9796 [80]. L'autre problème présent dans les algorithmes de signature est le problème du logarithme discret. Ceux-ci sont définis par la norme ISO/IEC 9796 – 3 : 2006 [81]. La signature El-Gamal est basée sur celui-ci. Elle implique de vérifier certaines conditions lors de la réception de la signature afin de ne pas être dupée par un attaquant. Pour une sécurité identique et avec les mêmes conditions de vérification, la signature Digital Signature Standart (DSS, ancien DSA, [82]) est une amélioration de la signature précédente car son résultat est plus court. Une variante de cette signature est le protocole $EC-DSA$ ([82]), c'est-à-dire une signature DSA adaptée aux courbes elliptiques. Enfin, le protocole considéré comme étant le plus simple des algorithmes de signature prouvés sûrs dans un modèle d'oracle aléatoire est le protocole de Schnorr ([83]). Il est en effet efficace et gère des signatures courtes.

Afin d'avoir un bon niveau de sécurité, certaines recommandations du $NIST$ et des courbes spécifiques sont indiquées dans [84]. La norme [81] indique également six autres schémas de signatures numériques de récupération de données : l'algorithme Nyberg-Rueppel (NR, [85]) travaillant sur les corps finis et les processus Elliptic Curve Nyberg-Rueppel ($ECNR$, [85]), Elliptic Curve Abe-Okamoto ($ECAO$, [86]), Elliptic Curves Message Recovery ($ECMR$), Efficient Certificate Path Validation ($ECPV$, [87]) et Elliptic Curve KCDSA-Nyberg-Rueppel ($ECKNR$) qui sont définis sur une courbe elliptique sur corps finis. L'$ANSSI$ (Agence Nationale de la Sécurité des Systèmes d'Information) considère quand à elle les schémas de signature asymétrique suivants comme conformes : le schéma RSA-SSA-PSS (RSA Signature Scheme with Appendix – Provably Secure encoding method for digital Signatures, Section 8.1 de [88]) sous certaines conditions indiquées dans [59], ainsi que $EC-DSA$ si l'une des courbes est : $P-256$, $P-384$, $P-521$, $B-283$, $B-409$ ou $B-571$ (respectivement dans [84] pages 8-10, 17-18 et 21).

Propriétés des signatures numériques

Étant donné qu'il est impossible de modifier un message signé, les procédés de signature permettent l'*intégrité des données*. De plus, leur premier usage est de permettre au destinataire de connaître de façon certaine l'identité de l'expéditeur. L'*authentification du signataire* est donc assurée et une *répudiation* de la part de l'auteur du message évitée.

Cependant, les protocoles de signature ne garantissent pas la *confidentialité* du message, ni son *anonymat*. De plus, ils ne permettent de signer généralement que des tailles de données restreintes (typiquement 128 ou 160 bits) impliquant, dans le cas de données plus longues, le découpage du message en blocs puis leurs signatures successives.

Infrastructure à clés publiques

Une infrastructure à clé publique (*PKI* pour Public Key Infrastructure), décrite par la norme $ISO/IEC\ 9594-8:2008$ ([89]), permet de gérer la distribution des clés publiques et est composée des trois entités distinctes. L'**autorité d'enregistrement** gère les certificats et contrôle l'identité de l'utilisateur final ; l'**autorité de validation** vérifie les informations des certificats, ainsi que le **répertoire** qui stocke leahs certificats numériques et les listes de révocation. Le principe de l'autorité de certification a été introduit par Kohnfelder en 1978 [90]. Il est utilisé pour attester de la liaison entre la clé publique et un identifiant. Par conséquent, une signature numérique de l'autorité de certification est utilisée pour relier une clé publique au titulaire de cette clé. Ainsi, en prenant la clé publique de l'autorité de certification, tout acteur de la *PKI* peut vérifier le lien entre une identité et une clé de n'importe quel certificat. En France, une *PKI* est également composée d'une autorité de séquestre qui stocke de manière sécurisée les clés de chiffrement générées par la *PKI* afin de permettre aux autorités de déchiffrer, si nécessaire, les données chiffrées lors d'une création de certificat.

Certificat numérique

Le certificat numérique est un objet numérique permettant d'associer à une entité (client ou serveur) sa clé publique. Il s'agit d'un document électronique signé qui permet de prouver l'identité de l'entité auprès d'un tiers de confiance, appelé Autorité de Certification.

Pendant son cycle de vie, un certificat est géré par une *PKI*. Cette infrastructure se sert de mécanismes de signature afin de certifier des clés publiques qui permettent de chiffrer et de signer des messages ou des flux de données. En matière de normes, les certificats doivent respecter de manière rigoureuse certains standards. La principale norme de définition de ces certificats est la norme *X*.509.

2.2.4 Discussion

Nous venons de décrire de nombreux protocoles cryptographiques classiques permettant d'assurer la sécurité des transactions, notamment la confidentialité des messages, ainsi que l'authentification des entités. Cependant, nous ne pouvons nous contenter de ces propriétés et donc de ces protocoles. Effectivement, la confidentialité d'un message ne suffit pas à protéger les données personnelles d'un individu. De plus, même si le fait de garantir l'identité de l'utilisateur est rassurant, cela peut entraîner une divulgation d'information trop importante. Certains mécanismes ont alors été créés afin de protéger spécifiquement les données personnelles des individus.

2.3 Protocoles cryptographiques de protection de la vie privée

2.3.1 Preuves de connaissance

Motivations

Le développement de nouvelles technologies, notamment dans le domaine des télécommunications, a fait prendre de l'ampleur aux procédés d'authentification. Ces derniers font tous intervenir une donnée secrète, connue uniquement de la personne autorisée, qui entraine alors l'apparition de nombreux protocoles appelés preuves de connaissance (*proof of knowledge*, en anglais). Certains utilisent un unique mot de passe, d'autres deux, comme par exemple l'authentification par signature et par déchiffrement. En cryptographie, ces preuves de connaissance particulières sont appelées des preuves d'identités. Il s'agit d'une généralisation des protocoles d'authentification.

Protocoles Zero-knowledge

Un nouveau concept permettant l'authentification et l'identification est développé en 1985 par Goldwasser et coll. [91] : les protocoles Zero-Knowledge (ou ZK) ou preuves à divulgation nulle de connaissance [92]. Il est désormais possible de prouver qu'un mot appartient à un langage sans révéler ce mot. Ces protocoles sont basés sur une entité, le Prouveur, qui souhaite prouver mathématiquement à une autre entité, le Vérifieur, la véracité d'une proposition sans jamais révéler cette propriété secrète. La norme ISO/IEC 9798 − 5 : 2009 [93] explicite ces mécanismes. Certains sont basés sur le problème de la factorisation des nombres premiers ou du logarithme discret. Dans [91], les auteurs donnent en exemple une preuve ZK du problème des non-résidu quadratique modulo m.

Un protocole simple permettant de comprendre le principe de ZK est le protocole de Fiat-Shamir ([94, 95, 96]) expliqué par la Figure 2.2. Le but du Prouveur est de prouver au Vérifieur qu'il connait la racine carré d'un entier modulo un autre entier, sans jamais lui dévoiler cette racine carré.

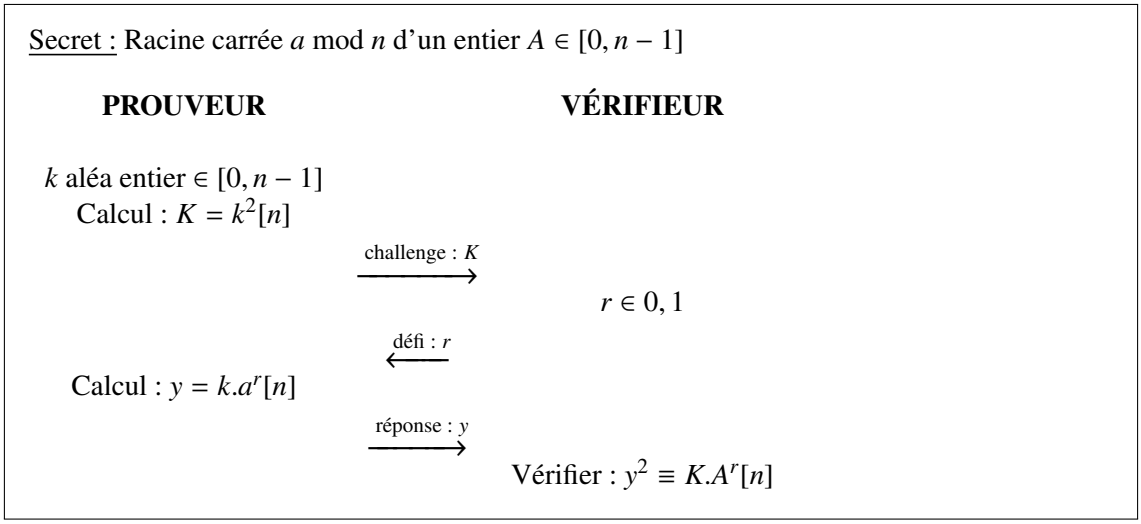


FIGURE 2.2 – Protocole Zero-knowledge de Fiat-Shamir

On attend d'un protocole d'identification qu'il respecte trois propriétés :

1. **Consistance** : La connaissance du secret permet au Prouveur de répondre au défi du Vérifieur ;
2. **Significatif** : Si la proposition est fausse, aucun Prouveur malicieux ne peut convaincre le Vérifieur que la proposition est vraie, ceci avec une forte probabilité. Le Prouveur doit donc connaître le secret ;
3. **Sans apport de connaissance (ou ZK)** : Le Vérifieur peut valider l'identification du Prouveur sans apporter de connaissance sur son identifiant autre que sa possession.

Le papier [97] de Goldreich, Micali et Wigderson a ainsi montré que, sous l'hypothèse de l'existence de chiffrement inviolable, il était possible de créer un système de preuve Zero-Knowledge pour le problème NP-Complet de 3-colorabilités des graphes : *"Un graphe G est-il coloriable avec trois couleurs sans que deux nœuds reliés portent la même couleur ?"*. Or, étant donné qu'un problème NP-complet peut être réduit à celui de coloration, tous les problèmes NP-Complet possède une preuve Zero-Knowledge.

Un autre protocole *ZK* nous intéressant particulièrement dans ce manuscrit est le protocole *ZK* de Schnorr [98] basé sur le problème du logarithme discret. Celui-ci est décrit par la Figure 2.3. Remarquons que p, q et g doivent être choisis par un tiers de confiance de telle sorte que le problème du logarithme discret dans $(\mathbb{Z}/p.\mathbb{Z})^*$ soit difficile.

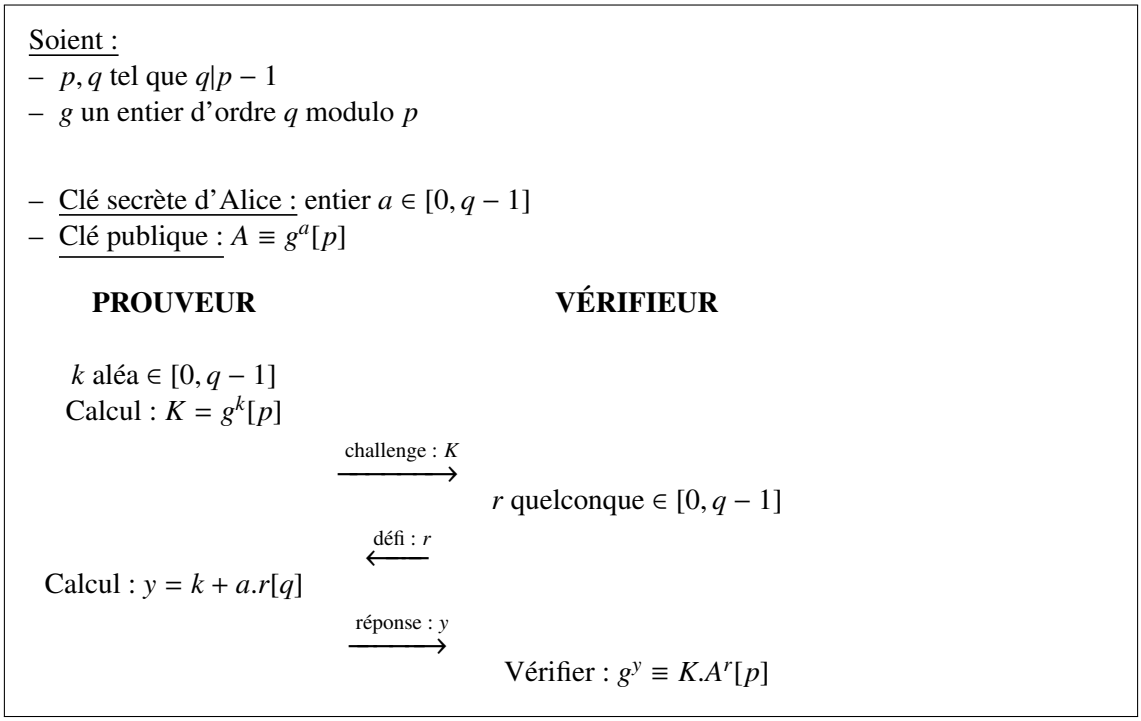


FIGURE 2.3 – Protocole Zero-Knowledge de Schnorr

Application

Les protocoles Zero-Knowledge peuvent être utilisés dans le cas où une personne veut prouver à une autre son identité sans la lui révéler. Ainsi, le fait de prouver qu'elle connait un secret particulier peut permettre de prouver son identité sans pour autant divulguer le secret ou l'identité précise.

Preuve du "OU"

La preuve du "OU" fait partie des preuves de connaissance potentiellement utiles pour prouver que l'on connait des valeurs sans les dévoiler. En effet, Cramer, Damgard and Schoenmakers décrivent dans [99] une méthode permettant de prouver la connaissance de d données parmi un ensemble de $n > d$ données, sans révéler celles-ci et en utilisant les matrices ou polynômes. Ce problème est également décrit dans la thèse de E. Hufschmitt [100], ainsi que l'exemple de la preuve de connaissance d'un logarithme discret parmi l relativement à une même base g_1.

Application

Soit un groupe de 100 personnes partageant des informations communes. Seul l'administrateur du groupe connait l'ensemble des i logarithmes discrets relativement à une base g. Ainsi, lorsqu'une personne souhaite avoir accès à une information concernant le groupe, elle prouve sa connaissance de j ($j < i$) logarithmes discrets relativement à la base g. Cette preuve lui permet de prouver son identité et ainsi d'accéder aux informations du groupe.

Preuve d'inégalité

Un autre type de preuve pouvant être très utile afin, par exemple, d'éviter la divulgation d'une information, est la preuve d'inégalité. Il s'agit d'une solution permettant à une entité de vérifier si la donnée de l'utilisateur est bien supérieure (ou inférieure) à la donnée requise par le vérifieur, sans pour autant que l'utilisateur dévoile cette information précise.

Ces preuves d'inégalité ont été introduites par Brickell et coll. [101]. Il s'agissait, dans un premier temps, du problème d'appartenance à un idéal. De nombreuses autres preuves ont découlé de cet article, notamment [102] qui concerne les systèmes anonymes de monnaie électronique. Ce protocole permet de travailler avec des intervalles plus petits par rapport aux autres schémas existants. D'autres preuves d'inégalités classiques sont décrites par Boudot dans son article [103] et reprises par Hufschmidt dans [100]. À titre indicatif, l'article [104] permet de prouver l'égalité entre deux secrets possédés par deux protagonistes différents.

Application

Lors de l'inscription d'un utilisateur à un site internet, il lui est souvent demandé un grand nombre d'informations et, notamment, sa date de naissance. Ainsi, bien que souvent demandée à des fins publicitaires, elle permet également de calculer votre âge et donc de vérifier vos droits d'accès à ce serveur. La date de naissance constitue une donnée extrêmement sensible au vue du rôle majeur qu'elle tient lors de l'établissement du passeport électronique. Ainsi, dans le but de préserver au maximum la vie privée de l'utilisateur et plus particulièrement, sa date de naissance, il est possible de fournir une preuve d'inégalité au site web prouvant que l'âge du client est bien supérieur à l'âge attendu, sans divulguer votre âge réel et ainsi sans fournir votre date de naissance.

2.3.2 Signature aveugle

Description

La signature aveugle est le procédé dédié à la vie privée qui permet de signer un document préalablement masqué et de garantir l'anonymat, le signataire n'a donc pas connaissance de son contenu. Le signataire est souvent une autorité de confiance différente des auteurs du message. Elle est notamment utilisée pour le vote électronique ou le porte monnaie électronique et a été présentée par David Chaum [105] en 1983 dans le cadre des systèmes de paiement. Un protocole de signature aveugle facilement compréhensible est le protocole de Schnorr [106, 107] décrit dans les encadrés 2.4 et 2.5.

Paramètres :

- G un groupe Schnorr ;
- $g \in G$ d'ordre q premier tel que le problème du logarithme discret soit difficile ;
- H une fonction de hachage (voir la Section 2.2.2) ;

Clés :

- Privée : $x \in]0, q[$ et $k \in (\mathbb{Z}/q.\mathbb{Z})^*$
- Publique : $y = g^{-x}[p]$ où p premier

Protocole :

BANQUE		**ALICE**
$r \equiv g^k[p]$		
	$\xrightarrow{r}$	
		$\alpha, \beta \in (\mathbb{Z}/q.\mathbb{Z})$
		$r' = r.g^{\alpha}.y^{\beta}$
		$e' = H(m\|\|r')$
		$e = e' - \beta[q]$
	$\xleftarrow{e}$	
$s = k + e.x[q]$		
	$\xrightarrow{s}$	
		Vérifier : $g^s.y^r \equiv r[p]$
		Calculer $s' = s + \alpha[q]$

FIGURE 2.4 – Protocole de signature aveugle de Schnorr (1/2)

Signature :

$$(r', s', e')$$

Finalement, lorsque Bob reçoit la signature aveugle du message d'Alice signé par l'autorité de confiance Banque, il n'a plus qu'à vérifier les équations suivantes :

$$\begin{cases} g^{s'}.y^{e'} = r' \\ e' = H(m\|r') \end{cases}$$

FIGURE 2.5 – Protocole de signature aveugle de Schnorr (2/2)

Une autre forme de signature se basant sur le protocole de signature aveugle de Schnorr en remplaçant notamment g par $g.h^n$ est le schéma de signature partielle proposé par Jean-Claude Pailles dans [108]. Ce protocole permet de lancer une opération de signature avec un partage de clés de session sans divulguer d'informations personnelles du client.

Application

Comme explicité par Chaum [105], la signature aveugle peut être un excellent outil pour les systèmes de paiement automatisés. Une tierce partie peut signer un document sans connaître son contenu, par exemple le montant de la transaction, leur bénéficiaire et la date du paiement. Le client obtient alors une preuve de paiement. Il ne s'agit cependant pas d'une architecture de paiement mais d'un protocole pouvant amener un niveau supérieur en termes de protection des informations d'achat du client. Ce protocole peut également être utilisé lors d'un vote électronique où le vote du citoyen doit être caché à l'autorité de signature. En ce qui concerne la signature partielle, il est possible d'imaginer le cas où un client désire obtenir un certificat anonyme afin d'accéder de façon sécurisée et anonyme à un serveur pour ensuite échanger une clé de session avec celui-ci.

2.3.3 Signature de groupe

Description

Chaum poursuit son travail avec Van Heyst [109] en 1991 sur les signatures en proposant un nouveau type : les signatures de groupe. Ils présentent quatre schémas cryptographiques qui varient en fonction du type de signature, du type de groupe (fixé ou non) et du nombre de calculs et/ou bits transmis. Certains se basent sur le problème du logarithme discret, d'autres sur le problème RSA. Cependant, ces schémas ont une longueur de clé

proportionnelle au nombre d'individus dans le groupe. Ainsi, en 1997, Camenisch et Stadler présentent dans [110] le premier protocole de signature de groupe dont les signatures et la clé publique ne dépendent pas du nombre de membres du groupe. Ce schéma est basé sur le logarithme discret et permet au gestionnaire du groupe d'ajouter de nouveaux membres sans changer la clé publique. La seule personne pouvant ensuite déterminer le signataire est le responsable du groupe. Deux autres types de signature de groupe existent :

- La signature de cercle est une signature de groupe sans possibilité d'identifier le signataire ;
- La signature "K parmi N" est une signature valable uniquement si au moins K membres de l'entreprise parmi les N signent le document.

Application

Ce genre de signature peut être intéressant lorsqu'un responsable d'une entreprise veut signer des informations au nom de tous les responsables de l'entreprise. L'information peut ainsi être vérifiée comme venant d'un responsable sans dévoiler l'identité du signataire.

Propriétés des signatures numériques dédiées à la protection de la vie privée

De la même façon que pour la signature numérique classique, les procédés de signature de groupe, aveugle et partielle assurent l'*intégrité des données*, l'*authentification du signataire*, ainsi que la *non-répudiation* du message. De plus, dans le cas des signatures aveugles et partielles, le signataire n'est pas en mesure d'accéder au contenu du message, la *confidentialité* de celui-ci est donc assurée. Les signatures partielles et de groupe permettent quant à elles de conserver l'*anonymat* du signataire et de l'expéditeur. Ainsi, alors que la signature de groupe permet de masquer l'identité du signataire, la signature aveugle permet de masquer le contenu du message et donc de garantir la **confidentialité des données**. Les différentes propriétés de ces signatures sont décrites en détails dans la thèse [100].

2.3.4 Accréditations anonymes

Description

Les accréditations anonymes, introduites par Chaum en 1985 [111, 112], sont des outils utilisant des données certifiées, appelées attributs, comme le nom, la ville, la date de naissance, etc... Ces systèmes impliquent trois acteurs : les **organisations** délivrent des accréditations à des **utilisateurs** qui peuvent ensuite les utiliser de façon anonyme pour prouver aux **vérifieurs** la validité des informations certifiées. Ainsi, les seules données révélées au vérifieur sont limitées aux informations pertinentes et nécessaires pour obtenir

service souhaité. De tels systèmes impliquent l'utilisation de différents schémas de signature dédiés à la protection de la vie privée telles que ceux détaillé dans la section précédente. Par exemple, le système de Camenisch et Lysyanskaya [113] est basé sur les signatures de groupe et est utilisé dans le système de gestion d'identité d'IBM, Idemix [54]. La technologie U-Prove de Microsoft [114, 115] utilise quant à elle le système d'accréditation anonyme de Brands basé sur les signatures aveugles [116].

Application

Un étudiant souhaite obtenir une place de cinéma à prix réduit. En montrant sa carte d'étudiant, il pourra obtenir une telle réduction. Cependant, cette carte contient actuellement des informations permettant d'identifier très clairement l'étudiant en question même si une telle précision n'est pas nécessaire pour obtenir le tarif préférentiel. Avec les accréditations anonymes, le jeune homme peut prouver qu'il est étudiant sans révéler aucune autre information le concernant et donc en restant anonyme.

Discussion

Les différents protocoles dédiés à la protection de la vie privée permettent donc d'assurer l'authentification des messages, leur intégrité, la non-répudiation et dans certains cas l'anonymat du signataire et de l'expéditeur, ainsi que la confidentialité des données. Cependant, même s'ils gèrent de nombreux principes et protègent davantage la vie privée des individus, peu de ces mécanismes sont actuellement utilisés et aucun protocole n'a encore été développé à notre connaissance pour les transactions électroniques sécurisées.

2.3.5 Partage de secret de Shamir

Description

Le principe de partage de secret de Shamir a pour but de diviser un secret en plusieurs parties qui seront distribuées aux différents acteurs du système [117]. L'idée est qu'il suffit de deux points pour déterminer l'équation d'une droite, de trois pour une parabole et de k points pour définir un polynôme de degré $k-1$ à l'aide de l'interpolation de Lagrange. Le polynôme ainsi retrouvé permet de recalculer le secret qui est alors l'ordonnée à l'origine. Ainsi, afin de retrouver le polynôme et donc le secret, un certain nombre $(k-1)$ de participants doivent s'unir. La Figure 2.6 décrit ce procédé.

Objectif :

– Partager un secret $S \in \mathbb{F}$, $\mathbb{F}$ corps fini

Détermination des parties à distribuer :

– Choisir : $k-1$ coefficients : $a_1, ..., a_{k-1} \in \mathbb{F}$, $a_0 = S$
– Poser : $f(X) = a_0 + a_1X + a_2X^2 + ... + a_{k-1}X^{k-1}$
– Distribuer les couples aux acteurs : $(x_i, f(x_i))$, $\forall x_i \in \mathbb{N}$
– Le secret est alors l'ordonnée à l'origine $S = a_0$.

Récupération du secret :

– Paramètres : k couples $(x_i, f(x_i))$
– Interpolation de Lagrange :
$f(X) = \sum_{j=0}^{n} f(x_j)(\prod_{i=0, i\neq j}^{n}(\frac{X-x_i}{x_j-x_i}))$
– Calcul de a_0
– $a_0 = S$

FIGURE 2.6 – Partage de secret de Shamir

Application

Supposons qu'un groupe de cinq associés décident qu'il est nécessaire de réunir au moins trois d'entre eux pour ouvrir un dossier délicat. Le partage de secret de Shamir peut gérer ce problème. En effet, en choisissant un polynôme de degré deux et en fournissant un point de ce polynôme à chacun, il suffit que trois d'entre eux se réunissent pour obtenir le secret, c'est-à-dire l'ordonnée à l'origine du polynôme et ainsi ouvrir/déchiffrer le document demandé.

2.3.6 Proxy de rechiffrement

Description

Le proxy de re-chiffrement est un outil de chiffrement permettant la protection de données personnelles lors d'un stockage sur une base de données. Il s'agit d'un outil cryptographique assez récent (1998, [118, 119]) qui permet de diffuser des informations uniquement entre différentes entités autorisées et ceci sans que l'infrastructure en charge de leur stockage et de leur diffusion n'ait la connaissance en clair de ces données. En d'autres termes, un proxy qui possède des informations spécifiques peut transformer un texte chiffré pour Alice en ce même texte chiffré pour Bob, sans jamais connaître le texte clair comme l'explique la Fig.2.7.

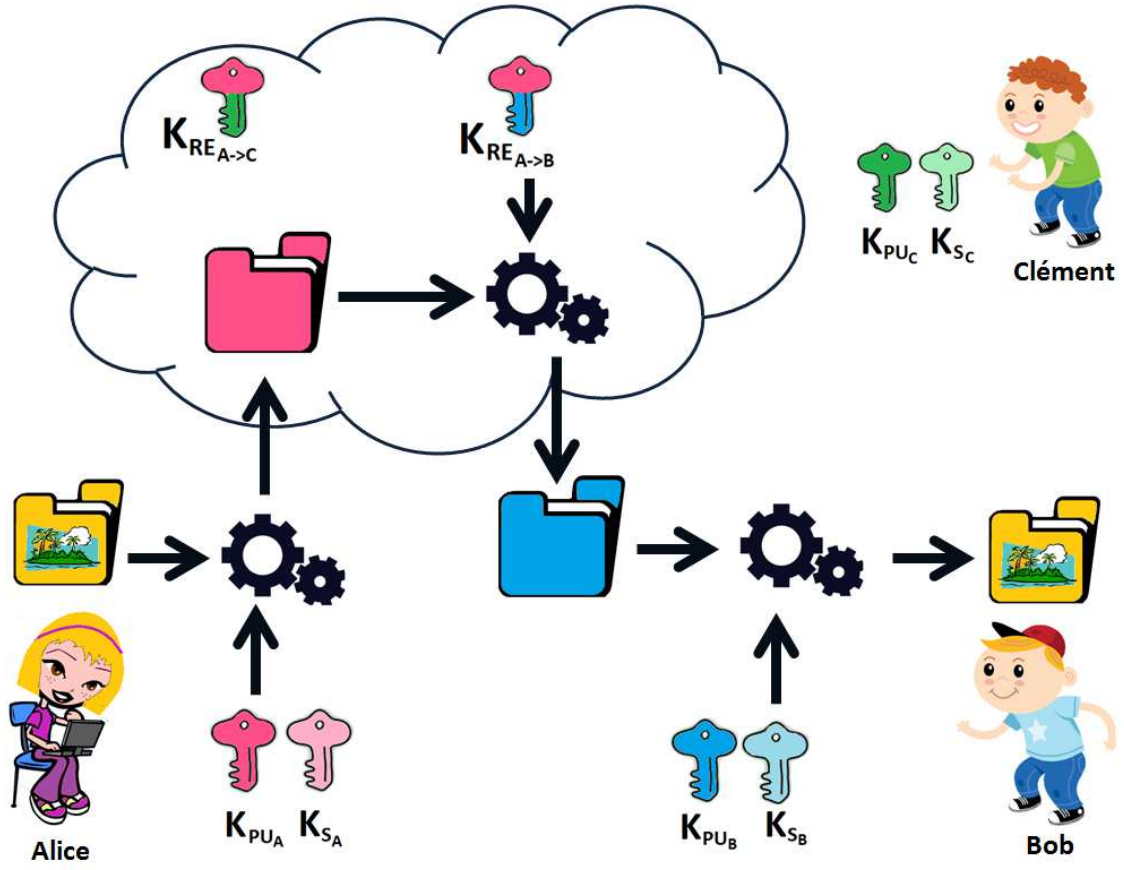


FIGURE 2.7 – Illustration de l'application utilisant le *PRE*

Application

Alice souhaite partager ses photos de vacances avec ses amis Bob et Clément. Elle chiffre ses images avec sa clé publique et les dépose sur le proxy. À partir de sa clé secrète et de la clé publique de Bob puis de Clément, deux clés de rechiffrement sont générées. Une est associée à Bob, l'autre à Clément. Si Bob veut accéder aux images, il contacte le proxy. Ce dernier rechiffre le dossier avec la clé de rechiffrement associée à Bob. Bob récupère ce nouveau chiffré qu'il peut finalement déchiffrer avec sa clé privée. Il obtient ainsi les photos d'Alice stockées grâce au proxy sans que ce dernier n'ait eu accès aux photos en clair.

2.3.7 Protection de la biométrie

Les systèmes d'authentification biométrique sont de plus en plus déployés pour remplacer les systèmes d'authentification classiques [120]. La sécurité de ces systèmes est nécessaire dans les applications du monde réel en biométrie et constitue un défi majeur.

Contexte

Les mécanismes d'authentification biométrique sont largement utilisés dans les systèmes de sécurité et offrent de nombreuses applications telles que l'e-gouvernement ou l'e-commerce. Les données biométriques sont considérées comme uniques pour chaque personne et sont directement liées à son propriétaire. Ces caractéristiques biométriques sont des données personnelles et extrêmement sensibles. Ainsi, dans le cas où les données biométriques originales seraient volées ou compromises, elles ne pourraient être révoquées. De récents résultats proposent ainsi une nouvelle approche pour ce problème en utilisant un modèle biométrique différent pour chaque application. La Figure 2.8 illustre ces différentes approches de protection des données biométriques.

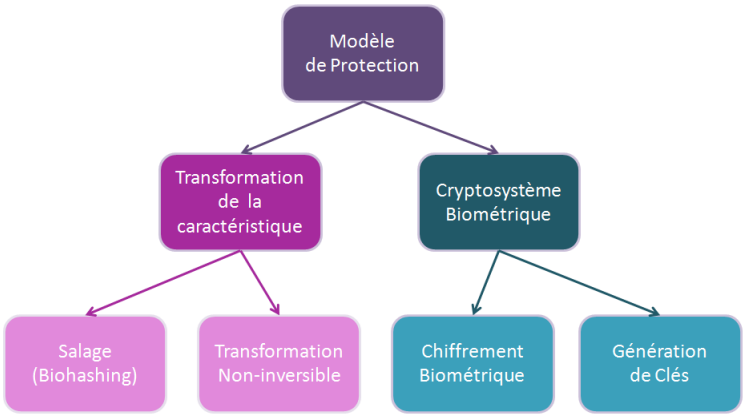


FIGURE 2.8 – Les différentes approches de protection de la biométrie (Figure issue de [121])

Le principe de la biométrie révocable, illustré par la Fig.2.9, a été introduit par Ratha et coll. dans [122] puis détaillé dans de nombreux articles, tels que [123], [124], [125] ou [126]. Ces derniers tentent de proposer une alternative au système biométrique classique en prenant en compte spécifiquement la protection de la vie privée. Cette méthode consiste à n'utiliser que des modèles révocables pour le stockage et la phase de vérification. Ceux-ci sont calculés à partir du modèle biométrique d'origine et d'une valeur aléatoire supplémentaire, appelé graine. Ainsi, si le modèle révocable est divulgué ou volé, un nouveau modèle est calculé à partir du modèle d'origine et d'une nouvelle graine aléatoire. La biométrie révocable permet également la génération de modèles différents pour une

personne et propose donc un modèle différent pour des applications différentes grâce à la génération d'une nouvelle graine. Cette transformation aléatoire doit cependant vérifier plusieurs critères comme la résiliabilité, la diversité et la non-inversibilité [127], [128].

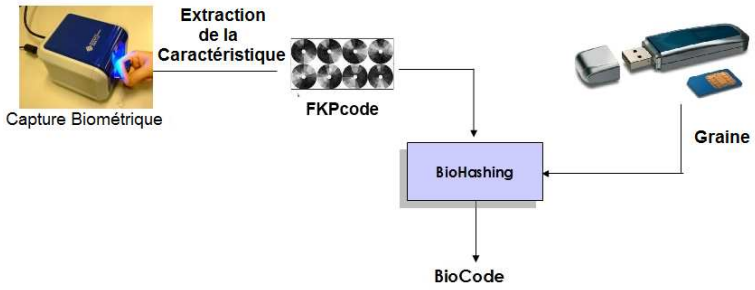


FIGURE 2.9 – Principe de biométrie révocable

Description

Le mécanisme d'authentification biométrique nécessite, au plus haut niveau, deux opérations : une phase d'enrôlement afin d'ajouter l'utilisateur au système biométrique qui inclut une acquisition et un prétraitement d'un modèle biométrique, ainsi qu'une phase de vérification comparant une nouvelle donnée biométrique présentée par l'utilisateur à celle stockée, appelée référence. Concernant la phase d'enregistrement, la caractéristique biométrique est calculée via un processus d'extraction et de discrétisation. Puis, par un procédé de biohashing, elle est transformée en un biocode à l'aide d'une fonction à sens unique comme le détaille la Fig. 2.10. Ce procédé utilise un secret aléatoire qui doit être stocké afin d'être réutilisé lors de la phase vérification.

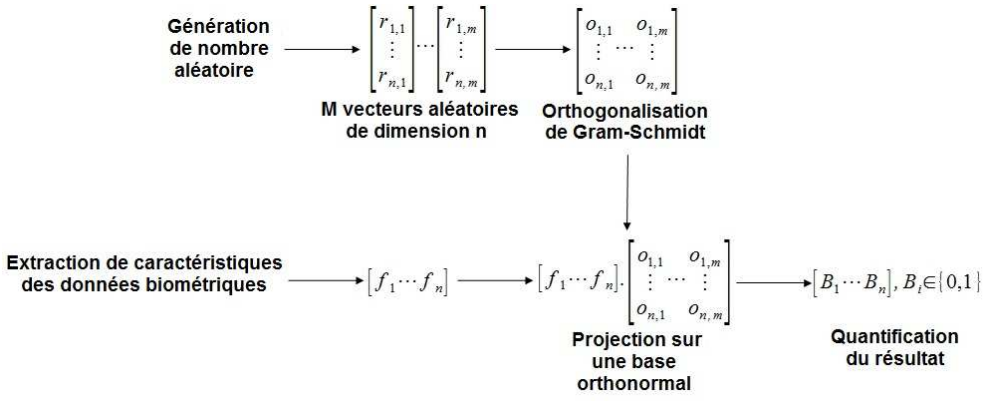


FIGURE 2.10 – Calcul du biohashing

De nombreux protocoles sont donc disponibles pour protéger les informations des différents acteurs d'une transaction. Cependant, lors d'une TES, l'échange de données est généralement nécessaire. Leur stockage dans des outils sécurisés est donc important.

2.4 Éléments sécurisés

Un Secure Element SE est un élément sécurisé, inviolable et permettant le stockage de clés ou d'informations sensibles. Un SE intègre également des applications capables de gérer un haut niveau de sécurité et donc de résister à de nombreuses attaques. Ils peuvent être intégrés dans des formes diverses. Ainsi, afin de protéger les données personnelles, de tels éléments sécurisés de stockage peuvent être utilisés.

2.4.1 Carte à puce

Une carte à puce est une carte plastique (parfois papier ou carton) qui comporte au moins un circuit intégré capable de contenir de l'information et donc des données personnelles. Il s'agit d'un système embarqué qui ne comporte actuellement ni écran, ni clavier mais une interface de communication simple. Son but est d'être portable, personnalisé et sécurisé. Il en existe deux types : avec ou sans contact. Les caractéristiques physiques, les dimensions et positions des contacts [129], les signaux électroniques et protocoles de transmissions [130], les commandes inter-industries pour l'échange [131], les identificateurs d'applications [132], les éléments de données inter-industries [133], les commandes inter-industries pour les $SCQL$ (Structured Card Query Language, [134, 135]) et enfin l'organisation des données et leur sécurisation [136] sont régis par la norme $ISO-7816$ [137]. De plus, depuis 2002, l'ISO $7816-11:2004$ traite de l'utilisation de la biométrie dans une carte. Les Sous-comités (SC) de l'ISO concernant les cartes sont le $SC17$ pour les cartes d'identification et le $SC27$ pour la sécurité IT et la cryptographie. Plus précisément, chaque SC est divisé en groupes de travail (WG). Au sein du $SC17$, le $WG4$ s'intéresse aux cartes avec contact et $WG8$ aux cartes sans contact. Qui plus est, le sans contact est normalisé essentiellement par l'ISO/IEC 14443 [138].

Ces cartes à puce permettent par exemple de sécuriser une clé privée et de stocker des certificats numériques. De plus, elles peuvent être un outil d'authentification efficace. Dans ce cas, l'utilisateur possède la carte et connaît le Numéro d'Identification Personnel (NIP ou code PIN). Si ce numéro est remplacé par de la biométrie, le niveau d'authentification est très fort. Plus particulièrement, dans le cas où la technologie requiert l'empreinte digitale du porteur de la carte, on parle de **Match On Card** (MOC).

Dans le domaine bancaire, le Chip Authentication Program (*CAP*, [139]) est un exemple d'authentification par carte à puce. À l'initiative de MasterCard, ce programme authentifie l'utilisateur d'une carte à puce bancaire *EMV* (Europay Mastercard Visa). Son concurrent direct, Visa, détient le programme appelé Dynamic Passcode Authentication (*DPA*). Le *CAP* est défini par un appareil portatif contenant : un lecteur de carte, un clavier décimal et un écran permettant d'afficher au plus 12 caractères. Ainsi, un client désirant s'authentifier doit insérer la carte puis entrer le code *PIN* correspondant. Il existe différentes applications une fois l'authentification terminée :

- Code/Identify : Cette fonction permet de générer un mot de passe à usage unique (OTP pour One Time Password en anglais) pour se connecter au site de la banque ;
- Response : Ce mode permet une authentification par challenge-réponse. Le site de la banque demande d'entrer un challenge dans le lecteur *CAP*. L'utilisateur reçoit une réponse sur ce lecteur qu'il entre dans le site web ;
- Sign : Il s'agit du mode précédent étendu. Effectivement, l'utilisateur doit également entrer dans le lecteur *CAP* : le challenge, le montant, la devise et son numéro de compte.

Par ces protocoles, les banques espéraient réduire les pertes liées à la fraude bancaire en ligne. Cependant, la carte à puce n'est pas sans contrainte. Les données sensibles ne doivent jamais sortir de leur environnement et la taille de stockage est restreinte du fait d'un système embarqué.

De nombreuses attaques ont également été dénombrées. Les attaques physiques reposent sur l'observation ou la modification du support d'exécution du programme. Certaines sont invasives et touchent directement les composants. Elles détruisent en général la carte ou la rendent inutilisable. Heureusement, elles sont coûteuses car elles nécessitent un matériel onéreux et des compétences pointues. À l'opposé, les attaques non-invasives ne détruisent pas la carte. Ces dernières [140, 141] se font par lecture ou par modification : Differential Power Analysis (*DPA*), Simple Power Analysis (*SPA*), timing-attack ([142]), Electro-Magnetic Analysis (*EMA*). Afin de s'y protéger, différentes implémentations ont été ajoutées au niveau matériel et logiciel : horloges désynchronisées, masque sur les données secrètes, vérification d'intégrité ou encore répétition des calculs. Les vulnérabilités peuvent être évitées grâce à des protocoles cryptographiques protégés [143] de plus en plus complexes et des vérifieurs de codes. En termes de normalisation, l'*ANSSI* recommande ([144]) le niveau de certification *EAL*4 avancé (Evaluation Assurance Level, [49]) comme niveau minimum de sécurité incluant une protection contre ce type d'attaques.

2.4.2 Carte SIM

La carte SIM (Subscriber Identity Module) en est un exemple. Il s'agit d'une puce contenant un microcontrôleur et de la mémoire. Elle contient les informations relatives à un abonné dont son numéro de téléphone cellulaire et est également conforme à la norme ISO/IEC 7816 . Cette carte peut servir d'élément sécurisé et est utilisée dans les téléphones portables pour l'identification du propriétaire et la sauvegarde d'informations diverses. C'est le microcontrôleur qui gère le droit d'accès à ces données et les fonctions de cryptographie. Ces dernières permettent d'avoir des algorithmes de chiffrement spécifiques pour la génération de clés et des algorithmes d'authentification.

2.4.3 Carte SD

Une carte SD (Secure Digital) est une carte mémoire utilisée pour le stockage de fichiers dans des appareils numériques. Contrairement aux autres outils sécurisés dont les normes sont régies par l'ISO/IEC, le standard de ces cartes mémoires avec contact est donné par la SDA (SD Card Association, [145]). Ce format inclut quatre familles de cartes : les $SDSC$ (original SD Standard-Capacity), les $SDHC$ (SD High-Capacity), les $SDXC$ (SD eXtended-Capacity) et les $SDIO$ qui combinent les fonctions Entrée/Sortie avec le stockage de données. Un dispositif de sécurité est disponible sur de nombreuses cartes SD afin de les ouvrir uniquement en lecture seule. Cependant, il s'agit souvent uniquement d'une simple encoche à déplacer. Il est également possible pour certaines cartes d'ajouter un mot de passe ou de gérer les droits d'accès à celles-ci.

2.4.4 Clé USB

Une clé USB (Universal Serial Bus) est un petit média amovible pouvant se brancher sur un ordinateur, une chaîne Hi-Fi, un lecteur de DVD,... La norme la concernant est la norme ISO/IEC 7816 – 12 : 2005. Une clé USB contient une mémoire flash mais aucun élément mécanique. Elle est ainsi très résistante au choc et portative. On peut considérer que les clés USB font partie de la famille des cartes à puce. Cependant, il n'y a qu'une minorité de ces clés qui protègent l'accès à la mémoire et évitent donc les intrusions. Ces dernières peuvent ainsi être utilisées pour stocker des données sensibles.

2.4.5 HSM

Un HSM (Hardware Security Module, Fig. 2.11) est un matériel qui offre des services de sécurité permettant de gérer (générer, stocker et protéger) des clés cryptographiques. Ce matériel est très utile pour gérer les clés privées. De plus il dispose de nombreux moyens

de protection matérielle, comme la détection d'ouverture ou d'arrachage, ainsi que des capteurs de déplacement, de température et de tension. Le *HSM* possède également une double sécurité : en cas de manipulation physique, il autodétruit ses données, et donc ces clés privées, et empêche ainsi le vol de ces données. Ainsi, afin de garantir la non compromission des données, toutes les informations sensibles sont manipulées uniquement dans l'enceinte sécurisée du *HSM*. En ce qui concerne ces moyens de protection logique, le *HSM* est capable de détecter une intrusion et les tentatives de fraude. De plus, il contrôle chaque signature des logiciels qu'il embarque et il nécessite la présence d'une majorité de personnes pour activer un secret. Cette caractéristique assure l'intégrité des données et empêche l'utilisation du secret par un seul individu, même s'il s'agit de l'administrateur. Les *HSM* sont omniprésents pour sécuriser des transactions de paiement ou encore des transactions distantes. Ils sont standardisés par la norme *FIPS* 140 ([146]) et Critères communs *EAL*4+ ([49, 147]). De plus, ils supportent des *API* (Application programming interface) cryptographiques majeures comme : *PKCS*♯11 − *CryptoAPI* ([148]) ou encore *Java JCA*/*JCE* ([148]).

FIGURE 2.11 – Aperçus d'un *HSM* BULL (Ancien et Nouveau modèles)

Dans cette thèse, nous utilisons en priorité le *HSM* de BULL : *CRYPT*2*Pay*. Il apporte la sécurité indispensable aux transactions de paiement. Son déploiement est ainsi lié à 95% au monde bancaire. Ce *HSM* permet également de fournir de nombreux services cryptographiques pour le chiffrement des données, l'authentification des acteurs, la signature ou encore le stockage sécurisé de clés. De plus, il offre de nombreuses fonctionnalités comme le montre la Fig.2.12 : Acquisition de transactions, autorisation EMV, gestion de PIN, impression de PIN, expertise de carte et déblocage de PIN, autorisation 3D-Secure, préparation des données EMV et piste, ainsi qu'un centre de gestion des clés combinant ces différentes fonctionnalités. Plus précisément, CRYPT2Pay fournit les schémas de chiffrement et de déchiffrement DES, Triple-DES, RSA (jusqu'à 2048 bits), AES (avec clés de 128, 192 ou 256 bits), les fonctions de hachage SHA-1 et SHA-256, les codes d'authentification de type HMAC et des générations de clés ECDSA. De plus, CRYPT2Pay est conforme au standard international de sécurité des cartes de paiement *EMV* 4.2 *CPA* [149] ou encore à 3D-Secure [150].

*CRYPT*2*PAY* a également une option intéressante : il est programmable. Ainsi, il est

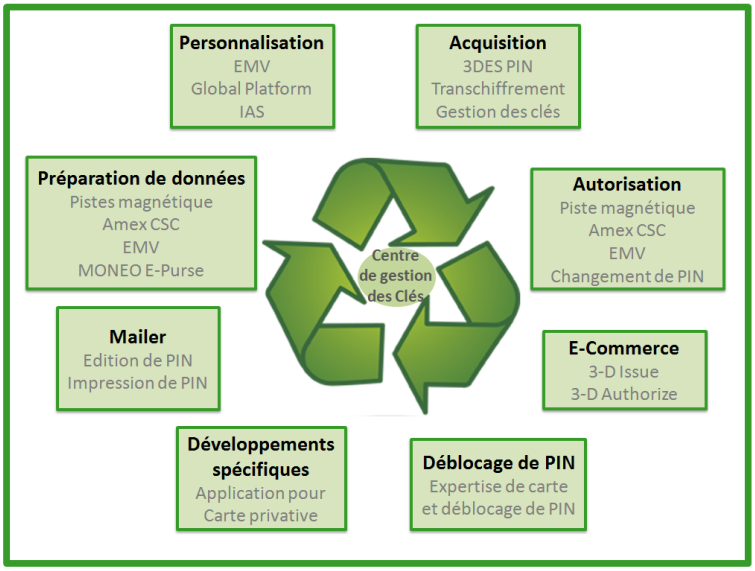


FIGURE 2.12 – Fonctionnalités offertes par *CRYPT2Pay* (Figure issue de [151])

possible d'y importer de nouvelles fonctions implémentées par nos soins afin d'améliorer les architectures proposées par la suite ou encore d'apporter un service supplémentaire à ce *HSM*.

2.5 Conclusion

Ce chapitre a présenté les principaux algorithmes de chiffrement existants, ainsi que leurs propriétés et les normes nationales et internationales les régissant. La seconde section de ce chapitre a énoncé plusieurs mécanismes à même de protéger les données sensibles des utilisateurs. Différents protocoles Zero-Knowledge et de nombreux types de signature ont été explicités, ainsi que des procédés de gestion de clés avec les *PKI*s. D'autres méthodes d'authentification ont également été exposées et une liste non exhaustive d'éléments sécurisés souvent utilisés pour stocker les données sensibles a été dressée.

Malheureusement, ces mécanismes sont actuellement peu utilisés et il existe, à notre connaissance, peu de protocole développé pour les transactions électroniques sécurisées qui les utilise.

La suite de cette thèse se concentre ainsi sur la mise en place d'architectures respectueuses de la vie privée employant des *TES*. Les modifications nécessaires à apporter aux algorithmes et aux outils existants afin de garantir l'ensemble de ces propriétés sont également explicitées.

Deuxième partie

Contributions de la thèse : Vers le développement d'architectures respectueuses de la vie privée

Chapitre 3

Système de gestion de la sécurité et de la protection des données personnelles sur Internet

Les fournisseurs de service en ligne font partie des domaines en plein essor sur Internet. Cependant, de nombreuses informations, souvent personnelles, sont échangées puis stockées dans d'immenses bases de données par les sites web et les commerçants en ligne. Malheureusement, ces fournisseurs de services ne proposent pas toujours une politique de sécurité claire sur la collecte et le stockage des informations demandées. De plus, les utilisateurs ont besoin de comprendre la façon dont elles sont collectées et stockées. Dans ce chapitre, nous proposons un système de gestion de ces données pour la navigation sur Internet avec plusieurs fonctionnalités. Cette solution se veut facile d'utilisation, centrée sur l'utilisateur et respectueuse de la vie privée.

Sommaire

3.1 Introduction

En 10 ans, le nombre d'utilisateurs connectés dans le Monde est passé de 394 millions en 2000 à 2044 millions en 2010, soit presque 30% de la population mondiale connectée et 67% de la population européenne [152]. De nombreuses possibilités s'offrent à ces internautes tels que l'utilisation de réseaux sociaux ou le e-commerce. Dans la plupart des cas, ils doivent fournir diverses informations personnelles afin d'accéder à un service en ligne. Ces informations sont parfois stockées sans réel contrôle, ni limitation de temps ce qui est contraire à la *LIL*. En outre, la plupart des renseignements demandés est uniquement utilisée pour des études de marché ou à des fins publicitaires sans réelle utilité pour le *SP*.

L'utilisation des données personnelles à des fins promotionnelles est un problème important en termes de vie privée car les utilisateurs sont rarement conscients des enjeux lorsqu'ils divulguent de telles informations. Ces données à caractère personnel doivent être protégées. Il est donc nécessaire de fournir des outils aux utilisateurs afin de leur permettre d'être sensibilisés à cette divulgation et ainsi de la limiter.

La législation en matière de protection de la vie privée sur Internet est différente pour chaque pays. Les États-Unis protègent ses jeunes internautes grâce au texte [153] et ses consommateurs par la loi Gramm-Leach-Bliley [154]. L'Union européenne dispose de la directive 95/46 sur la protection des données [155], et le Canada de la Loi PIPEDA [156]. Malheureusement, un grand nombre de publications se concentre sur les protocoles d'authentification sur Internet ou sur une technologie spécifique améliorant la protection de la vie privée des utilisateurs. Dans ce chapitre, nous proposons de rassembler plusieurs protocoles dans un logiciel de façon à offrir à l'utilisateur un outil pour protéger ses données personnelles sur Internet. Cette application possède plusieurs fonctionnalités : analyseur de conditions d'utilisation, analyseur de conditions d'accès d'un fournisseur d'accès, authentification forte par biométrie révocable, gestion des données de l'utilisateur ou encore gestionnaire d'identités. Ce système a pour objectif d'aider l'utilisateur à garantir des principes fondamentaux de protection de la vie privée décrits dans le chapitre 1.

Ce chapitre commence par la définition des exigences à prendre en compte pour protéger les données personnelles de l'utilisateur sur Internet. Ensuite, nous explicitons les technologies spécifiques au Web existantes en termes de protection de la vie privée. La section suivante détaille le système de gestion proposé et les fonctionnalités dont il dispose. Pour finir, une analyse de la solution est donnée dans la section 3.5.

3.2 Exigences de sécurité et de protection de la vie privée

Comme nous venons de le rappeler, les internautes doivent fournir un grand nombre d'informations. Il est alors nécessaire que les utilisateurs prennent en compte plusieurs exigences afin de protéger leurs données personnelles. Nous définissons ici huit exigences nécessaires à la réalisation de cet objectif :

• E_1 : L'**usage de mots de passe différents** ajoute un niveau de sécurité, notamment en permettant de ne pas être capable de lier les transactions venant d'un même utilisateur et ainsi éviter l'associabilité des données.
• E_2 : L'**authentification du SP** par l'utilisateur ou par une partie de confiance garantit l'identité du SP auprès de l'utilisateur.
• E_3 : La **confidentialité de l'identité de l'utilisateur vis-à-vis du SP**.
• E_4 : L'**authentification forte de l'utilisateur** afin d'éviter l'usurpation d'identité.
• E_5 : Le **stockage sécurisé des données** par l'utilisateur.
• E_6 : La prise en compte du **principe de minimisation des données** par l'utilisateur.
• E_7 : La prise en compte du **principe de souveraineté des données** implique que les données personnelles de l'internaute lui appartiennent avec son contrôle et son consentement sur leur utilisation.
• E_8 : La prise en compte du **principe de sensibilité des données** implique la prise en compte du degré de sensibilité des différentes données.

Actuellement, il n'est pas facile pour un utilisateur de gérer ces différentes exigences par lui-même. Plusieurs travaux dans la littérature permettent d'en assurer un certain nombre mais jamais la totalité.

3.3 État de l'art sur la protection des données personnelles sur Internet

À notre connaissance, aucun système complet de gestion des données de l'utilisateur sur Internet n'a encore été proposé. Les architectures connues et soulevées dans cette section ne permettent d'assurer en effet qu'une partie des exigences énoncées précédemment.

La plate-forme P3P [157], déjà énoncée dans ce rapport, en est un exemple. Elle définit la procédure de traitement et de récolte des données personnelles du site Web mais ne vérifie pas si le site la respecte réellement. De plus, il ne s'agit pas d'un outil de gestion

des données de l'utilisateur. Cette plate-forme s'intéresse uniquement à la problématique de politique de sécurité sans notion d'intrusion et de degrés de sensibilité des données. De la même façon, afin d'aider les internautes à gérer leurs informations personnelles et contrôler les politiques de vie privée des sites en ligne, IBM a proposé Tivoli [158], une infrastructure de gestion des données. Cependant, les règles de gestion pour celle-ci ne sont pas extrêmement précises. Cette architecture se concentre principalement sur la gestion du stockage des données de l'utilisateur et les contrôles d'accès. De plus, et contrairement à la plate-forme P3P, ce système a principalement été créé pour des entreprises. La Table 3.1 présente une analyse de ces deux plates-formes en fonction des exigences définies dans ce chapitre.

$\mathbf{E_i}$	**Propriétés**	**P3P**	**Tivoli**
E_1	Mots de passe différents	-	Non
E_2	Authentification du SP	-	**Oui**
E_3	Anonymat vis-à-vis du SP	Non	-
E_4	Authentification forte vis-à-vis de la banque du SP	-	**Partiel**
E_5	Stockage sécurisée des données	-	**Oui**
E_6	Principe de minimisation	Partiel	Partiel
E_7	Principe de souveraineté	**Oui**	Partiel
E_8	Principe de sensibilité	**Oui**	Non
	Score /8	2.5	**3.5**

TABLE 3.1: Comparaison des protocoles existants

En ce qui concerne les réseaux sociaux, les technologies permettant de protéger les données personnelles ne manquent pas. Cependant, elles n'assurent jamais la totalité des propriétés nécessaires à une complète protection. Nous pouvons par exemple citer les architectures suivantes : Diaspora [159], Peerson [160] et SAFEBOOK [161]. Diaspora permet aux utilisateurs de créer leurs propres serveurs afin d'héberger des données de façon décentralisée et sécurisée mais ne fournit aucune possibilité pour le chiffrement des données. À l'inverse, Peerson est une infrastructure peer-to-peer qui inclut des protocoles de chiffrement et qui permet aux utilisateurs de contrôler leurs données. En outre, les utilisateurs peuvent utiliser ce réseau social sans avoir accès à Internet. Enfin, SAFEBOOK est un réseau social en ligne qui se base sur la confiance pour préserver la vie privée des utilisateurs. Sa politique est donc loin d'être suffisamment précise. Jahid et coll. offrent dans [162] une étude critique sur les performances en termes de protection de la vie privée de ces différentes infrastructures. Ils proposent également une nouvelle architecture décentralisée pour les réseaux sociaux. Malheureusement, ces architectures sont souvent complexes pour des utilisateurs novices.

En ce qui concerne les technologies de l'information ciblant principalement le problème d'identité et de vie privée, de nombreux projets ont vu le jour. Dans le cadre de la volonté de l'Europe de prendre le leadership dans les matières liées aux technologies de l'information et à l'identité numérique, le projet académique **FIDIS** ([163]) se base sur la question de la diffusion de données à caractère personnel et déclare des objectifs liés à l'identité et à l'identification, ainsi qu'au vol d'identité numérique et à la sécurité. En termes d'amélioration de l'identification, les risques pour la vie privée de l'utilisateur et son autonomie croissent. Le projet **PRIME** (Privacy and Identity Management in Europe, [2]) donne alors des solutions pour lutter contre ces menaces. Il se consacre à une série de questions sur la gestion d'identité. Par exemple, il s'interroge aussi bien sur la normalisation que sur le développement de prototypes permettant le contrôle des données privées. **PICOS** (Privacy and Identity Management for Community Services, [164]), un autre projet international sur la gestion d'identité et la protection de la vie privée, a vu le jour en 2008. Celui-ci a pour objectifs de rechercher, développer, construire et évaluer une plate-forme ouverte de gestion d'identité respectant la vie privée pour les fournisseurs de services de communications mobiles. Le projet **ABC4TRUST** ([165]) doit quant à lui permettre de définir une architecture commune aux systèmes de certification basés sur les attributs (ABC pour Attribute-based Credentials). Il s'adresse notamment aux technologies préservant la vie privée. Le standard OpenId [166] implémente quant à lui un modèle en ligne de gestion d'identités centrée sur l'utilisateur. Cependant, bien que ce modèle soit simple d'utilisation pour l'internaute en lui présentant clairement les données allant être échangées durant la transaction, la sécurité de l'authentification est laissée aux développeurs et les attaques par phishing sont courantes. De la même manière, le modèle de client intelligent U-Prove permet de ne révéler que le minimum d'informations nécessaires en fonction du contexte [115, 114]. De plus, comme le montre le rapport [19], cette technologie résoud le problème de traçabilité non pris en compte par les autres systèmes de gestion d'identité. Cependant, celle-ci se concentre majoritairement sur le principe de minimisation des données de l'utilisateur, par exemple, à l'aide d'accréditations anonymes dans le cas de U-Prove.

Ainsi, à notre connaissance, il n'existe pas d'infrastructure complète permettant de protéger les données personnelles des utilisateurs sur Internet en divulguant un minimum d'informations et en respectant l'ensemble des principes de protection de la vie privée. Nous proposons ainsi dans ce chapitre un système simple de gestion des données pour l'utilisateur en utilisant plusieurs protocoles décrits dans le chapitre 2.

3.4 Gestion de sécurité et protection de la vie privée

L'application décrite dans ce chapitre doit être vue comme un logiciel installé sur l'ordinateur d'un utilisateur et disposant d'un certain nombre de fonctionnalités que nous détaillons par la suite. Ce système est centré sur l'utilisateur et lui permet de naviguer sur Internet en divulguant un minimum d'informations. Il a également pour objectif d'aider l'utilisateur à prendre en compte le degré de sensibilité de ses données et à gérer ses informations personnelles. L'application peut ainsi jouer différents rôles comme le montre la figure 3.1.

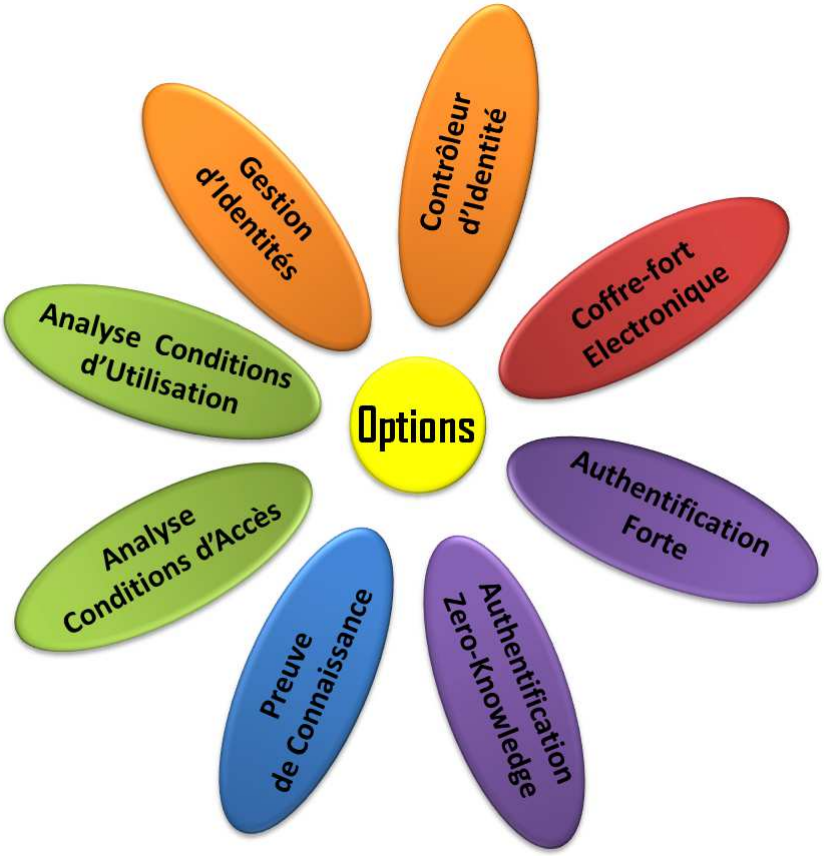


FIGURE 3.1 – Fonctionnalités proposées par l'application

Dans un premier temps, l'utilisateur naviguant sur Internet peut souhaiter avoir un aperçu des données que le site Web stocke lors d'une navigation et ainsi obtenir le degré de sensibilité des données demandées. L'utilisateur fait alors appel à l'analyseur de conditions d'accès et/ou à l'analyseur de conditions d'utilisation fournis par l'application. Ensuite, s'il ne souhaite pas fournir certaines informations, il peut demander à l'application de lui générer une preuve de connaissance. Afin de gérer l'ensemble de ses mots de passe,

l'utilisateur peut également compter sur le gestionnaire d'identités de son application. De plus, si le *SP* le permet, l'utilisateur peut demander à l'application de s'authentifier de manière forte ou encore de façon anonyme à l'aide d'un protocole Zero-knowledge. Pour finir, les différentes informations stockées par l'application se trouvent toutes dans le coffre-fort électronique fourni par celle-ci. Ce système se veut ainsi respectueux des principes fondamentaux en termes de vie privée et des exigences précédemment définies.

Ainsi, plus l'utilisateur fait appel aux fonctionnalités disponibles dans l'application, plus le niveau de protection de ses données personnelles sera important.

3.4.1 Fonctionnalités de l'application

Comme expliqué précédemment, l'application est installée sur l'ordinateur de l'utilisateur et permet de gérer et d'analyser les données de celui-ci pouvant être partagées. Le client décide ou non d'utiliser les options disponibles. Certaines de ces fonctionnalités entrainent le stockage de données personnelles de l'utilisateur ou encore de couple de login/mot de passe dans le coffre-fort électronique de l'application.

Coffre-fort Électronique

De nombreuses fonctionnalités sont disponibles dans cette application et un certain nombre d'entre elles nécéssite un stockage de données de façon sécurisée. La présence d'un coffre-fort électronique dans notre application est donc nécessaire. Ce coffre permet de créer une base de données chiffrée à l'aide d'un protocole de chiffrement symétrique, comme l'*AES*. Afin d'y accéder, il est possible d'imaginer une authentification forte comme décrit dans la section 3.4.1 ci-dessous. Un exemple de coffre-fort électronique possible est le logiciel de chiffrement TrueCrypt [167] qui met en œuvre différents algorithmes de chiffrement dont l'*AES* et qui, pour la version 6.0, a recu une certification de l'*ANSSI*.

Contrôleur d'identité

Si le fournisseur de service, avec lequel l'utilisateur veut converser, possède un certificat de signature, l'application peut avoir un rôle de vérifieur de signature. En effet, rares sont les utilisateurs qui vérifient l'identité du site Web sur lequel ils naviguent. Dans notre cas, il pourrait s'agir de l'envoi d'un challenge et la vérification de la signature de ce challenge par le *SP* qui seraient réalisés par l'application afin de contrôler l'identité du *SP*.

Authentification forte

Dans le cas où l'utilisateur doit être authentifié pour accéder à un site Web (réseaux sociaux, site marchand...), il est recommandé à l'utilisateur de faire appel à un protocole d'authentification forte afin d'éviter un grand nombre d'attaques, comme celles décrites dans le chapitre 1. En fonction des possibilités offertes par le fournisseur de services et des souhaits de l'utilisateur, différentes authentifications sont possibles car intégrées dans l'application générale.

Dans un premier temps et du fait que de nombreux sites Web se contentent de cette authentification, il est possible d'utiliser une solution simple de login/mot de passe. Ainsi, afin d'éviter les problèmes liés à l'utilisation multiple d'un même mot de passe, l'application peut générer un couple différent de login/mot de passe pour chaque site consulté. Elle stocke ensuite ces différents couples de façon sécurisée dans son coffre-fort électronique. Dans un second temps, si l'utilisateur possède un lecteur *CAP*, l'application peut permettre une authentification via ce lecteur comme nous l'avons détaillé dans le chapitre 2. L'application permet également de faire des authentifications à l'aide de biométrie révocable. Il s'agit alors de combiner un mot de passe avec une caractéristique biométrique : une empreinte digitale ou encore la dynamique de frappe au clavier. Comme nous l'avons décrit dans le chapitre 2, l'intérêt principal de cette méthode est qu'elle est révocable. Ainsi, si le modèle est divulgué ou volé, un nouveau modèle peut être calculé par l'application à partir du modèle d'origine et d'un nouveau mot de passe. L'application supprime alors l'ancien modèle de son coffre-fort électronique et stocke le nouveau modèle ainsi calculé. De plus, il est important de noter que, contrairement au modèle de dynamique de frappe au clavier, l'utilisation d'une empreinte digitale nécessite la possession d'un lecteur d'empreinte. La Figure 3.2 illustre l'implémentation d'une authentification par biométrie révocable dont le mot de passe est "azerty" pour l'utilisateur. Ce mot de passe, combiné à l'empreinte digitale de l'utilisateur extraite immédiatement, permet d'authentifier l'utilisateur à partir de 95% de similarité entre le biocode calculé et celui stocké. Si le site Web l'autorise, l'application peut offrir un autre moyen d'authentification à l'utilisateur : l'authentification Zero-Knowledge que nous avons détaillée dans le chapitre 2 section 2.3. La preuve d'authentification est alors réalisée par l'application.

Preuve de connaissance

Afin d'éviter de fournir certaines informations au fournisseur de service, le client peut également faire appel à l'application pour qu'elle génère une preuve de connaissance de l'information et prenne la responsabilité de cette preuve. Par exemple, l'application peut fournir une preuve que l'âge du client est supérieur à l'âge requis pour accéder au site au

FIGURE 3.2 – Authentification avec la biométrie révocable

lieu de donner sa date de naissance précise. Les encadrés 3.3 et 3.4 détaillent le cas où l'âge doit être supérieur à 16 ans et se base sur la preuve de Boudot dans son article [103].

Analyseurs de conditions d'utilisation

Le client peut demander ensuite une analyse des conditions d'utilisation du site consulté et ainsi de sa politique de sécurité. Ces conditions sont en effet trop rarement lues par les utilisateurs alors qu'elles explicitent clairement la politique du site vis-à-vis de la gestion des données enregistrées. L'application passe ainsi en revue ces conditions et met en lumière les termes importants en ce qui concerne la protection de la vie privée ou la divulgation de certaines informations personnelles, par exemple les informations bancaires.

Analyseurs de conditions d'accès

L'application peut également analyser les conditions d'accès au site. Pour chaque donnée demandée au client, le niveau de sensibilité de l'information attendue est ainsi indiqué. Ces deux options permettent au client de mieux gérer ses données et de minimiser leur divulgation.

Afin d'optimiser cette dernière option et de faciliter l'usage des clients, ces derniers peuvent pré-remplir un formulaire proposé par l'application. Ce formulaire permet au client de choisir les informations qu'il accepte de révéler sans condition, celles pour lesquelles il émet des réserves et celles qu'il refuse de transmettre. De la même façon que lors de l'analyse, le client est aidé lors du remplissage de ce formulaire afin d'évaluer le niveau de sensibilité de chaque type de donnée.

Cette preuve permet au Prouveur P de prouver au Vérifieur V qu'il connait une valeur $x = 19$ (l'âge du client calculé à partir de sa date de naissance, ici 19 ans), gardée secrète, supérieure à une valeur connue $a = 16$ de taille $2^l = 2^4$.

Soient :
- k le paramètre de sécurité
- G groupe cyclique d'ordre q premier tel que $q > 2^k$
- g générateur de G
- $h \in G$ tel que $log_g(h)$ non connu
- $x = 19$, $a = 16$ entiers de longueur $l = 4$

La preuve est basée sur le problème du logarithme discret. Une astuce pour cette preuve est d'utiliser la décomposition binaire de $a = 16$ et de $x = 19$.

PROUVEUR **VÉRIFIEUR**

Décomposition :

$1.a = a_0 + a_1.2^1 + a_2.2^2 + ... + a_{l-1}.2^{l-1}$
où $a_0 = 0, a_1 = 0, a_2 = 0, a_3 = 0, a_4 = 1$
$2.x = x_0 + x_1.2^1 + x_2.2^2 + ... + x_{l-1}.2^{l-1}$
où $x_0 = 1, x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 1$

Choix :

$r, r_0, r_1, r_2, r_3, r_4 \in \mathbb{Z}_p$

Calculs :

$C = g^x.h^r$
$C_i = g^{x_i}.h^{r_i}, \forall i \in 0, ..., 4$
$\tilde{C} = \prod_{0 \leq k \leq 4}(C_i^{2^i})$

$$\xrightarrow{C, C_0, C_1, C_2, C_3, C_4}$$

Preuve interactive entre le Prouveur et le Vérifieur

$$\text{POK } [\ \alpha, \beta, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \delta :$$

$$(C_0 = h^{\gamma_0} \vee \frac{C_0}{g} = h^{\gamma_0}) \bigwedge ... \bigwedge (C_4 = h^{\gamma_4} \vee \frac{C_4}{g} = h^{\gamma_4})$$

$$\bigwedge (C = g^{\alpha}.h^{\beta}) \bigwedge (C.\tilde{C}^{-1} = h^{\delta})$$

$$\bigwedge$$

$$[(\frac{C_4}{g} = h^{\gamma_4} \wedge a_4 = 0) \vee$$

$$(\frac{C_4}{g^{a_4}} = h^{\gamma_4} \wedge \frac{C_3}{g} = h^{\gamma_3} \wedge a_3 = 0)$$

$$\vee ... \vee$$

$$(\frac{C_4}{g^{a_4}} = h^{\gamma_4} \wedge ... \wedge \frac{C_1}{g^{a_1}} = h^{\gamma_l} \wedge \frac{C_0}{g} = h^{\gamma_0})] \ ]$$

FIGURE 3.3 – Preuve d'inégalité : $x > a$ connu (1/2)

C représente l'engagement de P et les différentes sous-preuves permettent à P de prouver l'ensemble de son affirmation :

- $(C_i = h^{\gamma_i} \vee \frac{C_i}{g} = h^{\gamma_i})$ indique que x_i vaut soit 0 (et donc $g^{x_i} = 1$) soit 1 (et donc $g^{x_i} = 0$).
  Ainsi, la partie $(C_0 = h^{\gamma_0} \vee \frac{C_0}{g} = h^{\gamma_0}) \bigwedge ... \bigwedge (C_{l-1} = h^{\gamma_{l-1}} \vee \frac{C_4}{g} = h^{\gamma_{l-1}})$ prouve à V que P connait les x_i et donc possède la bonne décomposition de x.
- $(C = g^{\alpha}.h^{\beta})$
- $(C.\tilde{C}^{-1} = h^{\delta})$ montre que $C.\tilde{C}^{-1}$ est bien une puissance de h.
  On peut en effet vérifier que :
  $C.\tilde{C}^{-1} = g^{x}.h^{r}.\prod_{0\leq k\leq 4}(C_i^{2^i})^{-1}$
  $= g^{x}.h^{r}.(C_0^{2^0}.C_1^{2^1}.C_2^{2^2}.C_3^{2^3}.C_4^{2^4})^{-1}$
  $= g^{x}.h^{r}.((g^{x_0}.h^{r_0})^{2^0}.(g^{x_1}.h^{r_1})^{2^1}..(g^{x_2}.h^{r_2})^{2^2}.(g^{x_3}.h^{r_3})^{2^3}.(g^{x_4}.h^{r_4})^{2^4})^{-1}$
  $= g^{x}.g^{-x_0}.g^{-2^1x_1}.g^{-2^2x_2}.g^{-2^3x_3}.g^{-2^4x_4}.h^{r}.h^{-r_0}.h^{-2^1r_1}.h^{-2^2r_2}.h^{-2^3r_3}.h^{-2^4r_4}$
  $= g^{x-x_0-2^1x_1-2^2x_2-2^3x_3-2^4x_4}.h^{r-r_0-2^1r_1-2^2r_2-2^3r_3-2^4r_4}$
  $= g^{x-(x_0+2^1x_1+2^2x_2+2^3x_3+2^4x_4)}.h^{r-r_0-2^1r_1-2^2r_2-2^3r_3-2^4r_4}$
  $= g^{x-(x)}.h^{r-r_0-2^1r_1-2^2r_2-2^3r_3-2^4r_4}$
  $= h^{r-r_0-2^1r_1-2^2r_2-2^3r_3-2^4r_4}$
  On a alors : $\delta = r - r_0 - 2^1r_1 - 2^2r_2 - 2^3r_3 - 2^4r_4$

- $[(\frac{C_4}{g} = h^{\gamma_4} \wedge a_4 = 1)$
  $\vee(\frac{C_4}{g^{a_4}} = h^{\gamma_4} \wedge \frac{C_3}{g} = h^{\gamma_3} \wedge a_3 = 0)$
  $\vee(\frac{C_4}{g^{a_4}} = h^{\gamma_4} \wedge \frac{C_3}{g^{a_3}} = h^{\gamma_3} \wedge \frac{C_2}{g^{a_2}} = h^{\gamma_2} \wedge a_2 = 0)$
  $\vee(\frac{C_4}{g^{a_4}} = h^{\gamma_4} \wedge \frac{C_3}{g^{a_3}} = h^{\gamma_3} \wedge \frac{C_2}{g^{a_2}} = h^{\gamma_2} \wedge \frac{C_1}{g^{a_1}} = h^{\gamma_1} \wedge a_1 = 0)$
  $\vee(\frac{C_4}{g^{a_4}} = h^{\gamma_4} \wedge \frac{C_3}{g^{a_3}} = h^{\gamma_3} \wedge \frac{C_2}{g^{a_2}} = h^{\gamma_2} \wedge \frac{C_1}{g^{a_1}} = h^{\gamma_l} \wedge \frac{C_0}{g} = h^{\gamma_0} \wedge a_0 = 0)]$
  prouve que le bit de poids fort de x est de rang supérieur à celui de a.

FIGURE 3.4 – Preuve d'inégalité : $x > a$ connu (2/2)

Ainsi, lors d'un accès à un site sur Internet, si une information présente dans le formulaire est requise, l'application peut directement pré-remplir la condition sans rien demander à l'internaute. Ceci simplifie la connexion au site web pour le client. Cependant, si une exigence n'est pas contenue dans le formulaire, ou si elle contient une condition (comme pour l'adresse postale), le consentement du client est automatiquement requis. Pour finir, avant d'envoyer ces conditions d'accès remplies, le client doit les vérifier et ainsi donner explicitement son consentement à leur divulgation. La Figure 3.8 de la section 3.6 montre une illustration de cette fonctionnalité.

Une solution est de créer un formulaire de conditions d'accès commun à tous les SP avec uniquement les informations adéquates et nécessaires pour le site Web. De plus, il est possible d'assurer l'authenticité de ce formulaire une fois rempli en le faisant certifier par une autorité de certification ou une autorité de confiance. Cependant, lors de

chaque changement, le client doit refaire certifier son formulaire. Contenant de nombreuses informations personnelles sur le client, le formulaire doit ensuite être stocké dans le coffre-fort électronique de l'application.

Gestionnaire d'identités

Dans le but d'éviter les problèmes d'oubli de mot de passe, il est commun pour le client d'utiliser le même mot de passe pour les différents sites consultés, ce qui pose des problèmes de sécurité. Cependant, l'utilisation d'un même couple permet d'associer des données de différentes bases de données, et donc augmente le risque d'attaques, telles que les usurpations d'identité. Le client peut alors utiliser l'application comme un Gestionnaire d'identités qui génère un couple différent pour chaque SP consulté et le stocke dans son coffre-fort électronique sécurisé.

3.4.2 Application à l'enregistrement en ligne

L'enregistrement en ligne est un cas d'usage parfaitement adapté à notre application où de nombreuses fonctionnalités peuvent être utilisées. La Figure 3.5 illustre le procédé d'enregistrement en ligne et les différentes fonctionnalités de l'application entrant en jeu. Lors d'un tel enregistrement, les utilisateurs doivent fournir diverses informations au prestataire de service, le SP.

Les informations attendues sont présentes dans les conditions d'accès de ce SP. Elles incluent en général : une adresse email valide, un nom d'utilisateur et un mot de passe associé au site Web. Malheureusement, de nombreuses informations supplémentaires et personnelles sont souvent demandées et obligatoires : adresse personnelle, numéro de téléphone, date de naissance, occupations favorites et loisirs. À la fin de ces conditions, l'utilisateur doit généralement répondre à une question en cochant une case souvent invisible : "J'accepte de recevoir des emails avec des offres sélectionnées pour moi par la SP et ses partenaires." Cependant, afin de permettre au SP et à ses partenaires de sélectionner de telles préférences pour l'utilisateur, ils utilisent ses renseignements personnels et suivent ses allées et venues sur le site. En outre, dans de nombreux cas, le consentement de l'utilisateur n'est pas explicitement demandé. Par exemple, les sites français décochent automatiquement cette case, il s'agit d'un consentement appelé opt-in, alors que les sites anglais préfèrent le consentement opt-out. Ainsi, si l'utilisateur ne refuse pas explicitement ces offres en décochant la case, des e-mails publicitaires lui sont envoyés automatiquement. Cette phase d'enregistrement peut donc s'avérer très envahissante pour la vie privée.

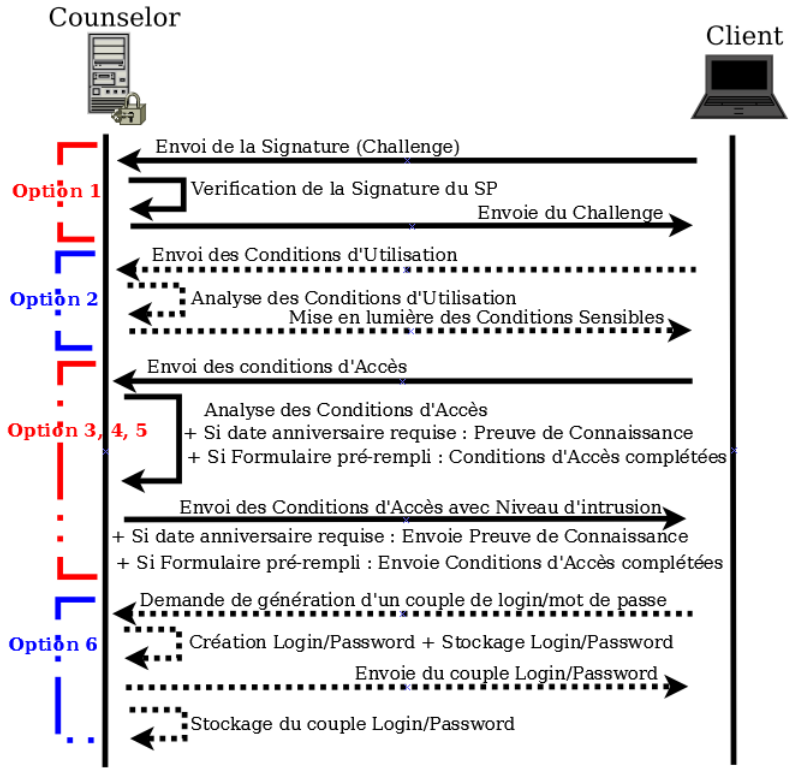


FIGURE 3.5 – Fonctionnalités entrant en jeu lors d'un enregistrement en ligne

Une fois ces conditions d'accès remplies, l'utilisateur doit s'authentifier auprès de ce fournisseur de service. Comme dans une transaction classique, la mise en place d'un canal sécurisé entre le client et le marchand à l'aide d'un protocole tel que SSL/TLS [168, 169] est réalisée. L'usage d'un tel protocole implique que le client ait confiance en son marchand et qu'il accepte les attaques référencées sur un tel protocole [170, 171, 172]. SSL/TLS n'est donc pas suffisant et il est nécessaire de recourir à une authentification multi-facteurs du marchand (ou service provider, SP) et du client. L'anonymat ou le pseudonymat permettent davantage de préserver la vie privée de l'utilisateur durant la connexion à un site commercial. Le "German Act" [173] requiert par exemple de ces fournisseurs de service qu'il puisse offrir à leur client la possibilité de se connecter de façon anonyme ou à l'aide d'un pseudonyme lors du paiement en ligne. Cependant, à notre connaissance, les authentifications de ce type ne sont actuellement pas supportées par les SP français.

L'application proposée dans ce chapitre peut parfaitement être utilisée lors d'un tel enregistrement en ligne. Il est cependant nécessaire d'apporter quelques modifications à l'enregistrement classique. Tout d'abord, le client, qui navigue sur le site Web du marchand,

demande au *SP* de s'identifier à l'aide d'un protocole traditionnel de challenge/réponse. L'application étant ouverte sur son ordinateur, le client peut choisir de lancer la vérification de l'authentification du *SP*. Après cette vérification d'authenticité, le client doit prendre en compte les conditions d'utilisation et d'accès à ce fournisseur. S'il les accepte, il doit renvoyer les informations demandées au *SP*. L'analyseur de conditions d'accès et d'utilisation de l'application entre en jeu en mettant en lumière les informations sensibles demandées par le *SP*, ainsi que les usages non adéquats des données stockées. De plus, si le client a rempli préalablement le formulaire auprès de l'application, celle-ci pré-remplit les conditions d'accès du client. Le formulaire peut dans ce cas contenir certaines informations avec des conditions, par exemple :

- Une adresse email est utile en cas de perte du mot de passe ;
- L'adresse postale est pertinente uniquement pour une livraison à domicile. Ainsi, elle doit être fournie uniquement à la fin de la transaction lors du choix de la livraison ;
- Le numéro de carte est stocké uniquement si le client l'exige étant donné qu'il s'agit d'une information extrêmement sensible et que ce stockage est contraire à *PCI DSS* [174] ;
- L'âge du client est utile si le site requiert un âge minimum. La date de naissance n'est cependant pas nécessaire, excepté, là encore, si le client l'exige. Dans les autres cas, une preuve de connaissance réalisée à partir de la date de naissance du client connue par l'application, est suffisante ;
- Une option peut être cochée si le client souhaite recevoir des emails publicitaires du site.

Si certaines de ces conditions d'accès ont une sensibilité importante en termes d'intrusion de la vie privée, une preuve de connaissance mathématique peut être délivrée au *SP* à la place de certaines informations précises, par exemple, une preuve d'âge supérieur à 16 ans au lieu de la date de naissance, ou encore le numéro de carte. Cette preuve de connaissance est alors générée par l'application. Une fois certain de traiter avec le *SP*, le client doit s'authentifier auprès du *SP*. Suivant le type d'authentification demandée par le *SP*, l'application peut générer pour le client un couple de login/mot de passe associé à ce *SP* et qui sera stocké dans le coffre fort électronique de l'application. Le client s'authentifie alors auprès de ce *SP* et fournit les conditions d'accès à ce site, ainsi que les preuves potentiellement générées. Dans le cas où aucune information particulière n'est exigée par le *SP* et si le *SP* le permet, il est préférable d'utiliser une authentification de type Zero-Knowledge qui peut être réalisée par l'application.

Finalement, si le client respecte l'ensemble des conditions demandées, le service peut être rendu, après une phase de paiement si nécessaire.

3.5 Sécurité et protection de la vie privée de l'application proposée

Grâce à l'application proposée, le client ne fournit aucune donnée, ni information sur ses attentes et besoins avant que l'application lui confirme l'authenticité du SP. La sécurité de cette application est liée à la possession d'un certificat par le SP, aux différents algorithmes d'authentification, ainsi qu'au canal sécurisé utilisé pour chaque transaction permettant d'assurer l'exigence E_2. L'application est capable de générer et de stocker de façon sécurisée les couples de login/mot de passe spécifiques aux différents sites Web où le client est enregistré. Ainsi, l'utilisateur évite les problèmes de corrélation de mots de passe ou d'usurpation d'identité. L'exigence E_1 est ainsi respectée. De plus, si le SP l'accepte, le client a la possibilité d'être anonyme à l'aide d'un protocole d'authentification Zero-Knowledge (E_3). Dans le cas contraire, afin d'éviter de nombreuses attaques, le client peut utiliser l'application pour lui fournir une authentification forte auprès du SP (E_4).

De plus, l'application est entièrement centrée sur le client et permet d'analyser les conditions d'utilisation et d'accès des différents sites consultés. Ensuite, un formulaire peut être complété par le client avec l'aide de l'application. Cette dernière indique alors le degré de sensibilité des renseignements demandés et, si un formulaire a été préalablement complété, elle pré-remplit les conditions d'accès au SP. Le principe de sensibilité des données E_8 est ainsi respecté. De la même façon, le principe de minimisation E_6 des données est assuré. En effet, les seules données divulguées sont les données nécessaires pour le SP. De plus, elles sont révélées uniquement après avoir eu la confirmation de la bonne identité du SP. Qui plus est, si le SP l'accepte, une preuve de connaissance peut être fournie à la place de certaines informations, notamment la date de naissance. L'ensemble des données stockées par l'application le sont dans un coffre-fort électronique, ce qui permet de respecter l'exigence E_5. Finalement, le client a toujours le dernier mot en ce qui concerne ses données, leur gestion au sein de l'application et leur divulgation au SP, le principe de souveraineté des données E_7 est donc également assuré. La Table 3.2 donne une comparaison de notre architecture par rapport aux protocoles évalués précédemment.

3.6 Preuve de concept

L'application n'a pas encore été implémentée dans sa globalité. Cependant, certaines fonctionnalités ont fait l'objet d'étude de projets étudiants et ont ainsi pu être implémentées séparément.

E_i	Propriétés	P3P	Tivoli	Application
E_1	Mots de passe différents	-	Non	**Oui**
E_2	Authentification du SP	-	**Oui**	**Oui**
E_3	Anonymat vis-à-vis du SP	Non	-	**Oui**
E_4	Authentification forte vis-à-vis de la banque du SP	-	Partiel	**Oui**
E_5	Stockage sécurisée des données	-	**Oui**	**Oui**
E_6	Principe de minimisation	Partiel	Partiel	**Oui**
E_7	Principe de souveraineté	**Oui**	Partiel	**Oui**
E_8	Principe de sensibilité	**Oui**	Non	**Oui**
	Score /8	2.5	3.5	**8**

TABLE 3.2: Comparaison des protocoles existants avec l'application proposée

3.6.1 Authentification via lecteur CAP

Les trois méthodes disponibles via le protocole *CAP* : Code, Challenge/Response et Sign ont été implémentées. L'interface graphique simulant le lecteur *CAP* a été développée en Java. La Fig.3.6 représente une illustration de la méthode Challenge/Response. Après avoir entré son identifiant *UserID* sur le site Web du fournisseur de service, le client doit entrer le challenge 45242622 dans son lecteur *CAP*. Ce dernier calcule alors le code réponse associé, ici 15230. L'utilisateur entre ce code sur son site et clique sur "Verify". Le fournisseur de service calcule à son tour le code réponse. S'il arrive au même résultat que le lecteur *CAP*, le client est authentifié.

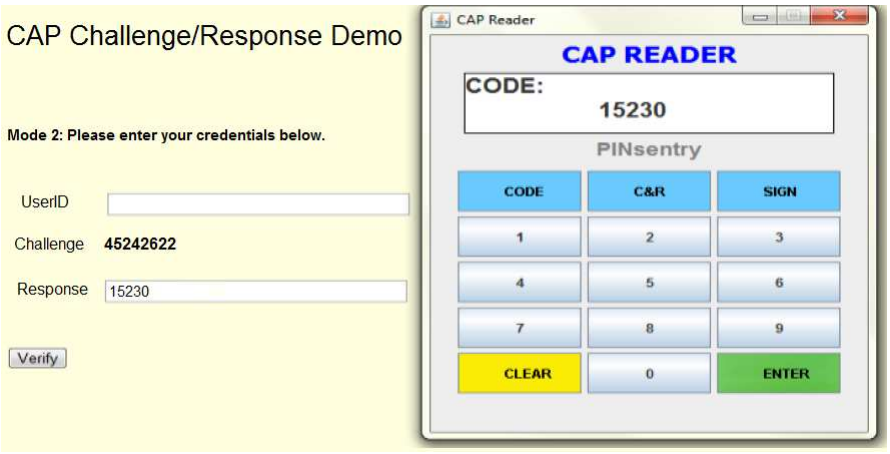


FIGURE 3.6 – Page de réglement par Challenge/Response

3.6.2 Authentification Zero-Knowledge

Le protocole Zero-Knowledge de Shnorr a été implémenté sur carte à puce sous la forme d'une applet Java Card en version 2.2.2 afin d'utiliser un environnement standardisé et certifié. Il fait appel à des paramètres de taille 512 bits. La carte à puce utilisée lors de cette implémentation est la carte NXP J3A080 dont les caractéristiques principales sont la double interface contact et sans-contact et la présence de plusieurs coprocesseurs pour l'exécution de protocoles tels que RSA ou DES.

3.6.3 Analyse des conditions d'utilisation

Lors de l'enregistrement en ligne d'un client, les conditions d'utilisation du SP sont analysées et certaines conditions allant à l'encontre de la vie privée (ici quatre) sont mises en lumière. L'application se contente de lire les conditions d'utilisation et de repérer des termes tels que "vie privée", "données personnelles stockées", "coordonnées bancaires", etc... Cette partie du logiciel est implémentée en Java. De plus, comme le montre la Fig.3.7, l'application affiche ensuite les phrases des conditions d'utilisation contenant les différents termes indiqués. Cette analyse permet au client de se rendre compte du niveau de sensibilité de la politique de sécurité du SP.

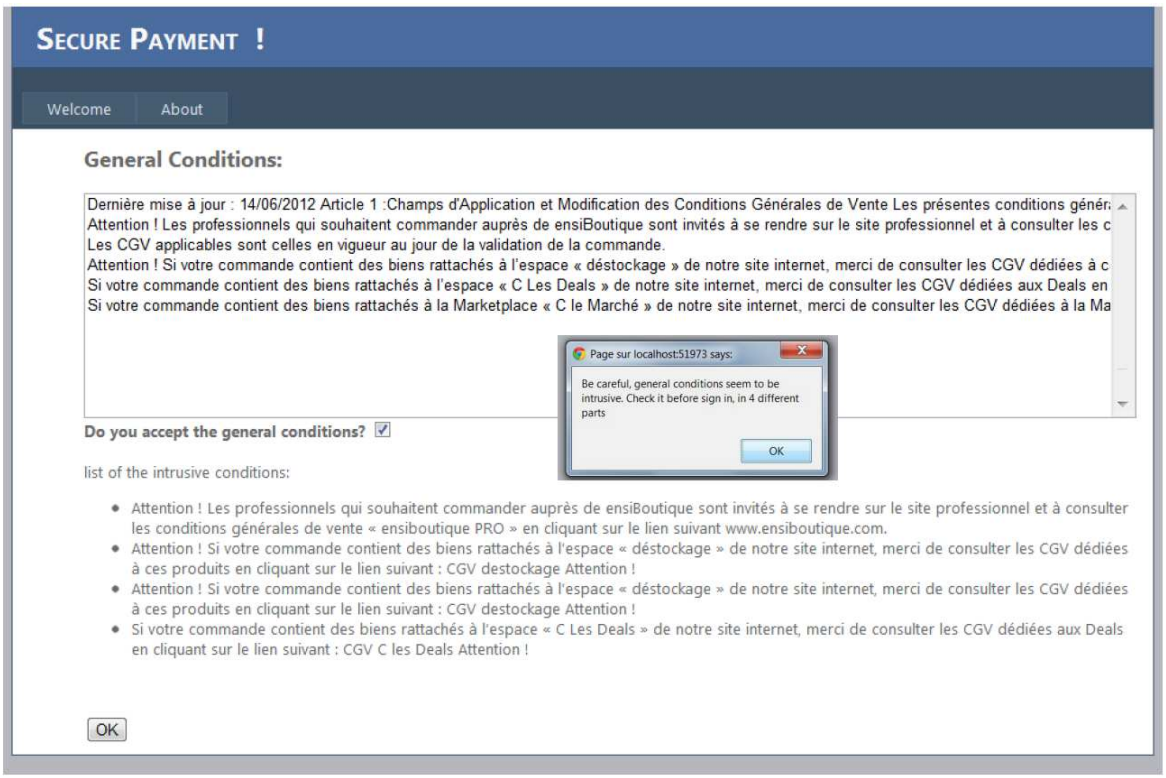


FIGURE 3.7 – Résultat de l'analyse des conditions d'utilisation

3.6.4 Formulaire et conditions d'accès

Dans le cas où l'utilisateur a pré-rempli des conditions d'accès, l'application peut les remplir pour lui. À tout moment, le client peut modifier les conditions qu'il accepte ou non de divulguer. La Figure 3.8 illustre l'implémentation du remplissage du formulaire par le client. Dans notre cas, les informations requises par le SP sont l'âge et la somme d'argent du pour le service demandé. Si le client accepte de les dévoiler, il clique sur "accept" et la transaction peut continuer. Sinon, le client peut modifier les données qu'il accepte de partager en cliquant sur "Modify shared data". Il est alors redirigé vers une autre page où il peut ajouter ("insert") et/ou supprimer ("delete") des données dont il autorise la divulgation.

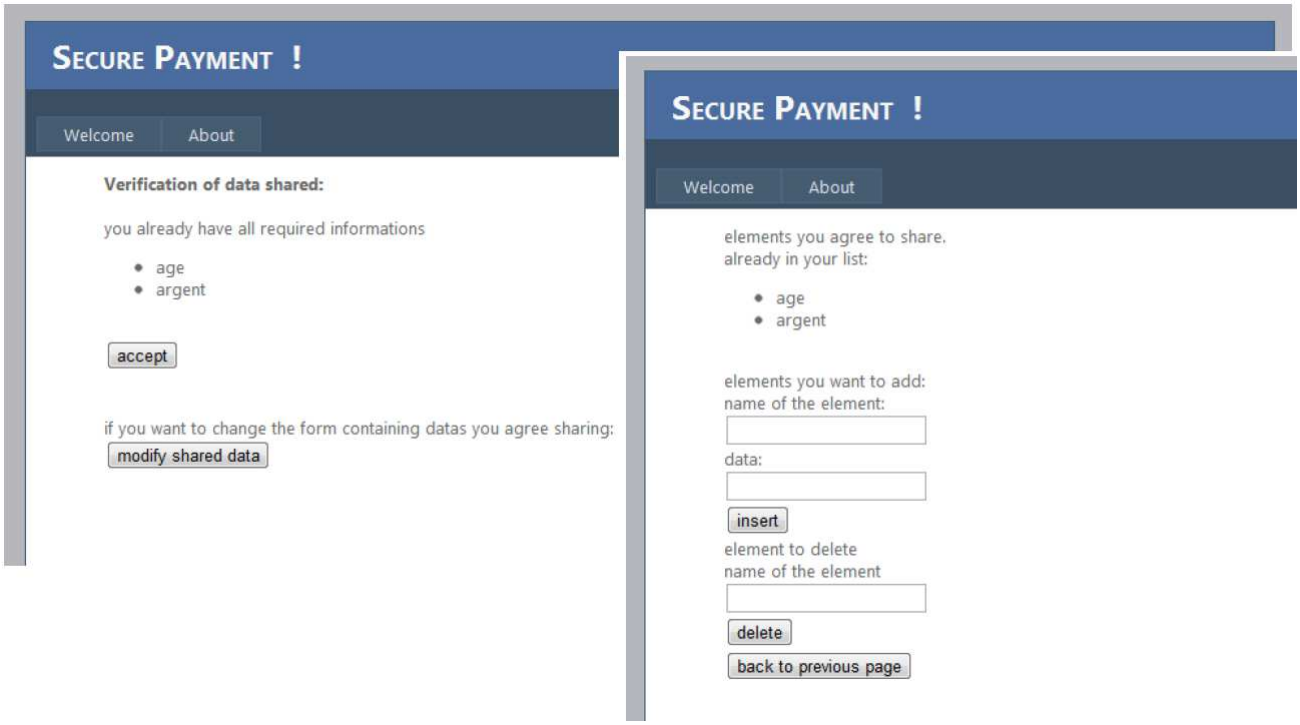


FIGURE 3.8 – Gestion du formulaire

3.7 Conclusion

Le grand nombre d'informations personnelles stockées dans les bases de données des fournisseurs de service requiert une attention particulièrement importante. Les informations de l'utilisateur sont effectivement souvent réutilisées à des fins statistiques et publicitaires, et donc rarement considérées comme des données sensibles appartenant au client.

L'application proposée utilise un formulaire pré-rempli, des outils cryptographiques bien connus tels que les certificats, les protocoles d'authentification, et les preuves de connaissance. Cette application est capable de jouer plusieurs rôles. Elle peut ainsi vérifier des signatures, analyser les conditions d'utilisation et d'accès au site concerné, fournir une preuve de connaissance, gérer un coffre-fort électronique et enfin, gérer les différentes identités du client. Ces fonctionnalités guident l'utilisateur lors de sa navigation sur Internet et le tient informé de l'importance de ses informations. Cette proposition offre ainsi une solution sécurisée, facile d'utilisation et donne un exemple d'application respectueuse de la vie privée qu'il est possible d'obtenir en utilisant certains des différents outils cryptographiques détaillés dans le chapitre 2.

À terme, nous souhaitons réunir l'ensemble des fonctionnalités programmées afin d'obtenir une unique application (un plugin sur son navigateur) pouvant être installée directement sur l'ordinateur des utilisateurs. D'autres domaines plus sensibles méritent également notre attention et nécessitent l'usage d'outils cryptographiques afin de protéger davantage les données personnelles entrant en jeu. C'est le cas notamment des données médicales des patients qui sont le thème du chapitre suivant.

Travail de l'auteur sur ce thème

- **A. Plateaux**, P. Lacharme, K. Murty, C. Rosenberger, *Online user's registration respecting privacy*, World Congress on Computer and Information Technologies (WCCIT), Sousse, Tunisie, Juin 2013.
- P. Lacharme, **A. Plateaux**, *PIN-based cancelable biometrics*, International Journal of Automated Identification Technology (IJAIT), 2011.

Chapitre 4

Protection des données personnelles médicales

De nombreuses infrastructures gérant les problèmes relatifs aux dossiers médicaux émergent dans plusieurs pays. Nous nous sommes intéressés dans cette thèse à la protection des données personnelles dans de tels systèmes utilisant de nombreuses TES, où la vie privée n'est que partiellement traitée, et où les concepts de minimisation et de souveraineté des données sont souvent négligés. Ce chapitre présente une infrastructure de e-santé visant à protéger les informations personnelles et à minimiser leur divulgation. De plus, le principe de souveraineté des données est assuré en conformité avec les contraintes médicales. Une implémentation logicielle permet de démontrer la faisabilité des concepts définis.

Sommaire

4.1 Introduction

Beaucoup d'informations personnelles et sensibles, ainsi que de nombreux acteurs (médecins, infirmières, patients ...) entrent en jeu dans les applications médicales. Des données médicales et administratives (nom, prénom, adresse...) y sont traitées par différents établissements (clinique standard, hôpital psychiatrique, cabinet ...). Le développement de l'e-santé est inéluctable. Par conséquent, le cas des dossiers médicaux est directement concerné par les problèmes de protection de données et leur sécurité est une priorité renforcée au vue de la grande sensibilité des données traitées.

De nombreuses réglementations existent et sont spécifiques au contexte de l'e-santé et au pays considéré. Ainsi, l'*HIPAA* (Health Insurance Portability and Accountability Act, [175]) est une loi de 1996 qui régit la gestion de l'assurance maladie aux Etats-Unis. La recommandation [176] de l'Union européenne définit, quant à elle, deux expressions importantes dans le domaine de la vie privée au sein d'un système de santé. Les *données personnelles* couvrent les informations relatives à une personne identifiable, alors que les *données médicales* se réfèrent à toutes les données personnelles relatives à la santé d'un individu. Par conséquent, ces données appartiennent au patient qui dispose de droit d'accès et de rectification sur ces informations. En France, la loi du 4 mars 2002 [177] régit les droits des malades et le code de santé publique [178] régit les différentes lois concernant ce contrôle d'accès aux informations. Plus précisément, l'article $L1111-7$ en ce qui concerne les principes généraux, les articles $R1111-1$ et $R1111-8$ sur l'accès aux informations de santé à caractère personnel. Les termes du dossier médical personnel (ou DMP) et dossier pharmaceutique (ou DP) sont quant à eux regroupés entre l'article $L1111-14$ et $L1111-24$ de ce même code.

Cependant, comme nous le voyons dans la suite, le DMP ne garantit pas la totalité des exigences attendues par un tel système de e-santé et beaucoup de publications sont centrées sur la sécurité des systèmes d'information de e-santé et sont généralement limitées aux hôpitaux seuls, sans aucune interaction entre eux. Le principe de minimisation des données et la notion d'intraçabilité sont au plus partiellement traitées et le principe de souveraineté des données est parfois affirmé sans vraiment être développé.

L'architecture proposée est une solution décentralisée d'un système d'informations de e-santé respectueuse de la vie privée. Elle traite aussi bien le problème de gestion d'identités à l'intérieur d'une institution médicale, que celui de chiffrement des dossiers

médicaux. L'approche est voulue simple et utilise uniquement des outils bien connus de cryptographie, comme une PKI pour l'authentification du personnel médical, l'AES pour la gestion des identités anonymes des patients, ainsi que le principe de partage de secret de Shamir pour une gestion des données chiffrées plus respectueuse de la vie privée des patients. Le principe de la souveraineté des données est pris en compte, et ceci même dans des contextes spécifiques des systèmes médicaux comme un événement d'urgence.

Ce chapitre commence par préciser les exigences spécifiques à respecter dans un contexte de e-santé. La section 4.3 continue sur un état de l'art concernant la protection des données médicales. La section 4.4 décompose et explique ensuite la nouvelle architecture, au sein d'un organisme de santé dans un premier temps puis entre deux institutions. L'analyse de la solution proposée est présentée dans la section 4.5. Les deux sections suivantes donnent un aperçu du démonstrateur de l'application mise en place et des perspectives de ce système.

4.2 Exigences de sécurité et de protection de la vie privée

Les problèmes de sécurité et de protection de la vie privée dans les dossiers médicaux sont considérés depuis le milieu des années 90, notamment par Anderson dans [179]. Il a proposé un ensemble de principes sur les traitements de données pour la e-santé qui doivent être vérifiés par toute politique de sécurité clinique. Plus récemment, l'étude EPHR (European Privacy and Human Rights, [180]) a présenté en 2010 un cadre de règles pour la vie privée dans l'Union européenne pour de nombreux cas. L'analyse des dossiers médicaux y été réalisée selon différents critères prenant en compte la vie privée des patients : contrôle de données par le patient, protection contre les utilisations secondaires, mesures de sécurité lors de l'utilisation et de la collecte des données et enfin utilisation de registres décentralisés. La création d'une base de données nationale centralisée contenant les dossiers médicaux a été envisagée dans de nombreux pays (comme pour le système britannique NHS) afin d'améliorer la disponibilité des dossiers médicaux. Néanmoins, de fortes critiques des praticiens, confondues avec une réaction négative de l'opinion publique, a émergé et dans plusieurs pays cette idée a été abandonnée. En effet, les données centralisées peuvent causer une perte totale des informations médicales en cas de problème du système, et une divulgation de ces informations aurait de graves conséquences nationales pour la vie privée des citoyens ([181] et [182]) et malades. Un système décentralisé est donc une solution logique et une conséquence directe du principe de sensibilité des données.

Afin d'accéder à ces informations stockées de manière décentralisée, la personne (ou le groupe de personnes) doit fournir les **droits appropriés** à la lecture ou à l'écriture de ce dossier médical. Le nom de la (ou des) personne(s) doit ensuite être noté dans le fichier afin de pouvoir vérifier à tout moment l'**historique des accès**. Ces contrôles permettent d'assurer une certaine confidentialité et intégrité des données en vérifiant que seules les personnes autorisées et ayant les bons droits, accèdent aux fichiers. Cependant, ces contrôles ne sont pas suffisants et il est nécessaire d'ajouter un niveau à la protection de ces données, par exemple, en les chiffrant. La **disponibilité des données** au sein d'une même institution, tout comme lors d'un **transfert entre deux organismes**, doit également être prise en compte dans un contexte de e-santé. Ces données peuvent en effet être partagées entre plusieurs prestataires de santé, un hôpital public et une clinique spécialisée par exemple.

Ensuite, au sein d'un même hôpital, les droits sont différents en fonction des employés. Un médecin peut accéder aux informations médicales de ses patients alors qu'une infirmière aura uniquement accès aux ordonnances. Dans les deux cas, et par le principe de minimisation des données, ils n'ont pas besoin de connaître les informations administratives du patient, à l'exception de son âge. Une **politique d'accès précise**, prenant en compte le rôle des acteurs et leurs identités, doit également être utilisée pour garantir la minimisation des données. Ce principe prévoit également le pseudonymat des données médicales et ainsi : le principe de **non-associabilité des données** et la non-corrélation des identifiants. Tous les dossiers médicaux sont donc anonymisés, sauf dans des cas exceptionnels, comme pour la fusion de plusieurs institutions. Toutefois, dans un cadre d'enquête ou d'urgence, l'anonymat doit pouvoir être levé. La propriété de **réversibilité** doit donc être ajoutée à notre proposition. Par ailleurs, des mesures efficaces afin d'empêcher l'agrégation de données de santé sont réalisées au sein d'un même hôpital et entre deux institutions.

Il faut rappeler que l'ensemble des informations médicales appartiennent au patient et non au médecin qui les crée, ou à l'hôpital qui les stocke. Cela signifie que les patients ont le **contrôle de leurs données** et peuvent y accéder librement. Cependant, la relation de confiance entre un médecin et son patient donne implicitement au docteur l'accord du malade pour accéder à son dossier médical. De même, le transfert entre deux établissements doit être réalisé avec le consentement du patient. Anderson rappelle, dans [182] et [183], que 50% des médecins n'aimeraient pas télécharger les détails cliniques du patient sans leur consentement spécifique et que de nombreux flux illégaux d'informations ont été découverts dans le NHS britannique [181, 184]. Ceci résulte du conflit entre le *consentement* du patient et la *nécessité de connaissance* des médecins. Par ailleurs, le principe de souveraineté des données doit prendre en compte les scénarii où la vie du patient dépend de

l'information externe, notamment dans les **cas d'urgence**. Ainsi, tout comme le principe de minimisation, la propriété de réversibilité des données est nécessaire. Nous avons alors extrait de ces analyses, un ensemble de seize exigences devant être respectées afin de protéger au maximum les données des patients au sein de divers hôpitaux :

• E_1 : La **confidentialité des transactions** exige que chaque information échangée soit chiffrée lors de leur transfert.
• E_2 : La **confidentialité des données stockées** exige que chaque information stockée soit chiffrée.
• E_3 : L'**intégrité** de l'information transmise permet d'assurer l'exactitude du contenu et donc la non-altération des données lors de leur transmission.
• E_4 : La **disponibilité** des données permet d'accéder en temps réel à l'information.
• E_5 : L'**authentification du patient** pour l'accès à ses données.
• E_6 : L'**authentification des employés** permet de contrôler l'appartenance d'un employé à un institut médical particulier.
• E_7 : Le **contrôle d'accès des employés** permet de contrôler que l'employé est en droit d'accèder à un dossier.
• E_8 : Des **politiques de sécurité différentes** en fonction des employés, du dossier et de la donnée demandée.
• E_9 : La **gestion séparée** du stockage des données et des identifiants afin d'éviter un lien entre les données correspondant aux différents identifiants d'un même patient.
• E_{10} : La **non-associabilité** des informations **entre les différentes bases de données** au sein d'un (ou entre plusieurs) établissement(s) de santé. Cette exigence implique l'usage d'identifiants différents et la séparation des données médicales et d'identité.
• E_{11} : L'**anonymat du patient** est garanti si les exigences E_3, E_4, E_9 et E_{10} sont assurées.
• E_{12} : Le **principe de souveraineté des données** implique que les données personnelles du patient lui appartiennent avec son contrôle et son consentement sur leur utilisation.
• E_{13} : Le **principe de sensibilité des données** implique que les données personnelles nécessitent une structure décentralisée pour leur stockage.
• E_{14} : La prise en compte des **cas d'urgence** implique la possibilité, dans un cas extrême, d'avoir accès à certaines données importantes sans l'accord préalable du patient.
• E_{15} : La gestion des **transferts entre différents établissements de santé** permet d'assurer également l'ensemble des exigences entre établissements.
• E_{16} : La **sauvegarde de l'historique** d'accès à des données afin de détecter, par exemple, les cas d'usurpation d'identité acceptés (ou non) par l'employé en question.
• E_{17} : La **facilité d'utilisation** pour le patient. En effet, dans le cas d'un système trop complexe pour le patient, ce dernier pourrait se contenter du minimum à gérer et donc ne pas utiliser l'ensemble des possibilités offertes pour protéger ses données.

Les exigences pour le respect des données des patients sont donc nombreuses. Cependant, à notre connaissance, et comme le confirme la section suivante, aucune architecture de e-santé dans la littérature ne respecte la totalité d'entre elles.

4.3 État de l'art sur la protection des données personnelles médicales

4.3.1 Le Dossier Médical Patient

En France, le système de e-santé le plus connu (et reconnu) est le *DMP*. Il s'agit d'un dossier informatisé permettant de regrouper les informations médicales de tout bénéficiaire de l'assurance maladie et de les partager entre les professionnels et établissements de santé. Il a été créé par la loi du 13 août 2004 relative à l'assurance maladie [185] mais n'a cependant été autorisé qu'en décembre 2010 par la *CNIL* [186]. Le *DMP* permet de respecter un certain nombre des exigences soulignées précédemment. En effet, ce dossier est uniquement consultable par les personnels de santé ayant directement obtenu les droits par le patient (E_7). Ces droits se matérialisent par la remise de la carte vitale par le patient aux personnels de santé. Les données y sont stockées chiffrées (E_2), les cas d'urgence pris en compte via le mode "bris de glace" (E_{14}) et le consentement du patient est considéré (E_{12} partielle). De plus, un historique d'accès est conservé et le transfert de donnée assurée (E_{15} et E_{16}). Il est possible d'accéder à tout moment à ce *DMP* via la plate-forme de santé et après une authentification, ou encore de la supprimer si le patient le souhaite (E_4, E_5, E_6). Cependant, certains points ne sont pas assurés ou restent très vagues. En effet, la CNIL suggère une séparation des données d'identités et des données médicales du patient, mais il n'existe qu'un identifiant national de santé *INS* unique pour chaque patient et les méthodes de séparation de ces informations ne sont pas détaillées (E_9 partielle, E_{10} et E_{11} non assurées). De plus, l'identifiant est connu à la fois des professionnels de santé autorisés et de l'hébergeur des données. Qui plus est, l'hébergeur stocke les données de façon centralisée (E_{13} non assurée) et, lorsque le patient donne son consentement à un établissement de santé, tous les professionnels de santé ont accès à ces données (E_8 et E_9 partielles). Afin de contrer ce dernier point, le patient doit refuser l'accès à toutes les personnes de l'hôpital qui ne le suivraient pas. Le *DMP* se révèle donc très lourd pour le patient qui doit gérer l'ensemble de ces données et des droits d'accès (E_{17} non assurée). Pour finir, il s'avère couteux comme le montre le rapport récent de la Cours des Comptes datant du 19 février 2013 [187].

4.3.2 Autres systèmes de la littérature

Une approche décentralisée des systèmes de e-santé a récemment été suggérée par les auteurs de [188]. Cette solution donne lieu à une gestion d'identité basée sur un identifiant local par patient. Ces identifiants permettent d'éviter l'agrégation des informations personnelles de santé. Une première solution est la gestion de ces identifiants locaux par une autorité centrale de confiance. Ce tiers possède une table globale des identifiants locaux qui permet de transférer les identités nécessaires d'un service à un autre. Néanmoins, cette approche présente des problèmes similaires à l'approche centralisée. Effectivement, la table est vulnérable à la suppression et à la divulgation de données. Ghindici dans sa thèse [189] a proposé une approche cryptographique en fournissant aux trois acteurs une clé privée. Cependant, cette approche étant un exemple d'application, elle n'a pas été proposée pour être réalisable dans une situation réelle.

Afin d'éviter de possibles agrégations, le gestionnaire d'identité ne doit pas stocker de liens entre les différents identifiants d'un patient sur une même base de données. Cette solution est proposée par Deng et coll. dans [190] et [191]. Cependant, ni les droits d'accès du personnel médical, ni l'authentification entre différents hôpitaux n'est pris en compte. Un autre problème vient du principe de non-associabilité. En effet, un médecin peut demander l'identifiant local au fournisseur d'identité grâce à une simple description du patient. Cela signifie qu'il y a une table d'information reliant les deux types de données. Cependant, la description du patient peut être une alternative à un identifiant global et offrir ainsi une possibilité d'agrégation.

D'autres alternatives pour le respect de la vie privée dans un contexte médical américain sont proposées par Ateniese et Medeiros ([192]) et par De Decker et coll. pour le système belge, [193]. Dans le premier cas, un schéma de signature de groupe est proposé, alors que dans le second, il s'agit de certificats préservant la vie privée. Ce dernier utilise de nombreux principes tels que la minimisation des données ou la non-associabilité entre organisations (pharmaciens, organisme de sécurité sociale,..). Malheureusement, ces différents protocoles ne considère pas le transfert d'informations entre institutions de santé, ni le principe de souveraineté des données.

4.3.3 Discussion

Une étude plus détaillée du *DMP* et de l'architecture de Mina Deng en fonction des exigences est décrite dans le Tab. 4.3. Cette étude confirme le manque de protection des données personnelles du patient au sein d'un établissement de santé et entre différents

établissements, et ceci même lorsqu'il s'agit d'architecture ayant pour but de protéger les données personnelles. De plus, ces méthodes restent difficiles à utiliser pour le patient et le partage entre différents établissements pose souvent problème.

$\mathbf{E_i}$	**Propriétés**	**DMP**	**Architecture de Mina Deng**
E_1	Confidentialité des transactions	-	**Oui**
E_2	Confidentialité des données	**Oui**	**Oui**
E_3	Intégrité	**Oui**	-
E_4	Disponibilité	Partiel	**Oui**
E_5	Authentification de Patient	**Oui**	**Oui**
E_6	Authentification des employés	**Oui**	**Oui**
E_7	Contrôle d'accès des employés	**Oui**	**Oui**
E_8	Politiques de sécurité différentes	Partiel	Non
E_9	Gestion séparée des données	**Oui**	Non
E_{10}	Non-associabilité	Non	**Oui**
E_{11}	Anonymat des bases de données	Non	**Oui**
E_{12}	Souveraineté des données	**Partiel**	Non
E_{13}	Sensibilité des données	Non	**Oui**
E_{14}	Cas d'urgence	**Oui**	Non
E_{15}	Transfert entre établissements	**Oui**	**Oui**
E_{16}	Historique sauvegardé	**Oui**	Non
E_{17}	Usage facile pour le patient	Non	Non
	Score /17	**10.5**	10

TABLE 4.1: Synthèse des analyses du *DMP* et de l'architecture de Mina Deng.

Ainsi, au vu des exigences fixées, nous proposons dans ce chapitre une nouvelle architecture plus efficace par son utilisation de la cryptographie symétrique et surtout plus respectueuse de la vie privée du patient.

4.4 Architecture de e-santé proposée

4.4.1 Portée de la protection

Dans le schéma proposé, nous utilisons une *PKI*, décrite par la Fig. 4.1, comme une infrastructure hiérarchique avec une autorité de certification. Cette autorité génère et stocke des paires de clés publique/privée utilisées pour signer les certificats des différents fournisseurs d'e-santé $H_1, \ldots, H_n$ afin de créer une relation de confiance entre eux. Ceux-ci génèrent et stockent à leur tour des paires de clés utilisées pour signer les certificats des membres du personnel M_i. En plus de ce certificat contenant nom, prénom, hôpital d'affi-

liation, droit d'accès, profession, date de validité du certificat, clé publique et algorithme utilisé, le personnel possède une identité numérique sous forme de login et mot de passe [1].

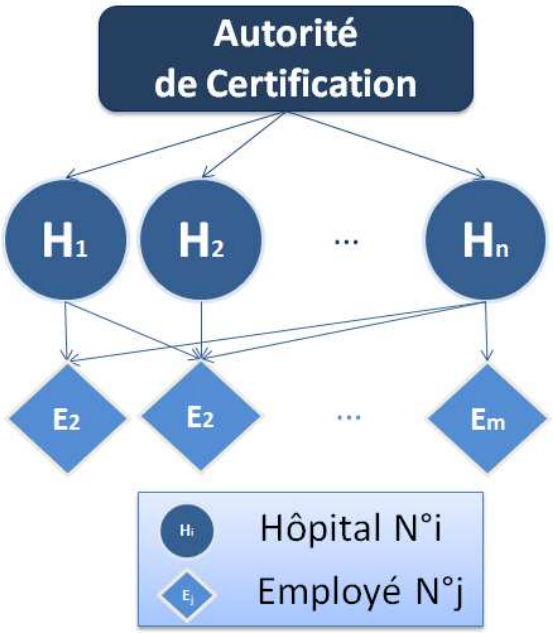


FIGURE 4.1 – Infrastructure à clé publique pour le système e-santé

À l'intérieur d'un système d'e-santé, les données personnelles médicales et administratives sont stockées dans autant de bases de données différentes nécessaires afin d'éviter un possible lien entre ces informations. Dans ce chapitre, nous nous contentons, afin de simplifier les explications, de deux bases de données. Le procédé s'itère facilement pour un plus grand nombre de bases de données. Par ailleurs, celles-ci sont pseudonymisées, ce qui entraîne que seules les personnes autorisées peuvent récupérer la véritable identité du patient. Un chiffrement de ces bases de données apporte également une meilleure sécurité.

Après une authentification réussie et avec le consentement du patient, un employé peut accéder aux données médicales du malade. S'il s'agit d'un médecin ayant eu accès aux données de ce patient dans un autre établissement qui possède d'autres informations, il doit mettre à jour l'ensemble des données via un canal chiffré sécurisé. Cependant, l'authentification ne suffit pas à la protection des données à caractère personnel. En effet, les droits d'accès spécifique de la personne authentifiée doivent être vérifiés. Les certificats peuvent par exemple permettent de gérer cette politique d'accès.

1. Le fait d'utiliser une carte à puce sécurisée pour le personnel médical permettrait d'accroître la sécurité de cette authentification. Dans ce cas, la carte pourrait contenir le certificat numérique du titulaire, avec la signature de l'autorité de certification.

À l'intérieur de l'hôpital et dans l'architecture proposée, le gestionnaire d'identité représente l'élément de confiance. Son rôle est de créer et gérer les identifiants locaux du patient à partir de son identifiant global *IdG*. L'identifiant global est l'unique numéro d'identification du patient utilisé dans toutes les structures de santé. En France, il peut s'agir par exemple du numéro de Sécurité Sociale qui, bien que quasiment publique, est en réalité une donnée unique et donc particulièrement sensible. Son stockage dans une table comportant tous les identifiants serait risqué pour la vie privée des patients. Afin de simplifier le système, nous utilisons directement l'identifiant global. Toutefois, afin d'éviter la lecture ou le stockage de celui-ci, l'utilisation d'une fonction de hachage, calculée par la carte du patient (telle que la carte vitale en France), est recommandée.

Deux contrôleurs d'identité sont ensuite utilisés pour vérifier les identités et droits d'accès aux deux bases de données à l'intérieur d'un établissement de santé. Le contrôleur d'accès spécifique aux informations médicales est nommé *CAM* (Contrôleur d'Accès Médical), celui correspondant aux données administratives est le *CAA* (Contrôleur d'Accès Administratif). La Figure 4.2 donne un aperçu de la gestion du contrôle d'accès.

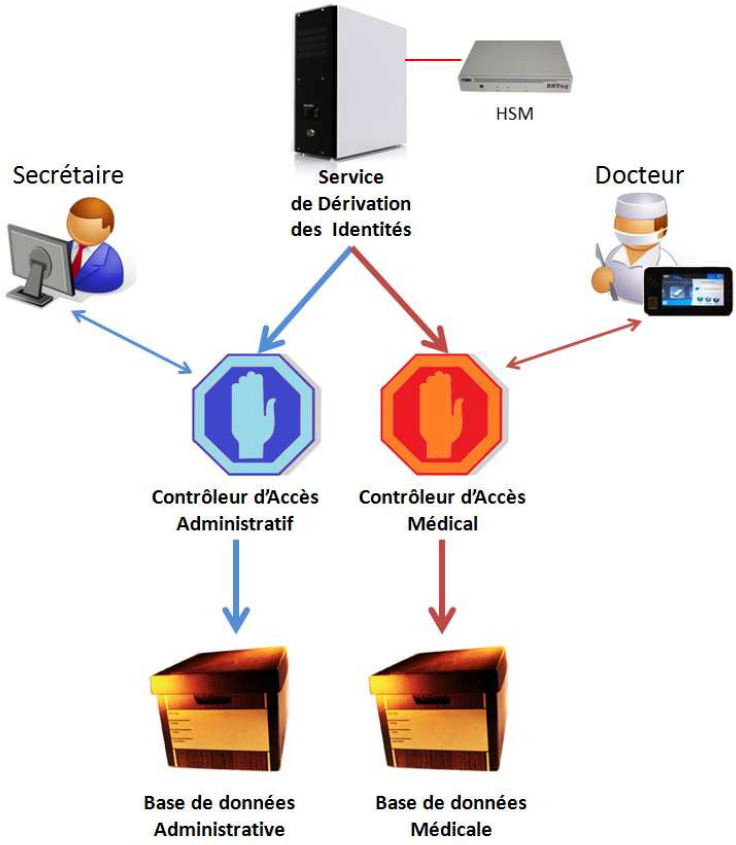


FIGURE 4.2 – Organisation de l'hôpital

4.4.2 Gestion d'identité à l'intérieur d'un hôpital

Dans cette architeture, un patient possède deux identifiants locaux, IdL_1 et IdL_2 calculés à partir de son identifiant global IdG. Ces identifiants locaux sont différents pour chaque base de données afin d'éviter la traçabilité des données et donc leur associabilité. Dans la Fig. 4.3, deux bases de données sont donc considérées : une contient les données médicales associées à l'identifiant IdL_2, l'autre les renseignements administratifs associés à IdL_1.

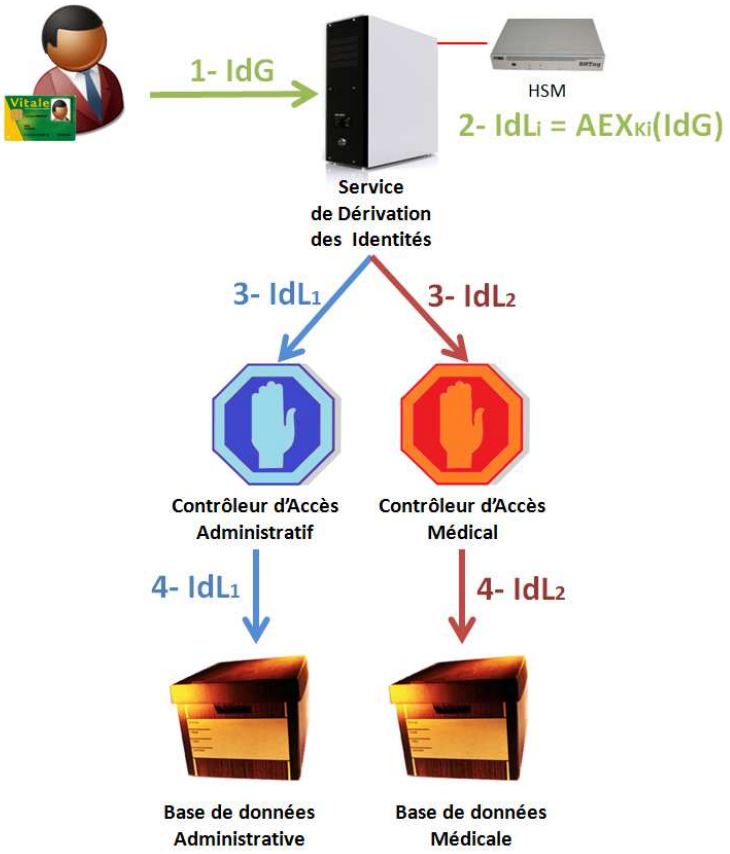


FIGURE 4.3 – Gestion des identités et contrôle d'accès pour une architecture e-santé

Nous notons que l'âge du patient, et non sa date de naissance, est présent dans son dossier médical. En effet, il est nécessaire pour un docteur de connaître l'âge de ses patients pour des raisons de prescriptions et de traitements. Durant la phase d'enregistrement d'un patient P, décrite dans la Fig. 4.4, le service de dérivation d'identité SDI utilise deux clés secrètes K_1 et K_2. Il applique ensuite l'algorithme AES à l'identifiant global IdG du patient avec K_1 pour obtenir l'identifiant local IdL_1 puis K_2 pour IdL_2. Ces clés sont gérées par l'institution indépendamment du IdG.

1. P donne son IdG au gestionnaire d'identité SDI.
2. SDI utilise K_1 et K_2 pour calculer les identifiants locaux : $IdL_1 = AES_{K_1}(IdG)$ et $IdL_2 = AES_{K_2}(IdG)$.
3. SDI supprime le IdG (qui ne doit jamais être stocké) et retourne IdL_1 et IdL_2 au patient (identifiant pour accéder au données administratives et médicales).
4. SDI retourne IdL_1 à CAA et IdL_2 à CAM.
5. Le patient donne son IdL_2 à son docteur.
6. Le docteur ajoute IdL_2 à sa liste de patients.

FIGURE 4.4 – Phase d'enregistrement

Le gestionnaire d'identité SDI n'a jamais accès aux données et est le seul à pouvoir récupérer l'identifiant global du patient avec la connaissance des clés secrètes K_1 et K_2. La gestion de ces clés secrètes doit être réalisée de manière sécurisée et à l'intérieur d'un élément de sécurité, comme dans un HSM (Hardware Security Module). L'identifiant global doit également être supprimé afin d'éviter la divulgation accidentelle de cette information très sensible : une base de données liant identifiants globaux et locaux des patients serait aussi risquée qu'un système centralisé. Ainsi, les contrôleurs d'accès CAA et CAM ne connaissant pas respectivement les identifiants IdL_2 et IdL_1, aucun lien ne peut donc être fait entre données médicales et administratives.

Dans l'étape suivante, le patient reçoit un formulaire médical avec ses nom, prénom et identifiant local IdL_2, utilisé comme pseudonyme dans son dossier médical. Ce formulaire peut prendre la forme d'un bracelet avec un tag $RFID$ (Radio Frequency IDentification) qui contiendrait les informations précédemment citées. Lors de la consultation, il donne cet identifiant au médecin comme une preuve de son consentement pour l'accès et la mise à jour de son dossier médical. Le médecin doit ainsi gérer, dans son ordinateur personnel, une liste de ses patients avec leurs identifiants locaux associés.

Dans chaque service consulté, le patient présente son formulaire médical ainsi que son identifiant local IdL_2, et permet ainsi au service d'accéder à son dossier médical. L'accès aux données requiert l'identifiant local du patient, ainsi qu'une authentification login/mot de passe. Le contrôleur d'accès vérifie ensuite la profession médicale et l'hôpital d'affiliation de l'employé. Un enregistrement de l'identité du demandeur, de la date et de l'heure est également réalisé par le gestionnaire d'identité. Malgré toutes ces vérifications, si un médecin obtient frauduleusement IdL_2 et accède aux données médicales d'un patient

qu'il ne suit pas, ce docteur pourra être retrouvé grâce à l'historique de consultations des dossiers. Ainsi, un membre du personnel ne connaissant pas l'identifiant local du patient sera considéré comme n'ayant pas le consentement du patient et ne pourra donc ni lire, ni modifier le fichier de données du patient. La procédure d'accès au dossier médical, qui est identique pour chaque employé, est détaillée dans la Fig. 4.5.

1. Le docteur fournit son login, son mot de passe et l'identifiant local IdL_2 de son patient à *CAM*.
2. *CAM* contrôle l'identité du docteur, son certificat personnel et ses droits d'accès.
3. *CAM* autorise ou non l'accès au dossier médical du patient puis répond à la demande de l'employé en utilisant l'identifiant local IdL_2. L'identité du demandeur, la date et l'heure sont enregistrées.

FIGURE 4.5 – Accès aux données médicales connaissant l'identifiant local

Dans un cas d'urgence où le patient est inconscient et où son identifiant local est inconnu, un médecin agréé ou le service ayant besoin d'accéder aux dossiers médicaux du patient, peut se contenter de donner l'identité du patient. Si le patient est déjà enregistré, le gestionnaire récupère alors l'identifiant local IdL_1 du patient, puis son identificateur global IdG à partir de K_1 et Id_1, et enfin l'identifiant local IdL_2 à partir de IdG et K_2. Ce cas est décrit dans la Fig. 4.6. Dans le cas où le patient ne serait pas enregistré, son dossier sera complété avec ses informations personnelles lorsqu'il en sera capable.

Dans ce cas, les contrôleurs d'accès enregistrent l'identité du demandeur, la date, l'heure et le type de renseignements médicaux demandés. Une fois le patient conscient, il peut être informé du déroulement des opérations. Ajoutons que les nom, prénom et année de naissance figurant dans le dossier administratif ne doivent pas être chiffrés afin d'assurer une telle admission au urgence.

4.4.3 Chiffrement des bases de données

Afin d'accéder aux dossiers médicaux, l'employé fournit un certain nombre de renseignements sur ses droits et sur son patient. S'il s'agit d'un employé autorisé, le contrôleur d'accès *CAM* lui transfère le dossier médical demandé. Cependant, afin d'apporter une meilleure sécurité, nous considérons le fait que les employés n'ont pas les mêmes droits au sein d'un hôpital. Dans notre cas, les droits d'accès simplifiés pour la compréhension du chapitre sont résumés par le Tableau 4.2. Un patient a tous les droits sur ses données.

1. Le docteur fournit : son login, son mot de passe et une identité (le nom) du patient au *CAA*.
2. *CAA* contrôle l'identité du docteur, son certificat et ses droits d'accès.
3. *CAA* retrouve l'identifiant local administratif IdL_1 avec l'identité du patient. Il l'envoie à *SDI*.
4. *SDI* retrouve alors l'identifiant global *IdG* en appliquant l'AES à IdL_1 avec la clé secrète K_1 : $IdG = AES_{K_1}^{-1}(IdL_1)$
5. *SDI* calcule l'identifiant local médical IdL_2 à l'aide de *IdG* et de K_2 :$IdL_2 = AES_{K_2}(IdG)$. Il supprime *IdG* et envoie IdL_2 au *CAM*.
6. *CAM* accepte ou refuse l'accès au dossier et répond à la demande. L'identité du demandeur, la date et l'heure sont enregistrées.

FIGURE 4.6 – Accès aux données médicales sans la connaissance de l'identifiant local

Le médecin a les droits sur l'ensemble des données médicales du patient, alors que l'infirmière peut uniquement accéder aux prescriptions. Finalement, la secrétaire a des droits sur les données administratives mais pas médicales. Il est également possible avec notre architecture de préciser davantage les droits de chaque employé en prenant en compte les chefs de service, les infirmières en chef, etc... Cependant, afin de simplifier les explications dans ce chapitre, nous nous contentons de ces quatre acteurs.

Acteur	Données administratives	Diagnostic	Prescription
Patient	X	X	X
Docteur		X	X
Infirmière			X
Secrétaire	X		

TABLE 4.2: Détails des droits d'accès des acteurs du système médical

Nous pourrions utiliser l'approche cryptographique de Ghindici dans sa thèse [189] en fournissant aux trois acteurs une clé privée : $K_{patient}$, $K_{docteur}$ et $K_{infirmire}$ où $K_{docteur}$ est chiffré par $K_{patient}$ et $K_{infirmire}$ par $K_{docteur}$. Ainsi, le patient a accès à la clé du docteur et le docteur à la clé de l'infirmière. La partie diagnostic est alors chiffrée par la clé du docteur, et peut donc également être lue par le patient, alors que les prescriptions et autres données, comme l'âge, sont chiffrées avec la clé de l'infirmière afin d'être lisible par les trois acteurs. La Fig. 4.7 permet de visualiser ces chiffrements.

FIGURE 4.7 – Organisation de l'hôpital

Partage de secret de Shamir

Dans une institution médicale, il est nécessaire de pouvoir ajouter des acteurs aux systèmes. La solution précédente n'est donc pas opérationnelle dans notre cas. Une autre idée venant rapidement à l'esprit est d'utiliser une signature de groupe. Ainsi, les employés qui suivent le même patient font partie d'un même groupe centré autour de ce patient. Cela implique donc que les employés possèdent une clé par groupe de patient. Cependant, dans une situation réelle, les employés sont amenés à suivre un très grand nombre de patients. Par conséquent, si cette solution était mise en place, elle nécessiterait un très grand nombre de clés à gérer par les employés. Cette option n'est donc pas exploitable. Nous nous sommes donc tourner vers le principe de partage de secret de Shamir, [117]. Son but est de diviser le secret en plusieurs parties distribuées aux participants. L'idée est qu'il suffit de n points pour définir un polynôme de degré $n-1$. Ainsi, afin de retrouver le secret, un certain nombre $(n-1)$ de participants doivent s'unir.

Dans notre proposition, le secret est la clé de déchiffrement des données. Ainsi, afin d'obtenir la clé permettant de déchiffrer les données du patient (prescriptions, diagnostics ou données administratives), le patient, le docteur et l'infirmière doivent utiliser leur clé. Cependant, il n'est pas envisageable que les différents acteurs se réunissent pour récupérer le secret. En effet, cela nécessiterait qu'à chaque demande d'accès à un dossier patient, il faille réunir un certain nombre d'employés qui suivent le patient afin de retrouver le secret. Nous imaginons très facilement qu'une telle situation n'est pas concevable dans un établissement de santé où chaque employé a un emploi du temps différent et suit de nombreux patients. Ainsi, afin de permette à une unique personne de retrouver le secret, des clés de serveurs sont nécessaires pour remplacer les clés des autres employés.

Supposons désormais qu'une infirmière souhaite accéder aux prescriptions d'un patient, celles-ci étant uniquement accessibles par le patient, l'infirmière et le docteur. Sachant que deux points, ici sous forme de clés, suffisent à définir une droite, la clé de déchiffrement des prescriptions se cachera dans une équation de degré un. L'infirmière combine alors sa clé avec celle du serveur associée aux prescriptions de ce patient. Avec ces deux clés, l'équation passant par ces deux points est retrouvée à l'aide de l'interpolation de Lagrange. L'ordonnée à l'origine peut être calculée. L'infirmière obtient ainsi la clé de déchiffrement des prescriptions du patient à l'aide de son unique clé sans avoir à contacter un autre employé pour réunir leurs parties du secret. De la même façon, la clé du docteur se trouve sur la même équation que celles des prescriptions. En effet, le docteur a un droit d'accès et de modifications de ces données.

Cependant, le docteur a également accès aux diagnostics. L'équation des diagnostics ne peut donc pas passer par la même droite (des prescriptions) étant donné que l'infirmière n'a pas de droit sur ceux-ci. Il faut alors déterminer une équation passant par les clés du médecin et du patient mais qui soit différente de celle des prescriptions. Dans notre cas, comme l'illustre la Fig. 4.8, il peut s'agir d'une parabole. Le déchiffrement des diagnostics est donc assimilé à une équation de degré deux et nécessite le stockage pour le serveur de deux clés supplémentaires. Ainsi, un médecin souhaitant accéder aux diagnostics d'un patient réunit sa clé avec les deux clés du serveur afin d'obtenir trois points permettant de retrouver l'équation de degré deux. L'ordonnée à l'origine de l'équation est alors retrouvée et ainsi la clé de déchiffrement des diagnostics du patient.

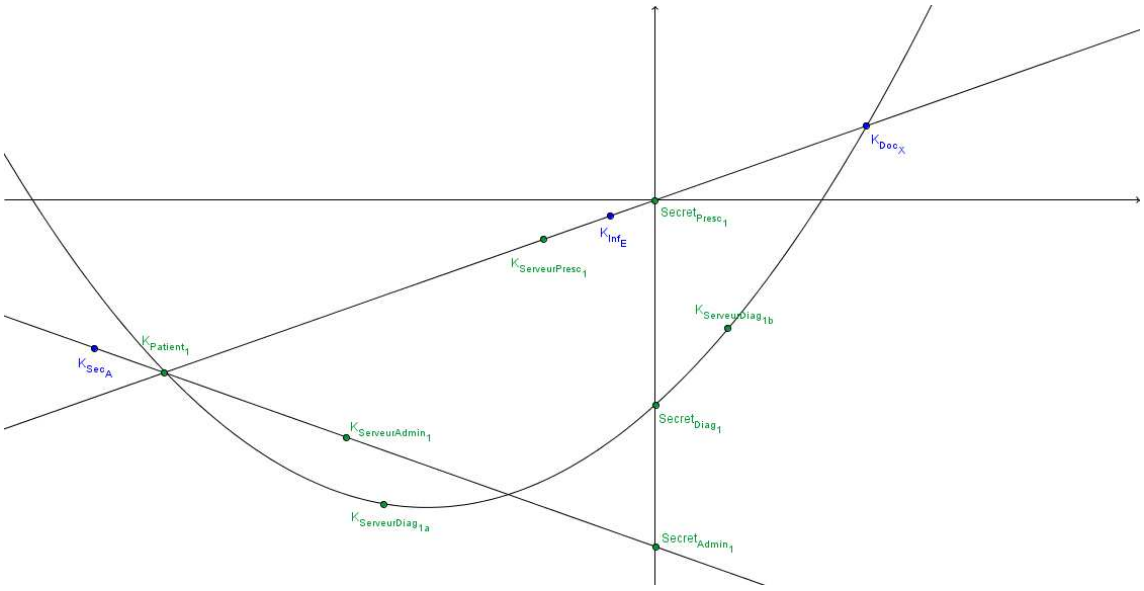


FIGURE 4.8 – Partage de secret de Shamir dans le système e-santé

La Figure 4.9 illustre la gestion des clés dans le cas d'un établissement de santé avec deux patients qui seraient suivis par les mêmes employés. Alors que la clé de déchiffrement des prescriptions du premier patient se cache dans une équation de degré un, celles du deuxième nécessite une équation de degré deux afin de passer par les clés du médecin, de l'infirmière et du second patient. Deux clés de serveurs sont alors nécessaires. On constate également qu'il est alors possible de combiner des clés de serveurs ($K_{ServeurPresc_1} = K_{ServeurDiag_{2a}}$ afin de réduire le nombre de clés de serveur à stocker. En effet, cette situation est clairement envisageable dans un établissement de santé gérant des centaines de patients.

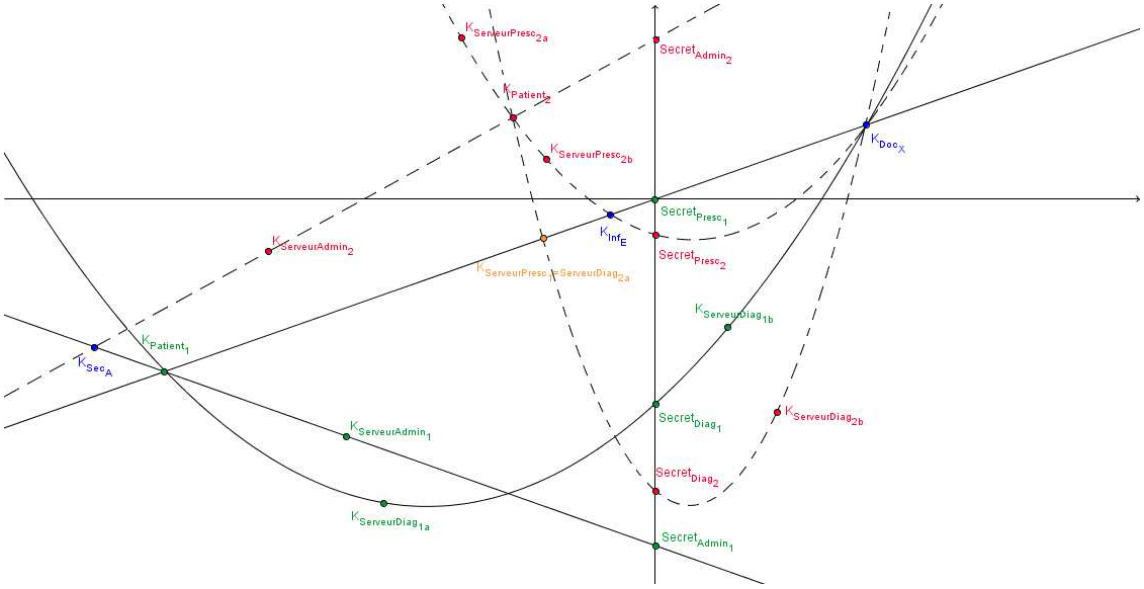


FIGURE 4.9 – Gestion des clés pour deux patients suivis par les mêmes employés

La Figure 4.10 illustre un cas très courant dans un établissement de santé : deux patients suivis par la même secrétaire mais un médecin et une infirmière différents.

Proxy de rechiffrement

Bien qu'une telle architecture soit viable, le contrôle d'accès aux informations médicales ou administratives nécessite le stockage d'un grand nombre de clés de serveur. Une nouvelle approche pour la gestion des droits d'accès peut alors être étudiée. L'utilisation d'un schéma de proxy de rechiffrement (comme expliqué dans le chapitre 2) ou d'un autre algorithme de partage de secret en ligne nous semble être une possibilité. En attendant de réaliser des études comparatives plus précises, nous avons pu constater que l'architecture avec le partage de secret de Shamir était plus rapide lorsqu'elle ne nécessite pas un trop grand nombre de secrets à partager. En effet, le *PRE* utilise de la cryptographie asymétrique

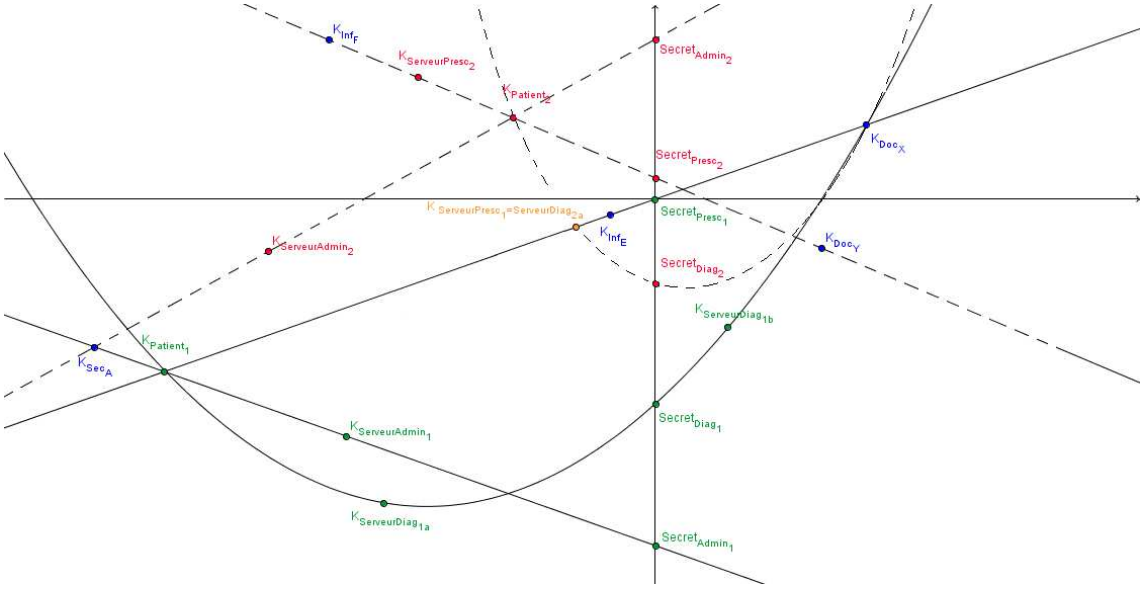


FIGURE 4.10 – Gestion des clés pour deux patients suivis par la même secrétaire

alors que l'architecture avec le partage de secret de Shamir utilise une clé symétrique. Cependant, plus le nombre d'acteurs de l'architecture (et donc de l'établissement de santé) augmente, plus le *PRE* se révèle intéressant à utiliser. En effet, alors que le nombre de calculs n'augmente pas pour le *PRE*, il augmente avec le nombre de clés à gérer pour la première architecture. . Ainsi, à l'aide des deux implémentations de cette architecture (voir section 4.7), nous pouvons constater qu'il sera nécessaire d'étudier davantage l'architecture avec le *PRE* qui semble plus durable et opérationnel.

4.4.4 Gestion d'identité entre hôpitaux

Les dossiers médicaux étant parfois croisés entre différents hôpitaux, ces-derniers doivent collaborer pour assurer un bon suivi médical au patient et une disponibilité des données. Ces transactions sont réalisées via un canal sécurisé et avec le consentement du patient, excepté dans les cas exceptionnels vus auparavant. Rappelons que pour lutter contre l'agrégation de données de différents hôpitaux, les lieu et date de naissance ou numéro de téléphone sont suffisants.

Le transfert est réalisé en plusieurs étapes et utilise l'infrastructure *PKI* décrite précédemment. Le protocole est le suivant : Un personnel médical M_1 de l'hôpital H_1 a besoin de l'ensemble des informations médicales d'un patient. Cependant, celles-ci sont à la fois stockées dans l'hôpital H_1 sous l'identifiant local IdL_{2,H_1} et dans l'hôpital H_2 sous l'identifiant IdL_{2,H_2}.

Tout d'abord, M_1 s'authentifie auprès de son contrôleur d'accès CAM_{H_1} et fournit l'identifiant local IdL_{2,H_1} du patient dont il désire le dossier. Si l'authentification et les droits d'accès sont valides, CAM_{H_1} contacte son service de dérivation d'identité IDS_{H_1}. Après avoir calculé IdG à partir de IdL_{2,H_1}, il contacte IDS_{H_2} via un canal sécurisé qui leur permet de communiquer en toute confiance. Ce dernier calcule IdL_{2,H_2}, joint son CAM qui lui fournit les informations demandées. Pour finir, les données sont transférées à M_1 via IDS_{H_2}, puis IDS_{H_1} et enfin CAM_{H_1}. L'identité de l'employé, ainsi que la requête sont enregistrées. La Fig. 4.11 explique en détail cette procédure.

1. M_1 envoie son login, son mot de passe et l'identifiant local IdL_{2,H_1} du patient à CAM_{H_1}.
2. CAM_{H_1} contrôle l'identité de M_1, son certificat et ses droits d'accès.
3. Si l'authentification est réussie, CAM_{H_1} transmet IdL_{2,H_1} à SDI_{H_1}.
4. SDI_{H_1} calcule l'IdG du patient à l'aide du IdL_{2,H_1}.
5. SDI_{H_1} s'authentifie auprès de SDI_{H_2} avec son certificat.
6. Après authentification, SDI_{H_1} et SDI_{H_2} ouvrent un canal sécurisé.
7. SDI_{H_1} envoie une demande à SDI_{H_2}, avec le nom de M_1, ses droits d'accès et l'IdG.
8. SDI_{H_2} calcule l'identifiant local correspondant IdL_{2,H_2} avec IdG et sa propre clé secrète, supprime IdG. Il envoie à CAM_{H_2} l'identifiant IdL_{2,H_2}
9. CAM_{H_2} vérifie les droits d'accès et retrouve les informations médicales qu'il envoie à SDI_{H_2}.
10. SDI_{H_2} fournit les informations à SDI_{H_1}
11. SDI_{H_1} fournit les informations à CAM_{H_1}
12. CAM_{H_1} transfère les données à M_1. L'historique est mis à jour.

FIGURE 4.11 – Communication entre hôpitaux

Transférer des données médicales entre deux hôpitaux implique que l'hôpital H_1 connaisse l'endroit où les informations sont stockées et dans notre cas la localisation de l'hôpital H_2. Le patient a donc besoin de déclarer à son médecin ses autres établissements de soins, ainsi que son consentement pour le transfert de ces données. L'accord du patient est ainsi obtenu par la connaissance de deux valeurs : IdL_{2,H_1} et H_2. Un employé ne possédant pas l'une des deux données ne peut donc pas facilement accéder à ces données. En cas d'urgence, le médecin ou service agréé peut ne pas avoir cette information. Une solution possible pourrait être le stockage, sur la carte à puce du patient par exemple, de la liste des établissements où les données médicales du patient sont enregistrées ou la gestion de ces différents établissements pas le médecin traitant du patient.

4.5 Sécurité et protection de la vie privée du modèle proposé

4.5.1 Sécurité des données

La confidentialité E_2, l'intégrité E_3 et la disponibilité E_4 des données à caractère personnel (administratives ou médicales) sont assurées en partie par les contrôleurs d'accès *CAM* et *CAA*. L'autorisation d'accès est contrôlée par une vérification des droits d'accès (E_7). La confidentialité des données exige que seul le personnel autorisé puisse lire le dossier médical. De plus, afin d'assurer un meilleur examen des accès, le nom du demandeur, la date et l'heure de l'opération sont enregistrés. L'exigence E_{16} est donc respectée. De la même façon, l'intégrité des données nécessite le même type de contrôle. Il faut alors ajouter les droits d'accès correspondant aux droits de création, modification ou destruction des données à caractère personnel. L'authentification du personnel médical est réalisée avec un système classique de login, mot de passe et grâce à un certificat spécifique à chaque employé qui contient leurs droits. De même, le patient est authentifié à chaque demande d'accès à son dossier. Par conséquent, seules les personnes autorisées accèdent ou mettent à jour ces données à l'intérieur de l'institution. Les exigences E_8 et E_6 sont donc assurées.

Durant le transfert de données médicales entre institutions, c'est l'infrastructure *PKI*, ainsi que le protocole d'authentification, qui assurent les différents principes de sécurité. En effet, l'hôpital H_1 étant certifié par une autorité de confiance, H_2 peut entrer en contact avec H_1 sans risque. H_1, quant à lui, vérifie l'authentification et le droit d'accès de ses propres employés afin de garantir l'intégrité des données. De plus, le canal sécurisé utilisé assure la confidentialité des données échangées entre les deux hôpitaux au cours de la communication et ainsi les exigences E_1 et E_{15}.

L'architecture proposée est simple d'utilisation pour le patient pour lequel il est facile d'accéder et de gérer ses informations. L'exigence E_{17} ainsi assurée permet d'améliorer la gestion par le patient de ses données et ainsi leur niveau de protection et de sécurité.

4.5.2 Minimisation des données

Étant donné que l'identifiant local est la seule valeur donnée relative à l'identité du patient, toutes les données médicales sont anonymes (E_{11}). La minimisation des informations est également réalisée via les droits d'accès. Ainsi, seuls les renseignements pertinents sont fournis au personnel médical. L'identifiant local est calculé par la méthode cryptographique *AES* et grâce à une clé secrète ce qui assure la divulgation des informations

uniquement aux personnes connaissant cet identifiant. De plus, les professionnels de santé n'ont jamais accès à l'identifiant global du patient étant donné qu'ils ne possèdent pas les clés secrètes. La seule entité les possédant est le gestionnaire d'identité SDI qui est l'élément de confiance du régime défini et qui assure la sécurité des clés en les stockant dans un élément de sécurité de type HSM.

L'exigence E_{10} de non-associabilité qui en découle est vérifiée dans le schéma proposé à l'intérieur de l'hôpital et entre les différents hôpitaux. Effectivement, si le contenu des dossiers médicaux et des dossiers administratifs sont volés, il n'est pas possible d'agréger ces informations étant donné que les identifiants locaux sont différent et non utilisables sans la connaissance des clés secrètes. La séparation des données médicales et administratives est donc nécessaire pour éviter une telle corrélation. De même, la possession de deux dossiers médicaux d'hôpitaux différents ne permet pas de lier des informations entre elles étant donné que les clés secrètes utilisées diffèrent pour chaque établissement.

De plus, la séparation (E_9) entre le calcul des identifiants locaux par le service de dérivation de clé et le stockage des données gérés par les deux contrôleurs d'accès, permet d'éviter le lien entre les données administratives et données médicales. Seul le service de dérivation d'identité SDI est capable de calculer des identifiants locaux et connaît ces deux identifiants. Toutefois, le SDI n'a pas accès aux bases de données (administratives et médicales). Par ailleurs, le contrôleur d'accès médical accède aux dossiers médicaux avec sa connaissance de l'identifiant médical local, mais, ne connaissant pas l'identifiant administratif, il ne peut lire les dossiers associés. La même contrainte est observée pour le contrôleur d'accès administratif.

Finalement, l'usage du partage de secret de Shamir permet de gèrer les droits d'accès aux informations du patient et assure que seule les données nécessaires soient fournies aux employés (E_8). En effet, en ce qui concerne le partage de secret de Shamir, chaque employé possède un point secret qui, réunit avec les $(n-1)$ points du serveur, permet de trouver le polynôme de degré $(n-1)$. De plus, si le serveur est endommagé ou corrompu, les $(n-1)$ points du serveur ne permettent pas de retrouver ce polynôme. Par conséquent, dans les deux cas, le principe de minimisation des données est assuré. Cette gestion du contrôle d'accès permet également de modifier la politique des droits d'accès à tout moment dans le cas où les droits d'un employé seraient changés. Par exemple, si une infirmière n'a plus accès aux données d'un patient. Comme nous l'avons vu précédemment, l'usage du proxy de rechiffrement permet également de gèrer les droits d'accès aux informations du patient. De plus, il ne nécessite pas un stockage trop important de clés ou de données.

4.5.3 Souveraineté et sensibilité des données

Dans l'architecture définie, le consentement du patient (E_{12}) pour l'accès à son dossier médical est réalisé par la connaissance de son identifiant local IdL_2. Ce dernier est obtenu à partir du patient par un médecin agréé, comme dans un scénario classique de e-santé. De plus, ses données sont stockées dans plusieurs bases de données afin d'éviter la faiblesse d'un système centralisé et de gérer les différents niveaux de sensibilité de celles-ci (E_{13}).

Dans certains cas exceptionnels où l'hôpital doit avoir accès aux données médicales sans le consentement du patient, par exemple en cas d'urgence, le calcul de l'identifiant global du patient pourra être réalisé par le service de dérivation d'identité et ceci sans le consentement du patient. L'enregistrement de l'identité du demandeur est alors considéré comme une protection supplémentaire de souveraineté des données. Le patient est informé dès que possible de la demande ainsi traitée. Les exigences E_{14} et E_{16} sont ainsi assurées.

Le Tableau 4.3 compare l'architecture du DMP et celle de Mina Deng avec notre proposition en fonction des exigences décrites précédemment. Comme le score le confirme, alors que le DMP et le protocole de Mina Deng ne prennent pas en compte toutes les exigences, le système proposé assure leur totalité.

E_i	Propriétés	DMP	Architecture de Mina Deng	Nouvelle Architecture
E_1	Confidentialité des transactions	-	**Oui**	**Oui**
E_2	Confidentialité des données	**Oui**	**Oui**	**Oui**
E_3	Intégrité	**Oui**	-	**Oui**
E_4	Disponibilité	Partiel	**Oui**	**Oui**
E_5	Authentification de Patient	**Oui**	**Oui**	**Oui**
E_6	Authentification des employés	**Oui**	**Oui**	**Oui**
E_7	Contrôle d'accès des employés	**Oui**	**Oui**	**Oui**
E_8	Politiques de sécurité différentes	Partiel	Non	**Oui**
E_9	Gestion séparée des données	**Oui**	Non	**Oui**
E_{10}	Non-associabilité	Non	**Oui**	**Oui**
E_{11}	Anonymat des bases de données	Non	**Oui**	**Oui**
E_{12}	Souveraineté des données	Partiel	Non	**Oui**
E_{13}	Sensibilité des données	Non	**Oui**	**Oui**
E_{14}	Cas d'urgence	**Oui**	Non	**Oui**
E_{15}	Transfert entre établissements	**Oui**	**Oui**	**Oui**
E_{16}	Historique sauvegardé	**Oui**	Non	**Oui**
E_{17}	Usage facile pour le patient	Non	Non	**Oui**
	Score /17	10.5	10	**17**

TABLE 4.3: Synthèse des analyses des différents protocoles étudiés

4.6 Éléments de validation métier

Afin de nous rendre compte des réels besoins attendus par les hôpitaux, nous avons pris contact avec un ancien Professeur du CHU de Caen, le Professeur Blanchère [2], ainsi qu'avec la référente en dermatologie de la région Haute-Normandie, le Docteur Dompmartin Anne. Suite à nos échanges, nous avons pu constater que notre architecture était réellement intéressante pour ce domaine et obtenir des précisions sur l'organisation au sein des hôpitaux et sur les attentes des employés. Concernant la hiérarchie entre soignants, quatre niveaux nous ont été indiqués. La personne disposant d'un plus grand nombre de droits est le médecin généraliste suivi des médecins spécialistes, de l'infirmière experte et enfin de l'infirmière libérale, le médecin traitant ayant tous les droits sur les données médicales du patient. Il nous a également été conseillé de prendre en compte le cas des maladies longue durée, souvent indiquées sous la mention *prise en charge à* 100% et dévoilant ainsi une information sensible concernant ce patient : le patient souffre d'une maladie dite "régulière". Enfin, lors de l'entrée d'un patient à l'hôpital, le médecin recevant ce patient doit pouvoir modifier les droits d'accès en fonction de son service et des personnes qui le suivront.

Dans la suite de nos travaux, nous prendrons en compte toutes ces contraintes afin de rendre notre architecture utilisable rapidement dans les hôpitaux.

4.7 Preuve de concept

Dans le cadre de stages étudiants à l'ENSICAEN, une preuve de concept a été déployée. L'implémentation en Java est une architecture classique de type Client-Serveur. Le Client se compose en différentes classes définissant les différents acteurs et les interfaces graphiques. Le serveur gère quant à lui les différents services et contient les fichiers utiles au fonctionnement de la plate-forme de santé et les données transmises sont chiffrées via une socket SSL. Les Figures 4.13, 4.15, 4.16, 4.21, 4.19, 4.20, 4.14 et 4.17 donnent une idée de l'architecture implémentée.

2. Directeur du Réseau de Télémédecine appliqué aux Plaies (TELAP), Membre du Comité stratégique restreint de Télémédecine et Systèmes d'Information de Santé de l'*ARS* (Agences Régionales de Santé) de Basse Normandie, Directeur du département de e-santé et domotique au pôle de compétivité *TES* (Transactions Électroniques Sécurisées).

4.7.1 Connexion en tant que médecin traitant

Supposons que je sois le médecin traitant du patient, je me connecte en tant que membre du personnel du service Pédiatrie. J'entre alors mes identifiants et peut choisir si je souhaite voir les données médicales ou administratives du client. Notons que le médecin traitant est le seule à pouvoir accéder à ces deux types d'informations.

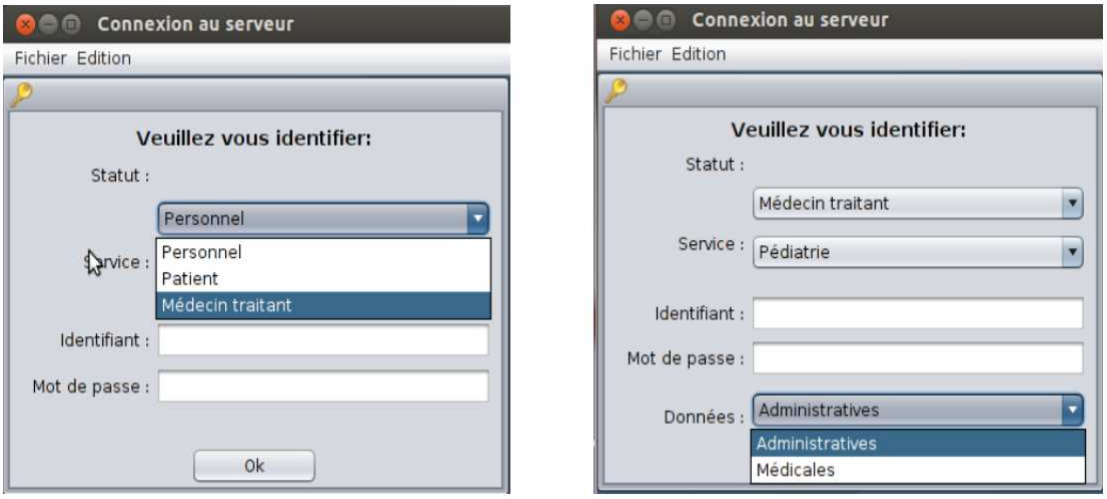


FIGURE 4.12 – Fenêtre principale du logiciel.

4.7.2 Connexion en tant que médecin

Supposons que je sois un médecin quelconque de l'établissement de santé, je me connecte en tant que membre du personnel du service Pédiatrie. J'entre alors mes identifiants.

FIGURE 4.13 – Fenêtre principale du logiciel.

J'accède à mon espace personnel dans lequel je peux retrouver l'ensemble des informations de mes patients à l'aide de leur identifiant local IdL_2.

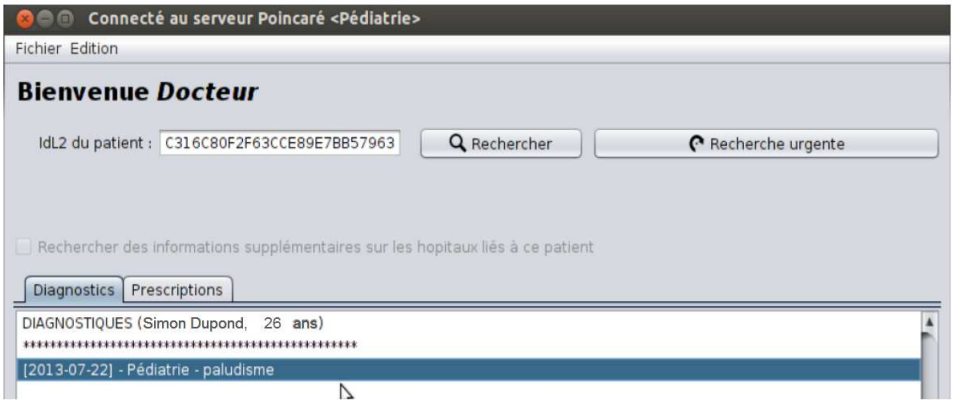


FIGURE 4.14 – Interface du logiciel pour le docteur.

4.7.3 Connexion en tant que patient

Supposons que je sois le patient, j'accède de la même façon que le médecin à mes informations personnelles.

FIGURE 4.15 – Interface du logiciel pour le patient.

Avec la touche "Donner les droits au personnel", je peux modifier, dans le cas d'un conflit, les droits d'accès des membres du personnel qui me suivent.

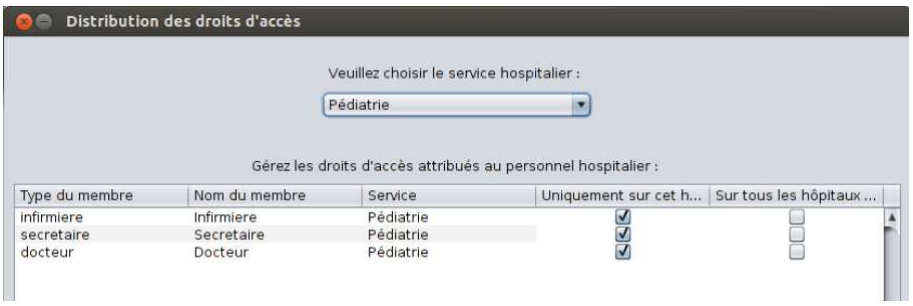


FIGURE 4.16 – Interface du logiciel de modification des droits d'accès.

4.7.4 Connexion en tant que secrétaire

Supposons que je sois une secrétaire, j'ai uniquement accès aux données administratives du patient. Je peux ainsi, grâce à la connaissance de son identifiant local IdL_1, consulter ou mettre à jour ses informations.

FIGURE 4.17 – Interface du logiciel pour une secrétaire.

4.7.5 Connexion en tant qu'administrateur

Enfin, supposons que je sois l'administrateur de l'hôpital, je peux ajouter un membre du personnel et indiquer si nécessaire qu'il s'agit du chef d'équipe.

FIGURE 4.18 – Création d'un membre du personnel.

Je peux également ajouter ou supprimer des services, des employés...

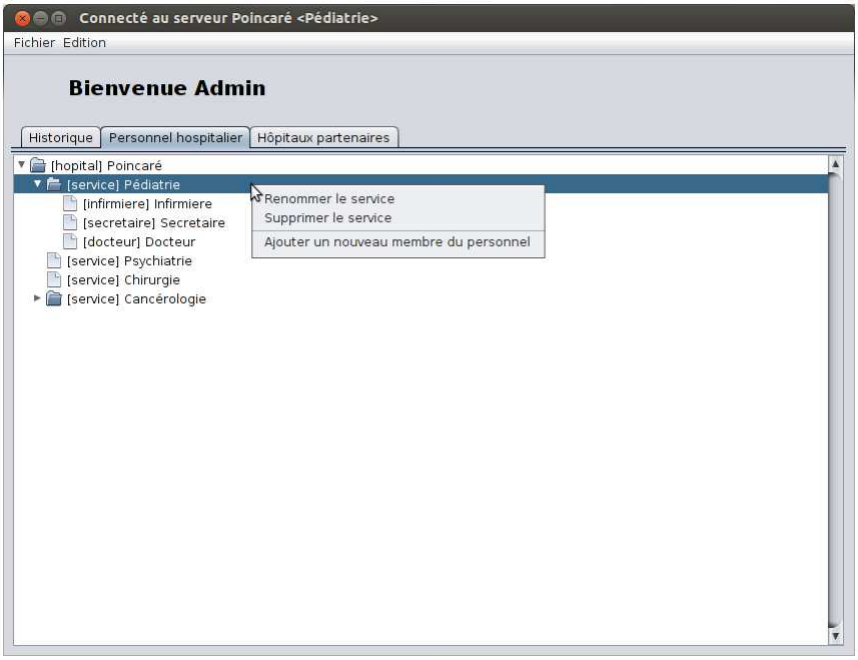


FIGURE 4.19 – Interface du logiciel pour l'administrateur - onglet "personnel hospitalier".

... ou des hôpitaux partenaires.

FIGURE 4.20 – Interface du logiciel après connexion en tant qu'administrateur - onglet "Hôpitaux partenaires".

De plus, je peux à tout moment accéder à l'historique de l'hôpital dans le cas d'un doute sur le personnel ayant accédé à des informations précises.

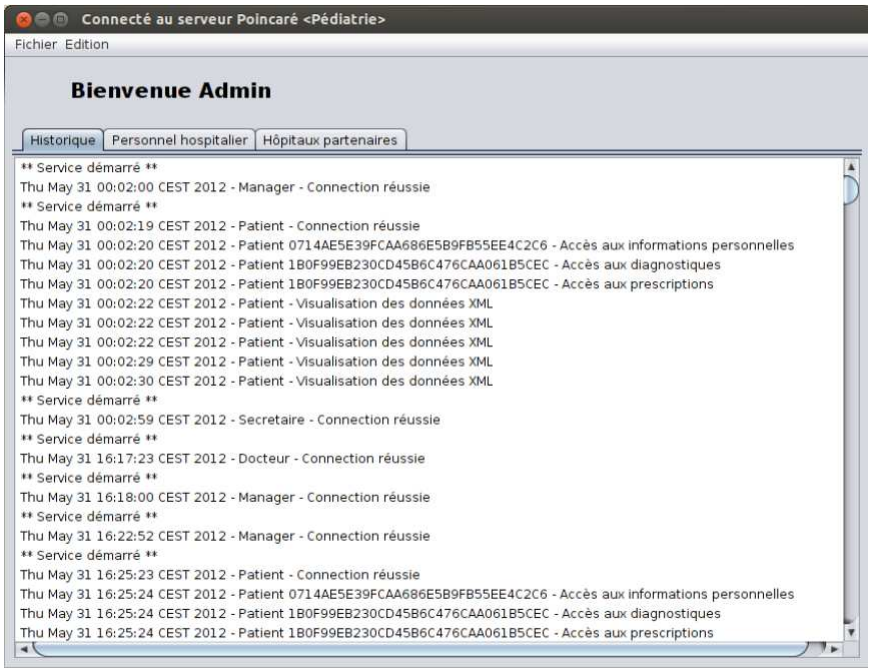


FIGURE 4.21 – Interface du logiciel pour l'administrateur - onglet "Historique"

4.8 Conclusion

La grande quantité d'informations sensibles présentes dans un système de e-santé et le transfert de ces données entre les institutions sont des défis complexes pour les technologies protégeant la vie privée. Les extensions possibles des systèmes de e-santé, comme une utilisation secondaire des données médicales pour la recherche, ont besoin d'attentions supplémentaires.

La minimisation des données et leur souveraineté sont également des principes nécessaires dans un système d'informations décentralisé de e-santé. Cependant, très peu de protocoles présents dans la littérature prennent en compte les exigences de sécurité et de vie privée en totalité. La solution que nous proposons dans ce chapitre a pour objectif d'assurer la totalité des propriétés définies ultérieurement, ainsi que les contraintes liées au système médical, tel que le transfert de données médicales entre le personnel autorisé et divers établissements de santé.

Rappel des contributions sur cette partie

- **A. Plateaux**, P. Lacharme, C. Rosenberger, K. Murty, *A Contactless E-health Information System with Privacy*, 9th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), Cagliari, Sardaigne, Juillet 2013.
- **A. Plateaux**, P. Lacharme, K. Murty, C. Rosenberger, *Minimisation des données de e-santé*, 4ème Atelier sur la Protection de la Vie Privée (APVP), Les Loges en Josas, Juin 2013.
- **A. Plateaux**, P. Lacharme, *Organisation d'une architecture de santé respectueuse de la vie privée*, 7ème Conférence sur la Sécurité des Architectures Réseaux et des Systèmes d'Information (SARSSI), Cabourg, France, Mai 2012.

Chapitre 5

Architecture de paiement en ligne

Internet est de plus en plus utilisé pour faire des achats en ligne au travers de différents systèmes de paiement. Les utilisateurs se sentent concernés par les problèmes de sécurité et d'atteinte à leur vie privée qui en résultent. Nous nous sommes ainsi concentrés pour ce dernier chapitre aux transactions électroniques sécurisées dans le cadre d'un paiement en ligne. Nous proposons une liste d'enjeux nécessaires à la sécurité et à la protection de la vie privée des utilisateurs et des marchands lors d'un paiement en ligne. À l'aide de ces exigences, nous analysons le niveau de protection de la vie privée du protocole actuel 3D-Secure et du protocole d'Ashrafi et Ng. Nous appuyons ces observations par les résultats d'une étude réalisée sur 354 internautes. Enfin, nous proposons plusieurs nouveaux protocoles respectant davantage les données personnelles des individus à partir des protocoles existants. Nous analysons la protection de la vie privée des protocoles proposés en les comparant à ceux présents dans l'état de l'art. Une implémentation permet finalement de démontrer la faisabilité des concepts définis.

Sommaire

5.1 Introduction

5.1.1 Quelques chiffres...

Le commerce électronique a considérablement augmenté avec la démocratisation d'Internet. On dénombre 100400 sites commerciaux français en 2011 [194] et 31 millions d'acheteurs en ligne à la fin du premier trimestre 2012, soit 77% des internautes. Le montant de la fraude dans les paiements électroniques augmente avec la même régularité et devient aujourd'hui une préoccupation majeure pour les institutions financières et les utilisateurs [195]. En effet, alors que les paiements à distance ne représentent que 8.4% de la valeur des transactions nationales, ils représentent 61 % du montant de la fraude [196].

Plusieurs directives liées à la sécurité des paiements en ligne ont vu le jour, comme par exemple la directive européenne $2000/31/CE$ sur la sécurité du e-commerce [197]. De la même manière, la directive sur les services de paiement [198] offre un vaste marché européen ainsi qu'une plate-forme juridique pour la SEPA (Single Euro Payment Area, [199]). Néanmoins, la sécurité et l'authentification dans le commerce électronique ne doivent pas être renforcées au détriment de la vie privée des utilisateurs [200]. De nombreuses informations personnelles sont impliquées lors d'un paiement sur Internet et doivent être protégées. Ainsi, nombreux acteurs et organismes de différents domaines acceptent les principes d'architecture centrée sur l'utilisateur et l'approche *privacy by design*. Il est de plus très important de prendre en compte les grands principes concernant la protection de la vie privée pour de tels systèmes définis dans le chapitre 1.

Malheureusement, l'industrie bancaire se concentre principalement sur l'usurpation d'identité et l'authentification des utilisateurs et délaisse la confidentialité des données personnelles des utilisateurs, ainsi que les principes nécessaires à la protection de la vie privée. Qui plus est, seulement 23% des transactions sont considérées comme sécurisées [196].

En mars 2013, dans le cadre d'un projet étudiant que j'ai encadré avec Sylvain Vernois, une enquête sur la protection de la vie privée lors d'un paiement sur Internet a été réalisée sur un panel de 354 individus. Cette population inclut des étudiants de la faculté de Caen, les étudiants de l'ENSICAEN (chimie, électronique, informatique, image), des salariés du laboratoire GREYC, des employés de divers organismes publics et privés, ainsi que des personnes extérieures. Le questionnaire se trouve dans l'AnnexeA. Nous sommes évidemment conscients des biais induits par un tel panel non représentatif de la population, cependant, les résultats obtenus n'en restent pas moins très intéressants.

On dénombre 42% des participants entre 16 et 25 ans, 36% entre 26 et 40 ans et 22% plus de 40 ans. 34% sont des femmes et donc 66% des hommes. Sur l'ensemble du panel, seulement 2% déclare ne jamais avoir fait d'achat sur Internet, 18% en font rarement, alors que 40% et 39% en font respectivement parfois et régulièrement. Ainsi, à la question *"Vous sentez vous concernés par les problèmes de respect de la vie privée sur Internet ?"*, 87% d'entre eux répondent positivement et 69% déclarent avoir des appréhensions lors d'un tel achat. En tête des craintes (avec environ 23% chacune) se trouve : la peur de se faire voler des informations personnelles et la peur qu'elles soient stockées. La Figure 5.1 donne des précisions sur ces appréhensions.

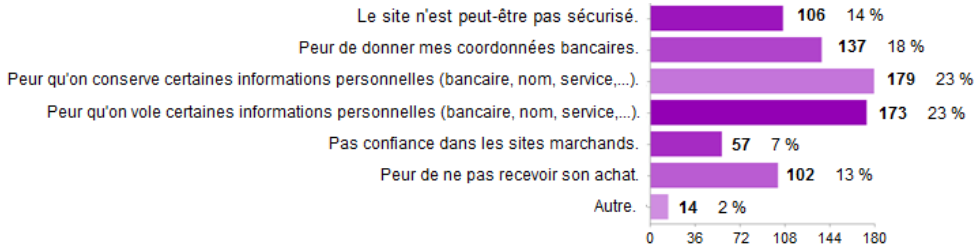


FIGURE 5.1 – Appréhensions du panel lors d'un paiement en ligne

5.1.2 État de l'art sur la protection des données personnelles bancaires

Quatre acteurs sont toujours présents lors d'un paiement électronique. Le **client** navigue sur le site Internet du **fournisseur de service** SP pour acheter un service en ligne. Ces deux acteurs ont chacun un fournisseur de paiement, la **banque du compte à débiter** et **la banque du compte à créditer**, respectivement appelés dans ce chapitre *banque du client* et *banque du SP*. Dans la plupart des architectures de paiement en ligne, un cinquième acteur entre en jeu, souvent il s'agit d'un tiers de confiance ou d'un système d'interopérabilité comme dans 3D-Secure. Le rôle de celui-ci est différent en fonction du protocole, bien qu'il permette généralement d'authentifier ou d'identifier les fournisseurs de paiement.

Dans l'objectif d'une acquisition, l'utilisateur doit s'enregistrer auprès du SP puis procéder au paiement. Ainsi, durant cette transaction, de nombreuses informations personnelles transitent, comme celles demandées durant l'enregistrement auprès du SP. Le SP peut également demander des informations optionnelles comme le nombre d'enfants ou la profession du client. La plupart de ces informations sont rarement utilisées pour la transaction de paiement, mais sont réclamées à des fins purement statistiques ou marketing.

Finalement, en plus de ces données d'enregistrement, dont le problème a été traité dans le chapitre 3, les informations bancaires du client, comme le *PAN* (Primary Account Number), la date de validité et le cryptogramme visuel ($CVX2$), sont également requises lors de la phase de paiement.

En France, les deux principaux services de paiement en ligne sont Sips [201] (Secure Internet Payment Services) d'Atos Worldline et PAYBOX [202]. Le premier permet à un de ces 18000 commerçants en Europe d'accepter des paiements en ligne et de les gérer sans pour autant conserver les données sensibles de ses clients. La seconde, similaire à Sips, se veut multi-canal et est le premier opérateur monétique multi-banques.

Cependant, le premier protocole proposé pour sécuriser les transactions électroniques est le protocole SET (Secure Electronic Transactions [203]). Plus tard, les protocoles standards de paiement en ligne sont renforcés par l'ajout d'un secret envoyé par mobile, comme avec le protocole 3D-Secure [150] ou, par l'utilisation d'un appareil supplémentaire, par exemple, un lecteur *CAP* (Chip Authentication Protocol [139]). Cependant, les résultats en termes de sécurité de ces réponses sont mitigés comme le soulignent [204, 205]. En outre, alors que le protocole SET a été largement étudié [206, 207, 208, 209], le protocole 3D-Secure est étonnamment négligé, à l'exception des analyses de Murdoch et Anderson [204] et de Pasupathinathan et coll. [210].

Néanmoins, bien que la norme PCI/DSS soit une première étape pour la protection de la vie privée, son champ d'application est essentiellement limité aux données bancaires [174] et la protection de la vie privée de l'utilisateur a complètement disparu des protocoles de paiement en ligne lors du passage de SET à 3D-Secure [204]. Des protocoles alternatifs sont proposés dans la littérature, mais généralement axés sur la sécurité des SP et non sur la protection de la vie privée de l'utilisateur, à l'exception des propositions partielles d'Ashrafi et Ng dans [211] et d'Antoniou et Batten dans [212].

La section 5.2 définit les exigences requises à la protection de la vie privée et à la sécurité de tels paiements. Les travaux existants, dont le protocole 3D-Secure et le protocole d'Ashrafi et Ng sont détaillés dans la section 5.3. Ces deux derniers protocoles y sont également comparés au protocole SET. Trois nouveaux protocoles, dont une amélioration de 3D-Secure et du protocole d'Ashrafi et Ng, sont proposés dans la section 5.4. Enfin, un aperçu du démonstrateur de l'architecture, basée sur la dernière amélioration, est détaillé dans la section 5.5. Pour finir, les perspectives et la conclusion sont données dans les sections 5.7 et 5.8.

5.2 Exigences de sécurité et de protection de la vie privée

Nous établissons dans cette section un ensemble d'exigences importantes à la sécurité et à la protection de la vie privée lors d'un achat sur Internet. Ces exigences complètent le travail d'Antoniou et Batten [212] qui ont regroupé et divisé en deux catégories une liste des risques présents dans la littérature. La première catégorie est composée de six risques liés à l'information révélée par un client au SP. La seconde représente sept risques liés au modèle traditionnel de commerce électronique et sont indépendants de l'information divulguée :

- R_{11} : Information de paiement interceptée entre le client et le SP.
- R_{12} : Informations de paiement volées au SP, avant ou après l'utilisation de ces informations par le SP.
- R_{13} : Détournement des informations de paiement par le SP :
  - avec double frais de la carte de crédit du client..
  - avec la vente de ces informations de paiement à une autre entité.
  - avec retrait sur la carte de crédit du client d'un montant supérieur à celui annoncé.
- R_{14} : Détournement des informations personnelles par le SP :
  - avec vente/divulgation des informations personnelles à une autre entité.
  - avec divulgation sans le consentement du client.
  - avec envoi de spam au client.
- R_{15} : Vol d'informations personnelles depuis le SP.
- R_{16} : Mise à jour des données personnelles ou de paiement sans accord du client.
- R_{21} : Identité réelle du SP cachée au client par le SP.
- R_{22} : Politique de sécurité publiée par le SP non respectée par le SP.
- R_{23} : Politique de retour du SP non respectée par le SP.
- R_{24} : Lieu du SP et législation du pays inconnu par le client.
- R_{25} : Produit attendu par le client non envoyé par le SP.
- R_{26} : Aucun produit du client ne peut être envoyé par le SP.
- R_{27} : Actes malveillants du SP et impossibilités pour le client de trouver et d'accuser le SP.

Lors d'un paiement en ligne, il est nécessaire de prendre en compte ces risques, ainsi que les principes fondamentaux en termes de protection de la vie privée et les propriétés de sécurité traitées dans le chapitre 1.

En fonction de l'information, les niveaux de protection et les droits d'accès aux données personnelles sont différents. Nous choisissons alors de diviser les données personnelles en trois parties :

1. Les informations d'identité Id incluent les informations permettant de connaître l'identité du client, par exemple, son nom.
2. Les informations de commande IC incluent le panier détaillé et les informations liées au service attendu, comme le nom du SP. Ces données sont connues du SP.
3. Les informations bancaires IB, comme par exemple le nom de la banque du client, la valeur du PAN ou du $CVX2$. Ces données sont connues par la banque du client.

Ainsi, bien que les informations bancaires, notamment le PAN, soient depuis 2003 considérées par la CNIL comme un réel moyen d'identification du client, nous choisissons de différencier les données de type état civil, que nous appelons informations d'identité, des données bancaires.

Une liste de dix-sept exigences, E_i, de sécurité et vie privée est établie. Elles prennent en compte l'ensemble des principes de protection de la vie privée, ainsi que les risques regroupés dans la littérature. Ces exigences devraient être respectées par n'importe quel système de paiement en ligne afin de protéger au maximum les données personnelles du client et du SP :

• E_1 : La **confidentialité des transactions** exige que chaque information échangée soit chiffrée afin de protéger ces données contre les entités exterieures au système.

• E_2 : L'**intégrité** de l'information transmise permet d'assurer l'exactitude du contenu et donc la non-altération des données lors de leur transmission.

• E_3 : La **confidentialité de l'identité du client vis-à-vis du SP** veille à ce que le client puisse accéder à un service sans divulguer son identité au SP. Cette exigence est levée si le client souhaite une livraison du service à domicile.

• E_4 : La **confidentialité de l'identité du client vis-à-vis de la banque du SP** veille à ce que le client puisse accéder à un service sans divulguer son identité à cette banque.

• E_5 : L'**authentification du client** par une partie de confiance garantit l'identité du client. La partie de confiance peut être par exemple la banque du client ou une autre partie de confiance *où le client est déjà enregistré*.

• E_6 : L'**authentification du SP** par le client ou par une partie de confiance garantit l'identité du SP auprès du client.

• E_7 : L'**authentification des banques** par une partie de confiance garantit l'identité de la banque du SP et de la banque du client.

• E_8 : La **non-associabilité** des différentes transactions réalisées permet de ne pas être capable de lier les commandes venant d'un même client.

• E_9 : La **non-réutilisabilité** des informations transmises (bancaires ou autres) permet d'avoir des transactions uniques et non rejouables.
• E_{10} : La **confidentialité des informations de commande** IC garantit que seules les personnes autorisées ont accès aux informations de commande. En particulier, le panier du client est inconnu de la banque du client et de la banque du SP.
• E_{11} : La **confidentialité des informations bancaires** IB (ou principe de minimisation des données du client) veille à ce que seules les personnes autorisées aient accès aux données bancaires. Cette exigence inclut le fait que le SP ne connait pas les coordonnées bancaires du client.
• E_{12} : L'**anonymat du client** est garanti si les exigences E_3, E_4, E_9 et E_{10} sont assurées. En effet, les données IC et IB permettent partiellement d'identifier le client.
• E_{13} : Le **principe de minimisation des données du SP** inclut le fait que la banque du SP ne connaît pas le client (E_4) et que le client ne connaît pas la banque du SP. Cette dernière condition est très importante lorsque le SP est une petite organisation et sa banque est le même fournisseur de paiement que la banque personnelle du gestionnaire.
• E_{14} : Le **principe de souveraineté des données** implique que les données personnelles du client lui appartiennent avec son contrôle et son consentement sur leur utilisation.
• E_{15} : Le **principe de sensibilité des données** implique que les données personnelles nécessitent une structure décentralisée pour leur stockage.
• E_{16} : La **possession d'un certificat par le client** ne devrait **pas** être **obligatoire** afin de faciliter le paiement électronique au client.
• E_{17} : La **non-répudiation** de la commande par le client afin d'ajouter une sécurité supplémentaire au fournisseur de service.

Dans la suite du chapitre, nous faisons état des travaux existants en termes de systèmes de paiement en ligne et étudions plus en détails le protocole 3D-Secure actuellement déployé et le protocole d'Ashrafi et Ng [211] avancé comme respectant la vie privée des utilisateurs. Ces deux derniers protocoles sont ensuite analysés et comparés au protocole SET à l'aide des exigences ainsi définies.

5.3 État de l'art de la protection des données personnelles et bancaires

Une transaction commence par une authentification du client et du SP et par l'établissement d'une connexion sécurisée à l'aide d'un protocole tel que SSL/TLS (Secure Sockets Layers/Transport Layer Security, [213, 214]). Ce protocole implique que le client ait confiance dans le SP par lequel ses informations bancaires transitent. De plus, les

protocoles de type SSL/TLS permettent de sécuriser la transaction entre les clients et le SP mais ne fournissent pas de véritable authentification du client. Le client peut aussi faire appel à un tiers de confiance, comme Paypal [215]. Ce dernier est un des protocoles d'intermédiation de paiement les plus connus sur Internet et comptabilise plus de 5 millions de comptes en France. Cependant, l'ouverture d'un compte Paypal implique l'envoi de nombreuses informations personnelles (nom, prénom, adresse mail, *PAN*, date d'expiration de la carte, $CVX2$, etc...). L'utilisateur peut ensuite réaliser des paiements par Internet sans avoir à (re-)communiquer ses informations bancaires à chaque transaction. Toutefois, bien que Paypal spécifie dans sa politique de confidentialité *ne pas vendre ou louer ces informations* [216], celles-ci peuvent être *partagées avec des parties tiers* dans le monde. Il est donc nécessaire de s'interroger sur la protection de données personnelles et sur l'utilisation d'un tel service.

Un des premiers protocoles visant à protéger la vie privée des clients est proposé par Chaum en 1990 [217], basé sur l'utilisation de signature aveugle. Plusieurs systèmes de paiement ont été proposés afin d'améliorer la sécurité, mais sans se focaliser sur la protection de la vie privée. Dans [218], un protocole de paiement sécurisé permet de gérer différents aspects telle que la carte à puce avec des fonctionnalités réseaux ou la multiplicité des entités. Cependant, ces scénarios gèrent la sécurité des transactions mais ne s'intéressent pas à la notion de vie privée de l'utilisateur. Antoniou et Batten se sont intéressés à la confiance au sein des systèmes de commerce électronique [212] et proposent quatre modèles avec quatre niveaux différents de protection des données personnelles. Cependant, leurs solutions sont centrées autour d'un libérateur qui connaît tous les acteurs du processus. Une méthode de calcul est également proposée afin de mesurer le degré de vie privée, ainsi qu'une liste de risques identifiés dans la littérature vis-à-vis de la vie privée. Le protocole d'Ashrafi et Ng [211] fait quant à lui appel à une authentification basée sur un mot de passe non-réutilisable. Ce processus garantit une authentification du client et minimise les risques pour le SP. Cependant, il a la même complexité que le protocole 3D-Secure et utilise la société de cartes de crédit, ainsi qu'une passerelle de paiement optionnelle. La sécurité du protocole est donc basée sur la société de cartes qui stocke toutes les informations de paiement du client dans une base de données locale et centralisée.

Un nouveau type de monnaie électronique a vu le jour en 2009, le Bitcoin. Il s'agit à la fois d'une devise monétaire et d'un système de paiement dans cette devise. La valeur de la devise est déterminée par son usage économique et le marché des changes, alors que le système de paiement fait appel à un ordinateur du réseau choisi de façon aléatoire. Bien qu'un tel système permette l'anonymat durant la transaction de paiement, il ne s'agit

pas d'une infrastructure centralisée. Le Bitcoin est donc totalement différent des systèmes bancaires internationaux classiques. Nous ne nous intéresserons donc pas davantage à une telle infrastructure dans cette thèse.

5.3.1 Le protocole SET

Un consortium de sociétés de cartes de crédit, incluant VISA et MasterCard, ainsi que des sociétés de logiciels, ont développé le protocole *SET* (Secure Electronic Transactions, [203]) pour sécuriser les transactions électroniques de paiement par carte de crédit. Celui-ci est partiellement décrit par la Fig5.2.

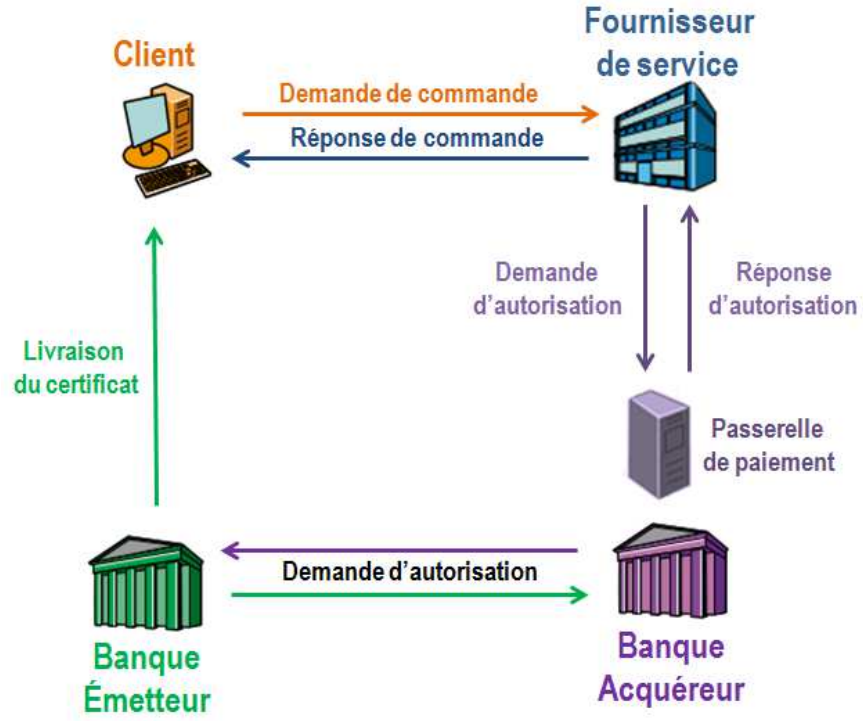


FIGURE 5.2 – Le protocol SET [219]

Le protocole *SET* se déroule en deux étapes :

1. Étape d'enregistrement :
   - Le client contacte sa banque qui supporte le protocole SET ;
   - Après authentification de son client, la banque joue le rôle d'autorité de certification. Le client reçoit alors son certificat signé par sa banque ;

2. Étape d'achat :
   - Le client contacte le fournisseur de service qui possède son propre certificat ;
   - Le SP envoie au client sa clé publique et une copie de son certificat pour vérification de sa validité ;
   - Après vérification, le client envoie au SP :
     - Son certificat pour permettre au SP de vérifier son identité ;
     - Son panier chiffré avec la clé publique du SP ;
     - Ses coordonnées bancaires chiffrées avec la clé publique de la passerelle de paiement (la banque du SP) ;
   - Le SP demande une autorisation de paiement à la passerelle de paiement en lui envoyant :
     - Les coordonnées bancaires du client chiffrées avec la clé publique de la passerelle ;
     - Les détails du paiement chiffrés avec la clé publique de la passerelle ;
   - La passerelle envoie au SP une confirmation chiffrée avec la clé publique du SP ;
   - Le SP envoie au client la réponse de la passerelle chiffrée avec la clé publique du client ;
   - Le SP fournit le service demandé par le client ;
   - Le SP envoie à la passerelle une demande de transaction chiffrée avec la clé publique de la banque ;
   - La banque (passerelle) transfère le paiement au SP.

Ainsi, le protocole SET assure un certain nombre de propriétés en intégrant la protection de la vie privée du client et du SP. La confidentialité et l'intégrité des données sont ainsi respectées et le protocole fournit une authentification mutuelle par un tiers de confiance (la banque du SP) entre le SP et le client. De plus, le SP ne connaît pas les coordonnées bancaires du client et la banque du client ne connaît pas le contenu de la commande. Finalement, le client ne connaît pas l'identité de la banque du SP. Cependant, la banque du SP est l'autorité de confiance du protocole. De plus, bien que la banque du client n'ait pas connaissance du contenu de la commande du client, elle connait l'identité du SP et la double signature garantissant l'authenticité n'est pas suffisamment précise. Qui plus est, le consentement du client lors de l'envoi de ses coordonnées bancaires ne peut pas être prouvé [208] et de nombreuses attaques de SET peuvent être citées [209]. Les auteurs dans [207] mettent également en lumière la complexité et la contradiction présentes dans les spécifications de SET. Ce protocole a donc été analysé en détail dans le début des années 2000 et des améliorations, ainsi que des simplifications, ont été proposées [207, 220, 221].

Pour finir, l'installation par le client d'un logiciel spécifique, ainsi que la distribution de lecteurs de carte et de certificats par le SP, rendent le protocole complexe pour le client et coûteux pour le SP. Ainsi, comme précisé dans [204], toutes ces contraintes ont entrainé l'abandon du protocole SET en faveur du protocole 3D-Secure et les quelques parties de SET concernant la vie privée ont tout simplement été supprimées.

5.3.2 Le protocole 3D-Secure

Développé par Visa en 2001, le protocole 3D-Secure [150] est architecture sécurisée utilisée fréquemment pour les paiements électroniques sur Internet. D'autres organismes financiers ont également développé leurs propres implémentations, comme MasterCard avec MasterCard SecureCode et American Express avec SafeKey. Une comparaison entre 3D-Secure et MasterCard SecureCode est donnée dans [210].

Le protocole est composé de neuf étapes échangées entre les cinq acteurs du système. Comme le montre la Fig. 5.3, le cinquième acteur est le serveur Visa, appelé serveur directory. Il a pour but de "rediriger" les communications entre le SP et la banque du client. De plus, afin de mettre en place ce protocole, un module dédié appelé MPI (Merchant Plug In) doit être importé sur le site du SP.

A. Le client envoie au SP son intention d'achat accompagnée de son PAN, du $CVX2$ et de la date d'expiration de sa carte.

B. Le MPI envoie une requête de verification (VEReq) au serveur directory.

C. Le serveur directory contrôle l'identité du SP, le numéro de carte et la banque du client puis retrouve le serveur de contrôle d'accès, ou ACS (Access Control Server), de la banque émettrice où est enregistrée la carte du client.

D. L'ACS contrôle si le client est enregistré dans le programme 3D-Secure de la banque émettrice et envoie l'URL d'authentification du client au MPI (VERes - Verify Enrollment result).

E. Le MPI envoie une requête PAReq à travers cette URL. Le message contient les détails de l'achat autorisé et demande à l'ACS d'authentifier son client. Le protocole d'authentification dépend de la banque du porteur.

F. Le client fournit les informations nécessaires pour s'authentifier auprès de sa banque.

G. L'ACS envoie au MPI une confirmation de l'authentification du client (message PARes), enregistré par le MPI.

H. Le SP s'authentifie à sa banque, qui vérifie la nature de la transaction auprès de la banque du client et confirme l'autorisation de paiement au SP. La banque du client stocke les informations de paiement pour assurer la non-répudiation de la transaction.

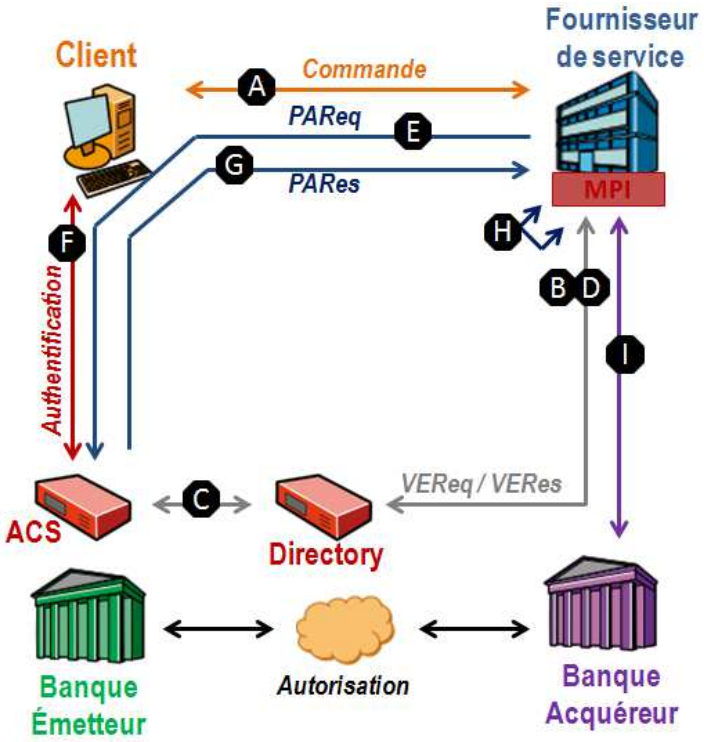


FIGURE 5.3 – Le protocole 3D-Secure

La principale faille de sécurité des implémentations 3D-Secure, soulignée dans [204], a été corrigée par de nombreuses banques. L'authentification du client à l'aide de sa date de naissance (ou d'autres secrets triviaux) est depuis remplacée par un mot de passe à usage unique, un OTP (One Time Password), envoyé au téléphone mobile de l'utilisateur. Notons qu'étant donné que la phase de paiement complète n'est pas décrite, le protocole 3D-Secure ne contient ici que neuf étapes.

Discussion sur le protocole 3D-Secure

Dans un premier temps (étape A), le client envoie ses informations bancaires au SP, qui peut l'identifier, en contradiction avec les exigences E_3 et E_{11}. Les exigences E_4 et E_{13} ne sont pas respectées car la banque du client connaît l'identité du SP et, la banque du SP l'identité du client. Ensuite, la banque du client n'étant pas la seule à authentifier son client (entre D et G), l'exigence E_5 n'est pas respectée. De la même façon, l'authentification du SP n'est ni réalisée par le client, ni par une autre partie dans laquelle le client a confiance (étape C et H), ainsi l'exigence E_6 n'est pas assurée. Il en est de même pour l'authentification des banques qui est réalisée par le serveur directory, l'exigence E_7 est donc uniquement partiellement respectée. Étant donné que les données de commande du client (E_{10}) sont contenues dans le message PAReq envoyé à l'ACS (étape E), elles ne

peuvent être confidentielles. Ainsi, les exigences E_3, E_4, E_{10} et E_{11} n'étant pas respectées, l'exigence d'anonymat E_{12} ne peut l'être également. Finalement, l'exigence E_{14} ne peut être assurée étant donné que les informations du client passent par plusieurs entités sans que le client en soit informé. De plus, la sensibilité des informations qui transitent n'est pas suffisamment prise en compte et donc ne permet pas d'assurer l'exigence E_{15}. Le protocole 3D-Secure garantit donc seulement quatre des seize exigences décrites précédemment, dont une partiellement.

De plus, et comme souvent dans les architectures de paiement en ligne, le cinquième acteur entre toujours en jeu au milieu de la transaction, comme pour le serveur directory dans 3D-Secure ou la société de carte dans le protocole de Ashrafi et Ng [211], comme nous le détaillons dans la section suivante.

5.3.3 Le modèle d'Ashrafi et Ng

Ashrafi et Ng dans [211] proposent un schéma préservant la vie privée et faisant appel à la société de cartes. Ce protocole utilise en option une passerelle de paiement, que nous ne développons pas dans ce chapitre, et se déroule, comme le montre la Fig.5.4, de la manière suivante :

A. Le client demande une facture ainsi que les clés publiques du SP et de la société émettrice de cartes de crédit.

B. Le SP envoie au client un identifiant de transaction, le certificat du SP et celui de la société émettrice. Ce message est signé par le SP.

C. Si la vérification des deux certificats est correcte, le client génère deux paquets :
- Les informations de paiement : le haché du détail de la carte, du mot de passe et du détail de la commande, l'horodatage et la période de validité. Ce paquet sera chiffré avec la clé publique de la société émettrice de cartes.
- Les informations de commande : l'identifiant de la transaction, le montant de la transaction, l'horodatage et la période de validité. Ce paquet sera chiffré par la clé publique du SP.

D. Les deux paquets chiffrés sont envoyés au SP.

E. Le SP déchiffre avec sa clé privée le deuxième paquet et le vérifie. Si la vérification est correcte, le SP génère un identifiant de paiement unique.

F. Les deux paquets sont envoyés à la société émettrice des cartes.

G. La société émettrice des cartes déchiffre les informations de paiement avec sa clé privée. Elle vérifie l'horodatage et la date d'expiration. Le haché des données de la

carte et le haché du mot de passe du client sont également vérifiés via la table de hachage stockée localement par la société.

H. La société émettrice des cartes envoie le message à la banque du client pour qu'elle vérifie le solde du client.

I. La banque du client vérifie le solde de son client.

J. La banque du client approuve ou non la transaction et fait connaître sa décision à la société de cartes. Ce dernier envoie cette réponse au SP.

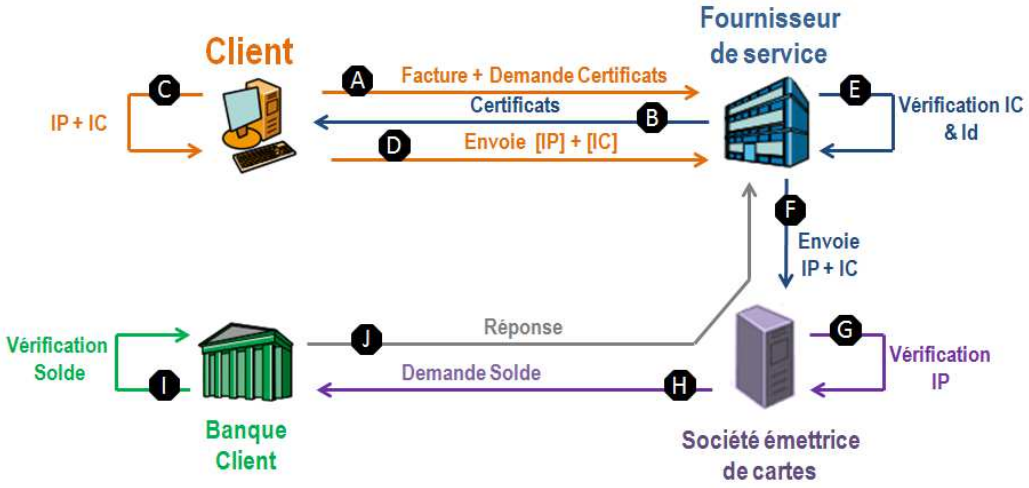


FIGURE 5.4 – Le protocole d'Ashrafi et Ng

Discussion sur le schéma d'Ashrafi et Ng

Le modèle d'Ashrafi et Ng permet d'assurer certaines exigences sur la vie privée et possède un niveau supérieur en termes de protection de la vie privée que le protocole 3D-Secure. Le chiffrement des données, la signature et les éléments d'horodatage permettent d'assurer les exigences E_1 et E_2. Mais, la présence du haché de la commande permet uniquement de respecter partiellement l'exigence E_{10}. Ensuite, l'anonymat du client vis-à-vis du SP et la banque du SP permet de respecter les exigences E_3 et E_4. De plus, le client authentifie la société de cartes, ainsi que le SP grâce à leur certificat, ce qui permet à l'exigence E_6 d'être respectée. Cependant, l'authentification du client par la société de cartes, et non par sa propre banque, ne permet pas d'assurer l'exigence E_5. De la même façon, la confidentialité des données bancaires (E_{11}) et de commande (E_{10}) n'est pas entièrement respectée. En effet, la société de cartes possède beaucoup de ces informations via le deuxième paquet du client transféré par le SP. Ainsi, l'exigence d'anonymat E_{12} ne peut être assurée. Enfin, étant donné que le client calcule et fournit volontairement les

deux paquets de la transaction, le principe de souveraineté des données peut être respecté (E_{14}). Néanmoins, du fait du stockage centralisé de nombreuses données par la société de cartes, l'exigence E_{15} ne peut que partiellement être assurée. Ce protocole permet donc de respecter entièrement sept des quinze exigences et d'en assurer partiellement trois autres.

5.3.4 Discussion

Nous pouvons résumer les exigences respectées par les trois protocoles définis ci-dessus par le Tab.5.1. Nous prendrons dans la suite du chapitre le protocole SET comme référent. En effet, c'est le premier protocole à avoir été proposé comme architecture de paiement en ligne prenant en compte la protection des données personnelles des internautes et à avoir été réellement utilisé.

E_i	Propriétés	SET	3DS	A & N
E_1	Confidentialité des transactions	**Oui**	**Oui**	**Oui**
E_2	Intégrité	**Oui**	**Oui**	**Oui**
E_3	Confidentialité de l'identité du C vis-à-vis du SP	Non	Non	**Oui**
E_4	Confidentialité de l'identité du C vis-à-vis de la banque du SP	Non	Non	**Oui**
E_5	Authentification de C	Non	Non	Non
E_6	Authentification du SP	**Oui**	Non	**Oui**
E_7	Authentification des banques	Non	**Partiel**	Non
E_8	Non-associabilité	-	-	-
E_9	Non-réutilisabilité	-	-	-
E_{10}	Confidentialité des IC	**Partiel**	Non	**Partiel**
E_{11}	Confidentialité des IB	**Partiel**	Non	**Partiel**
E_{12}	Anonymat du C	Non	Non	Non
E_{13}	Minimisation des données du SP	Non	Non	Non
E_{14}	Souveraineté des données	Partiel	**Oui**	Partiel
E_{15}	Sensibilité des données	**Partiel**	Non	**Partiel**
E_{16}	Certificat non obligatoire	Non	**Oui**	**Oui**
E_{17}	Non-répudiation	**Oui**	Non	Non
	Score /17	6	3.5	**7.5**

TABLE 5.1: Analyse des protocoles 3D-Secure, d'Ashrafi et Ng par rapport à SET

Grâce à cette analyse, nous constatons clairement la faiblesse des protocoles de paiement en ligne actuels. Nous proposons dans la suite de ce chapitre trois architectures d'e-paiement respectant davantage la vie privée sur Internet des clients.

5.4 Propositions d'architectures de paiement

Les deux premières architectures de cette section sont des améliorations de deux architectures existantes. Nous détaillons ainsi une modification du protocole d'Ashrafi et Ng, ainsi qu'une amélioration simple du protocole utilisé actuellement par de nombreux fournisseurs de service : le protocole 3D-Secure.

5.4.1 Proposition 1 : Amélioration de 3D-Secure

Présentation

Afin d'améliorer le protocole 3D-Secure, on peut utiliser le certificat de la banque du SP. De plus, dans le protocole 3D-Secure, le $CVX2$ et la date d'expiration sont présents uniquement pour la compatibilité avec les systèmes de paiement existants. Ces données ne sont donc pas nécessaires. Ainsi, étant donné que l'authentification du client à sa banque est forte, ces deux éléments sont inutiles. On utilise alors le certificat de la banque du SP contenant en plus des informations standards, la clé publique du serveur directory. Deux étapes du protocole doivent alors être modifiées :

A. Le client envoie au SP son intention d'achat accompagné de son PAN **chiffré avec la clé publique du serveur directory**. Cette clé est présente dans le certificat de la banque du SP fourni initialement par le SP au client.

C. Le serveur directory contrôle l'identité du SP, puis, **déchiffre le** PAN **avec sa clé privée**. Il peut alors vérifier le numéro de carte et la banque du client puis retrouver le serveur de contrôle d'accès, ou ACS, de la banque émettrice où est enregistrée la carte du client.

Analyse

Cette amélioration permet ainsi de minimiser la connaissance du SP concernant les informations bancaires du client, puisqu'il n'a plus accès au $CVX2$ et à la date d'expiration de la carte, celles-ci étant chiffrées par la banque du client (E_{11} assuré). De plus, étant donné que le PAN est encore connu du serveur directory mais que l'authentification du client est gérée par une unique partie de confiance qui est sa propre banque, l'exigence E_5 est partiellement respectée. Finalement, avec ces améliorations, la sensibilité des données est davantage prise en compte et donc l'exigence E_{15} est partiellement assurée.

Nous constatons donc que ces modifications n'entraînent pas de changement majeur dans le protocole et peuvent donc facilement être déployées. La Figure 5.2 montre l'augmentation du niveau de protection de vie privée du protocole 3D-Secure une fois amélioré.

Néanmoins, toutes les exigences en matière de protection de la vie privée ne sont pas remplies, par exemple, la banque du client connaît ses achats. De plus, de manière générale, le client n'est pas anonyme vis-à-vis du cinquième acteur (le serveur directory). Cet acteur est nécessaire pour authentifier les banques entre elles, mais il est possible de mettre en oeuvre une architecture où le client est anonyme vis-à-vis de ce cinquième acteur. D'autres améliorations sont aussi possibles au niveau de la souveraineté des données ou au niveau de l'authentification des différents acteurs. Une critique plus générale sur l'authentification utilisée dans le protocole 3D-Secure concerne la réalité des deux canaux distincts lors de l'authentification du client, lorsque l'achat est réalisé par le smartphone qui reçoit le SMS.

Tableau de synthèse

La comparaison rapide du niveau de protection de la vie privée du protocole 3D-Secure avec sa modification peut être résumée par le Tab.5.2. On constate qu'à l'aide d'une simple adaptation, il est possible de gagner 2 points en termes de protection des données personnelles et d'approcher le protocole SET.

E_i	**Propriétés**	**SET**	**3DS**	**3DS Amélioré**
E_1	Confidentialité des transactions	**Oui**	**Oui**	**Oui**
E_2	Intégrité	**Oui**	**Oui**	**Oui**
E_3	Confidentialité de l'identité du C vis-à-vis du SP	Non	Non	Non
E_4	Confidentialité de l'identité du C vis-à-vis de la banque du SP	Non	Non	Non
E_5	Authentification de C	Non	Non	**Partiel**
E_6	Authentification du SP	**Oui**	Non	Non
E_7	Authentification des banques	Non	**Partiel**	**Partiel**
E_8	Non-associabilité	-	-	-
E_9	Non-réutilisabilité	-	-	-
E_{10}	Confidentialité des IC	**Partiel**	Non	Non
E_{11}	Confidentialité des IB	Partiel	Non	**Oui**
E_{12}	Anonymat du C	Non	Non	Non
E_{13}	Minimisation des données du SP	Non	Non	Non
E_{14}	Souveraineté des données	**Partiel**	Non	Non
E_{15}	Sensibilité des données	**Partiel**	Non	**Partiel**
E_{16}	Certificat non obligatoire	Non	**Oui**	**Oui**
E_{17}	Non-répudiation	**Oui**	Non	Non
	Score /17	**6**	3.5	5.5

TABLE 5.2: Comparaison du protocole 3D-Secure initial avec sa version modifiée

5.4.2 Proposition 2 : Amélioration du modèle d'Ashrafi et Ng

Présentation

Une modification du protocole d'Ashrafi et Ng permet d'éviter le stockage de toutes les informations bancaires du client dans la base de données centralisée de la société de cartes et ainsi d'assurer complètement E_5, E_{10}, E_{11}, ainsi que le principe de sensibilité des données. Dans notre version, la société de cartes agit comme un relais et ne procède plus aux nombreuses vérifications des données du client. Celles-ci sont déléguées à la banque du client qui en a déjà connaissance. Seules trois étapes ont alors besoin d'être modifiées (les sept autres étapes étant les mêmes que précédemment) :

C. Le client génère deux paquets : Les informations de commande sont chiffrées avec la clé publique du SP, et les informations de paiement sont séparées en deux parties :
- le haché du détail de la carte, le haché du mot de passe du client et le haché du détail de la commande sont **chiffrés avec la clé publique de la banque du client** ;
- **le nom de la banque du client**, l'horodatage et la période de validité sont chiffrés avec la clé publique de la société émettrice de cartes.

G. La société de cartes **déchiffre la première partie des informations de paiement** et vérifie l'horodatage, la période de validité, ainsi que le nom de la banque du client. Le paquet est ensuite chiffré avec la clé publique de la banque du client.

I. La banque du client **déchiffre et vérifie les informations de paiement**, ainsi que le solde du client.

Analyse

Ces modifications du protocole d'Ashrafi et Ng permettent d'assurer entièrement trois exigences supplémentaires. En effet, l'exigence E_5 est assurée grâce à l'authentification du client par sa banque. De plus, l'information n'a pas besoin d'être stockée une seconde fois par la société de cartes étant donné qu'elle est déjà connue d'une partie de confiance, la banque du client. L'exigence E_{11} est donc également respectée. Enfin, la création d'une base de données contenant tous les détails de paiement des clients est évitée, le principe de sensibilité des données peut donc être assuré (E_{15}).

Tableau de synthèse

La comparaison du niveau de protection de la vie privée entre le protocole initial d'Ashrafi et Ng et son amélioration peut se faire rapidement sur le Tab.5.3. On constate qu'à l'aide de simples modifications, il est possible de gagner 3 points en termes de

protection des données personnelles et de largement dépasser les deux versions de 3D-Secure, ainsi que SET, considéré comme référence.

E_i	Propriétés	SET	A & N	A & N Amélioré
E_1	Confidentialité des transactions	**Oui**	**Oui**	**Oui**
E_2	Intégrité	**Oui**	**Oui**	**Oui**
E_3	Confidentialité de l'identité du C vis-à-vis du SP	Non	**Oui**	**Oui**
E_4	Confidentialité de l'identité du C vis-à-vis de la banque du SP	Non	**Oui**	**Oui**
E_5	Authentification de C	Non	Non	**Oui**
E_6	Authentification du SP	**Oui**	**Oui**	**Oui**
E_7	Authentification des banques	Non	Non	Non
E_8	Non-associabilité	-	-	-
E_9	Non-réutilisabilité	-	-	-
E_{10}	Confidentialité des IC	**Partiel**	**Partiel**	**Partiel**
E_{11}	Confidentialité des IB	Partiel	Partiel	**Oui**
E_{12}	Anonymat du C	Non	Non	Non
E_{13}	Minimisation des données du SP	Non	Non	Non
E_{14}	Souveraineté des données	Partiel	**Oui**	**Oui**
E_{15}	Sensibilité des données	Partiel	Partiel	**Oui**
E_{16}	Certificat non obligatoire	Non	**Oui**	**Oui**
E_{17}	Non-répudiation	**Oui**	Non	Non
	Score /17	6	7.5	**10.5**

TABLE 5.3: Comparaison du protocole initial d'Ashrafi et Ng avec la version modifiée

5.4.3 Proposition 3 : Nouvelle architecture de paiement en ligne

Présentation

La nouvelle architecture de paiement que nous proposons combine les avantages des systèmes de chèques électroniques et la facilité d'usage des systèmes de paiement en ligne décrits et analysés dans les sections précédentes. Cependant, l'architecture n'est pas considérée comme un schéma de chèque électronique [222] qui est souvent difficile à utiliser pour un utilisateur lambda. Ces systèmes nécessitent effectivement l'utilisation d'un certificat pour le client et d'une carte de chéquier électronique. De plus, de nombreux calculs et stockages par la banque du client sont nécessaires, même si des améliorations existent [223]. Finalement, ces schémas ne prennent généralement pas en compte la protection de la vie privée, à l'exception de celui présenté dans [224].

L'architecture proposée se concentre sur la phase de paiement et nous supposons que la phase d'authentification entre le SP et le client s'est déroulée correctement. Par conséquent, si le SP accepte l'anonymat du client ou une authentification Zero-Knowledge de celui-ci, le client peut avec cette nouvelle architecture faire un paiement de manière anonyme vis-à-vis du SP et de la banque du SP. L'architecture implique cinq acteurs : le client, le fournisseur de service ou marchand SP, les deux banques associées à ces deux acteurs, ainsi qu'un cinquième acteur, le système interbancaire SI. Ce dernier est le tiers de confiance de l'architecture et son rôle est détaillé par la suite. Chaque banque génère une paire de clé, dont la clé publique est certifiée par le système interbancaire. Ce dernier publie ces certificats contenant : le nom de la banque, la clé publique, l'algorithme de hachage associé, l'algorithme de signature associé et le nom de l'autorité de certification.

De la même manière, le SP possède une paire de clé dont la clé publique est certifiée par le tiers de confiance contractualisé, par exemple le SI. Ce certificat est composé des mêmes informations que ceux des banques : nom, clé publique, fonction de hachage, algorithme de signature et nom de l'autorité de certification, mais également des paramètres décrivant le schéma de paiement reconnu par le SP et permettant de sécuriser le futur paiement, par exemple, American Express, VISA ou MasterCard. Afin de permettre au SP de ne pas révéler l'identité de sa banque au client lors de la transaction, il est préférable de ne pas confier la génération du certificat du SP à la banque du SP. Le système interbancaire pourrait très bien jouer ce rôle.

Dans ce chapitre, nous utilisons des notations particulières afin de détailler notre protocole de paiement :

- $Sign_X(m)$: Signature du message m par l'auteur X avec récupération du message ;
- $[m]_{K_{PU_X}}$: Chiffrement du message m par la clé publique K_{PU} de l'acteur X ;
- $[m]_{K_{S_X}}$: Chiffrement du message m par la clé de session K_S de l'acteur X ;
- A_n : n_{ime} nombre aléatoire utilisé pour garantir la fraicheur du message ;
- $H(m)$: Hash du message m.

L'architecture de paiement respectant la vie privée de l'utilisateur proposée est basée sur la génération de deux documents : un contrat entre le SP et le client, ainsi que la création d'un document de banque électronique ou chèque de banque électronique, appelé pour simplifier chèque. Comme expliqué précédemment, nous rappelons que ce dernier document est différent d'un chèque généré dans les architectures de chèque électronique classique.

Le système interbancaire SI joue le rôle de partie de confiance de l'architecture et

permet la communication entre les deux banques sans révéler d'information sur les autres acteurs. Bien que ce cinquième acteur puisse être supprimé de l'architecture et qu'il a le plus petit rôle possible, il n'est pas judicieux de s'en passer. En effet, il permet de vérifier l'authentification des deux banques et ainsi d'éviter le blanchiment d'argent. Un tel système interbancaire se doit de répondre à des problématiques de réglementations internationales afin de gérer des transferts entre pays. Il faut alors imaginer une sorte de PKI de SI avec différentes couches de réglementations afin de résoudre les éventuels problèmes entre les différents pays du Globe.

Des sociétés internationales comme Visa ou MasterCard pourraient actuellement jouer le rôle du SI. Cependant, dans un cadre européen et en raison de sa neutralité, la société belge SWIFT (Society for Worldwide Interbank Financial Telecommunication, [225]) serait un bon candidat pour un système interbancaire. SWIFT est en effet un réseau interbancaire offrant divers services tels que des transferts interbancaires entre centaines de pays différents. Ainsi, la nouvelle architecture proposée est composée de quatorze étapes illustrées par la Fig.5.5 et détaillées ci-dessous.

Dans un premier temps, le client remplit son panier *Basket* et l'envoie au SP avec l'aléa A_1, ainsi qu'une clé de session K_{S_1} (Étape A). Ces informations sont chiffrées avec la clé publique du SP. A_1 permet d'assurer la fraîcheur du message et K_{S_1} de chiffrer les données entre le client et le SP étant donné que le client n'a pas de certificat et qu'il ne possède donc pas de clé publique. Dans le cas où il en possède une, la clé de session est remplacée par sa clé publique.

$$Client \rightarrow SP : [Basket, A_1, K_{S_1}]_{K_{PU_{SP}}} \qquad (A)$$

Ensuite, le SP génère un contrat avec son client à l'aide des informations recueillies à l'étape (A). Le contrat contient (Étape B).

- Le montant total *Amount* de l'achat ;
- Un numéro de commande aléatoire *Order* généré par le SP. Ce numéro ne doit pas avoir de lien avec l'identité du SP mais permet de lier la commande au futur chèque généré par la banque du client ;
- Une clé symétrique aléatoire K_{S_2} chiffrée avec la clé publique de la banque du SP $K_{PU_{Bank_{SP}}}$;
- Le nom du bénéficiaire du chèque $Benef$ chiffré par la clé précédemment générée K_{S_2}. On utilise une clé de session pour chiffrer le nom du bénéficiaire afin de réduire la complexité des calculs ;
- L'URL du SP afin de retourner à la page paiement le moment venu ;

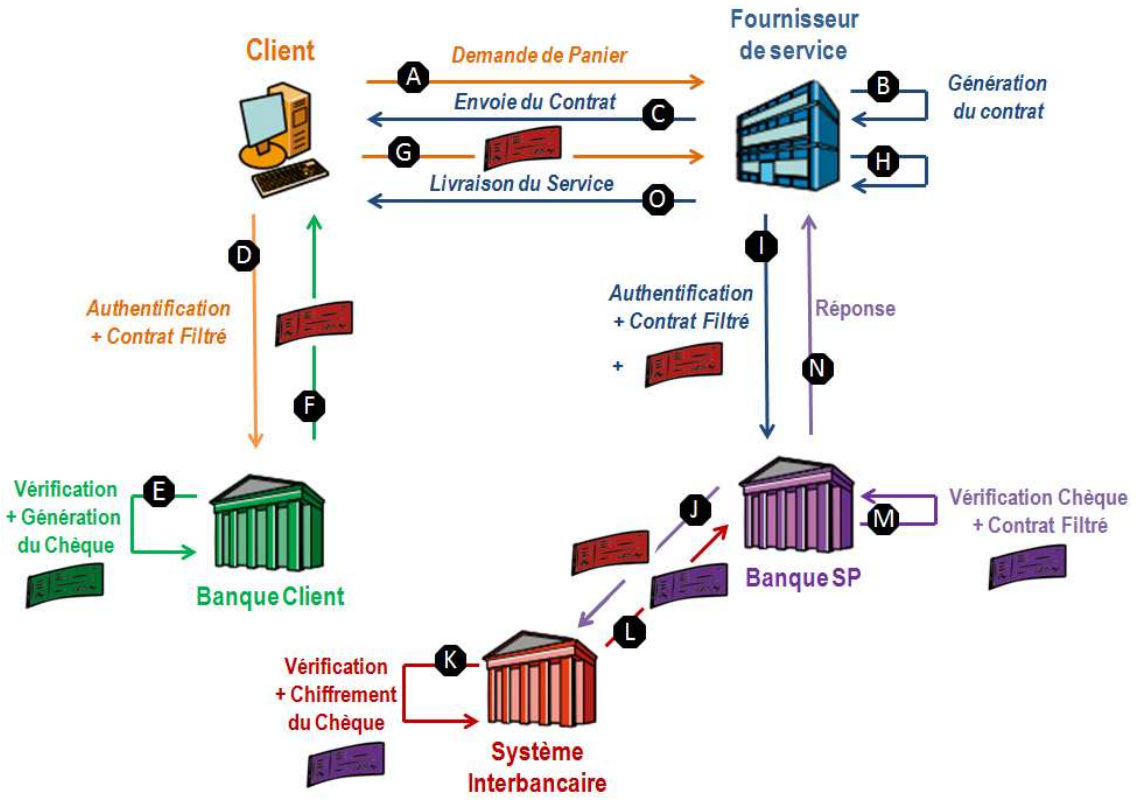


FIGURE 5.5 – Architecture de paiement en ligne proposée.

– Les détails du panier *Basket* : liste des achats, quantité, prix unitaire.

$$SP : Contract = \{Amount, Order, [K_{S_2}]_{K_{PU_{Bank_{SP}}}}, [Benef]_{K_{S_2}}, \\ URL, Basket\} \qquad (B)$$

Afin d'éviter la non-répudiation et d'assurer l'authenticité du SP, le SP doit signer ce contrat avec sa clé privée. Celui-ci est ensuite envoyé signé au client, ainsi que le hash du panier et de l'aléa A_1 précédemment envoyé par le client (Étape C).

$$Client \leftarrow SP : [Sign_{SP}(Contrat, H(N_1, Basket))]_{K_{S_1}} \qquad (C)$$

L'étape D consiste en l'authentification du client à sa banque en utilisant un plugin dédié à son navigateur Internet. Le plugin établit une connexion HTTPS et joint le contrat filtré à la banque du client. Le nombre de transactions entre le client et sa banque est réduit à une unique transaction d'authentification et de transfert du contrat filtré afin d'augmenter le taux d'acceptation de la banque et de diminuer le montant de la commission. Le protocole d'authentification dépend de la banque du client. Cependant, une authentification forte,

comme celles détaillées dans le chapitre 2, est recommandée. Le contrat filtré contient uniquement les informations nécessaire du contrat précédemment généré par le SP : le montant total, la devise, la clé symétrique chiffré, le nom du bénéficiaire du chèque chiffré par cette dernière clé, ainsi que le numéro de commande. Ainsi, la banque du client ne connait pas l'identité du bénéficiaire et donc du SP. De plus, l'aléa A_2 assure la fraîcheur du message. Étant donné que le client ne possède pas en général de certificat, la banque du client utilise une clé de session K_{S_3} afin de chiffrer les messages échangés avec son client.

$$Client \rightarrow Bank_C : [Amount, Order, [K_{S_2}]_{K_{PU_{Bank_{SP}}}}, [Benef]_{K_{S_2}}, A_2, K_{S_3}]_{K_{PU_{Bank_C}}} \quad (D)$$

Ensuite, si l'authentification du client est réussie et que celui-ci est solvable, la banque répond positivement à la requête du client générant le chèque de banque associé au contrat filtré (Étape E). Ce chèque comprend : le montant total, la devise, le numéro de commande aléatoire, le nom du bénéficiaire chiffré, la clé symétrique de chiffrement de ce nom, une durée de validité du chèque *DueTimeDate*, les informations sur la banque du client *BankDetails*, ainsi que la signature de la banque du client. Ainsi, aucune information bancaire du client n'est transmise.

$$Bank_C : Cheque = Sign_{Bank_C}(Amount, Order, [K_{S_2}]_{K_{PU_{Bank_{SP}}}}, [Benef]_{K_{S_2}}, DueTimeDate, BankDetails) \quad (E)$$

La banque du client signe le chèque et le chiffre avec la clé publique du système interbancaire SI. Ainsi, SI pourra vérifier l'identité des banques et la validité du chèque. Ce chèque est ensuite envoyé au client (Étape F) qui le transfère au SP (Étape G). Les nombres aléatoires A_2 et A_3 assurent la fraîcheur des transactions. De plus, A_2 permet d'identifier la précédente requête. Cette réponse étant chiffrée par la clé publique du SI, le SP ne peut pas connaître les informations de la banque du client.

$$Client \leftarrow Bank_C : [[Cheque]_{K_{PU_{IS}}}, A_2]_{K_{S_3}} \quad (F)$$

$$Client \rightarrow SP : [[Cheque]_{K_{PU_{IS}}}, A_3]_{K_{PU_{SP}}} \quad (G)$$

Pour l'étape H, le SP obtient le chèque chiffré $[Cheque]_{K_{PU_{IS}}}$ et A_3 en déchiffrant avec sa clé privée. A l'étape I, le SP s'authentifie à sa banque et lui fournit le contrat filtré, ainsi que le chèque chiffré par la clé publique de sa banque et précédemment signé par la banque du client. Comme pour l'authentification du client à sa banque, le protocole d'authentification du SP à sa banque dépend de la banque du SP mais une authentification forte est conseillée (voir chapitre 2). Le contrat filtré envoyé contient : le montant total, la devise, le nom du bénéficiaire, le numéro de commande et le nombre aléatoire A_4.

$$SP \rightarrow Bank_{SP} : [Amount, Order, Benef, [Cheque]_{K_{PU_{IS}}}, A_4]_{K_{PU_{Bank_{SP}}}} \quad (I)$$

Afin de confirmer la validité des banques et du chèque, la banque du SP s'authentifie au SI et lui transfère le chèque chiffré avec la clé publique du SI (Étape J) en utilisant A_5 pour la fraîcheur de la transaction.

$$Bank_{SP} \rightarrow IS : [[Cheque]_{K_{PU_{IS}}}, A_5]_{K_{PU_{IS}}} \quad (J)$$

Le système interbancaire vérifie l'identité de la banque du SP et déchiffre le chèque électronique avec sa clé privée (Étape K). La validité du chèque, sa signature et ainsi l'identité de la banque du client peuvent ensuite être contrôlées. Ensuite, si les authentifications des banques sont correctes, le SI rechiffre le chèque avec la clé publique de la banque du SP et le lui transfère (Étape L). A_5 est à nouveau utilisé pour identifier la requête.

$$Bank_{SP} \leftarrow IS : [Cheque, A_5]_{K_{PU_{Bank_{SP}}}} \quad (L)$$

La banque du SP déchiffre alors le chèque avec sa clé privée (Étape M). Elle vérifie d'abord que le montant et la devise indiqués sur le chèque sont les mêmes que ceux indiqués sur le contrat filtré du SP. Elle déchiffre ensuite la clé symétrique avec sa clé privée et peut ainsi déchiffrer le nom du bénéficiaire du chèque. Elle le compare alors à celui présent dans le contrat filtré. La banque du SP peut également faire la vérification de la signature de la banque du client. Cependant, cette vérification n'est pas obligatoire étant donné qu'elle a déjà été réalisée par le système interbancaire. La banque du SP peut ainsi directement utiliser les informations de la banque du client.

Finalement, si toutes les informations coïncident, la banque du SP contacte le SP et confirme l'authenticité et la validité du chèque. Cette confirmation permet au SP de délivrer le service à son client (Étape N, O). Les nombres aléatoires A_3 et A_4 permettent là encore d'identifier la requête et de garantir la fraîcheur des transactions. Dans le cas où une des vérifications ne serait pas correcte, la transaction est évidemment interrompue.

$$SP \leftarrow Bank_{SP} : [Response, Amount, Order, A_4]_{K_{PU_{SP}}} \quad (N)$$
$$Client \leftarrow SP : [Service, Amount, Order, A_3]_{K_{S_1}} \quad (O)$$

La banque du SP contacte enfin la banque du client qui a pu être localisée grâce aux informations présentes dans le chèque. La procédure de débit/crédit entre ces banques complète cette nouvelle architecture de paiement où le chèque est utilisé comme une preuve de paiement.

Analyse de cette nouvelle architecture

Cette proposition, bien que remettant en cause les protocoles de paiement actuels comme 3D-Secure, permet de ne faire transiter aucune information de paiement concernant le Client via le SP et d'assurer la majorité des exigences lors des neuf premières étapes du protocole. Les cinq dernières étapes permettent d'assurer uniquement l'authentification des banques par le système interbancaire (A_7) et ainsi permettent d'éviter le blanchiment d'argent. De plus, le protocole proposé ne contient pas plus de transactions que le protocole complet de 3D-Secure qui n'est pas détaillé dans ce chapitre.

Sécurité des données et authentification

La création d'un canal sécurisé entre les différents auteurs et l'utilisation de schémas de chiffrement permettent de gérer la confidentialité des données échangées durant la totalité du protocole. L'exigence E_1 est ainsi assurée. L'usage de nombres aléatoires évite l'associabilité des données et le rejeu d'une transaction et permet ainsi d'assurer les exigences E_8, E_9 et également d'assurer l'intégrité des données E_2. De plus, les deux banques et le SP possèdent des certificats signés pour ces deux premières par le système interbancaire. De plus, contrairement au protocole SET, le tiers de confiance n'est pas la banque du SP mais le système interbancaire. La possession de tels certificats permet aux acteurs d'avoir les clés pour signer, chiffrer et déchiffrer les informations qu'ils souhaitent transférer. Le système interbancaire gère donc les certificats des banques et authentifie la banque du client et la banque du SP. De plus, il vérifie les informations contenues dans les chèques électroniques signés et confirme la validité de celui-ci auprès de la banque du SP. Le contrat signé par le SP permet quant à lui d'obtenir le service souhaité avec les conditions indiquées. Finalement, la validation de l'identité de la banque du client par la banque du SP, ainsi que la vérification des informations de transaction assure au SP d'être payé une fois le service fourni. Ainsi, les exigences E_6 et E_7 peuvent être assurées. De plus, ces vérifications par le système interbancaire permettent d'éviter le blanchiment d'argent par un SP malveillant ou un banque malveillante.

Analyse de la protection de la vie privée

Á notre connaissance, l'architecture proposée est plus respectueuse de la vie privée de l'utilisateur que l'ensemble des protocoles de paiement en ligne existants et dans la littérature. L'authentification du SP par le client et par sa banque assure la validité du SP et permet donc au client de ne fournir aucune information personnelle avant d'être certain de demander un service du SP. De plus, l'identité du client n'est jamais révélée au SP, si celui-ci permet un enrôlement anonyme. De même, elle n'est jamais dévoilée à la banque du SP. La banque du client est la seule à authentifier son client. Ainsi, les

exigences E_3, E_4 et E_5 sont respectivement assurées. Plus précisément, dans le but de respecter davantage la vie privée du client durant le transfert de données aux différentes banques, le numéro de commande, utilisé dès l'étape C, ne doit pas contenir d'information concernant le SP, tel que son numéro commercial. Il est par conséquent aléatoire ou non-identifiable. Dans le cas où les deux banques seraient les même, toutes les exigences seraient assurées de la même façon, à l'exception de la connaissance du SP et de son client. De plus, la banque du client ne connait ni le contenu du panier de son client, ni le SP avec lequel son client traite. En effet, à partir du moment où le client est identifié et solvable, il peut faire une transaction de paiement. De plus, le nom du bénéficiaire du chèque est chiffré et non déchiffrable par la banque du client. Ainsi, l'exigence E_{10} est ainsi assurée.

Cette nouvelle architecture résout également les problèmes du protocole 3D-Secure. Les informations bancaires du client sont en effet ignorées du SP et même de la banque du SP assurant ainsi l'exigence E_{11}. Le chèque chiffré avec la clé publique du système interbancaire permet au SP de ne pas connaître la banque du client. De plus, contrairement à l'ensemble des protocoles existant, aucune information bancaire du client ne transite lors de ce protocole. Ainsi, les exigences E_3, E_4, E_{10} et E_{11} sont respectées et donc, si le SP le permet, le client peut être anonyme (E_{12}).

Finalement, le chèque chiffré avec la clé publique du système interbancaire évite au client de connaître la banque du SP. De plus, la banque du SP ne connait pas le client avec lequel le SP traite. Le principe de minimisation des données du SP est donc respecté (E_{13}). Cette exigence est notamment importante lorsque le SP est une petite organisation et par conséquent lorsque la banque du SP est également la banque personnelle du marchand. Qui plus est, une sécurité supplémentaire pour le SP est possible si le client possède un certificat. En effet, l'exigence E_{17} de non-répudiation pourra ainsi être assurée.

Au vu des nombreuses analyses ci-dessus, nous constatons que les données sensibles du client et du SP sont protégées et que le client ne fournit que les données nécessaires, appropriées et pertinentes. Les principes de minimisation et de sensibilité sont donc respectées (E_{15}). De plus, contrairement à l'ensemble des protocoles existants, le cinquième acteur de l'architecture joue un rôle à la fin de la transaction et a donc une connaissance restreinte des données transitant. La protection totale de la vie privée du client n'est ainsi jamais rendu impossible. De plus, une fois le contrat signé par le SP, le client doit cliquer deux fois pour accepter la transaction : une fois pour la confirmation de son panier et une autre fois pour la validation du paiement. Ainsi, le fait de devoir lire et cliquer deux fois pour valider une même information assure le consentement du client à payer pour ce service

(E_{14}). Ces actions pourront être remplacées par la signature du client si celui-ci possède un certificat. Dans le futur, le certificat pourrait se trouver dans la carte d'identité du client ou dans son passeport par exemple. Cependant, pour le moment, le fait de posséder un certificat n'est pas nécessaire pour le client (E_{16}), contrairement au protocole SET.

Tableau de synthèse

Ce nouveau protocole ayant était construit de façon à respecter un maximum d'exigences, il est ainsi logique qu'il assure la totalité de celles-ci comme le résume le Tab.5.4. L'architecture est donc clairement supérieure en termes de protection des données personnelles au protocole SET.

E_i	Propriétés	SET	Nouveau Protocole
E_1	Confidentialité des transactions	**Oui**	**Oui**
E_2	Intégrité	**Oui**	**Oui**
E_3	Confidentialité de l'identité du C vis-à-vis du SP	Non	**Oui**
E_4	Confidentialité de l'identité du C vis-à-vis de la banque du SP	Non	**Oui**
E_5	Authentification de C	Non	**Oui**
E_6	Authentification du SP	**Oui**	**Oui**
E_7	Authentification des banques	Non	**Oui**
E_8	Non-associabilité	-	**Oui**
E_9	Non-réutilisabilité	-	**Oui**
E_{10}	Confidentialité des IC	Partiel	**Oui**
E_{11}	Confidentialité des IB	Partiel	**Oui**
E_{12}	Anonymat du C	Non	**Oui**
E_{13}	Minimisation des données du SP	Non	**Oui**
E_{14}	Souveraineté des données	Partiel	**Oui**
E_{15}	Sensibilité des données	Partiel	**Oui**
E_{16}	Certificat non obligatoire	Non	**Oui**
E_{17}	Non-répudiation	**Oui**	(Option)
	Score /17	6	**16**

TABLE 5.4: Comparaison du nouveau protocole de paiement en ligne avec le protocole référent SET

Traitement des litiges

Contrairement à 3D-Secure où tous les acteurs se connaissent, la nouvelle architecture se base sur les connaissances réciproques des acteurs. Cela n'empêche en rien la prise en compte des problèmes de litiges liés, par exemple, à une rupture de dialogue entre deux

acteurs, à une erreur de transaction, ou encore à un problème quelconque, et ceci même lors du paiement. En effet, dans le cas d'un litige, le système interbancaire peut retrouver les banques ayant participées à la commande dont le numéro est *Order*. Á l'aide de ce numéro de commande, chacune des banques peut retrouver son client (et donc également le SP), ainsi que le chèque associé à la transaction. Le client et le SP sont alors avertis de l'annulation de la transaction.

Dans le cas où le problème est lié à la livraison du service par le SP et donc que le client a déjà été débité, le procédé inverse se met en place. Le client contacte sa banque qui contacte le système interbancaire. L'ensemble des cas de litiges peut ainsi être réglé à l'aide du numéro de commande et des différents niveaux d'acteurs de la transaction.

Discussion

Dans ce chapitre, nous avons proposé trois nouvelles architectures de paiement en ligne. Deux d'entres elles sont des améliorations d'algorithmes présents dans la littérature, dont un est implémenté et utilisé actuellement par de nombreuses banques et nombreux fournisseurs de service. La troisième architecture proposée est basée sur la génération d'un document électronique, appelé chèque, fourni par la banque du client et par un contrat généré par le SP. Comme le résume la Fig.5.5, les analyses des différents protocoles traités dans ce chapitre montre clairement la robustesse en termes de sécurité et de protection des données personnelles de la nouvelle proposition. Ces analyses ont pu être précisément réalisées à l'aide des exigences définies dans ce chapitre et prenant en compte les différentes contraintes présentes dans la littérature en termes de protection de la vie privée lors un paiement en ligne.

Le protocole d'Ashrafi et Ng n'ayant jamais été exploité, il est peu probable que son amélioration que nous proposons dans ce chapitre le soit également. En effet, au vu des nombreuses modifications à apporter aux protocoles actuels pour le mettre en place, il serait alors judicieux de privilégier la troisième architecture qui obtient un score nettement supérieur en termes de vie privée et de sécurité. À contrario, les petites modifications à apporter à 3D-Secure, afin de le rendre plus respectueux des données personnelles des utilisateurs, rendent l'amélioration proposée tout à fait envisageable et exploitable rapidement. Afin de nous rendre compte de la faisabilité de la troisième proposition, une preuve de concept a également été développée et permet de confirmer la possibilité de mettre en place une telle architecture.

E_i	Propriétés	SET	3DS	3DS Amélioré	A & N	A & N Amélioré	Nouveau Protocole
E_1	Confidentialité des transactions	**Oui**	**Oui**	**Oui**	**Oui**	**Oui**	**Oui**
E_2	Intégrité	**Oui**	**Oui**	**Oui**	**Oui**	**Oui**	**Oui**
E_3	Confidentialité de l'identité du C vis-à-vis du SP	Non	Non	Non	**Oui**	**Oui**	**Oui**
E_4	Confidentialité de l'identité du C vis-à-vis de la banque du SP	Non	Non	Non	**Oui**	**Oui**	**Oui**
E_5	Authentification de C	Non	Non	Partiel	Non	**Oui**	**Oui**
E_6	Authentification du SP	**Oui**	Non	Non	**Oui**	**Oui**	**Oui**
E_7	Authentification des banques	Non	Partiel	Partiel	Non	Non	**Oui**
E_8	Non-associabilité	-	-	-	-	-	**Oui**
E_9	Non-réutilisabilité	-	-	-	-	-	**Oui**
E_{10}	Confidentialité des IC	Partiel	Non	Non	Partiel	Partiel	**Oui**
E_{11}	Confidentialité des IB	Partiel	Non	**Oui**	Partiel	**Oui**	**Oui**
E_{12}	Anonymat du C	Non	Non	Non	Non	Non	**Oui**
E_{13}	Minimisation des données du SP	Non	Non	Non	Non	Non	**Oui**
E_{14}	Souveraineté des données	Partiel	Non	Non	**Oui**	**Oui**	**Oui**
E_{15}	Sensibilité des données	Partiel	Non	Partiel	Partiel	**Oui**	**Oui**
E_{16}	Certificat non obligatoire	Non	**Oui**	**Oui**	**Oui**	**Oui**	**Oui**
E_{17}	Non-répudiation	**Oui**	Non	Non	Non	Non	(Option)
	Score /17	6	3.5	5.5	7.5	10.5	**16**

TABLE 5.5: Synthèse des analyses des différents protocoles étudiés

5.5 Éléments d'acceptabilité de l'architecture proposée

Lors de l'étude réalisée durant un projet étudiant de l'ENSICAEN, nous avons posé un certain nombre de questions aux personnes sondées sur leur connaissance en termes de fichiers et de certificats électroniques. Notre objectif était de connaître le degré d'acceptation d'une telle nouveauté pour des personnes lambda. Ainsi, différentes questions ont été posées à ce sujet dont les réponses confirment la possibilité d'un tel protocole respectueux de la vie privée et ceci, même si la transaction devait durer un peu plus longtemps qu'un des protocoles existants :

- À la question "Seriez vous prêt à attendre moins d'une minute supplémentaire si cela vous permettait de protéger davantage vos données personnelles lors du paiement en ligne ?", la réponse était "oui" à 97%, soit 268 sondés sur 277 réponses.
- À la question "Si vous deviez faire quelques manipulations lors d'un paiement sur internet (comme enregistrer un fichier puis le mettre en pièce jointe) et si cela vous permettait de faire votre achat sur internet de façon sécurisée sans fournir aucune de vos informations personnelles, ni aucune de vos coordonnées bancaires au marchand,

seriez-vous prêt à le faire", la réponse était également positive à 89%, soit 315 sondés sur 354 réponses.

– En ce qui concerne l'usage d'un certificat pour le client et sur un panel regroupant beaucoup de scientifiques, 64% savent ce qu'est un certificat numérique. On peut donc en déduire que sur un panel clairement représentatif de la population, le pourcentage serait nettement moins important. Il était donc important de proposer une architecture de paiement en ligne ne nécessitant pas la possession d'un certificat mais pouvant être facilement modifiée lors d'une démocratisation de tels outils.

5.6 Preuve de concept

Afin de nous rendre compte de la faisabilité de notre concept et de sa facilité d'utilisation, il était nécessaire de proposer une preuve de concept. Plusieurs étudiants de l'ENSICAEN ont ainsi implémenté l'architecture proposée. La plate-forme réalisée est découpée en cinq parties distinctes représentant les cinq acteurs de l'architectures. La partie client correspond à un plugIn développé pour le navigateur Firefox. Les différents langages utilisés sont le C#, le langage HTML et le JavaScript. Un aperçu du rendu actuel est donné par les Fig. 5.6, 5.7, 5.8, 5.9, 5.10 et 5.11.

Supposons qu'un client souhaite acheter un bien sur le site marchand (factice) de l'ENSICAEN. Il doit alors se connecter à son compte préalablement créé. Évidemment, lors de l'enregistrement, il aura uniquement fourni les informations nécessaires et pertinentes, à savoir un login et un mot de passe, les autres informations demandées étant optionnelles.

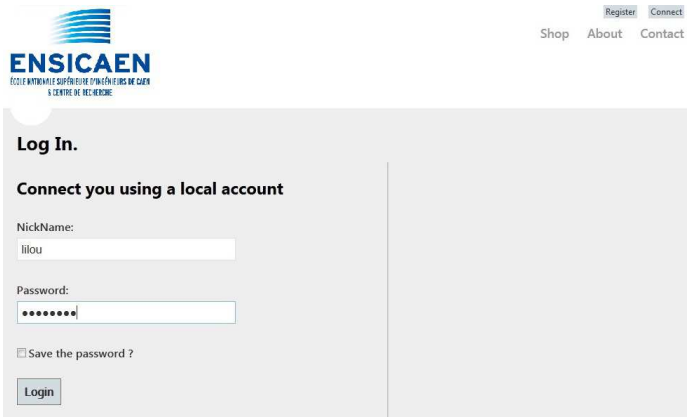


FIGURE 5.6 – Connexion au site marchand ENSICAEN

Une fois connecté, le client, qui souhaite acheter une paire de bottes, remplit son panier.

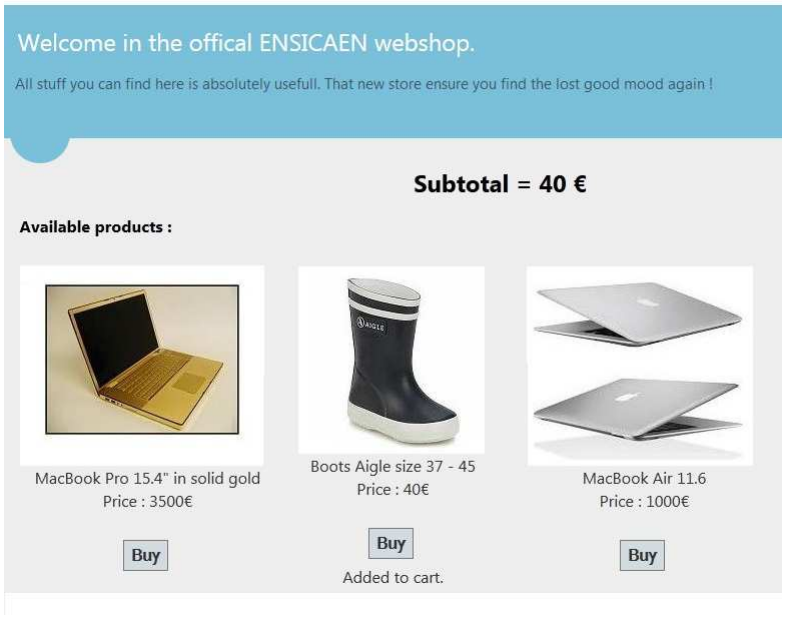


FIGURE 5.7 – Choix du panier.

Le panier ainsi rempli, le client doit le valider deux fois (procédure du double-clique) et accepter les conditions de vente du site.

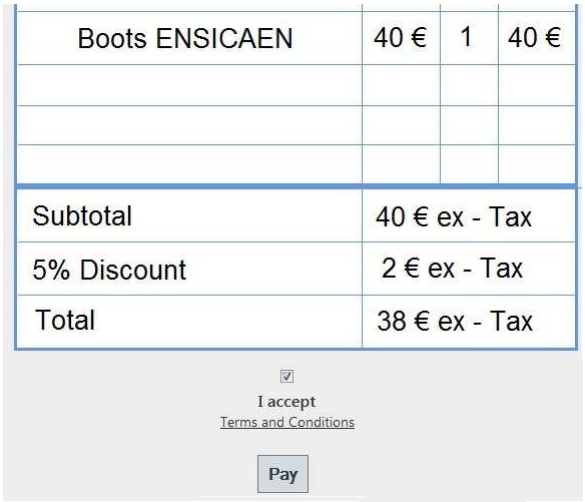


FIGURE 5.8 – Récapitulatif de commande.

Le client doit ensuite choisir son mode de livraison. Dans le cas où il choisit une livraison via un relais, aucune adresse personnelle ne lui est demandée. Dans notre cas, il choisit la livraison à domicile et entre son adresse postale qui ne sera pas conservée.

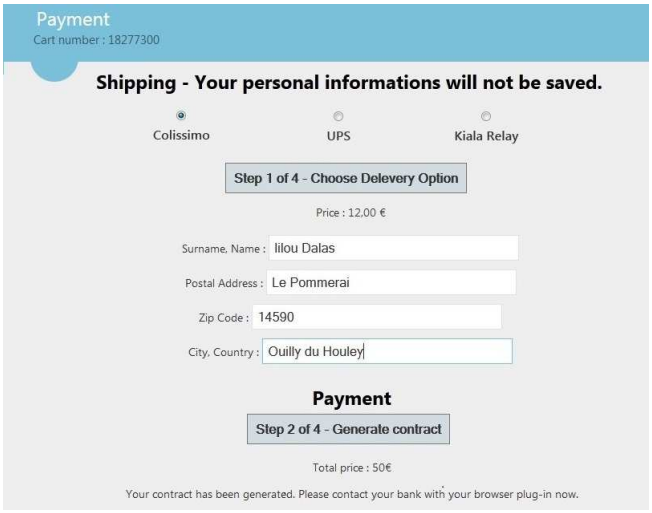


FIGURE 5.9 – Choix du mode de livraison

Le contrat est ensuite généré par le marchand à l'aide des informations du panier. Le chèque de banque est généré par la banque du client et transféré à la banque du marchand. Toutes ces transactions sont réalisées de manière transparente pour le client qui est averti de chacune d'entre elles.

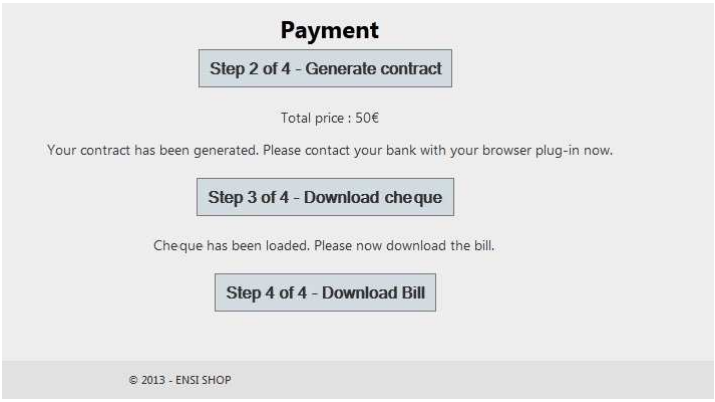
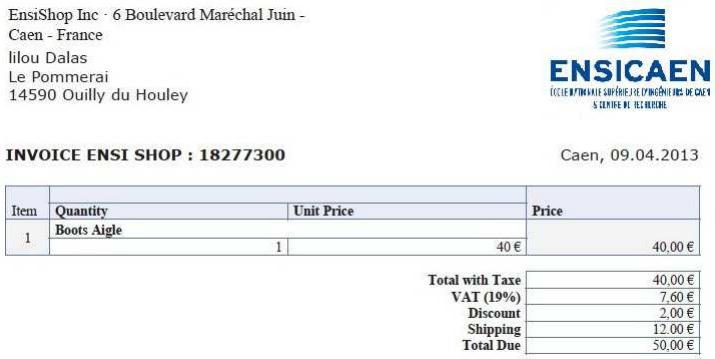


FIGURE 5.10 – Génération du contrat et du chèque.

Le client peut finalement télécharger sa facture avec les détails de son achat et/ou la recevoir par email s'il le souhaite. Il doit alors entrer son adresse mail à la fin de la transaction.



EnsiShop Inc - 6 Boulevard Maréchal Juin -
Caen - France
lilou Dalas
Le Pommerai
14590 Ouilly du Houley

ENSICAEN

INVOICE ENSI SHOP : 18277300 Caen, 09.04.2013

Item	Quantity	Unit Price	Price
1	Boots Aigle		
	1	40 €	40,00 €

Total with Taxe	40,00 €
VAT (19%)	7,60 €
Discount	2,00 €
Shipping	12,00 €
Total Due	50,00 €

FIGURE 5.11 – Génération de la facture.

Une fois la facture téléchargée, la transaction est terminée. Le client va pouvoir récupérer son service à la date prévue par le SP.

5.7 Conclusion

Un grand nombre d'informations sensibles est transféré lors d'un paiement en ligne, ce qui introduit de nombreuses failles en termes de protection de la vie privée des utilisateurs. Les systèmes de paiement électronique actuels, comme 3D-Secure, ne visent pas à assurer de tels principes de protection. De plus, les schémas visant à les améliorer, tel que celui d'Ashrafi et Ng ou SET, sont complexes et/ou ne prennent pas en compte un certain nombre de ces principes. Nous avons pu constater, grâce à l'étude statistique réalisée, que ces problèmes de protection de la vie privée sont pourtant bien présents dans l'esprit des utilisateurs. Une liste exhaustive d'exigences indispensables en termes de sécurité et de protection de la vie privée a été présentée. Ainsi, une amélioration, pouvant être déployée facilement, du protocole 3D-Secure a pu être proposée et également, de façon similaire, une amélioration du protocole d'Ashrafi et Ng a été avancée. La nouvelle architecture proposée permet de contrer ces faiblesses en respectant la vie privée du client contre les banques et le SP et améliore également la protection de la vie privée du SP. Cette solution est principalement basée sur la génération d'un chèque bancaire électronique.

Cette architecture est totalement compatible avec les principes de minimisation des données, de souveraineté de données et de sensibilité des données. Plus particulièrement, la transaction de paiement ne divulgue et n'utilise aucune coordonnée bancaire du client. De plus, ce dernier n'a pas besoin d'avoir de connaissances particulières en termes d'outils cryptographiques. Concernant la non-répudiation, elle pourrait être améliorée en fournissant au client un certificat. Cependant, comme nous avons pu le constater durant notre

étude, cette option ne sera pas viable avant quelques années. Par ailleurs, la preuve de concept ainsi développée nous permettra de lancer une autre étude auprès des utilisateurs. Nous pourrons ainsi tester le degré d'acceptation de cette nouvelle architecture.

5.8 Perspectives

Cette architecture peut également être envisagée dans le cas d'un transfert d'argent entre deux comptes clients. Dans cette configuration, le SP est remplacé par un autre client d'une banque quelconque.

Par la suite, nous aimerions étudier les chiffrements malléables. Ces algorithmes pourraient être une autre solution permettant de gérer différemment le chèque de banques. De même, nous pourrions envisager d'utiliser un algorithme de signature aveugle afin de faire signer le contrat non filtré par la banque du client et ainsi obtenir une preuve de paiement pour la commande associée.

Travaux de l'auteur sur ce thème

- **A. Plateaux**, P. Lacharme, V. Coquet, S. Vernois, K. Murty, C. Rosenberger, *An e-payment architecture ensuring a high level of privacy protection*, 9th International Conference on Security and Privacy in Communication Networks (SecureComm), Sydney, Australie, Septembre 2013.
- **A. Plateaux**, P. Lacharme, V. Coquet, S. Vernois, G. Frey, A. Gouriou, *Protection de la vie privée dans les modèles de paiement en ligne*, 8ième Conférence sur la Sécurité des Architectures Réseaux et des Systèmes d'Information (SARSSI), Mont de Marsan, France, Septembre 2013.
- **A. Plateaux**, P. Lacharme, C. Rosenberger, *Protection de la vie privée dans le système de paiement 3D-Secure*, 4ième Atelier Protection de la Vie Privée (APVP), Les Loges en Josas, France, Juin 2013.
- **A. Plateaux**, V. Coquet, S. Vernois, P. Lacharme, C. Rosenberger, K. Murty, *A Privacy Preserving E-Payment Architecture*, Financial Cryptography and Data Security (FC) - Session POSTER, Okinawa, Japon, Avril 2013.
- **A. Plateaux**, V. Coquet, P. Lacharme, S. Vernois, C. Rosenberger, *E-payment architecture preserving the privacy*, Société BULL SAS, Numéro du brevet déposé : *US* 04097.

CONCLUSIONS ET PERSPECTIVES

Bilan

Nous avons étudié dans ce mémoire trois domaines d'applications des transactions électroniques sécurisées (ou TES). Nous avons commencée par définir les termes principaux de la thèse que sont la vie privée et les transactions électroniques sécurisées. Afin de positionner la problématique, nous avons également répertorié les menaces existantes envers la vie privée des utilisateurs, ainsi que les propriétés attendues en termes de protection des données personnelles et de sécurité. Nous avons ensuite exposé les solutions existantes dans la littérature permettant d'assurer uniquement certaines exigences en termes de protection des données personnelles.

Le manuscrit s'articule ensuite autour de trois domaines : la gestion des données personnelles sur Internet, la e-santé, ainsi que le paiement en ligne. Pour chacun de ces domaines, nous avons proposé une ou plusieurs architectures complètes assurant la sécurité du système et permettant de protéger la vie privée des internautes, des patients et des clients.

Contributions

Pour la gestion des données personnelles d'un internaute, nous proposons une application utilisant un formulaire pré-rempli, des outils cryptographiques bien connus tels que les certificats, des preuves de connaissance, ainsi que des protocoles d'authentification tels que la biométrie révocable combinée à un mot de passe, l'authentification par lecteur CAP ou encore l'authentification de type Zero-Knowledge. Cette application possède différentes fonctionnalités : elle est capable de vérifier l'identité du SP sur lequel navigue l'utilisateur, d'analyser les conditions d'utilisation et d'accès du site concerné, de fournir une preuve de connaissance pour les données sensibles telles que la date de naissance, et enfin de gérer différentes identités de l'utilisateur. L'ensemble des données traitées sont évidemment

stockées dans un coffre-fort électronique. Ces différentes fonctionnalités permettent à l'utilisateur d'être guidé lors de sa navigation sur Internet et informé du degré de sensibilité de ses informations. Cette première architecture offre ainsi une solution sécurisée, centrée sur l'utilisateur et facile d'utilisation pour ce dernier.

En ce qui concerne notre deuxième domaine d'étude, nous construisons un système de e-santé permettant de protéger les dossiers médicaux (et administratifs) des patients au sein d'un établissement de santé mais également lors de transferts de ses données entre différentes institutions. L'architecture proposée est une solution décentralisée qui traite le problème de la gestion des identités des patients et des employés à l'intérieur d'une institution médicale et gère ainsi le contrôle d'accès aux données du patient. Le chiffrement des dossiers médicaux est également pris en compte dans notre approche qui est voulue facile d'utilisation et basée sur des outils bien connus de cryptographie. Cette architecture a pour objectif d'assurer la totalité des exigences nécessaires en termes de protection des données personnelles et de sécurité des systèmes de e-santé.

Les dernières contributions de cette thèse résolvent une partie des failles liées aux systèmes de paiement électronique actuels, comme 3D-Secure. Nous proposons ainsi une amélioration de ce système et du protocole d'Ashrafi et Ng. Ces deux améliorations permettent d'augmenter le degré de protection des données personnelles de ces deux systèmes de paiement en ligne. Nous détaillons pour finir une nouvelle architecture de paiement sur Internet permettant de gérer l'ensemble des failles des différents protocoles existants dans la littérature et d'assurer la totalité des propriétés attendues dans un tel système. Cette nouvelle architecture est principalement basée sur la génération d'un chèque de banque électronique et d'un contrat entre le client et son SP. De plus, elle ne divulgue et n'utilise aucune coordonnée bancaire du client qui n'a, quant à lui, besoin d'aucune connaissance cryptographique particulière. Cette architecture, dite de privacy by design, permet également d'augmenter la sécurité dans la mesure où la confidentialité des informations personnelles bancaires est d'autant plus garantie.

Perspectives

En ce qui concerne le système de gestion des données sur Internet, nous aimerions à terme réunir l'ensemble des fonctionnalités implémentées afin d'obtenir une unique application. Celle-ci aura alors la forme d'un plugin à installer directement sur le navigateur favori de l'utilisateur. Il serait également intéressant de pouvoir ajouter des fonctionnalités à cette application, par exemple, la possibilité d'obtenir une note de confiance du site en question en vérifiant qu'il ne s'agisse pas d'un site de phishing [226].

Concernant l'infrastructure de e-santé, nous aimerions réaliser sa mise en place au sein d'un service de santé d'un hôpital afin de constater sa faisabilité et de modifier en conséquence son implémentation. Nous devons également prendre en compte les différents cas particuliers liés notamment aux remboursements des soins qui donnent, entre autre, une indication sur le rang social du patient ou encore sa pathologie. De plus, il serait intéressant de se pencher sur la question des prescriptions médicales et des informations échangées avec les pharmaciens afin de sécuriser cette chaîne médicale.

Pour finir, l'architecture de paiement en ligne pourrait être décrite de façon à pouvoir réaliser un transfert d'argent entre deux banques. Le transfert ne se ferait plus entre le client et le SP mais entre deux clients. De plus, afin de gérer différemment le chèque de banque généré lors de la transaction, nous aimerions étudier les chiffrements malléables. Une autre solution serait d'utiliser un algorithme de signature aveugle afin de faire signer le contrat, cette fois de façon non filtré, par la banque du client et ainsi obtenir une preuve de paiement pour la commande associée.

Publications de l'auteur

Revue internationale

1. P. Lacharme, **A. Plateaux**, *PIN-based cancelable biometrics*, International Journal of Automated Identification Technology (IJAIT), 2011.

Conférences internationales avec comité de lecture et avec actes

1. **A. Plateaux**, P. Lacharme, V. Coquet, S. Vernois, K. Murty, C. Rosenberger, *An e-payment architecture ensuring a high level of privacy protection*, 9th International Conference on Security and Privacy in Communication Networks (SecureComm), Sydney, Australie, Septembre 2013.

2. **A**. Plateaux, P. Lacharme, C. Rosenberger, K. Murty, *A Contactless E-health Information System with Privacy*, 9th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), Cagliari, Sardaigne, Juillet 2013.

3. **A. Plateaux**, P. Lacharme, K. Murty, C. Rosenberger, *Online user's registration respecting privacy*, World Congress on Computer and Information Technologies (WCCIT), Sousse, Tunisie, Juin 2013.

4. **A**. Plateaux, V. Coquet, S. Vernois, P. Lacharme, C. Rosenberger, K. Murty, *A Privacy Preserving E-Payment Architecture*, Financial Cryptography and Data Security (FC) - Session POSTER, Okinawa, Japon, Avril 2013.

5. J. Vincent, V. Alimi, **A**. Plateaux, C. Gaber, M. Pasquet, *A Mobile Payment Evaluation Based on a Digital Identity Representation*, Collaboration Technologies and Systems (CTS), Denver, Colorado, Etats-Unis, Mai 2012.

Conférences nationales avec comité de lecture et avec actes

1. **A. Plateaux**, P. Lacharme, V. Coquet, S. Vernois, G. Frey, A. Gouriou, *Protection de la vie privée dans les modèles de paiement en ligne*, 8ième Conférence sur la Sécurité des Architectures Réseaux et des Systèmes d'Information (SARSSI), Mont de Marsan, France, Septembre 2013.
2. **A. Plateaux**, P. Lacharme, *Organisation d'une architecture de santé respectueuse de la vie privée*, 7ième Conférence sur la Sécurité des Architectures Réseaux et des Systèmes d'Information (SARSSI), Cabourg, France, Mai 2012.

Conférences nationales avec comité de lecture et sans actes

1. P. Lacharme, K. Murty, C. Rosenberger, *Minimisation des données de e-santé*, **A. Plateaux**, 4ième Atelier Protection de la Vie Privée (APVP), Les Loges en Josas, France, Juin 2013.
2. **A. Plateaux**, P. Lacharme, C. Rosenberger, *Protection de la vie privée dans le système de paiement 3D-Secure*, 4ième Atelier Protection de la Vie Privée (APVP), Les Loges en Josas, France, Juin 2013.

Dépôt de Brevet

1. **A. Plateaux**, V. Coquet, P. Lacharme, S. Vernois, C. Rosenberger, *E-payment architecture preserving the privacy*, Société BULL SAS, Numéro du brevet : *US* 04097.

Bibliographie

[1] Les transactions Électroniques sécurisées : Un enjeu clé de la société de lŠinformation de demain, Novembre 2006. [cité p. 10, 11, 169]

[2] Privacy and identity management for europe, May 2008. [cité p. 10, 21, 59]

[3] Information technology – security techniques – entity authentication – part 1 : General, 2010. [cité p. 11]

[4] WE. Burr, DF. Dodson, WT. Polk, et al. Information security : Electronic authentication guideline, September 2004. [cité p. 11]

[5] S. Krawczyk and AK. Jain. Securing electronic medical records using biometric authentication. In *Audio-and Video-Based Biometric Person Authentication*, pages 1110–1119. Springer, 2005. [cité p. 12]

[6] Code civil : Article 9. 1994. [cité p. 12]

[7] Assemblée nationale and Sénat. *Loi 78-17 : Lois Informatiques et Libertés*, Janvier, 6 1978. [cité p. 12]

[8] Cnil : Commission nationale informatique et libertés, 1978. [cité p. 13]

[9] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre ier - principes et définitions : Article 2. [cité p. 13]

[10] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre ier - principes et définitions : Article 3. [cité p. 13]

[11] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre 5 - obligations incombant aux responsables de traitements et droits des personnes : Section 1 : Obligations incombant aux responsables de traitements : Article 32. [cité p. 14]

[12] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre 5 - obligations incombant aux responsables de traitements et droits des personnes : Section 2 : Droits des personnes à l'égard des traitements de données à caractère personnel : Article 38. [cité p. 14]

[13] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre 5 - obligations incombant aux responsables de traitements et droits des personnes : Section 2 : Droits des personnes à l'égard des traitements de données à caractère personnel : Article 39. [cité p. 14]

[14] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre 5 - obligations incombant aux responsables de traitements et droits des personnes : Section 2 : Droits des personnes à l'égard des traitements de données à caractère personnel : Article 41. [cité p. 14]

[15] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre 5 - obligations incombant aux responsables de traitements et droits des personnes : Section 2 : Droits des personnes à l'égard des traitements de données à caractère personnel : Article 42. [cité p. 14]

[16] Loi 78-17 du 6 janvier 1978 modifiée : Chapitre 5 - obligations incombant aux responsables de traitements et droits des personnes : Section 2 : Droits des personnes à l'égard des traitements de données à caractère personnel : Article 40. [cité p. 14]

[17] Le Parlement Européen. Directive 2004/18/ce du parlement europeen et du conseil du 31 mars 2004 relative à la coordination des procédures de passation des marchés publics de travaux, de fournitures et de services. *Journal officiel de l'Union européenne FR*, 50(134/114), 2004. [cité p. 14]

[18] S Corone. L'école privée, prestataire de services. *Le Monde*, Economie(560 mots), October, 2 2007. [cité p. 14]

[19] J. Vincent. *Gestion d'identité en contexte télécom*. PhD thesis, Université de Caen, 2013. [cité p. 15, 59]

[20] Décret numéro 2007-451 du 25 mars 2007 modifiant le décret numéro 2005-1309 du 20 octobre 2005 pris pour l'application de la loi numéro 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, modifiée par la loi numéro 2004-801 du 6 août 2004, Mars, 28 2007. [cité p. 15]

[21] Partie législative : Livre iii des crimes et délits contre les bien : Titre ii des autres atteintes aux biens : Chapitre iii des atteintes aux systèmes de traitement automatisé de données. [cité p. 15]

[22] MSDN Microsoft. The stride threat model, 2005. [cité p. 15]

[23] D. Mina. *Privacy Preserving Content Protection*. PhD thesis, 2010. [cité p. 15, 19, 21]

[24] LM. LoPucki. Human identification theory and the identity theft problem. *Texas Law Review*, 80 :89–134, 2001. [cité p. 16]

[25] S. Bellovin. Defending against sequence number attacks. Technical report, RFC 1948, May 1996. [cité p. 16]

[26] N. Chou, R. Ledesma, Y. Teraguchi, D. Boneh, and JC. Mitchell. Client-side defense against web-based identity theft. In *Proc. NDSS*, 2004. [cité p. 17]

[27] R. Pries, W. Yu, X. Fu, and W. Zhao. A new replay attack against anonymous communication networks. In *Communications, 2008. ICC'08. IEEE International Conference on*, pages 1578–1582. IEEE, 2008. [cité p. 17]

[28] T. Aura. Strategies against replay attacks. In *Computer Security Foundations Workshop, 1997. Proceedings., 10th*, pages 59–68. IEEE, 2002. [cité p. 17]

[29] Y. Bar-Shalom. *Tracking and data association*. Academic Press Professional, Inc. San Diego, CA, USA, 1987. [cité p. 17]

[30] A. Srivastava and A. Eustace. Atom : A system for building customized program analysis tools. In *Proceedings of the ACM SIGPLAN 1994 conference on Programming language design and implementation*, pages 196–205. ACM, 1994. [cité p. 18]

[31] SZS. Idrus, E. Cherrier, C. Rosenberger, P. Bours, et al. A preliminary study of a new soft biometric finger recognition for keystroke dynamics. In *9th Summer School for Advanced Studies on Biometrics for Secure Authentication : Understanding Man Machine Interactions in Forensics and Security Applications*, 2012. [cité p. 18]

[32] I. Mann. *Hacking the human : social engineering techniques and security countermeasures*. Gower Publishing Company, 2008. [cité p. 18]

[33] KD. Mitnick and WL. Simon. *The art of deception : Controlling the human element of security*. John Wiley & Sons, Inc. New York, NY, USA, 2003. [cité p. 18]

[34] H. Hasle, Y. Kristiansen, K. Kintel, and E. Snekkenes. Measuring resistance to social engineering. *Information Security Practice and Experience*, pages 132–143, 2005. [cité p. 18]

[35] M. Egele, C. Kruegel, E. Kirda, H. Yin, and D. Song. Dynamic spyware analysis. In *USENIX Annual Technical Conference on Proceedings of the USENIX Annual Technical Conference*, pages 1–14. USENIX Association, 2007. [cité p. 18]

[36] S. Saroiu, SD. Gribble, and HM. Levy. Measurement and analysis of spywave in a university environment. In *Proceedings of the 1st conference on Symposium on Networked Systems Design and Implementation-Volume 1*, page 11. USENIX Association, 2004. [cité p. 18]

[37] R. Romain. *Contributions à la dynamique de frappe au clavier : multibiométrie, biométrie douce et mise à jour de la référence*. PhD thesis, Université de Caen, 2012. [cité p. 18]

[38] WE. Burr, DF. Dodson, and WT. Polk. *Electronic authentication guideline : Recommendations of the National Institute of Standards and Technology*. US Dept. of Commerce, Technology Administration, National Institute of Standards and Technology, 2006. [cité p. 19]

[39] R. Anderson. A security policy model for clinical information systems. In *IEEE*, Cambridge, 1996. University of Cambridge Computer Laboratory. [cité p. 19]

[40] B. Schneier. Protecting privacy and liberty. *Nature*, 413(6858) :773, 2001. [cité p. 19]

[41] AL. Allen. Constitutional law and privacy. *A companion to philosophy of law and legal theory*, pages 139–155, 1996. [cité p. 19]

[42] Guidelines for the regulation of computerized personal data files adopted, 14 December 1990. [cité p. 19]

[43] Information technology – security techniques – code of practice for information security management, 2005. [cité p. 20]

[44] Protection of individuals with regard to the processing of personal data and on the free movement of such data, 24 October 1995. [cité p. 20]

[45] On the processing of personal data and the protection of privacy in the electronic communications sector, 2002. [cité p. 20]

[46] Human rights, 1987. [cité p. 20]

[47] Y. Deswarte and S. Gambs. Towards a privacy-preserving national identity card. *Data Privacy Management and Autonomous Spontaneous Security*, pages 48–64, 2010. [cité p. 21]

[48] *European Commission : Communication from the Commission to the european parliament, the council, the economic and social committee and the commitee of the regions*. Nov. 4, 2010. [cité p. 21]

[49] *Common Criteria for Information Technology Security Evaluation*. Department of Health, july 2009. [cité p. 21, 49, 51]

[50] A. Pfitzmann and M. Hansen. Anonymity, unlinkability, unobservability, pseudonymity, and identity management - a consolidated proposal for terminology. Technical Report, 2008. v0.31. [cité p. 21]

[51] K. Cameron. The laws of identity. *Microsoft Corp*. [cité p. 21]

[52] N. FIPS. 197 : Announcing the advanced encryption standard (aes). *Information Technology Laboratory, National Institute of Standards and Technology, Nov*, 2001. [cité p. 22, 29]

[53] Liberty alliance project whitepaper : Personal identity, March 23 2006. [cité p. 22]

[54] Idemix : pseudonymity for e-transactions, 2008. [cité p. 22, 43]

[55] D. Chaum. Security without identification : Transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10) :1030–1044, 1985. [cité p. 22]

[56] J. Camenisch and E. Van Herreweghen. Design and implementation of the idemix anonymous credential system. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, page 30. ACM, 2002. [cité p. 22]

[57] Assemblé Nationale and Sénat. Loi 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique, 2004. [cité p. 29]

[58] PUB FIPS. 46-3, data encryption standard (des). *National Institute for Standards and Technology*, 25, 1999. [cité p. 29]

[59] ANSSI. Référentiel général de sécurité : Version 1.0 : Annexe b1 : Mécanismes cryptographiques : Règles et recommandations concernant le choix et le dimensionnement des mécanismes cryptographiques, Janvier, 26 2010. [cité p. 29, 31, 34]

[60] Nist : National institute of standarts and technology. [cité p. 29]

[61] J. Daemen and V. Rijmen. *The design of Rijndael : AES–the advanced encryption standard*. Springer Verlag, 2002. [cité p. 29]

[62] M. Dworkin. Recommendation for block cipher modes of operation : Methods and techniques, 2001. [cité p. 30]

[63] M. Dworkin. Special publication 800-38b : Recommendation for block cipher modes of operation : The cmac mode for authentication, 2005. [cité p. 30]

[64] NF. Pub. Secure hash standard, 2004. [cité p. 30]

[65] G. Bertoni, J. Daemen, M. Peeters, and G. Van Assche. Keccak sponge function family main document. *Submission to NIST (Round 2)*, 3, 2009. [cité p. 30]

[66] RL. Rivest, A. Shamir, and LM. Adleman. Cryptographic communications system and method, September 20 1983. US Patent 4,405,829. [cité p. 31]

[67] K. Aoki, J. Franke, T. Kleinjung, A. Lenstra, and D. Osvik. A kilobit special number field sieve factorization. *Advances in Cryptology–ASIACRYPT 2007*, pages 1–12, 2008. [cité p. 31]

[68] M. Fellows and N. Koblitz. Combinatorial cryptosystems galore. *Finite Fields : Theory, Applications, and Algorithms*, 168 :51–61. [cité p. 31]

[69] B. Buchberger and F. Winkler. *Gröbner bases and applications*. Cambridge Univ Pr, 1998. [cité p. 31]

[70] T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *Information Theory, IEEE Transactions on*, 31(4) :469–472, 2002. [cité p. 32]

[71] S. Pohlig and M. Hellman. An improved algorithm for computing logarithms over gf(p) and its cryptographic significance. *Information Theory, IEEE Transactions on*, 24(1) :106–110, 1978. [cité p. 32]

[72] V. Miller. Use of elliptic curves in cryptography. In *Advances in Cryptology-CRYPTO'85 Proceedings*, pages 417–426. Springer, 1985. [cité p. 32]

[73] N. Koblitz. Elliptic curve cryptosystems. *Mathematics of computation*, 48(177) :203–209, 1987. [cité p. 32]

[74] J. Stern. La cryptologie : des messages secrets aux transactions sécurisées. 40ème anniversaire du LAAS Toulouse, 20 juin 2008. [cité p. 32]

[75] P. Rogawayy. Optimal asymmetric encryption how to encrypt with rsa. 1995. [cité p. 32]

[76] H. Feistel. *Cryptography and computer privacy*. Scientific American, 1973. [cité p. 32]

[77] M. Hellman. New directions in cryptography. *IEEE transactions on Information Theory*, 22(6) :644–654, 1976. [cité p. 33]

[78] E. Caprioli. Ecrit et preuve électroniques dans la loi n 2000-230 du 13 mars 2000, 2000. [cité p. 33]

[79] I. de Lamberterie and JF. Blanchette. Le décret du 30 mars 2001 relatif à la signature électronique : lecture critique, technique et juridique. *JCP Entreprise et Affaires*, (30-26), 2001. [cité p. 33]

[80] Information technology – security techniques – digital signature schemes giving message recovery – part 2 : Integer factorization based mechanisms, 2010. [cité p. 34]

[81] Information technology – security techniques – digital signature schemes giving message recovery – part 3 : Discrete logarithm based mechanisms, 2006. [cité p. 34]

[82] P. Gallagher and CF. Director. Fips pub 186-3 federal information processing standards publication digital signature standard (dss). 2009. [cité p. 34]

[83] C. Schnorr. Efficient identification and signatures for smart cards. In *Advances in Cryptology-Crypto'89 Proceedings*, pages 239–252. Springer, 1990. [cité p. 34]

[84] Recommended elliptic curves for federal government use, July 1999. [cité p. 34]

[85] K. Nyberg and RA. Rueppel. Message recovery for signature schemes based on the discrete logarithm problem. *Designs, Codes and Cryptography*, 7(1) :61–81, 1996. [cité p. 34]

[86] M. Abe and T. Okamoto. A signature scheme with message recovery as secure as discrete logarithm. *Advances in Cryptology-ASIACRYPT'99*, pages 378–389, 2004. [cité p. 34]

[87] R. Mukkamamla and M. Halappanavar. Ecpv : Efficient certificate path validation in public-key infrastructure. In *Proceedings of 17th IFIP WG11*, volume 3. [cité p. 34]

[88] J. Jonsson et al. Public-key cryptography standards (pkcs)# 1 : Rsa cryptography specification version 2.1, 2003. [cité p. 34]

[89] Information technology : Open systems interconnection : The directory : Public-key and attribute certificate frameworks, 2008. [cité p. 35]

[90] LM. Kohnfelder. *Towards a practical public-key cryptosystem*. PhD thesis, Massachusetts Institute of Technology laboratory, 1978. [cité p. 35]

[91] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof-systems. In *Proceedings of the seventeenth annual ACM symposium on Theory of computing*, pages 291–304. ACM, 1985. [cité p. 36]

[92] D. Chaum. Zero-knowledge undeniable signatures. In *Proceedings of the workshop on the theory and application of cryptographic techniques on Advances in cryptology*, pages 458–464, 1991. [cité p. 36]

[93] Iso/iec 97/98-5 information technology – security techniques – entity authentication – part 5 : Mechanisms using zero-knowledge techniques. 2009. [cité p. 36]

[94] U. Feige, A. Fiat, and A. Shamir. Zero-knowledge proofs of identity. *Journal of cryptology*, 1(2) :77–94, 1988. [cité p. 37]

[95] A. Shamir and A. Fiat. Method, apparatus and article for identification and signature, May 31 1988. US Patent 4,748,668. [cité p. 37]

[96] A. Fiat and A. Shamir. How to prove yourself : Practical solutions to identification and signature problems. In *Advances in Cryptology-Crypto'86*, pages 186–194. Springer, 1987. [cité p. 37]

[97] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity or all languages in np have zero-knowledge proof systems. *Journal of the ACM (JACM)*, 38(3) :690–728, 1991. [cité p. 37]

[98] CP. Schnorr. Efficient signature generation by smart cards. *Journal of cryptology*, 4(3) :161–174, 1991. [cité p. 38]

[99] R. Cramer, I. Damgard, and B. Schoenmakers. Proof of partial knowledge and simplified designe of witness hiding protocols. *CRYPTO'94*, 1994. [cité p. 38]

[100] E. Hufschmitt. *Signatures pour l'anonymat fondées sur les couplages et applications*. PhD thesis, Université de Caen-Basse-Normandie, Spécialité Informatique, November 2007. [cité p. 38, 39, 42]

[101] E. Brickell, D. Chaum, I. Damgård, and J. van de Graaf. Gradual and verifiable release of a secret. In *Advances in Cryptology-CRYPTO'87*, pages 156–166. Springer, 2006. [cité p. 39]

[102] A. Chan, Y. Frankel, and Y. Tsiounis. Easy come-easy go divisible cash. *Advances in Cryptology-EUROCRYPT'98*, pages 561–575, 1998. [cité p. 39]

[103] F. Boudot. Efficient proofs that a committed number lies in an interval. In *Advances in Cryptology-EUROCRYPT 2000*, pages 431–444. Springer, 2000. [cité p. 39, 63]

[104] F. Boudot, B. Schoenmakers, and J. Traore. A fair and efficient solution to the socialist millionaires' problem. *Discrete Applied Mathematics*, 111(1-2) :23–36, 2001. [cité p. 39]

[105] D. Chaum. Blind signatures for untraceable payments. In *Advances in Cryptology : Proceedings of Crypto*, volume 82, pages 199–203, 1983. [cité p. 40, 41]

[106] C. Schnorr. Security of blind discrete log signatures against interactive attacks. *Information and Communications Security*, pages 1–12, 2001. [cité p. 40]

[107] D. Pointcheval and J. Stern. Security arguments for digital signatures and blind signatures. *Journal of cryptology*, 13(3) :361–396, 2000. [cité p. 40]

[108] JC. Pailles. Mobile transactions : trust and privacy aspects. In *C&ESAR 2008*. Orange Labs, Orange FT Group, 2008. [cité p. 41]

[109] D. Chaum and E. Van Heyst. Group signatures. In *Proceedings of the 10th annual international conference on Theory and application of cryptographic techniques*, pages 257–265. Springer-Verlag, 1991. [cité p. 41]

[110] J. Camenisch and M. Stadler. Efficient group signature schemes for large groups. *Advances in Cryptology-CRYPTO'97*, pages 410–424, 1997. [cité p. 42]

[111] D. Chaum. Blind signatures for untraceable payments. In *Crypto*, volume 82, pages 199–203, 1982. [cité p. 42]

[112] D. Chaum. Security without identification : Transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10) :1030–1044, 1985. [cité p. 42]

[113] J. Camenisch and A. Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Advances in Cryptology-EUROCRYPT 2001*, pages 93–118. Springer, 2001. [cité p. 43]

[114] S. Brands and C. Paquin. U-prove cryptographic specification v1. 0. Technical report, 0. Tech. rep., Microsoft Corporation (March 2010), 2010. [cité p. 43, 59]

[115] C. Paquin. U-prove technology overview v1. 2011. [cité p. 43, 59]

[116] Stefan A. Brands. *Rethinking Public Key Infrastructures and digital certificates : builgind in privacy*. The MIT Press, 2000. [cité p. 43]

[117] A. Shamir. How to share a secret. *Communications of the ACM 22*, pages 612–613, 1979. [cité p. 43, 89]

[118] M. Blaze, G. Bleumer, and M. Strauss. Divertible protocols and atomic proxy cryptography. In *Advances in Cryptology-EUROCRYPT'98*, pages 127–144. Springer, 1998. [cité p. 44]

[119] G. Ateniese, K. Fu, M. Green, and S. Hohenberger. Improved proxy re-encryption schemes with applications to secure distributed storage. *ACM Transactions on Information and System Security (TISSEC)*, 9(1) :1–30, 2006. [cité p. 44]

[120] Iso/iec jtc 1/sc 37 biométrie, 2002. [cité p. 46]

[121] AK. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP Journal on Advances in Signal Processing*, 2008 :113, 2008. [cité p. 46, 169]

[122] N. Ratha, J. Connell, and RM. Bolle. Enhancing security and privacy in biometrics-biased authentication systems. *IBM syst*, 40(3) :614–634, 2001. [cité p. 46]

[123] BJA. Teoh and CLD. Ngo. Cancellable biometrics featuring with tokenised random number. *Pattern recognition Letters*, 26(10) :1454–1460, 2005. [cité p. 46]

[124] A. Nagar, K. Nandakumar, and AK. Jain. Biometric template transformation : A security analysis. In *Media Forensics and Security*, 2010. [cité p. 46]

[125] N. Ratha, S. Chikkerur, J. Connell, and RM. Bolle. Generating cancelable fingerprint templates. *IEEE Transactions on pattern analysis and machine intelligence*, 29(4) :561–572, 2007. [cité p. 46]

[126] YS. Kim, ABJ. Teoh, and KO. Toh. A performance driven methodology for cancelable face templates generation. *Pattern recognition*, 43(7) :2544–2559, 2010. [cité p. 46]

[127] A. Teoh, D. Ngo, and A. Goh. Random multispace quantisation as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Trans. Pattern Anal Mach. Intell*, 28(12) :1892–1901, 2006. [cité p. 47]

[128] AK. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP J. Advances in signal processing*, 8(2) :1–17, 2008. [cité p. 47]

[129] Identification cards – integrated circuit(s) cards with contacts – part 1 : Physical characteristics, 1998. [cité p. 48]

[130] Identification cards – integrated circuit cards – part 2 : Cards with contacts – dimensions and location of the contacts, 2007. [cité p. 48]

[131] Identification cards – integrated circuit cards – part 3 : Cards with contacts – electrical interface and transmission protocols, 2006. [cité p. 48]

[132] Identification cards – integrated circuit cards – part 5 : Registration of application providers, 2004. [cité p. 48]

[133] Identification cards – integrated circuit cards – part 6 : Interindustry data elements for interchange, 2004. [cité p. 48]

[134] A. Hindle and DM. German. Scql : a formal model and a query language for source control repositories. *ACM SIGSOFT Software Engineering Notes*, 30(4) :1–5, 2005. [cité p. 48]

[135] Identification cards – integrated circuit(s) cards with contacts – part 7 : Interindustry commands for structured card query language (scql), 1999. [cité p. 48]

[136] Identification cards – integrated circuit cards – part 4 : Organization, security and commands for interchange, 2005. [cité p. 48]

[137] The iso 7816 smart card standard : Overview. [cité p. 48]

[138] Iso/iec 14443 identification cards – contactless integrated circuit cards – proximity cards. 2008. [cité p. 48]

[139] MasterCard International. Chip authentication program functional architecture, Sept., 2004. [cité p. 49, 108]

[140] K. Gandolfi, C. Mourtel, and F. Olivier. Electromagnetic analysis : Concrete results. In *Cryptographic Hardware and Embedded Systems-CHES 2001*, pages 251–261. Springer, 2001. [cité p. 49]

[141] JJ. Quisquater and D. Samyde. Electromagnetic analysis (ema) : Measures and countermeasures for smart cards. *Smart Card Programming and Security*, pages 200–210, 2001. [cité p. 49]

[142] P. Kocher. Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems. In *Advances in Cryptology-CRYPTO'96*, pages 104–113. Springer, 1996. [cité p. 49]

[143] M. Hendry. *Smart card security and applications*. Artech House Publishers, 2001. [cité p. 49]

[144] Atmel Smart Card ICs, Hitachi Europe Ltd., Infineon Technologies AG, and Philips Semiconductors. Platform protection profile, July 2001. [cité p. 49]

[145] Technical Commitee. Sd specification : Physical layer simplified specification version 4.10. *SD Card Association*, 2013. [cité p. 50]

[146] PUB FIPS. 140-1 : Security requirements for cryptographic modules, January, 11 1994. [cité p. 51]

[147] K. Nishilura, S. Ishikawa, K. Hirota, H. Aburatani, M. Hirose, T. Yoshihisa, M. Tsukamoto, S. Nishio, K. Shimizu, and T. Miura. Technologies de l'information – techniques de sécurité – critères d'évaluation pour la sécurité ti, 2005. [cité p. 51]

[148] Pkcs-11 - crypto api : Guide de programmation, Septembre 2007. [cité p. 51]

[149] LLC EMVCo. Integrated circuit card, specifications for payment systems. *EMV2000, Dec*, 2000. [cité p. 51]

[150] 3-d secure protocol specification - core functions, July 16, 2002. [cité p. 51, 108, 115]

[151] Bull SAS. Crypt2pay : Security services. 2009. [cité p. 52, 169]

[152] Union internationale des télécommunications. Itu - free statistics. 2011. [cité p. 56]

[153] Federal Trade Commission. Coppa – children's online privacy protection act. 2002. [cité p. 56]

[154] JC. Cuaresma. Gramm-leach-bliley act, the. *Berkeley Tech. LJ*, 17 :497, 2002. [cité p. 56]

[155] On the protection of individuals with regards to the processing of personal data and on the free movement of such data, 1995. [cité p. 56]

[156] Privacy Commissioner of Canada. The personal information protection and electronic documents act. *Available at SSRN 1403922*, 2009. [cité p. 56]

[157] R. Wenning, M. Schunter, L. Cranor, M. Marchiori, et al. The platform for privacy preferences 1.1 (p3p1. 1) specification. *W3C Working Group Note*, 2006. [cité p. 57]

[158] IBM. Tivoli security products. 1996. [cité p. 58]

[159] Project diaspora. 2010. [cité p. 58]

[160] S. Buchegger, D. Schiöberg, LH. Vu, and A. Datta. Peerson : P2p social networking : early experiences and insights. In *Proceedings of the Second ACM EuroSys Workshop on Social Network Systems*, pages 46–52. ACM, 2009. [cité p. 58]

[161] LA. Cutillo, R. Molva, and T. Strufe. Safebook : Feasibility of transitive cooperation for privacy on a decentralized social network. In *World of Wireless, Mobile and Multimedia Networks & Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a*, pages 1–6. IEEE, 2009. [cité p. 58]

[162] S. Jahid, S. Nilizadeh, P. Mittal, N. Borisov, and A. Kapadia. Decent : A decentralized architecture for enforcing privacy in online social networks. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 326–332. IEEE, 2012. [cité p. 58]

[163] European Academic institutions and Companies. Future of identity in the information society, 2009. [cité p. 59]

[164] Picos - privacy and identity management for community services, 2008. [cité p. 59]

[165] Abc4trust, 2010. [cité p. 59]

[166] D. Recordon and D. Reed. Openid 2.0 : a platform for user-centric identity management. In *Proceedings of the second ACM workshop on Digital identity management*, pages 11–16. ACM, 2006. [cité p. 59]

[167] TrueCrypt Foundation. Truecrypt : Free open-dource on-the-fly encryption. 1997. [cité p. 61]

[168] A. Freier, P. Karlton, and P. Kocher. The secure sockets layer (ssl) protocol version 3.0. 2011. [cité p. 67]

[169] T. Dierks. The transport layer security (tls) protocol version 1.2. 2008. [cité p. 67]

[170] D. Wagner and B. Schneier. Analysis of the ssl 3.0 protocol. In *The Second USENIX Workshop on Electronic Commerce Proceedings*, pages 29–40, 1996. [cité p. 67]

[171] E. Gabrilovich and A. Gontmakher. The homograph attack. *Communications of the ACM*, 45(2) :128, 2002. [cité p. 67]

[172] O. Aciiçmez, W. Schindler, and CK. Koç. Improving brumley and boneh timing attack on unprotected ssl implementations. In *Proceedings of the 12th ACM conference on Computer and communications security*, pages 139–146. ACM, 2005. [cité p. 67]

[173] L. Bygrave. Germanys teleservices data protection act. privacy law and policy reporter. 1998. [cité p. 67]

[174] PCI DSS. Payment card industry data security standard, 2006. https ://www.pcisecuritystandards.org/. [cité p. 68, 108]

[175] Health insurance portability and accountability act. hipaa administrative simplification : enforcement ; final rule, 2006. [cité p. 76]

[176] Recommendation of council of europe n. r(97)5 on the protection of medical data, February 1997. [cité p. 76]

[177] La loi n 2002-303 du 4 mars 2002, relative aux droits des malades et à la qualité du système de santé. *Parue au Journal officiel de la République Française*, 2002. [cité p. 76]

[178] Code de la santé publique, code. 2000. [cité p. 76]

[179] RJ. Anderson. A security policy model for clinical information systems. In *Security and Privacy, 1996*, pages 30–43, Cambridge, 2002. University of Cambridge Computer Laboratory, IEEE. [cité p. 77]

[180] European privacy and human rights (ephr), 2010. [cité p. 77]

[181] R. Anderson. Under threat : patient confidentiality and nhs computing. *Drugs and Alcohol Today*, 6(4) :13–17, 2006. [cité p. 77, 78]

[182] R. Anderson. Patient confidentiality and central databases. *Br J Gen Pract*, 58(547) :75–76, 2008. [cité p. 77, 78]

[183] Medix uk plc, November 2007. [cité p. 78]

[184] F. Caldicott and G. Britain. *Report on the review of patient-identifiable information*. Department of Health, 1997. [cité p. 78]

[185] Loi. 810 du 13 août 2004 relative à l'assurance maladie. *Parue au Journal officiel de la République Française (JORF) le*, 17, 2004. [cité p. 80]

[186] CNIL. Délibération numéro 2010-449 du 2 décembre 2010 portant autorisation des traitements de données personnelles mis en IJuvre par les professionnels et établissements de santé nécessaires à la première phase de déploiement généralisé du dossier médical personnel. 2010. [cité p. 80]

[187] Cours des Comptes. Rapport public annuel : Les téléservices publics de santé. 2013. [cité p. 80]

[188] C. Quantin, G. Coatrieux, M. Fassa, V. Breton, DO. Jaquet-Chiffelle, P. De Vlieger, N. Lypszyc, JY. Boire, C. Roux, and FA. Allaert. Centralised versus decentralised management of patientsŠ medical records. In IOS Press, editor, *Medical Informatics in a United and Healthy Europe K.-P. Adlassnig et al. (Eds.)*, 2009. [cité p. 81]

[189] D. Ghindici. *Information flow analysis for embedded systems : from practical to theoretical aspects*. PhD thesis, INRIA, Sophia-Antipolis et Univ. Laval, Canada, 2008. [cité p. 81, 88]

[190] M. Deng, D. De Cock, and B. Preneel. Towards a cross-context identity management framework in e-health. *Online Information Review*, 33(3) :422–442, 2009. [cité p. 81]

[191] M. Deng, R. Scandariato, D. De Cock, B. Preneel, and W. Joosen. Identity in federated electronic healthcare. In *Wireless Days, 2008. WD'08. 1st IFIP*, pages 1–5. IEEE, 2009. [cité p. 81]

[192] G. Ateniese and B. De Medeiros. Anonymous e-prescriptions. pages 19–31. ACM, 2002. [cité p. 81]

[193] B. De Decker, M. Layouni, H. Vangheluwe, and Verslype K. Anonymous e-prescriptions. pages 118–133. Public Key Infrastructure, 2008. [cité p. 81]

[194] FEVAD Fédération e-commerce et vente à distance. Chiffres clés 2012, vente à distance et e-commerce aux particuliers, 2012. http ://www.fevad.com/. [cité p. 106]

[195] Y. Espelid, LH. Netland, A. Klingsheim, and K. Hole. A proof of concept attack against norwegian internet banking systems. *Financial Cryptography*, pages 197–201, 2008. [cité p. 106]

[196] Observatoire de la sécurité des cartes de paiement. Rapport annuel 2011 de l'observatoire de la sécurité des cartes de paiement, 2011. [cité p. 106]

[197] European Commission. Directive 2000/31/EC of the european parliament and of the council of 8 june 2000 on certain legal aspects of information society services, in particular electronic commerce, in the internal market ('directive on electronic commerce'), 2000. [cité p. 106]

[198] European Commission. Directive 2007/64/EC of the european parliament and of the council of 13 november 2007 on payment services in the internal market, 2007. [cité p. 106]

[199] European Payments Council. Single euro payment area, 2007. http ://www.sepafrance.fr/. [cité p. 106]

[200] S. Katsikas, J. Lopez, and G. Pernul. Trust, privacy and security in e-business : Requirements and solutions. *Advances in Informatics*, pages 548–558, 2005. [cité p. 106]

[201] Atos Worldline. Sips e-payment, solution de paiement sécurisé, 2002. [cité p. 108]

[202] Paybox : solution de paiement sécurisé du e-commerce, 1999. [cité p. 108]

[203] S.E.T. Secure electronic transaction specification. *Book 1 : Business Description*, 2002. [cité p. 108, 113]

[204] S. Murdoch and R. Anderson. Verified by visa and mastercard securecode : or, how not to design authentication. *Financial Cryptography and Data Security*, pages 336–342, 2010. [cité p. 108, 115, 116]

[205] S. Drimer, S. Murdoch, and R. Anderson. Optimised to fail : Card readers for online banking. *Financial Cryptography and Data Security*, pages 184–200, 2009. [cité p. 108]

[206] C. Meadows and P. Syverson. A formal specification of requirements for payment transactions in the SET protocol. In *Proceedings of Financial Cryptography and Data Sceurity*, 1998. [cité p. 108]

[207] G. Bella, F. Massacci, L. Paulson, and P. Tramontano. Formal verification of cardholder registration in SET. *Computer Security - ESORICS 2000*, pages 159–174, 2000. [cité p. 108, 114]

[208] S. Bella, L. Paulson, and F. Massacci. The verification of an industrial payment protocol : The SET purchase phase. In *Proceedings of ACM CCS*, pages 12–20. ACM, 2002. [cité p. 108, 114]

[209] S. Brlek, S. Hamadou, and J. Mullins. A flaw in the electronic commerce protocol set. *Information Processing Letters*, 97(3) :104–108, 2006. [cité p. 108, 114]

[210] V. Pasupathinathan, J. Pieprzyk, H. Wang, and JY. Cho. Formal analysis of card-based payment systems in mobile devices. In *Proceedings of the 2006 Australasian workshops on Grid computing and e-research - Volume 54*, pages 213–220, 2006. [cité p. 108, 115]

[211] M. Ashrafi and S. Ng. Enabling privacy-preserving e-payment processing. In *Database Systems for Advanced Applications*, pages 596–603. Springer, 2008. [cité p. 108, 111, 112, 117]

[212] G. Antoniou and L. Batten. E-commerce : protecting purchaser privacy to enforce trust. *Electronic commerce research*, 11(4) :421–456, 2011. [cité p. 108, 109, 112]

[213] A. Freier, P. Kocher, and P. Karlton. RFC 6101 : The secure sockets layer (SSL) protocol version 3.0, 2011. [cité p. 111]

[214] T. Dierks. RFC 5246 : The transport layer security (TLS) protocol version 1.2, 2008. [cité p. 111]

[215] Paypal : Achetez, vendez et envoyez de l'argent en ligne, 1988. [cité p. 112]

[216] Paypal. Privacy policy for paypal services, 2012. [cité p. 112]

[217] D. Chaum, A. Fiat, and M. Naor. Untraceable electronic cash. In *Advances in Cryptology CRYPTO'88*, pages 319–327. Springer, 1990. [cité p. 112]

[218] M. Carbonell, J. Torres, A. Izquierdo, and D. Suarez. New e-payment scenarios in an extended version of the traditional model. *Computational Science and Its Applications–ICCSA 2008*, pages 514–525, 2008. [cité p. 112]

[219] M. Pasquet, C. Rosenberger, F. Cuozzo, et al. Security for electronic commerce. *Encyclopedia of Information Science and Technology,*, 4 :14, 2008. [cité p. 113, 170]

[220] G. Bella, F. Massacci, and L. Paulson. Verifying the SET purchase protocols. *Journal of Automated Reasoning*, 36(1) :5–37, 2006. [cité p. 114]

[221] A. Fioravanti and F. Massacci. How to model (and simplify) the SET payment phase for automated verification. In *IJCAR'01*, 2001. [cité p. 114]

[222] M. Anderson. The electronic check architecture. *Financial Services Technology Consortium*, 1998. [cité p. 123]

[223] TH. Chen, SC. Yeh, KC. Liao, and WB. Lee. A practical and efficient electronic checkbook. *Journal of Organizational Computing and Electronic Commerce*, 19(4) :285–293, 2009. [cité p. 123]

[224] V. Pasupathinathan, J. Pieprzyk, and H. Wang. Privacy enhanced electronic cheque system. In *E-Commerce Technology, 2005. CEC 2005. Seventh IEEE International Conference on*, pages 431–434. IEEE, 2005. [cité p. 123]

[225] Y. Shah, F. Vanbever, and G. Leibbrandt. Society for worldwide interbank financial telecommunication. 1977. [cité p. 125]

[226] S. Gastellier-Prevost, GG. Granadillo, and M. Laurent. Decisive heuristics to differentiate legitimate from phishing sites. In *Network and Information Systems Security (SAR-SSI), 2011 Conference on*, pages 1–9. IEEE, 2011. [cité p. 140]

Annexe

Annexe A

Étude : Paiement sur Internet et vie privée

Etude : Payer sur internet en préservant votre vie privée.

1.1- Votre tranche d'âge ?

10-15 16-20 21-25 26-30 30-35 36-40
41-45 46-50 51-55 56-60 +60

1.2- Votre sexe ? Homme Femme

2- Faites-vous des achats sur internet ? Jamais Rarement Parfois Régulièrement

3- Vous sentez vous concerné par les problèmes de protection de votre vie privée sur internet ?

Oui Non

4.1- Avez-vous des appréhensions quand vous faites un achat sur internet ? Oui Non

4.2- Si oui, pourquoi ?

Le site n'est peut être pas sécurisé
Peur de donner mes coordonnées bancaires
Peur qu'on conserve certaines informations personnelles (bancaire, nom, service…)
Peur qu'on vole certaines informations personnelles (bancaire, nom, service…)
Pas confiance aux sites
Peur de ne pas recevoir son achat
Autres : ……..
……..

5.1- Savez-vous ouvrir un fichier/document sur votre ordinateur ? Oui Non

5.2- Savez-vous enregistrer un fichier/document sur votre ordinateur ? Oui Non

5.3- Savez-vous mettre un document en pièce jointe ? Oui Non

5.4- Si non, pour (au moins) une des questions 5.1, 5.2, 5.3 :

Seriez-vous prêt à apprendre s'il faut moins d'une minute et si cela vous permet de protéger vos informations personnelles lors du paiement ? Oui Non

6.1- Savez-vous ce qu'est un certificat numérique ? Oui Non

Il s'agit d'une sorte de carte d'identité numérique. Il est utilisé principalement pour identifier une personne ou entité (site…) de façon sûre et pour chiffrer des échanges entres entités.

6.2- Aimeriez-vous apprendre à les utiliser ? Oui Non

7- Si vous deviez faire quelques manipulations lors d'un paiement sur internet (comme enregistrer un fichier puis le mettre en pièce jointe) et si cela vous permettait de faire votre achat sur internet de façon sécurisée sans fournir aucune de vos informations personnelles, ni aucunes de vos coordonnées bancaires au marchant, seriez-vous prêt à le faire ?

Oui Non

8- Remarques sur le problème de protection de la vie privée lors d'un achat sur Internet ?

……
……
……

FIGURE A.1 – Questionnaire globale

Liste des abréviations

ACS	Access Control Server
AES	Advanced Encryption Standard
ANSSI	Agence Nationale de la Sécurité des Systèmes d'Information
API	Application Programming Interface
CA	Autorité de Certification
CAA	Contrôleur d'Accès Médical
CAM	Contrôleur d'Accès Administratif
CAP	Chip Authentification Program
CBC	Cipher-Block Chaining
CE	Courbes Elliptiques
CNIL	Commission nationale de l'informatique et des libertés
*CVX*2	Cryptogram Value
DES	Data Encryption Standard
DMP	Dossier Médical Patient
DPA	Dynamic Passcode Authentication
DSA	Digital Signature Algorithm
DSS	Digital Signature Standart
DoS	Denial of Service
EAL	Evaluation Assurance Level
ECC	Elliptic Curve Cryptography
EC − *DSA*	Elliptic Curve Digital Signature Algorithm
EMV	Europay Mastercard Visa
FIDIS	Future of Identity in the Information Society
FIPS	Federal Information Processing Standards
FPGA	Field-Programmable Gate Array
GId	Global Identifier (Identifiant Global)
HSM	Hardware Security Module
IB	Informations Bancaire

IC	Informations de Commande
IdG	Identifiant Global
IdL	Identifiant Local
IEC	International Electrotechnical Commission
INS	Identifiant National de Santé
IP	Internet Protocol
ISO	International Organization for Standardization
LId	Local Identifier (Identifiant Local)
LIL	Loi Informatique et Libertés
	Detectability, Divulgation, Unawareness, Non-conformity
MAC	Message Authentification Code
MPI	Merchant Plug-In
NFC	Near Field Communication
NIST	National Institute of Standards and Technology)
OTP	One Time Password
P3P	Privacy Preferences Project
PAN	Primary Account Number
PCI − DSS	PCI Data Security Standard
PET	Privacy Enhancing Technologies
PIN	Personal Identification Number
PKI	Public Key Infrastructure
PRE	Proxy Re-Encryption
PRIME	PRivacy and Identity Management for Europe
RFC	Request For Comments
RFID	Radio Frequency IDentification
ROM	Real-Only Memory
RSA	Rivest Shamie Adleman
SC	Sous-Comités
SD	Secure Digital
SDI	Service de Dérivation d'Identité
SE	Secure Element
SET	Secure Electronic Transaction
SHA	Secure Hash Algorithm
SI	Système Interbancaire
SIM	Subscriber Identity Module
SP	Service Provider (ou Fournisseur de Service)
SSL	Secure Sockets Layers
	Denial of Service et Elevation of privilege
SWIFT	Society for Worldwide Interbank Financial Telecommunication
TES	Transactions Électroniques Sécurisées

TLS	Transport Layer Security
USB	Universal Serial Bus
WG	Work Group
ZK	Zero-Knowledge (ou à divulgation nulle de connnaissance)

Table des figures

Liste des tableaux

Avec l'utilisation de notre carte bancaire pour payer un achat sur Internet ou de notre téléphone portable pour nous connecter aux réseaux sociaux, les transactions électroniques font partie de notre quotidien et sont désormais incontournables. Malheureusement, lors de tels échanges, un grand nombre de données personnelles sont transférées et une telle informatisation n'est pas sans conséquence. Les problèmes de sécurisation et de protection de ces données sont bien présents.

Dans cette thèse, nous nous concentrons sur la problématique de la protection de la vie privée des utilisateurs dans des systèmes informatiques. Pour cela, nous nous intéressons à trois domaines d'actualité. Dans un premier temps, nous proposons un système de gestion des données centré sur l'utilisateur. Ainsi, lors de sa navigation sur Internet, l'internaute sera guidé et aura la possibilité de faire appel aux huit fonctionnalités offertes par l'application. Un second problème, sur lequel nous avons travaillé, est le cas des dossiers médicaux des patients et de l'accès à ces documents confidentiels. Nous proposons une architecture de e-santé permettant la protection des informations personnelles des patients au sein d'un établissement de santé et entre plusieurs établissements. Pour finir, nous avons travaillé dans le domaine de la monétique et plus précisément sur le paiement en ligne. Nous exposons ainsi trois nouveaux protocoles respectant davantage les données personnelles des internautes. Deux d'entre eux sont une amélioration de protocoles existants : 3D-Secure et le protocole d'Ashrafi et Ng. La dernière architecture, totalement nouvelle, permet de procéder à un paiement sur Internet sans fournir aucune information bancaire du client. Pour chacune de ces infrastructures, des exigences de sécurité et de protection de la vie privée sont décrites. Les solutions existantes, ainsi que celles proposées, sont détaillées et analysées en fonction de ces exigences. Les propositions d'architectures respectueuses de la vie privée ont toutes fait l'objet d'une preuve de concept avec une implémentation logicielle.

Operational solutions for secure electronic transactions ensuring the privacy.

By using one's credit card to make a purchase on the Internet or one's mobile phone to connect to social networks, electronic transactions have become part of one's daily routine, in a seemingly inescapable fashion. Unfortunately, these exchanges involve the transfer of a large amount of personal data. Such computerization is not without consequence. The issues of security and privacy protection are truly present.

In this thesis, we address the following issue : how to protect one's personal data in computer systems, focusing on three topical subjects. First, we propose a data management system centered on the user. Thus, when the user browses on the Internet, he/she will be guided and have the opportunity to refer to any of the eight features of the application. The second area deals with the managing of the patient's medical records and access control. We propose an e-health architecture in order to ensure the protection of the patient's personal data both within a health establishment and between separate institutions. Finally, we are interested in the field of electronic banking, and more specifically, online payment. We have suggested three new e-payment protocols ensuring the client's privacy. The first two protocols improve existing ones : 3D -Secure, Ashrafi and Ng. The last and completely new architecture allows to pay on the Internet without disclosing any of the user's banking information. With each of these architectures, come security and privacy requirements. The analysis of existing solutions and new propositions are carried out in accordance with these security requirements. Each architecture presented here ensures privacy and comes with a software proof of concept.

Indexation Rameau : Transaction électronique sécurisée, vie privée, sécurité des systèmes informatiques, Protection de l'information, Cryptographie. Indexation libre : Protection de la vie privée, E-santé, Paiement en ligne.

Spécialité Informatique et Applications

Laboratoire GREYC - UMR CNRS 6072 - Université de Caen Basse-Normandie - Ensicaen
6 Boulevard du Maréchal Juin - 14050 CAEN CEDEX