Organisation d'une architecture de santé respectueuse de la vie privée

Aude Plateaux* † ‡ §, Patrick Lacharme * † ‡

*Université de Caen Basse-Normandie, UMR 6072 GREYC, F-14032 Caen, France
†ENSICAEN, UMR 6072 GREYC, F-14050, Caen, France
‡CNRS, UMR 6072 GREYC, F-14032, Caen, France
§BULL SAS - Unité Monétique et PKI, Rue Jean Jaurès, 78340 LES CLAYES SOUS BOIS, France
aude.plateaux@ensicaen.fr, patrick.lacharme@ensicaen.fr

***Résumé*—De nombreuses infrastructures gérant les problèmes relatifs aux dossiers médicaux émergent dans plusieurs pays. La grande quantité des données stockées dans ces dossiers, ainsi que la sensibilité des informations traitées, donnent lieu à de nombreuses activités de normalisation et méritent donc une attention particulière. Cependant, la protection de la vie privée dans de tels systèmes n'est que partiellement traitée et les concepts de minimisation et de souveraineté des données sont souvent négligés.**

Cet article présente une infrastructure 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.

Mots clés : e-santé ; vie privée ; gestion d'identité ; dossiers médicaux ; chiffrement de bases de données.

I. Introduction

Les données personnelles sont particulièrement exposées depuis le développement des nouvelles technologies comme Internet, les réseaux sociaux et plus généralement avec l'augmentation des bases de données sensibles. De nombreux règlements tentent d'assurer la sécurité des systèmes d'information et la protection de la vie privée de l'utilisateur. Ainsi, la résolution 45/95, adoptée par l'Assemblée générale des Nations Unies [bGA90] dès 1990 présente plusieurs principes relatifs aux fichiers personnels informatisés, dont le principe de sécurité. Il décrit les mesures devant être prises pour protéger les fichiers contre les risques naturels et humains. 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) **Principe de la sensibilité des données** : les données personnelles traitées sont considérées comme sensibles et nécessitent une structure décentralisée pour leur stockage.
2) **Principe de souveraineté des données** : les données personnelles appartiennent à un particulier, avec un contrôle et une autorisation sur leurs utilisations et leurs finalités.
3) **Principe de minimisation des données** : la divulgation de données personnelles doit être limitée à des données adéquates, pertinentes et non excessives. Il comprend l'anonymat et l'intraçabilité des données.

La Commission européenne a examiné récemment les moyens de renforcer le principe de minimisation des données (novembre 2010, [ec10]). D'un point de vue juridique, plusieurs règlements concernant la protection des données personnelles et la confidentialité de l'utilisateur, tels que les directives de l'Union européenne (1995, [EUd95], 2002, [eu002]), ou l'article 8 de la Convention européenne [eu887] de sauvegarde des droits de l'homme et des libertés fondamentales, ont été créés.

Le cas des dossiers médicaux est particulièrement sensible au problème de protection de données. Effectivement, ils constituent de grandes bibliothèques d'information personnelles et sensibles où de nombreux acteurs (médecins, infirmières, patients ...) entrent en jeu. On y trouve aussi bien des données médicales, éventuellement utilisées par plusieurs établissements (clinique standard, hôpital psychiatrique, cabinet ...), que des données administratives (nom, prénom, adresse...).

Beaucoup de publications sont centrées sur la sécurité des systèmes d'information 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.

Contribution. Cet article propose une solution décentralisée d'un système d'information e-santé respectueuse de la vie privée. Cette architecture 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 proposée 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 afin de gérer le chiffrement des bases de données.

Organisation Ce papier commence par un bref état de l'art sur la gestion d'identité, les principes de sécurité et de vie privée et sur les exigences spécifiques au contexte de e-santé. La section III décompose et explique la nouvelle architecture, au sein d'un organisme de santé dans un premier temps puis entre deux institutions. Enfin, l'analyse de la solution proposée est présentée dans la section IV et des améliorations possibles sont données dans la conclusion.

II. Vie privée et système de e-santé

A. Identité numérique et vie privée

Dans un système d'information, les exigences pour le respect de la vie privée sont nombreuses. La classe FPR dans les exigences fonctionnelles des Critères Communs ([cc09]) décrit quatre contraintes : anonymat, pseudonymat, intraçabilité et non-observabilité. Les principes de minimisation des données, d'anonymat ou de pseudonymat sont également discutés dans le rapport technique de Pfitzmann et Hansen [PH08] et par Cameron avec ses sept lois de l'identité [Cam]. Plus précisément, l'*anonymat* garantit qu'un utilisateur peut accéder à une ressource sans divulguer son identité, alors que le *pseudonymat* exige que cette personne soit responsable de son utilisation. La notion de *non-associabilité* garantit que les données personnelles sont protégées contre une procédure d'agrégation. Cette notion est liée à l'anonymat. En effet, l'association de données peut permettre de récupérer l'identité d'un individu. Par ailleurs, il faut ajouter à ce principe la possibilité pour un attaquant de récupérer une donnée en dehors du système. La date de naissance d'un patient peut être retrouvée, par exemple, grâce aux réseaux sociaux. Ainsi, des données réparties dans différentes organisations peuvent être corrélées, l'*intraçabilité* n'est donc pas un principe élémentaire.

Une identité est représentée par le nombre d'attributs suffisant permettant d'identifier un individu dans une population donnée dont on connaît les caractéristiques générales. La gestion des identités consiste en *"des systèmes et processus qui gèrent et contrôlent ceux qui accèdent aux ressources et ce que chaque utilisateur est en droit de faire avec ces ressources, ceci en conformité avec les politiques de l'organisation"*, [PH08]. La personne liée à cette identité numérique est responsable de ses actes. Le vol d'identité est donc une menace importante pour les utilisateurs.

B. État de l'art des systèmes d'informations relatifs à la santé et à la protection de la vie privée

Le problème de la protection de la vie privée engendré par la numérisation des dossiers médicaux est considéré depuis le milieu des années 90, notamment par Anderson dans [And02]. Il a proposé un ensemble de principes sur les traitements de données pour l'e-santé qui doivent être vérifiés par toute politique de sécurité clinique. L'étude EPHR (European Privacy and Human Rights, [PE110]) a, quant à elle, présenté en 2010 un cadre de règles pour la vie privée dans l'Union européenne pour de nombreux cas. D'autres réglementations spécifiques au contexte de l'e-santé et à un pays existent. Ainsi, l'HIPAA (Health Insurance Portability and Accountability Act, [HIP06]) est une loi de 1996 qui régit la gestion de l'assurance maladie aux Etats-Unis. La recommandation [CE997] 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. Ainsi, 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. De plus, les informations médicales doivent être mises à disposition des acteurs autorisés. Il est nécessaire de protéger ces données contre les nombreuses attaques possibles, comme par exemple la divulgation d'information, et d'éviter les fuites d'information lors de transferts entre organismes (hôpitaux) ou lors d'utilisations secondaires.

Une approche décentralisée des systèmes de e-santé a récemment été suggérée par les auteurs de [QCF+09]. Cette solution donne lieu à une gestion d'identité basée sur un identifiant local par patient. Ces identifiants permettre 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.

Afin d'éviter de possibles agrégations, le gestionnaire d'identité ne doit pas stocker de relation entre les différents identifiants d'un patient sur une même base de données. Cette solution est proposée par Deng et coll. dans [DDCP09] et [DSDC+09]. 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 ([ADM02]) et par De Decker et coll. pour le système belge, [DDLVK08]. 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, il ne considère pas le transfert d'information entre institutions de santé, ni le principe de souveraineté des données. Le nouveau protocole proposé dans cet article est également plus efficace par son utilisation de cryptographie symétrique pour gérer le pseudo-anonymat des informations.

C. Exigences nécessaires pour un système e-santé respectueux de la vie privée

Exigences de sécurité. La personne (ou le groupe de personnes) désirant accéder à un dossier médical doit fournir les droits appropriés à la lecture ou à l'écriture de ce document. Le nom de la (ou les) 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 la confidentialité, l'intégrité et la disponibilité des données au sein d'une même institution, tout comme lors d'un transfert entre deux organismes. Ces données peuvent effectivement être partagées entre plusieurs prestataires de santé, un hôpital public et une clinique spécialisée par exemple.

Minimisation des données médicales. Un médecin peut accéder pleinement 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 est également utilisée pour la minimisation des données. Ce principe prévoit également l'anonymat des données médicales et donc le principe de non-associabilité des données. Tous les dossiers médicaux sont donc anonymisés, sauf dans certains cas exceptionnels, comme

pour la fusion de plusieurs institutions. Toutefois, dans un cadre d'enquête par exemple, l'anonymat peut être levé et la propriété de réversibilité ainsi 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.

Souveraineté des données. L'ensemble des informations médicales appartiennent au patient et donc ni au médecin qui les crée, ni 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 [And08] et [med07], 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, [And06], [CB97]. Ceci résulte du conflit entre *consentement* du patient et *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.

Système décentralisé. Une base de données nationale contenant les dossiers médicaux a été considérée 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, ont émergé. 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 ([And06] et [And08]) et malades. Un système décentralisé est donc une solution logique et une conséquence directe du principe de sensibilité des données.

III. Architecture de e-santé proposée

A. Approche globale de l'infrastructure

Dans le schéma proposé, la PKI, décrite par la Fig. 1, est 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'affiliation, 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].

À l'intérieur d'un système d'information e-santé, les données personnelles, médicales ou administratives, sont stockées dans deux bases de données différentes afin d'éviter un possible lien entre ces informations. Par ailleurs, elle sont pseudo-anonymisées, permettant ainsi aux seules personnes autorisées

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.

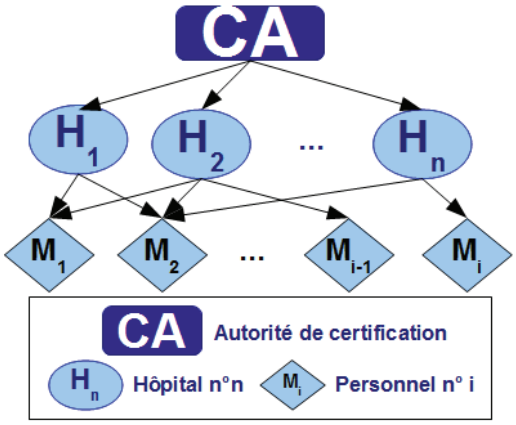

Figure 1. Infrastructure à clé publiques pour le système e-santé

de 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 crypté sécurisé. Cependant, l'authentification ne suffisant pas à la protection des données à caractères personnelles, les certificats permettent de gérer l'accès à celles-ci.

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 des patients 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 et lors d'un transfert de dossier. Par conséquent, c'est une donnée très 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, est recommandée.

Deux contrôleurs d'identité sont ensuite utilisés pour vérifier les identités et droits d'accès à l'intérieur des hôpitaux. Le contrôleur d'accès spécifique aux informations médicales est nommé CAM (Contrôle d'Accès Médical), celui correspondant aux données administratives est le CAA (Contrôle d'Accès Administratif).

B. Gestion d'identité à l'intérieur d'un hopital

Un patient possède deux identificants locaux, IdL_1 et IdL_2 calculés à partir de son identifiant global IdG. Ces identifiants locaux sont différents pour chaque répertoire afin d'éviter la traçabilité des données et donc leur associabilité. Dans la Fig. 2, 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.

Nous notons que l'âge du patient, et non sa date de naissance, est présent dans le dossier médical du malade. En effet, pour des raisons de prescriptions et de traitements, il est nécessaire pour un docteur de connaître l'âge de ses patient.

Durant la phase d'enregistrement d'un patient P, décrite dans le cadre III-B.1, le service de dérivation d'identité SDI utilise deux clés secrètes K_1 et K_2. Il applique ensuite

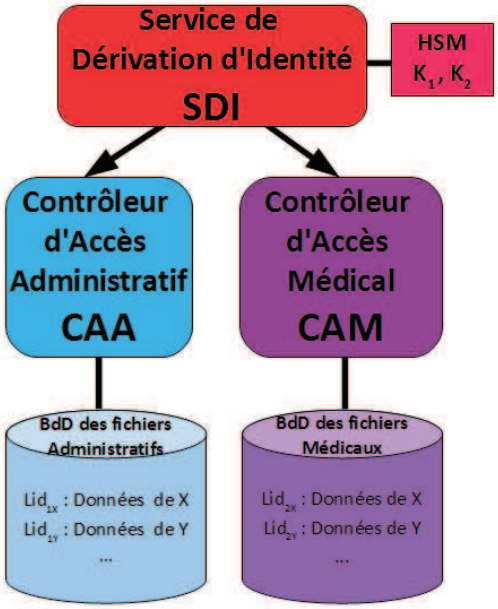

FIGURE 2. Gestion des identités et contrôle d'accès pour une architecture e-santé

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.

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. De plus, les contrôleurs d'accès CAA et CAM ne connaissent pas respectivement les identifiants IdL_2 et IdL_1, aucun lien ne peut donc être fait entre données médicales et administratives.

III.B.1 Phase d'enregistrement

1) P donne son IdG au gestionaire 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 stoqué) et retourne au patient ses identifiants locaux IdL_1 et IdL_2.
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.

Le patient reçoit ensuite un formulaire médical avec ses nom, prénom et identifiant local LId_2, utilisé comme pseudonyme dans son dossier médical. 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 donc 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 requière en effet 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. Malgrè 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 consultation 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 dossier. La procédure d'accès au dossier médical, qui est identique pour chaque employé, est détaillée dans l'encadré III-B.2.

III.B.2 Accès aux données médicales connaissant l'identifiant local

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.

Dans certains cas exceptionnels, comme en cas d'urgence où le patient est inconscient et donc où son identifiant local est inconnu, un médecin agréé, ou le service, ayant besoin d'accèder 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 l'encadré III-B.3.

III.B.3 Accès aux données médicales sans la connaissance de l'identifiant local

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 : $IdG = AES^{-1}_{K_1}(IdL_1)$
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.

Dans ce cas, les contrôleurs d'accès enregistrent l'identité du demandeur, avec la date, l'heure et le type de renseignements médicaux demandés. Une fois le patient à nouveau conscient, il peut être informé du déroulement des opérations.

C. Chiffrement des bases de données

Afin d'accéder aux dossiers médicaux, l'employé fournit un certains nombres 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, il est possible de considérer le fait que les employés n'ont pas les mêmes droits au sein d'un hôpital. En effet, alors que le docteur et le patient peuvent accéder au dossier médical complet du patient, l'infirmière a uniquement besoin d'accéder aux prescriptions et n'a donc aucun droit sur les diagnostics.

Nous pourrions utiliser l'approche cryptographique de Ghindici dans sa thèse [Ghi08] en fournissant aux trois acteurs une clé privée : $K_{patient}$, $K_{docteur}$ et $K_{infirmière}$ où $K_{docteur}$ est chiffré par $K_{patient}$ et $K_{infirmière}$ par $K_{docteur}$. Ainsi, le patient a accès à la clé du docteur et le docteur à la clé de l'infirmière. Les données composant le dossier sont alors chiffrées comme suit :

- La partie diagnostic est chiffrée par la clé du docteur, et peut donc également être lu par le patient ;
- Les prescriptions et autres données, comme l'âge, sont chiffrées avec la clé de l'infirmière afin d'être lisible par les trois acteurs.

Cependant, dans une institution médicale, il est nécessaire de pouvoir ajouter des acteurs aux systèmes. Cette solution n'est donc pas opérationnelle. Il est alors possible de se tourner vers le principe de partage de secret de Shamir, [Sha79]. 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 certains nombres $(n-1)$ de participants doivent s'unir. On a alors le Tableau 3.

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

FIGURE 3. Détails des droits d'accès des acteurs du système médical

Le patient, le docteur et l'infirmière doivent donc utiliser leur clé, ainsi que celle du serveur, pour obtenir la clé permettant de déchiffrer les données. Par conséquent, sachant que deux points, ici sous forme de clés, suffisent à définir une droite, le clé de déchiffrement des prescriptions, se cachera dans une équation de degré un. De la même façon, le déchiffrement des données administratives fait appel à la clé de la secrétaire et du serveur. La clé de déchiffrement des diagnostics quant à elle se dissimulera dans une équation de degré deux. La Fig. 4 donne un aperçu de la solution.

Par contre, les nom, prénom et année de naissance figurant dans le dossier administratif ne doivent pas être chiffrés afin d'assurer une admission potentielle au urgence.

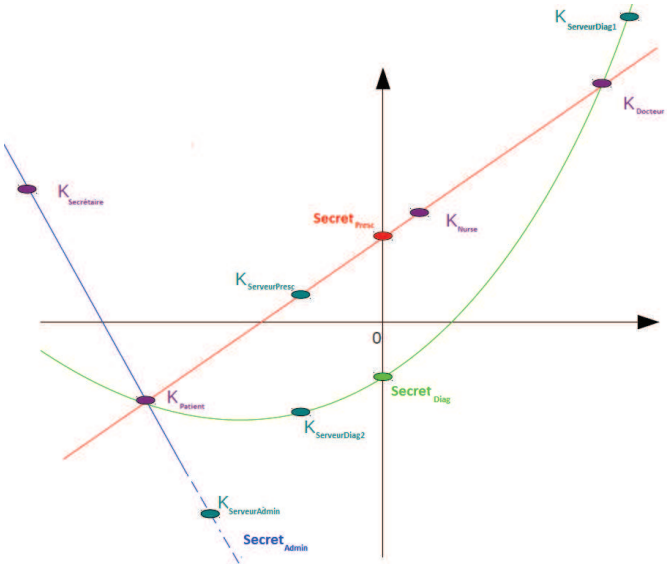

FIGURE 4. Partage de secret de Shamir dans le système e-santé

D. 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 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}. Ce transfert est réalisé en plusieurs étapes et utilise l'infrastructure PKI.

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 calculer 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 alors IdL_{2,H_2}, joint son contrôleur d'accès médical 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. L'encadré III-D.3 explique en détail cette procédure.

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.

III.D.3 Communication entre hôpitaux

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} ouvre 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'identifiant global 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} retrouve les informations médicales et vérifie les droits d'accès. Il les 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'identité du demandeur, la date et l'heure sont enregistrées.

IV. Analyse de sécurité du modèle proposée

A. Sécurité des données

La confidentialité, l'intégrité et la disponibilité des données à caractère personnel (administratives ou médicales) sont assurées par les contrôleurs d'accès CAM et CAA. L'autorisation d'accès est contrôlée par une vérification des droits d'accès. 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. 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. Par conséquent, seules les personnes autorisées accèdent ou mettent à jour ces données à l'intérieur de l'institution.

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.

B. Anonymat et intraçabilité 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. 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 donnée 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 élément de sécurité de type HSM.

Le principe de non-associabilité qui en découle est vérifié 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 différents 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.

Enfin, la séparation 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 de 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 accéde 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.

C. Souveraineté de données

Dans l'architecture définie, le consentement du patient 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.

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 considérée comme une protection supplémentaire de souveraineté des données. Le patient est informé dès que possible de la demande ainsi traitée.

V. 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. De plus, les extensions possibles de ce système, comme une utilisation secondaire des données médicales pour la recherche, ont besoin d'une attention supplémentaire. En effet, cette divulgation de données doit utiliser une anonymisation complète des informations qui seront, par exemple, uniquement des prescriptions et des diagnostics.

La minimisation des données et leur souveraineté sont des principes nécessaires dans un système d'information décentralisé de e-santé. La solution proposée, conçue avec des éléments simple, tente de préserver la vie privée du patient en prenant en compte des contraintes du système médical, tels que le transfert de données médicales entre le personnel autorisé et diverses institutions. La protection de la vie privée étant un enjeu majeur, la Convention Européenne souhaite établir des régimes de certification européen pour la création d'un label la concernant.

Remerciements

Les auteurs voudraient remercier Coline Migonney pour sa participation durant la phase *chiffrement des bases de données*, ainsi que Christophe Rosenberger et Wipa Chaisantikulwat pour leur relecture.

Références

[ADM02] G Ateniese and B De Medeiros. Anonymous e-prescriptions. pages 19–31. ACM, 2002.

[And02] R.J. 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.

[And06] R. Anderson. Under threat : patient confidentiality and nhs computing. *Drugs and Alcohol Today*, 6(4) :13–17, 2006.

[And08] R. Anderson. Patient confidentiality and central databases. *Br J Gen Pract*, 58(547) :75–76, 2008.

[bGA90] Adopted by General Assembly. Guidelines for the regulation of computerized personal data files. *resolution 45/95*, December 1990.

[Cam] K. Cameron. The laws of identity. *Microsoft Corp.*

[CB97] F. Caldicott and G. Britain. *Report on the review of patient-identifiable information.* Department of Health, 1997.

[cc09] *Common Criteria for Information Technology Security Evaluation.* Department of Health, july 2009.

[CE997] Recommendation of council of europe n. r(97)5 on the protection of medical data, February 1997.

[DDCP09] 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.

[DDLVK08] B. De Decker, M. Layouni, H. Vangheluwe, and Verslype K. Anonymous e-prescriptions. pages 118–133. Public Key Infrastructure, 2008.

[DSDC+09] 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.

[ec10] *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.

[eu002] On the processing of personal data and the protection of privacy in the electronic communications sector, 2002.

[eu887] European convention on human rights, 1987.

[EUd95] On the protection of individuals with regards to the processing of personal data and on the free movement of such data, 1995.

[Ghi08] Dorina Ghindici. *Information flow analysis for embedded systems : from practical to theoretical aspects.* PhD thesis, INRIA, Sophia-Antipolis et Univ. Laval, Canada, 2008.

[HIP06] Health insurance portability and accountability act. hipaa administrative simplification : enforcement ; final rule, 2006.

[med07] Medix uk plc, November 2007.

[PE110] European privacy and human rights (ephr), 2010.

[PH08] A. Pfitzmann and M. Hansen. Anonymity, unlinkability, unobservability, pseudonymity, and identity management - a consolidated proposal for terminology. Technical Report, 2008. v0.31.

[QCF+09] Catherine Quantin, Gouenou Coatrieux, Maniane Fassa, Vincent Breton, D-O Jaquet-Chiffelle, Paul De Vlieger, N Lypszyc, J-Y Boire, Christian Roux, and F-A Allaert. Centralised versus decentralised management of patients' medical records. In IOS Press, editor, *Medical Informatics in a United and Healthy Europe K.-P. Adlassnig et al. (Eds.)*, 2009.

[Sha79] Adi Shamir. How to share a secret. *Communications of the ACM 22*, pages 612–613, 1979.