Université Laarbi Tébessi – Tébessa
Faculté des Sciences Exactes, Science
de la Nature et de la Vie
Département des Mathématiques et Informatique

LAMIS
Laboratoire de Mathématiques,
Informatique et Systèmes

THÈSE

Pour obtenir le grade de
docteur 3ème cycle L.M.D. en
informatique

Option : Réseaux de Systèmes Intelligents

Thème

Sécurité de l'internet des objets en utilisant la biométrie

Présentée par : **BENTAHAR Atef**

Soutenue le : 26/05/2022, devant le jury composé de :

Pr. Laouar Med Ridda	Université Laarbi Tébessi- Tébessa	Président;
Pr. Derdour Makhlouf	Université Laarbi Ben Mhid- Oum El bouaghi	Examinateur;
Dr. Laimeche Lakhdar	Université Laarbi Tébessi- Tébessa	Examinateur;
Dr. Menassel Rafik	Université Laarbi Tébessi- Tébessa	Examinateur;
Pr. Meraoumia Abdallah	Université Laarbi Tébessi- Tébessa	Rapporteur;
Pr. Bendjenna Hakim	Université Laarbi Tébessi- Tébessa	Co-Rapporteur;

Année universitaire: 2021 / 2022

Sécurité de l'internet des objets en utilisant la biométrie

BENTAHAR Atef
atef.bentahar@univ-tebessa.dz

Laboratoire de Mathématiques, Informatique et Systèmes (LAMIS),
Université Larbi Tébessi - Tébessa, Algerie.

Laboratoire de Mathématiques,
Informatique et Systèmes


بسم الله الرحمن الرحيم

إلى العزيزة أمي التي وافتها المنية إبان كتابة هذه الاطروحة. أسأل الله ان يرحمها ويتغمد روحها الطاهرة.

À mon père; À mes frères et sœurs
À ma petite famille; en particulier ma compagne et mes chers enfants.
À mes amis ;
À tous ceux qui m'ont enseigné
À tous ceux qui me sont chers

Remerciements

*Loué soit "**Allah**" pour sa grâce par laquelle Les bonnes sont faites. Tout cela grâce à "**Allah**", avant et après, qui nous a aidés à cela, et nous n'avons réussi que par la permission "d'**Allah**".*

*Tout d'abord, je tiens à remercier mon directeur de thèse, **Pr. Meraoumia Abdallah**, Maître de conférences à l'Université Larbi Tébessi - Tébessa, pour sa confiance et son soutien absolu. Sans ses conseils avisés et ses précieuses observations, ce travail ne serait jamais achevé.*

*Je tiens également à remercier le **Pr Benjanna Hakim**, directeur du laboratoire LAMIS à l'Université Larbi Tébessi- Tébessa et le co-directeur de ma thèse, pour son assistant tout au long de la préparation du travail, et pour mettre à notre disposition toutes les ressources de laboratoire possibles.*

Un merci tout spécial aux membres du jury avec qui j'aimerais partager mon travail et qui nous aideront certainement à l'améliorer avec leurs commentaires de valeur et leurs suggestions utiles.

*Je tiens à remercier sincèrement mon frère **Dr. BentaharTarek**, Maître de conférences à l'Université Larby Tébessi, pour son précieuse aide dans la préparation, la finalisation et l'amélioration de la thèse.*

*Remerciement spécial à **Pr Amroune Mohamed** pour ses conseils et orientations.*

*Je remercie tous mes enseignants et collègues, notamment du **Laboratoire LAMIS**, pour l'aide qu'ils m'ont apportée, et merci aussi aux amis de m'avoir aidé de toutes leurs forces.*

*Je tiens également à remercier ceux qui étaient prêts à chaque fois que nous avions besoin d'eux. Particulièrement mon cher ami **Zeroual Abdelhakim**.*

Enfin, ma famille, pour leur indirecte soutien.

Résumé

Aujourd'hui, les réseaux répandus tels que l'Internet des objets (IoT) comportent des données de perception qui sont désormais disponibles en tant que service public. Ce nouveau modèle, appelé "Perception en tant que service (*Sensing as a Service* - S^2aaS) ", permet aux propriétaires de ces objets de vendre et d'acheter des "données" aux consommateurs intéressés dans un grand marché électronique ouvert. En effet, construire l'industrie des services utilisant le modèle S^2aaS basé sur l'IoT nécessite de trouver des solutions légères et précises pour que les êtres humains puissent pleinement bénéficier de ces services, tout en assurant une authentification précise et un échange de données sécurisé. À ces fins, les technologies biométriques sont considérées comme le moyen le plus fiable d'authentification de l'être humain en raison de ses résultats efficaces dans de nombreux domaines.

Cette thèse propose des solutions de sécurisation IoT utilisant la technologie biométrique. Pour que le niveau de sécurité soit à l'ampleur désiré, les objectifs de sécurité bien connus doivent être atteints par ces solutions. D'autre part, les solutions doivent respecter les contraintes et les limites de l'IoT et définir clairement « Où et Comment » pour inclure la technologie biométrique. Nos propositions fournissent des solutions pour les éléments IoT suivant, pour l'architecture bout-en-bout, pour l'accès au cloud, et pour l'architecture entière S^2aaS basée sur l'IoT.

Dans les solutions proposées, les données biométriques sont exprimées par un vecteur de caractéristiques qui est extrait et sécurisé par des algorithmes fiables et adaptés. Ces algorithmes biométriques sont utilisés conjointement avec des algorithmes de cryptographie pour partager une clé secrète et établir une session d'échange de données sécurisées. Pour améliorer la discrimination du vecteur de caractéristiques biométriques extrait, une nouvelle technique de quantification est utilisée. De plus, pour corriger certaines erreurs de transmission, des nombreux codes correcteurs sont incorporés.

Notre solution pour le modèle S^2aaS exploite le sous-système biométrique et permet ainsi à l'utilisateur d'accéder rapidement et en toute sécurité aux services publics via un simple site Web plutôt qu'une carte à puce spéciale. Le sous-système biométrique est utilisé pour authentifier les utilisateurs ainsi que pour partager la clé secrète afin gagner du temps et de l'énergie. Afin de respecter les contraintes de l'IoT, les processus lourds sont évités notamment lors des intervalles critiques et fréquents.

Afin de valider notre travail, la résilience de nos systèmes contre diverses attaques de transmission et de sauvegarde a été discutée et les solutions ont été formellement prouvées par un outil de sécurité spécial. De plus, les résultats obtenus ont été comparés à d'autres travaux récents en termes de sécurité, de coûts de communication, de coûts de calcul et de certaines métriques pertinentes, pour montrer que nos solutions sont légères et offrent plus de sécurité.

Mots clés:

Internet des Objets, Perception en tant que service, Sécurité de l'information, Reconnaissance biométrique, Cryptographie, Crypto-systèmes biométriques, Fusion de données.

Abstract

Today, widespread networks such as the Internet of Things (IoT) have sensing data now available as a public service. This new model, called "Sensing as a Service (S^2aaS)", allow owners of these things to sell and buy "data" to interested consumers in a large open electronic marketplace. Indeed, building the service industry using the IoT-based S^2aaS model requires finding lightweight and precise solutions so that human beings can fully benefit from these services, while ensuring precise authentication and secure data exchange. For these purposes, biometric technologies are considered to be the most reliable means of human being authentication due to its effective results in many fields.

This thesis proposes IoT security solutions using biometric technology. For the level of security to be at the desired scale, well-known security objectives must be met by these solutions. On the other hand, solutions must respect the constraints and limitations of IoT and clearly define “Where and How” to include biometric technology. Our proposals provide solutions for the following IoT elements, for end-to-end architecture, for cloud access, and for the fully IoT-based S^2aaS architecture.

In the proposed solutions, the biometric data is expressed by a feature vector which is extracted and secured by reliable and adapted algorithms. These biometric algorithms are used in conjunction with cryptography algorithms to share a secret key and establish a secure data exchange session. To improve the discrimination of the extracted biometric feature vector, a new quantification technique is used. In addition, to correct certain transmission errors, numerous correcting codes are incorporated.

Our solution for the S^2aaS model leverages the biometric subsystem and thus enables the user to quickly and securely access public services via a simple website rather than a special smart card. The biometric subsystem is used to authenticate users as well as share the secret key to save time and energy. In order to respect the constraints of the IoT, heavy processes are avoided especially during critical and frequent intervals.

In order to validate our work, the resilience of our systems against various transmission and storage attacks was discussed and the solutions were formally proven by a special security tool. Moreover, the results obtained were compared with other recent works in terms of security, communication costs, computational costs and some relevant metrics, to show that our solutions are lightweight and offer more security.

Keywords:

Internet of Things, Sensing as Service, Information Security, Biometric Recognition, Cryptography, Biometric Cryptosystems, Data Fusion

ملخص

اليوم، الشبكات العنكبوتية المنتشرة مثل إنترنت الأشياء (IoT) لديها بيانات استشعار متاحة كخدمة عامة. يسمح هذا النموذج الجديد المسمى "الاستشعار كخدمة (S^2aaS)" لأصحاب هذه الأشياء ببيع وشراء "البيانات" من و الى المستهلكين المهتمين في الأسواق الإلكترونية المفتوحة. في الواقع، بناء صناعة الخدمات باستخدام نموذج الاستشعار كخدمة القائم على إنترنت الأشياء يتطلب إيجاد حلول خفيفة ودقيقة حتى يتمكن العملاء من الاستفادة الكاملة من هذه الخدمات، مع ضمان المصادقة الدقيقة و التبادل الآمن للبيانات.لهذه الأغراض، تعتبر تقنيات القياسات الحيوية أكثر الوسائل موثوقية لمصادقة هويات العملاء نظرًا لنتائجها الفعالة في العديد من المجالات.

تقترح هذه الاطروحة حلولاً أمان إنترنت الأشياء باستخدام تقنية القياسات الحيوية. لكي يكون مستوى الأمان عند المستوى المطلوب، يجب أن تتحقق أهداف الأمان المعروفة من خلال هذه الحلول. من ناحية أخرى، يجب أن تحترم الحلول قيود ونقاط ضعف إنترنت الأشياء، وأن تحدد بوضوح "أين وكيف" يتم تضمين تكنولوجيا القياسات الحيوية. تقدم مقترحاتنا حلولا لعناصر إنترنيت الأشياء التالية، لإنترنت الأشياء طرف إلى طرف، وللوصول إلى السحابة، وللبنية الكاملة للاستشعار كخدمة القائمة على إنترنت الأشياء.

في الحلول المقترحة ، يتم التعبير عن البيانات البيومترية بواسطة متجه للخصائص يتم استخراجه وتأمينه بواسطة خوارزميات موثوقة ومكيفة .تُستخدم خوارزميات القياسات الحيوية هذه جنبًا إلى جنب مع خوارزميات التشفير لمشاركة مفتاح سري ومنه تبادل البيانات بصفة آمنة .لتحسين دقة متجه الخصائص البيومترية المستخرجة ، يتم استخدام تقنية جديدة للتكميم .بالإضافة إلى ذلك، لتصحيح بعض أخطاء الإرسال تم دمج العديد من خوارزميات التصحيح.

يعمل حلنا الخاص بنموذج S^2aaS على تعزيز النظام الفرعي للقياسات الحيوية، ومن ثم يمكّن المستخدم من الوصول بسرعة وأمان إلى الخدمات العامة عبر موقع ويب بسيط بدلاً من بطاقة ذكية خاصة. يتم استخدام النظام الفرعي للقياسات الحيوية لمصادقة المستخدمين وكذلك مشاركة المفتاح السري لتوفير الوقت والطاقة .من أجل احترام قيود إنترنت الأشياء، يتم تجنب العمليات الثقيلة خاصة خلال الفترات الحرجة والمتكررة.

من أجل التحقق من صحة عملنا، تمت مناقشة مرونة أنظمتنا ضد هجمات النقل والتخزين المختلفة وتم إثبات فعالية الحلول المقترحة بواسطة أداة أمان خاصة .علاوة على ذلك، تمت مقارنة النتائج التي تم الحصول عليها مع الأعمال الحديثة الأخرى من حيث الأمان وتكاليف الاتصال والتكاليف الحسابية وبعض المقاييس ذات الصلة، لإظهار أن حلولنا خفيفة وتوفر المزيد من الأمان.

الكلمات المفتاحية:

إنترنت الأشياء ، الاستشعار كخدمة ، أمن المعلومات ، التعرف على المقاييس الحيوية ، التشفير ، أنظمة التشفير البيومترية ، دمج البيانات.

Table des matières

Liste d’abréviations

AES	Advanced Encryption Standard
BCH	Bose–Chaudhuri–Hocquenghem
Com_i	Commitment de l’utilisateur i
CRC	Cyclic Redundancy Check
CS	Clé Secrète
DCT	Discrete Cosine Transform
Dec_{KEY}(Msg)	Message Msg déchiffré par la clé symétrique KEY
DES	Data Encryption Standard
DH	Diffie-Hellman Key Exchange
DoS	Denial-of-Service
DR-B	Direct Recognition- Binary
DR-NB	Direct Recognition-Non Binary
DWT	Discrete Wavelet Transform
EA	Ensemble d'apprentissage (Training Set)
ECC	Elliptic Curve Cryptography
ECDH	ECC Diffie-Hellman
$ECDH_{si}$	Diffie-Hellman Key exchanged (Service/User)
$ECDH_{sk}$	Diffie-Hellman Key exchanged (Service/Fog)
ECR	Excluded Client Rate
Een_{KEY}(Msg)	Message Msg chiffré par la clé symétrique KEY
EER	Equal Error Rate
EH	Extended Hamming
$E_q(a,b)$	Groupe en Courbe Elliptique, (a, b) est un ensemble de points et q est un nombre premier.
ESP	Extended Service Providers Layer
FAR	False Acceptance Rate
FC	Fuzzy Commitment
FC-H	Fuzzy Commitment avec code Extended Hamming
FC-RS	Fuzzy Commitment avec code Reed-Solomon
FE	Fuzzy Extracor
FN_k	Fog Node k
FRR	False Rejection Rate
FV	Fuzzy Vault
FV-H	Fuzzy Vault avec code Extended Hamming
FV-RS	Fuzzy Vault avec code Reed-Solomon
G	Point dans E_q dont l'ordre est une grande valeur
GAR	Genuine Acceptance Rate
Gen(.)	Fonction de générateur de Fuzzy Extracor
GF	Galois Field
H (.)	Fonction de hachage
ID_{FNk}	Fog identifier
ID_i	User identifier
ID_{SNj}	Sensor identifier
IE	Identification Error
IER	Identification Error Rate
IoT	Internet of Things (l'Internet des objets)

$\mathbf{K_{FNk}}$	Fog secret key
$\mathbf{K_i}$	User secret key
$\mathbf{K_{SNj}}$	Pre-shared key (Fog/Sensor)
$\mathbf{K_{url}}$	Service secret key
MK	Master key
MP	matrice de projection
$\mathbf{n_{FNk}}$	Fog private key
$\mathbf{n_i}$	User private key
$\mathbf{n_{url}}$	Service private key
PB	Paramètre public
PCA	Principal Component Analysis
$\mathbf{P_{FNk}}$	Fog public key
$\mathbf{P_i}$	User public key
Pub/Sub	Publish/Subscribe (Publier/S'abonner)
$\mathbf{P_{url}}$	Service public key
Rep(.)	Fonction de reproduction de Fuzzy Extracor
RFID	Radio Frequency IDentification
$\mathbf{R_i}$	Random challenger
ROI	Region of Interest
ROR	Rank One Recognition
RPR	Rank of Perfect Recognition
RS	Reed-Solomon
RSA	Rivest–Shamir–Adleman
S	Cloud Service
$\mathbf{S^2}$**aaS**	Sensing as a Service (Perception en tant que service)
$\mathbf{S^2}$**aaS-IoT**	Sensing as a Service based on Internet of Things
$\mathbf{SN_j}$	Sensor Node j
SP	Sensor Data Publishers Layer
SR	Success Rate
SS	Secure Sketch
$\mathbf{T_i}$	Current timestamp
TSCH	Time Synchronization and Channel Hopping
$\mathbf{U_i}$	User i
URL	Service identifier
$\mathbf{Vault_i}$	Vault de l'utilisateur i
VM	Virtual Machine
$\mathbf{W_i}$	le vecteur biométrique extrait de l'utilisateur i
WSN	Wireless Sensor Network

Liste des Tableaux

Table des Figures

Introduction Générale

Introduction Générale

1. Mise en contexte

Le domaine de l'internet et des réseaux en général connaît un développement étonnant et rapide. Après qu'internet n'était qu'une communication entre un groupe d'ordinateurs, puis l'implication d'un groupe de téléphones portables, il est devenu le rassemblement de tous les différents appareils électroniques. L'intégration des capteurs (*Sensors*), actionneurs (*Actuators*), et la radio-identification (*Radio Frequency IDentification* - RFID) dans presque toutes les entités physiques les rend connectables à Internet et connectables les unes aux autres. Ces entités physiques de toutes sortes sont appelés "objets". Au fil du temps, le nombre de ces objets augmente rapidement; il y a actuellement 30,7 milliards d'objets connectés, et d'ici 2025, il devrait y en avoir 75,44 milliards [Statista2020]. Cette grande quantité d'objets connectés est ce que l'on appelle aujourd'hui "l'internet des objets" (*Internet of Things* - IoT) [Vans2011].

Les évolutions continues du marché électronique et les tendances technologiques au cours de la dernière décennie ont modifié la notion et la valeur des objets interconnectés. À savoir, la combinaison de matériel à faible coût et le réseau à haut débit ; à la fois filaires et sans fil, a permis la création d'une nouvelle génération d'objets dotés d'une connectivité fiable et omniprésente sur l'ensemble de l'internet. Le bon exemple c'est le G5[1]. Ces systèmes facilitent la récolte et l'échange de données en temps réel, et offrent une visibilité et un contrôle sans précédent des opérations et des processus. L'utilisation accrue de l'informatique basées sur le cloud a introduit des capacités d'analyse de données encore plus avancées, inaugurant une nouvelle ère de prise de décision, de contrôle et d'automatisation intelligents. Cette tendance IoT basé sur le cloud est appelé « *cloud-based IoT*» [Al-Turjman2020].

De là, deux types d'architectures IoT peuvent être extraits, l'une est l'IoT conventionnel et l'autre est l'IoT basé sur le cloud. Dans le premier, chaque individu gère et bénéficie de ses objets privés. Cette architectures contient trois domaines, domaine de la perception (*Sensing Domain*), domaine réseau (*Network Domain*), et domaine d'utilisateur (*User Domain*) [Kumar2017]. Elle est également appelée architecture IoT de bout en bout (*End-to-End IoT*). La seconde architecture est composée des quatre domaines suivants, domaine de la perception (*Sensing*

[1] La 5G est le réseau mobile de 5e génération. Il s'agit d'une nouvelle norme mondiale sans fil après les réseaux 1G, 2G, 3G et 4G. La 5G permet un nouveau type de réseau conçu pour connecter pratiquement tout le monde et tous ensemble, y compris les machines, les objets et les appareils.

Domain), domaine du brouillard (*fog domain*), domaine de cloud (*Cloud Domain*), et domaine d'utilisateur (*User Domain*).

Selon l'IoT basé sur le cloud actuel, les données agrégées de divers objets interagissent au niveau du cloud pour fournir le service souhaité. Quand il y a plusieurs utilisateurs et une grande foule de propriétaires des objets, il y aura un grand nombre de services interconnectés. Il est donc possible de profiter des avantages de ces services en les combinant en un seul service unifié. La création et la configuration de ce service unifié nécessite certainement une solution pour le problème de l'hétérogénéité, de l'incohérence et de l'interopérabilité. Cette faiblesse a été surmontée grâce à nouveau paradigme appelé « *Sensing as a Service* (S^2aaS) » [Perera2019].

D'un autre point de vue, l'être humain faisant partie de ces objets connectés, l'incorporation des technologies biométriques pour intégrer l'élément humain dans ces architectures IoT présente un moyen de sécurité fiable en général, et un moyen d'authentification et d'identification en particulier. La dernière décennie a vu une augmentation spectaculaire du taux d'adoption des technologies biométriques. Ceci est motivé par une amélioration significative des technologies habilitantes et des capacités de traitement des données tels que la puissance de calcul, les développements algorithmiques, les entités de stockage et la maitrise de la consommation d'énergie. Les progrès de l'informatique mobile avec l'apparition d'appareils intelligents ont également une grande contribution dans l'amélioration des technologies biométriques.

Le domaine de la biométrie présente aujourd'hui un segment de marché bouillonnant qui concurrence les autres domaines [Global2021]. La biométrie est également marquée par sa diversité et son innovation dynamique. Depuis le début des premières formes de la biométrie avant des dizaines d'années, le domaine a rencontré une lente évolution où seule une pointe de technologie, principalement l'empreinte, était disponibles et utilisées. En revanche, au cours des trois dernières décennies, le domaine biométrique a connu une explosion d'avancés permettant l'émergence d'autres technologies qui ont ouvert plus de possibilité pour exploiter autres modalités biométriques. Leur intégration dans divers nouveaux appareils et leur mise en œuvre sont devenus maîtrisables. Mais faire répandre cette technologie n'est pas sans effets négatifs, effectivement, l'information biométrique émergées dans ces technologies est plus vulnérables aux imposteurs et aux attaques qui sont aussi développés.

A la première vue, il semble que nous ayons présenté deux concepts distincts, à savoir l'Internet des objets et les technologies biométriques. Mais en fait, ils sont fortement liés, ou plutôt ils sont complémentaires l'un vis-à-vis l'autre lorsque l'être humain est devenu une partie intégrante de

l'IoT. Ainsi, l'inclusion et l'implication de la technologie biométrique dans l'IoT est devenu impératif pour faciliter l'intégration des êtres humaines et améliorer l'interaction homme-objet.

Dans le modèle S^2aaS, L'être humain peut interagir avec IoT à travers trois situations. La première situation est en tant que consommateur de données. La deuxième est en tant que producteur de données, et la troisième en tant que propriétaire de données. Tout simplement, où l'être humain peut être intégré dans l'IoT, la technologie biométrique peut être impliquée.

2. Motivations

La cybersécurité est devenu un objectif pour toutes les institutions, quel que soit leur type. Les besoins croissants de la sécurité dans des services critiques tels que la santé, les finances, le gouvernement, la zone militaire, les réacteurs nucléaires, etc. nous conduit à développer des solutions puissante, efficace et rapide, car il existe des applications qui ne tolèrent pas les erreurs ou les retards. La solution de sécurité ; y compris la sécurité biométrique, est applicable dans deux axes, l'axe de communications, l'axe de sauvegarde (bases de données). La technologie biométrique pourrait permettre de contrôler et de sécuriser l'interaction l'homme et avec d'autres technologies de communication et de sauvegarde telles que l'IoT.

Les systèmes de sécurité basés seulement sur les mots de passe sont sujets à de nombreuses attaques menaçantes. Le stockage des fichiers contenant des mots de passe hachés est inutile, car ils sont sujets aux attaques par dictionnaire et aux attaques de table Arc-en-Ciel (*Rainbow Table Attacks*) [Beaver2016]. Un niveau de sécurité plus élevé est assuré par l'authentification multifactorielle. L'authentification multifactorielle offre au moins deux niveaux de sécurité en fonction de ce que vous possédez ou de ce que vous rappelez. L'inconvénient majeur de ces moyens d'authentification est qu'ils sont faciles à se perdre, se voler ou s'oublier. Faire recours à un autre moyen d'authentification qui ne peut pas être oublié, perdu, deviné ou falsifié est incontournable. La technologie biométrique est l'une des solutions pour remédier à ce problème. Alors que les technologies biométriques peuvent s'avérer être des outils relativement robustes et efficaces pour restreindre la fraude, ils sont eux-mêmes pas à l'abri des attaques frauduleuses et d'exploitation illégales, en particulier après l'extraction des vecteurs biométriques car une fois l'extraction effectuée, l'information biométrique est présentée par un vecteur qui se vole facilement. Par conséquent, il est nécessaire de fournir des moyens et des mécanismes de protection du vecteur des caractéristiques au niveau de la base de données et au niveau de transmission/traitement.

D'autre part, lorsque la technologie biométrique est intégrée à la technologie IoT basée sur les services publics (S^2aaS); où les données produites sont consommables par des différents utilisateurs, il est plus vulnérable aux risques de différents cyberattaques par rapport à l'ancien paradigme IoT. Cela est dû à l'absence de propriétaire unique des services, à la diversité des ressources et à la présence de transactions financières lourdes à ce modèle. S^2aaS présente des caractéristiques et des contraintes uniques concerne principalement la conception des mécanismes de défense efficaces qui sont: Interopérabilité des services, technologies multiples, applications multiples, évolutivité (*Scalability*), *Big Data*, disponibilité, limitations des ressources, emplacements distants, mobilité, sensibilité au délai [Ghorbani 2017], [Perera2013].

3. Questions de départ et problématique

La sécurité de l'IoT utilisant les technologies biométriques est abordée dans cette thèse, afin d'atteindre le niveau de sécurité requis, à savoir, authentification mutuelle, confidentialité, intégrité, autorisation, non-répudiation, disponibilité, fraîcheur (*Freshness*), confidentialité persistante. Les solutions proposées doivent respecter les contraintes susmentionnées concernant les limitations de l'IoT et répondre à la question « Où et Comment » inclure la technologie biométrique. À partir de ce point, notre problématique peut être résumée en général dans la question : " ***Comment sécuriser l'Internet des objets en utilisant la technologie biométrique ?***"

Comme nous l'avons précédemment mentionné, l'être humain peut interagir avec IoT basée sur les services à travers trois situations, en tant que consommateur de données, en tant que producteur de données, et en tant que propriétaire de données. Les mécanismes de sécurité utilisés dans les trois situations, y compris les mécanismes biométriques, sont les mêmes. Les protocoles de sécurité IoT et les algorithmes peuvent être mis en œuvre pour n'importe quel rôle humain, qu'il soit producteur, consommateur ou propriétaire. Pour cela, il suffit que nous nous concentrions et travaillons dans cette thèse sur le rôle du consommateur uniquement le faite qu'il puisse présenter les deux autres rôles. Cependant, nous avons traité ce rôle au niveau IoT de bout en bout, au niveau d'accès au cloud, et au niveau S^2aaS basé sur l'IoT. De ce point de vue, on peut diviser la problématique en trois sous-problématiques propres à chaque niveau. Ces sous-problématiques ont été révélées en passant en revue les solutions présentées dans les littératures actuelles. Il est de préférence de les appeler lacunes ou défauts plutôt que de sous-problématiques.

A) Lacunes communes au niveau IoT de bout en bout

Les solutions consacrées à l'IoT de bout en bout se caractérisent par les points communs suivants:

- Basées sur trois rôles: utilisateur, passerelle IoT (nœud de brouillard) et capteur.
- Utiliser des cartes à puce.
- Utiliser des canaux sécurisés.

Par conséquent:

- Les schémas présentés ne peuvent être appliqués que lorsque l'utilisateur communique directement avec le nœud de brouillard ou la passerelle appropriée, et cela ne semble pas être le cas pour l'IoT basé sur les services. Car le concept de service n'est pas supporté malgré il est la pierre angulaire du paradigme IoT actuel.
- L'utilisation d'une carte à puce spéciale pour chaque application IoT est une solution lourde, coûteuse, et pose le problème de garder autant de cartes que les services demandés.
- Les canaux disponibles ne sont pas toujours sécurisés sur les marchés en ligne.

B) Lacunes communes d'accès au cloud

Il est possible de noter quelques lacunes dans les crypto-systèmes biométriques présentés dans la littéraire pour accéder au cloud, notamment:

- Les solutions traitaient du problème de sauvegarde des données biométriques dans le cloud et n'abordaient pas le problème de la sécurité de transmission. Notamment ce que nous avons besoin aujourd'hui d'accès à distance aux services cloud. Les algorithmes à proposer doivent être bien exploités pour sécuriser la transmission des données biométriques.
- Puisqu'il existe de nombreuses techniques d'extraction biométriques et de nombreux types des codes correcteurs, et il existe aussi un grand espace de clés, nous pouvons concevoir un choix optimisé pour achever à un schéma plus efficace et plus léger.
- Les solutions proposées peuvent être élargies et complétés avec les travaux de l'implication biométrique à l'architecture IoT de bout en bout pour préparer et créé un schéma complet qui inclut tous les acteurs de l'IoT basé sur les services.

C) Lacunes communes au niveau S^2aaS basé sur l'IoT

- L'une des lacunes majeures des travaux connexes est qu'il n'existe pas un schéma de sécurité complet de l'implication des technologies biométriques pour S^2aaS basé sur l'IoT avec toutes

ses entités, telles que les utilisateurs, les services de cloud, les nœuds de brouillard, et les nœuds de capteurs.

- De même, la plupart de ces travaux contiennent au moins un canal sûr, en particulier lors de l'enregistrement des utilisateurs, et cela n'est pas valable lorsque le service est public et l'enregistrement est faisable via un site web et non via un serveur local ou via une carte à puce.

Il est à noter que les travaux connexes correspondant sont expliqués en détail dans la section de l'état de l'art dans le cinquième chapitre.

4. Objectifs de la thèse

Afin de résoudre les problèmes de sécurité dans les applications IoT, nous visons à proposer une solution en utilisant la biométrie pour chaque niveau mentionné ci-dessus, au niveau de l'IoT de bout-en-bout, au niveau d'accès au cloud, et au niveau S^2aaS basé sur l'IoT.

A) Solution proposée pour sécuriser l'IoT de bout en bout

Les principales contributions de notre travail à ce niveau peuvent être résumées comme suit:

- Utiliser la biométrie dans l'IoT de bout-en-bout pour l'authentification et pour le partage sécuritaire de la clé.
- Appliquer de nouveau crypto-système biométrique, combiné avec un protocole d'établissement de clé efficace dans un seul système pour économiser du temps et de l'énergie.
- Incorporer des codes correcteurs pour fournir une certaine tolérance aux erreurs.

B) Solution proposée pour sécuriser l'accès au cloud

Notre contribution pour sécuriser l'accès au cloud peut être résumée comme suit:

- Invertir le concept des crypto-systèmes biométriques pour sécuriser les transmissions et pour partager les clés plus tôt que sécuriser seulement la sauvegarde dans la base de données.
- Proposer des crypto-systèmes adaptés aux contraintes IoT basés sur des différents techniques d'extraction biométrique.
- Améliorer le processus d'authentification pour qu'elle soit plus légère et plus précise en :
    - ✓ Proposant une extraction appropriée par une nouvelle technique de quantification.
    - ✓ Utilisant des nombreuses modalités biométriques et des nombreuse codes correcteurs.

C) Solution proposée pour sécuriser S^2aaS

Notre contribution dans ce contexte peut se résumer dans les points suivants:

- Suggérer un schéma d'échange de clé simple, efficace et sécurisé pour le modèle S^2aaS basé sur l'IoT.
- Inclure tous les acteurs IoT possibles.
- Attribuer des tâches de la carte à puce aux terminaux des utilisateurs (généralement des smartphones). Par conséquent, ils peuvent accéder rapidement et en toute sécurité aux services publics via un simple site Web.
- Traiter le cloud comme un nœud non fiable, et tous les canaux de transmission ne sont pas sécurisés, car S^2aaS est un véritable environnement.
- Intégrer le partage de clé secrète dans le sous-système biométrique ainsi que l'authentification pour gagner du temps et de l'énergie.
- Eviter les processus lourds dans les intervalles critiques et fréquents.

5. Propriétés des techniques utilisées

Les solutions proposées reposent conjointement sur des technologies biométriques, des codes correcteurs et des techniques des cryptographies. Les solutions doivent être à faible coût de calcule tout en tenant en compte le niveau de sécurité requis. Pour cela, nos solutions répondent aux exigences suivantes:

- **Protocoles légers:** les solutions de sécurité doivent adopter des protocoles légers. Il existe des protocoles de communication spécifiquement conçus pour les réseaux sans fil à faible puissance qui comble les limitations des dispositifs IoT, telles que le faible débit, la faible puissance, la perte des paquets, et la transmission sur des courtes portées. Ces protocoles montrent leur efficacité dans des applications de vaste couverture comme les villes intelligentes [Khorov2015].
- **Petit vecteur de caractéristiques:** qu'il s'agisse d'une authentification biométrique, ou autre, les messages transmis sur le système IoT doivent être courts et de petite longueur.
- **Algorithmes légers:** Quel que soit l'algorithme utilisé, des algorithmes de chiffrement/dé chiffrement, des algorithmes d'extraction de caractéristiques biométriques, des crypto-systèmes biométriques, ou des codes correcteurs, tous doivent respecter les contraintes de l'IoT. Les algorithmes ne doivent inclure que des opérations simples avec de petites clés ; si l'algorithme est basé sur les clés. Cependant, ils doivent maintenir une sécurité élevée.

- **Défense collaborative (défense inter-domaines):** une solution collaborative où les différents domaines de l'IoT (Cloud, Brouillard, Perception, et Utilisateur) s'interagissent les uns avec les autres sera beaucoup plus efficaces qu'appliquer des contre-mesures à chaque domaine séparément.

6. Méthodes de validation

Pour valider nos solutions et vérifier l'utilité de nos systèmes nous avons suivi :

- Une analyse informelle pour montrer comment nos systèmes peuvent résister aux diverses attaques de transmission et du sauvegarde sans affecter l'agilité et la précision.
- Une analyse du coût de l'énergie et du coût de la communication.
- Une analyse formelle de la solution proposée pour S^2aaS par un outil de sécurité spécial (AVISPA) [AVISPA], par lequel les objectifs de sécurité souhaités sont vérifiés.
- Une validation des sous-systèmes biométriques en termes de temps de calcul et de taux de reconnaissance conventionnels ainsi que de nouveaux taux.
- Une comparaison avec d'autres travaux récents connexes en termes de sécurité, de coût de communication, et de coût de calcul et de certaines métriques pertinentes, pour montrer que nos schémas sont légers et offrent plus de protection.

7. Plan de la thèse

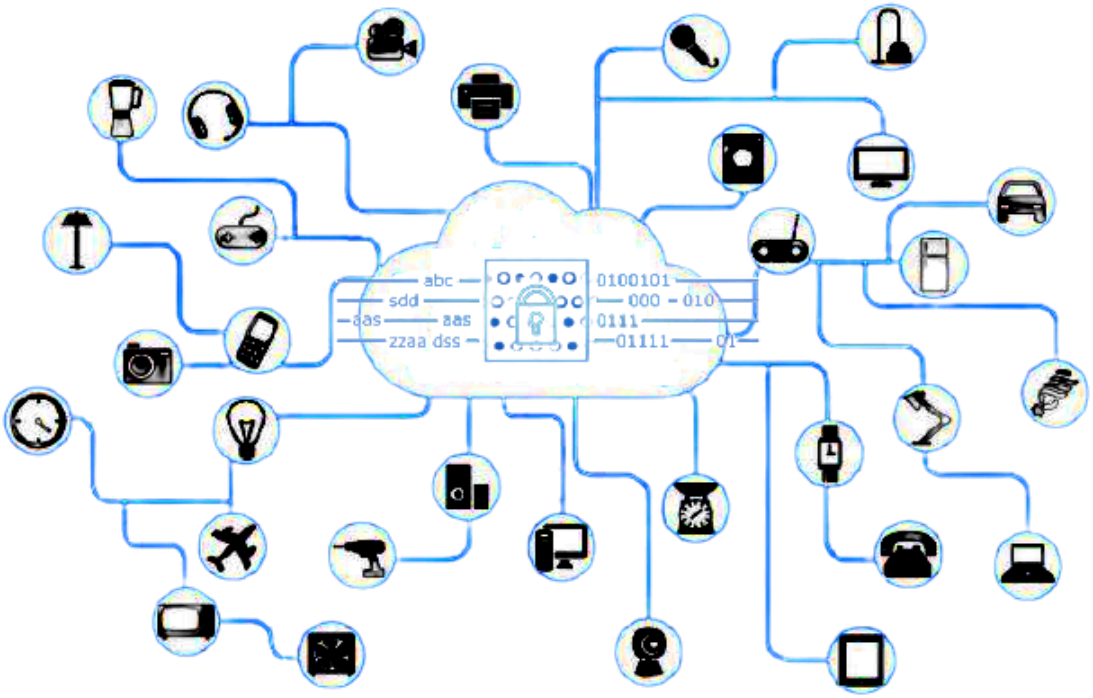
La thèse se compose de quatre chapitres après l'introduction générale et se termine par une conclusion générale comme suit:

- *L'introduction générale:* comprend le contexte de travail, les motivations de la recherche, les questions de départ et les problématiques incluant les lacunes des solutions associées, les objectifs souhaités, les solutions proposées, et un aperçu général sur les méthodes utilisées.
- *Le premier chapitre:* Ce chapitre est divisé en deux parties, la première est consacrée à tout ce qui concerne le paradigme IoT, son historique, la nature de ses éléments, ses architectures actuelles dont l'architecture S^2aaS. La deuxième partie concerne la sécurité IoT en général et en particulier, le modèle S^2aaS basé sur l'IoT.
- *Le deuxième chapitre:* Ce chapitre présente la technologie biométrique et son fonctionnement. Il fournit également les mécanismes de protection du vecteur de caractéristiques biométriques.
- *le troisième chapitre:* Ce chapitre traite de l'inclusion et de l'implication de la technologie biométrique dans l'IoT. Le chapitre est divisé en deux parties principales, la première présente

un état de l'art et les progrès menés dans cet axe. Tandis que la deuxième partie est un recueil de nos contributions à l'égard de l'intégration de la biométrie dans l'environnement IoT.

- *le quatrième chapitre:* Ce dernier chapitre est consacré aux résultats expérimentaux obtenus. Pour chacune de nos solutions, nous avons abordé les points suivants, la sélection des paramètres expérimentaux et de l'environnement, la discussion des résultats, la comparaison avec d'autres travaux récents pertinents, et un bref résumé des résultats.
- *Conclusion générale:* nous terminons cette thèse par une conclusion générale qui résume l'ensemble de nos travaux et décrit les perspectives attendues.

Chapitre 1 : L'Internet des Objets : Concepts, Défis, et Sécurité



Chapitre 1: L'internet des Objets : Concepts, Défis, et Sécurité

1.1 Introduction

L'intégration des capteurs (Sensors), actionneurs (actuators), et la radio-identification (*Radio Frequency IDentification* - RFID) dans presque toutes les entités physiques rends ces dernières connectables les unes aux autres et connectables à l'internet. Au fil du temps, le nombre de ces entités connectées (objets) augmente rapidement; il y a actuellement 30,7 milliards d'objets connectés, et d'ici 2025, il devrait y en avoir 75,44 milliards [Statista2020]. Les évolutions continues du marché électronique et les tendances technologiques au cours de la dernière décennie ont modifié la notion et la valeur des objets interconnectés. À savoir, la combinaison de matériel à faible coût et le réseau à haut débit ; à la fois filaires et sans fil, a permis la création d'une nouvelle génération de objets dotés d'une connectivité fiable et omniprésente sur l'ensemble de l'Internet. Le bon exemple c'est le G5.

Ces systèmes facilitent la récolte et l'échange de données en temps réel, et offrent une visibilité et un contrôle sans précédent des opérations et des processus. L'utilisation accrue de l'informatique basées sur le cloud a introduit des capacités d'analyse de données encore plus avancées, inaugurant une nouvelle ère de prise de décision, de contrôle et d'automatisation intelligents. En gros, ces nouveaux paradigmes sont appelés Internet des objets (*Internet of Things* - IoT) [Vans2011].

Dans l'IoT actuel, les données agrégées de divers objets interagissent au niveau du cloud pour fournir le service souhaité. Quand il y a plusieurs utilisateurs et une grande foule de propriétaires des objets, il y aura un grand nombre de services interconnectés. Par conséquent, le problème de l'inhomogénéité, de l'incohérence et de l'interopérabilité sera révélé. Cette faiblesse a été surmontée grâce à nouveau paradigme appelé *Sensig as a Service* (S^2aaS) [Perera2019]. Puisque S^2aaS présente des services dans un environnent ouvert sur le marché, alors il est plus vulnérable aux risques de différents cyberattaques par rapport à l'ancien paradigme IoT. Cela est dû à la diversité des ressources, des propriétaires et des services.

Ce chapitre est divisé en deux parties, la première est consacrée à tout ce qui concerne le paradigme IoT, son historique, la nature de ses éléments, ses architectures actuelles dont l'architecture S^2aaS. La deuxième partie concerne la sécurité IoT en général et en particulier, le modèle S^2aaS basé sur l'IoT.

1.2 Historique de l'internet des objets

En Californie, quelques mois seulement après que deux personnes ont mis le pied sur la lune pour la première fois, deux ordinateurs ont commencé à s'envoyer des messages en utilisant des protocoles conçus pour permettre à d'autres ordinateurs de se connecter facilement et de rejoindre la fête [Leiner2009]. Le 29 octobre 1969, un ordinateur en laboratoire à *University of California, Los Angeles* (UCLA) et un ordinateur en laboratoire au *Stanford Research Institute* (SRI) ont établi la première communication d'hôte à hôte dans ce qui allait devenir Internet [Savio2011]. *Vint Cerf* et deux collègues ont inventé le terme « Internet » comme une version abrégée de « *Inter-networking* » en décembre 1974 [Cerf1974]. Il n'a pas fallu longtemps pour que davantage d'ordinateurs et leurs périphériques, ainsi que davantage de réseaux d'ordinateurs et même d'équipements industriels se connectent et commencent à communiquer des messages.

Au début de 1982, une machine à soda est sans doute devenue le premier appareil connecté à Internet, annoncé par un courrier électronique largement diffusé qui partageait son histoire instrumentée et interconnectée avec le monde [Teicher2018]. Lorsque les ordinateurs coûtaient un million de dollars et que l'ARPANET (internet précédent) était encore la seule technologie de réseau en ville, *Kazar* a déclaré qu'un monde dominé par l'IoT semblait être un fantasme lointain. Il y avait une blague courante sur la façon dont votre grille-pain un jour va être sur l'internet, a-t-il déclaré, les gens ont ri de ça [Teicher2018]. En 1991, il était clair pour *Mark Weiser* que de plus en plus de choses auraient un jour des ordinateurs intégrés, y compris des téléphones portables, des voitures, même des poignées de porte et un jour même des vêtements [Weiser1991]. En 1999, *Kevin Ashton* est le premier à avoir inventé le terme « *Internet of Things* » [Zhang2020]. À ce stade, il considérait l'identification par radiofréquence (RFID) comme essentielle à l'Internet des objets, qui permettrait aux ordinateurs de gérer tous les objets individuels. Le concept principal de l'Internet des objets est d'intégrer des émetteurs-récepteurs mobiles à courte portée dans divers gadgets et nécessités quotidiennes pour permettre de nouvelles formes de communication entre les personnes et les objets, et même entre les objets eux-mêmes.

Aujourd'hui l'IoT continue de croître rapidement. En fait, IoT constitue la base de ce qui est devenu la quatrième révolution industrielle (4IR ou *Industry* 4.0[2]) et la transformation numérique des entreprises et de la société [Lee2014]. La première révolution industrielle était la machine à vapeur

[2] Industry 4.0 fait référence à la mise en réseau intelligente de machines et de processus pour l'industrie à l'aide des technologies de l'information et de la communication. Elles utilisent des systèmes de contrôle modernes, disposent de systèmes logiciels intégrés et disposent d'une adresse Internet pour se connecter et être adressée via l'IoT.

comme machine focale, la deuxième révolution comprenait les machines de production de masse, la troisième révolution était basée sur des machines avec des ordinateurs embarqués, et la quatrième révolution (aujourd'hui) des machines et des objets interconnectés, y compris des informations sur les matériaux et la consommation d'énergie entrant et sortant des systèmes cyber-physiques interconnectés à l'échelle mondiale.

1.3 Objets dans l'IoT

L'IoT est l'intersection d'Internet, des objets, et des données. Des processus et des normes ont également été ajoutés pour une définition plus complète de l'IoT. Les objets sont définis comme « Tout » ou même n'importe quoi, allant des appareils électroménagers aux bâtiments, aux voitures, aux personnes, aux animaux, aux arbres, aux plantes, etc.

La condition préalable aux objets IoT est la capacité de communiquer via Internet. Les objets dans l'IoT doivent être identifiables de manière unique au moins durant une session avec la capacité de capturer et d'actionner convenablement. Nous distinguons différents types d'objets connectés à l'IoT selon leur rôle, que nous mentionnons principalement, capteurs, RFID et actionneurs. En fait, les objets sont les éléments dans lesquels des capteurs, des identifiants, des actionneurs, des logiciels ou une connectivité Internet sont embarqués. Cependant, il est de coutume de se référer aux capteurs et aux actionneurs comme des objets eux-mêmes.

1.3.1 *Capteurs*

Un capteur est un dispositif (généralement électronique) qui détecte des événements ou des changements dans son environnement physique (par exemple, la température, le son, la chaleur, la pression, le débit, le magnétisme, le mouvement et les paramètres chimiques et biochimiques) et fournit une sortie correspondante sous forme électrique. Nous pouvons le définir tout simplement comme un convertisseur phénomène physique/électrique.

Les capteurs peuvent être très simples avec une fonction principale pour collecter et transmettre des données, ou intelligents en fournissant des fonctionnalités supplémentaires pour filtrer les données dupliquées et notifier la passerelle IoT uniquement lorsque des conditions très spécifiques sont remplies. Cela ne nécessite qu'une logique de programmation soit présente sur le capteur lui-même.

Les capteurs peuvent récolter de grandes quantités de données à tout moment et de n'importe quel endroit, et les transmettre sur un réseau IoT (*IoT Network*) en temps réel. Les données sont ensuite analysées et éventuellement corrélées avec autres bases de données pour fournir des informations commerciales utiles ou une meilleure connaissance de l'environnement. Cela peut créer des opportunités et/ou des gains d'efficacité et de productivité. Dans certains cas, les sorties des

capteurs sont traitées directement au niveau de capture lui-même par une application légère (*Lightweight Application*).

Les capteurs IoT devraient avoir de nombreuses caractéristiques telles que compactibilité, haute sensibilité, détection intelligente, filtrage des données, limitation de bruit, minimisation de l'interruption, etc. Peut-être le plus important, les capteurs doivent avoir une consommation électrique minimale, car plusieurs facteurs déterminent les exigences de faible consommation d'énergie dans l'IoT, example, des capteurs pour plusieurs verticaux IoT sont installés dans des endroits difficiles à atteindre pour remplacer les batteries. Autres soucis peuvent inclure le stockage de données, la capacité de calcul et l'auto-avertissement des symptômes anormaux [Rayes2016].

1.3.2 *Identification par radiofréquence*

RFID n'est pas un capteur mais un mécanisme permettant de capturer des informations pré-embarquées dans ce que l'on appelle l'étiquette d'un objet (*Tag*) à l'aide d'ondes radio. RFID se compose de deux parties : une étiquette et un lecteur. De plus, l'étiquette comporte deux parties : une micro-puce qui stocke et traite les informations, et une antenne pour recevoir et transmettre un signal. L'étiquette contient le numéro de série spécifique d'un objet spécifique. Le lecteur lit les informations codées sur une étiquette, à l'aide d'un émetteur-récepteur radio bidirectionnel, en émettant un signal vers l'étiquette à l'aide d'une antenne. L'étiquette répond avec les informations écrites dans sa mémoire. Le lecteur transmettra ensuite les résultats de la lecture à un programme informatique.

Il existe deux types d'étiquettes (*Tags*) RFID :

- Les étiquettes passives qui sont alimentées par l'énergie des ondes radio d'interrogation du lecteur RFID.
- Les étiquettes actives qui sont alimentés par une batterie et peuvent donc être lus à une plus grande distance du lecteur RFID, jusqu'à des centaines de mètres.

Contrairement à un code-barres, l'étiquette n'a pas besoin d'être dans la ligne de mire du lecteur, elle peut donc être intégrée dans l'objet suivi. Cependant, comme toute autre technologie, la RFID présente un certain nombre d'inconvénients, mais l'inconvénient potentiel est l'interférence entre plusieurs lecteurs et étiquettes si le système global n'est pas configuré de manière appropriée. Chaque lecteur RFID scanne essentiellement toutes les étiquettes qu'il récupère dans sa portée. Cela peut créer une confusion entre les informations d'étiquette (par exemple, facturer un client pour des articles dans le panier de quelqu'un d'autre dans la même portée.

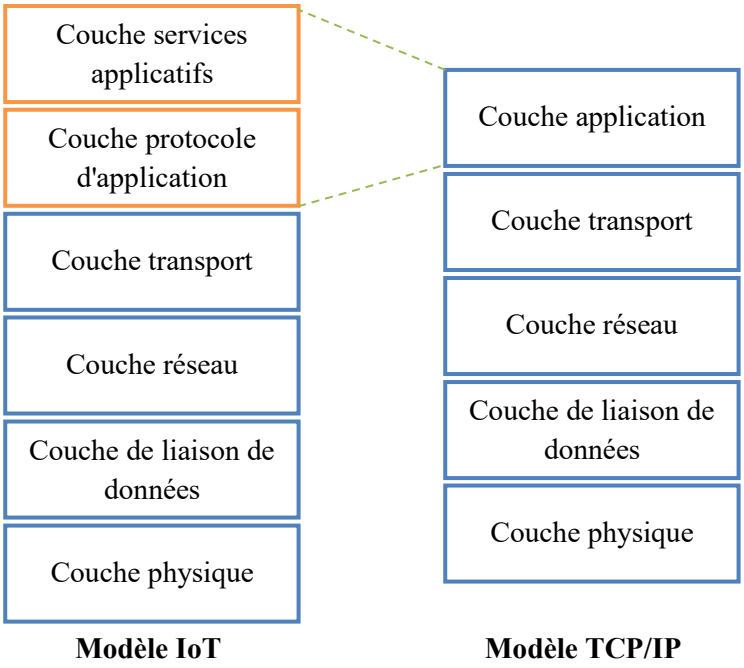


Figure 1.1– Pile de protocoles IoT (Inspiré de [Rayes2016a]).

La RFID est déjà utilisée par un grand nombre d'applications. Les meilleurs exemples sont les suivants, contrôle et gestion d'accès, commerce, transport et logistique, passeports, paiements de transport, suivi des animaux, implantation humaine, sport, et suivi des participants à la conférence [Rayes2016].

1.3.3 *Actionneurs*

La récolte et l'affichage de données par un système de surveillance sont inutiles à moins que ces données ne soient traduites en informations pouvant être utilisées pour contrôler ou gouverner un environnement avant qu'un service ne soit affecté. Les actionneurs utilisent des données récoltées et analysées par les capteurs pour faire l'action appropriée. Par exemple, l'arrêt du débit de gaz lorsque la pression mesurée est inférieure à un certain seuil.

Un actionneur est un type de moteur chargé de contrôler ou d'agir dans un système. Il prend une source de données ou d'énergie (par exemple, la pression du fluide hydraulique, d'autres sources d'énergie) et convertit les données/énergie en mouvement pour contrôler un système.

Il existe plusieurs types d'actionneurs :

- Actionneurs électriques sont des dispositifs entraînés par de petits moteurs qui convertissent l'énergie en couple mécanique.
- Actionneurs linéaires mécaniques : Les actionneurs mécaniques convertissent le mouvement rotatif en mouvement linéaire.

- Actionneurs hydrauliques : Les actionneurs hydrauliques sont des dispositifs simples avec des pièces mécaniques qui sont utilisées sur des vannes linéaires ou les quarts de tour. Ils sont conçus en basant sur la loi de Pascal : Lorsqu'il y a une augmentation de la pression en tout point d'un fluide incompressible confiné, alors il y a une augmentation égale en chaque point du récipient.
- Actionneurs pneumatiques : Les actionneurs pneumatiques fonctionnent sur le même concept que les actionneurs hydrauliques, sauf que le gaz comprimé est utilisé à la place du liquide.

Les actionneurs peuvent aussi être classifiés en actionneurs manuels et actionneurs automatiques. Les actionneurs manuels utilisent des leviers, des engrenages ou des roues pour permettre le mouvement demandé, tandis qu'un actionneur automatique dispose d'une source d'alimentation externe pour faire fonctionner un mouvement automatiquement.

1.4 Paradigmes d'interaction d'objets IoT

La pile des protocoles IoT peut être visualisée comme une extension du modèle de protocole en couches TCP/IP et est composée des couches suivantes (Voir Figure 1.1), couche physique, couche de liaison de données, couche internet, couche transport, couche de protocoles d'application et couche de services d'application [Rayes2016a]. Notez que la Couche Application de la pile des protocoles TCP/IP est étendue en deux couches dans la pile de protocoles IoT: protocole d'application et services d'application. Dans cette section, nous nous concentrons davantage sur la couche de protocole d'application. Parce qu'il différencie les protocoles IoT des autres protocoles TCP/IP d'une part, et fait partie de notre travail d'autre part.

Les protocoles d'application sont chargés de gérer la communication entre les entités d'application, c'est-à-dire les objets ou les passerelles, et les applications. Ils prennent généralement en charge le

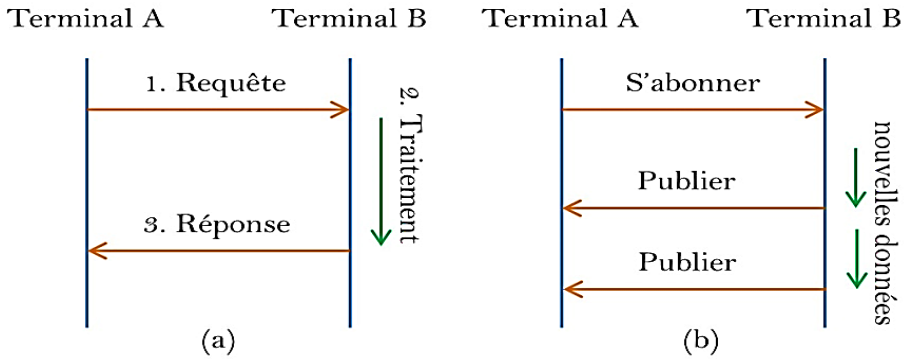


Figure 1.2 – Les différents paradigmes d'interaction, (a) Le paradigme Requête/Réponse, (b) Le paradigme Publier/S'abonner (Inspiré de [Rayes2016a]).

flux de données (par exemple, des lectures ou des mesures) des objets vers les applications et le flux d'informations de commande ou de contrôle (par exemple, pour déclencher ou actionner des dispositifs terminaux) dans le sens inverse. Ces protocoles définissent la sémantique et les mécanismes d'échange de messages entre les extrémités communicantes.

Les protocoles d'application prennent en charge différents modèles de communication. Ces modèles permettent différents paradigmes d'interaction entre les applications et les dispositifs IoT. Les deux paradigmes célèbres sont, Requête/Réponse et Publier/S'abonner (Voir Figure 1.2).

Le paradigme Requête/Réponse (*Request/Response*) permet une communication bidirectionnelle entre les terminaux. L'initiateur de la communication envoie un message de requête, qui sera reçu et exploité par le point terminal destinataire. Ce dernier envoie alors un message de réponse à l'initiateur d'origine. Ce paradigme est adapté aux déploiements IoT qui présentent une ou plusieurs des caractéristiques suivantes :

- Le déploiement suit une architecture client-serveur.
- Le déploiement nécessite une communication interactive : les deux terminaux ont des informations à envoyer à l'autre côté.

Cependant, tous les déploiements IoT n'ont pas les caractéristiques ci-dessus. En particulier, dans de nombreux scénarios, tout ce qui est requis est une communication unidirectionnelle d'un producteur de données (par exemple, un capteur) à une entité consommatrice (l'application).

Pour cela, le paradigme Requête/Réponse est sous-optimal en raison de la surcharge des messages inutiles s'exécutant dans le sens inverse. C'est là que le modèle de Publier/S'abonner (*Publish/Subscribe*) s'intervient.

Le paradigme Publier/S'abonner, souvent appelé Pub/Sub, permet une communication unidirectionnelle d'un publieur à un ou plusieurs abonnés. Les abonnés déclarent leur intérêt pour une classe ou une catégorie particulière de données au publieur. Lorsque le publieur dispose de nouvelles données à partir de cette classe, il les envoie dans des messages aux abonnés intéressés. Ce paradigme est optimal pour les applications IoT qui nécessitent une communication unidirectionnelle. Ainsi que, le modèle Pub/Sub est bien adapté aux déploiements IoT qui peuvent bénéficier des caractéristiques suivantes :

- Couplage lâche entre les terminaux de communication, en particulier par rapport au modèle client-serveur.
- Meilleure évolutivité en tirant parti du parallélisme et des capacités de multidiffusion du réseau de transport sous-jacent.

1.5 Architecture IoT

Avant d'introduire l'architecture selon les domaines IoT, nous allons décrire brièvement les composants clés des solutions IoT qui sont constitués d'entités de dispositifs IoT (*IoT devices*), d'entités de réseau IoT (*IoT network elements*), d'une plate-forme de services IoT (*IoT Services Platform*) et d'entités d'applications IoT (*IoT Applications*) [Rayes2016b].

- **Entités de dispositifs IoT:** Les dispositifs IoT comprennent les dispositifs de capture, les actionneurs et les passerelles. Les principales fonctions des passerelles sont (I) la récolte de données et la collecte des informations des dispositifs IoT, (II) le filtrage et simple corrélation des informations collectées, (III) le transfert des données corrélées vers la couche réseau, et (IV) la prise de mesures sur les dispositifs (par exemple, la mise hors tension) en fonction des commandes des couches supérieures.
- **Entités de réseau IoT:** Les entités du réseau IoT fournissent des services du réseau sous-jacent à la plate-forme de services. Ils comprennent des super-passerelles, des routeurs d'accès, des commutateurs et éventuellement des serveurs de gestion d'éléments avec des fonctions de gestion de réseau spécifiques.
- **Plateforme de services IoT:** La plateforme de services IoT est parfois appelée « plateforme IoT » ou « la plate-forme de services d'application IoT » de toute solution IoT. Il est responsable de la surveillance et du contrôle des éléments IoT dans la couche de dispositifs IoT et de réseau. Il est également responsable de la création d'une intégration directe entre les dispositifs physiques (par exemple, les capteurs, les actionneurs, les passerelles) et les systèmes d'applications informatiques pour améliorer l'efficacité, la précision et les avantages économiques. La plateforme de services IoT reçoit des informations des dispositifs IoT et des entités réseau et fournit des services aux entités d'application. Plus important encore, il fournit des fonctions de gestion au niveau du réseau et souvent au niveau du service.
- **Entités d'application IoT:** Les entités d'application reçoivent des informations de la plateforme de services et fournissent des services et des fonctions de niveau commercial. Ces fonctions dépendent généralement du type de domaine d'application.

Basé sur les éléments mentionnés ci-dessus, on peut dire qu'il existe deux principales architectures IoT qui sont, architecture IoT à trois domaines et architecture IoT à quatre domaines.

1.5.1 *Architecture IoT à trois domaines*

Cette architecture se compose de trois domaines qui sont, le domaine de la perception (*sensing domain*), le domaine réseau (*network domain*), et le domaine d'utilisateur (*user domain*). Cette

architecture est également appelée architecture IoT de bout en bout (*End-to-End*). Le composant manquant dans cette architecture est le cloud, où les objets du domaine perception sont utilisés, surveillés, contrôlés et gérés directement depuis l'application utilisateur [Kumar2017].

Cette architecture semble plus adaptée au réseau de capteurs sans fil (*Wireless Sensor Networks* - WSNs) lorsque l'utilisateur communique directement avec leurs propres objets à travers la passerelle responsable (*Sink node* dans *WSN, IoT gateway* dans *IoT*). Cependant, cela ne semble pas être le cas pour l'IoT basé sur les services. La raison en est que le concept de service n'est pas pris en charge malgré qu'il soit la pierre angulaire du paradigme IoT modern.

1.5.2 *Architecture IoT à quatre domaines:*

Aussi appelé « *cloud-based IoT*» [Al-Turjman2020], comme illustré dans la Figure. 1.3, l'architecture est composée des quatre domaines suivants, domaine de la perception (*sensing domain*), domaine du brouillard (*fog domain*), domaine de cloud (*cloud domain*), et domaine d'utilisateur (*user domain*).

- **Domaine de la perception :** Ce domaine comprend une partie des dispositifs IoT, et est composé de tous les objets intelligents qui ont la capacité de percevoir le milieu environnant et de rapporter les données récoltées à l'un des nouds du domaine du brouillard. De l'autre côté, ils reçoivent des commandes pour exécuter une action donnée comme dans le cas des actionneurs. Dans certains cas, les objets intelligents peuvent prendre des décisions et exécuter des actions automatiquement, comme pulvériser de l'eau dans le cas d'incendie. Les objets intelligents dans ce domaine devraient changer d'emplacement au fil du temps.
- **Domaine du brouillard:** ou « *fog domain* » en anglais, aussi appelé « domaine géodistribué». Ce domaine se compose d'un ensemble de dispositifs de brouillard situés dans des zones très peuplées de nombreux objets intelligents. Chaque dispositif de brouillard se voit attribuer un

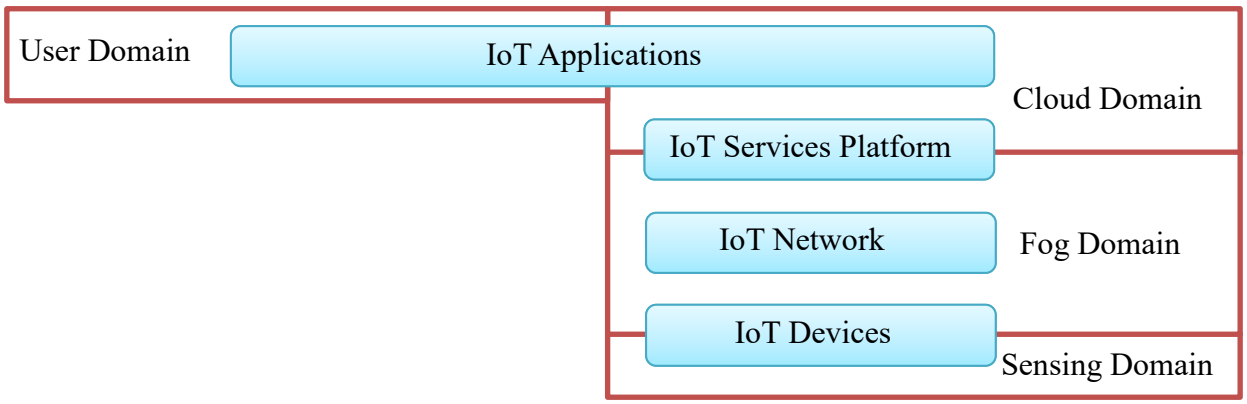


Figure 1.3– Architecture IoT à quatre domaines.

ensemble d'objets intelligents où les objets attribués rapportent leurs données capturées au dispositif de brouillard. Le dispositif de brouillard effectue des opérations sur les données récoltées, notamment l'agrégation, le prétraitement et le stockage. Les dispositifs de brouillard sont également connectés les uns aux autres afin de gérer la communication entre les objets intelligents et afin de coordonner quel dispositif de brouillard sera responsable de la gestion de quel objet car les objets changent d'emplacement au fil du temps. Chaque dispositif de brouillard est également connecté à un ou plusieurs serveurs dans le domaine de cloud. certains objets intelligents peuvent joués le rôle d'un nœud de brouillard comme le montre la Figure 1.4, s'ils ont suffisamment de capacité de calcul, de stockage, ou d'énergie.

- **Domaine de Cloud :** Ce domaine est composé d'un grand nombre de serveurs qui hébergent les applications chargées d'effectuer les opérations de traitement informatique lourdes sur les données rapportées par les dispositifs de brouillard. Des applications de calcul léger peuvent également résider dans le domaine du brouillard. Grâce à cette couche, les informations agrégées de différents nœuds de brouillard sont coopérées pour fournir un service public harmonieux.
- **Domaine d'utilisateur :** On entend par utilisateur chaque consommateur de données via une application IoT installée dans son terminal. Ce ne sont pas seulement l'humain et son terminal qui constituent le nœud utilisateur, mais tout dispositif qui consomme les données collectées à partir des dispositifs IoT.

Bien que le domaine du brouillard soit très similaire au domaine du cloud, il existe trois différences clés qui distinguent les entités de brouillard des serveurs cloud, (I) *Emplacement:* contrairement aux serveurs cloud qui sont généralement situés loin des objets intelligents, les dispositifs de brouillard sont placés à proximité des objets intelligents. (II) *Mobilité:* étant donné que l'emplacement de l'objet intelligent peut changer au fil du temps, les Machines virtuelles créées pour gérer ces objets dans le domaine du brouillard doivent être déplacées d'un dispositif de brouillard à un autre. (III) *Capacité de calcul inférieure*: les dispositifs de brouillard qui sont installés dans un certain emplacement devraient avoir une capacité de calcul inférieure par rapport aux capacités offertes par le cloud.

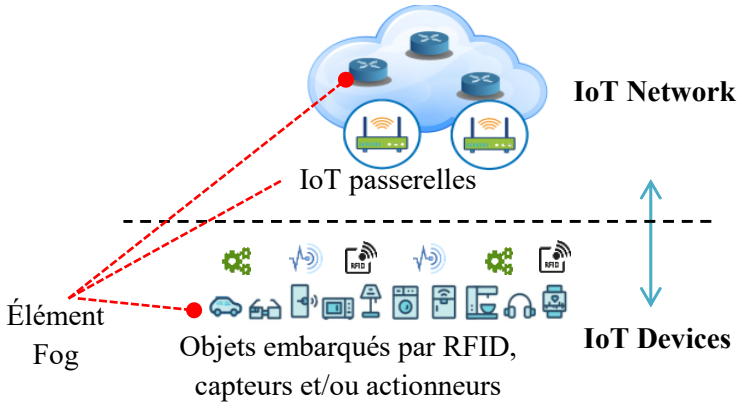


Figure 1.4 – Les objets que peuvent être joués le rôle d'un nœud du brouillard (*fog node*) (Inspiré de [Rayes2016c]).

1.6 Perception en tant que Service

Dans l'IoT actuel, les données collectées par différentes solutions IoT coincées dans des silos de données séparés, chaque propriétaire de données n'aura accès qu'à ses propres donnés à partir de plusieurs solutions IoT .Pour cela, il n'y a aucun moyen de partager ces données avec un tiers. Le téléchargement de ses propres données personnelles n'a aucune valeur pour les propriétaires de données à moins qu'il n'existe un moyen d'analyser et d'en extraire des informations utiles, qui peuvent être utilisées pour améliorer nos vies (y compris les comportements, les habitudes, les modes de vie, la consommation de ressources, etc.). C'est une tâche très difficile et très longue pour un propriétaire non technicien, même pour un expert, d'analyser et d'extraire des informations utiles à partir de données brutes. Du point de vue de l'analyse des données, afin d'effectuer des analyses avancées et d'extraire des informations utiles, les données d'un grand nombre de propriétaires de données doivent être traitées et analysées ensemble. Le modèle S^2aaS basé sur l'IoT intervient pour apporter des solutions à ces problèmes grâce à une architecture intégrée à quatre couches [Perera2013], (voir Figure 1.5), (I) les capteurs et les propriétaires « *Sensors and Sensor Owners Layer* ». (II) Publieurs de données de capteurs « *Sensor Data Publishers Layer* - SP». (III) Fournisseurs de services étendus «*Extended Service Providers Layer* - ESP». (IV) Consommateurs de données de capteurs « *Sensor Data Consumers Layer* ».

1.6.1 *Couche des capteurs et leurs propriétaires*

Cette couche se compose de capteurs et de propriétaires de capteurs. Comme nous l'avons vu précédemment dans le domaine de la perception, nombreux capteurs peuvent être attachés à un objet ou un dispositif qui percept, mesure ou capture un phénomène physique. Ces informations peuvent être utilisées pour mieux comprendre le comportement et les préférences des utilisateurs. De plus, ils ont la capacité d'envoyer des données vers le cloud. D'autre part, le propriétaire d'un

certain capteur à tout moment peut changer au fil du temps. La Couche des capteurs et les propriétaires est, plus ou moins, logiquement équivalent à la couche Perception de IoT conventionnel, La principale différence ici est que les propriétaires de capteurs sont ajoutés à cette couche dans le modèle S^2aaS.

1.6.2 *Couche des Publieurs de données de capteurs*

Les publieurs de données de capteurs (SP) sont des entités commerciales distinctes. La principale responsabilité de ces entités est de détecter les capteurs disponibles, de communiquer avec les propriétaires de capteurs et d'obtenir l'autorisation de publier les données dans le cloud. Lorsqu'un propriétaire de capteur enregistre un capteur spécifique, SP collecte des informations sur la disponibilité du capteur, les préférences et les restrictions du propriétaire, le retour attendu, etc. Toutes ces informations doivent être publiées dans le cloud. Une fois l'enregistrement effectué, un SP attend qu'un consommateur de données fasse une demande. Lorsqu'un SP reçoit une telle demande, il transmet tous les détails, y compris l'offre, au propriétaire de capteur correspondant pour l'acceptation ou le rejet. Si le propriétaire accepte l'offre, le consommateur correspondant pourra acquérir des données de ce capteur via le SP pendant la période mentionnée dans le contrat (offre).

1.6.3 *Couche des Fournisseurs de services étendus*

Cette couche peut être considérée comme la plus intelligente parmi les quatre couches qui intègrent l'intelligence à l'ensemble du modèle de service. Les services fournis par les ESP peuvent être très variés d'un fournisseur à l'autre. Cependant, il existe certaines caractéristiques fondamentales des ESP. Pour devenir un ESP, ils doivent fournir des services à valeur ajoutée [Nesse2013] aux consommateurs. Cependant, dans certains cas, une seule entité commerciale peut remplir à la fois les rôles de SP et deESP. Chaque SP a accès (uniquement) aux capteurs qui lui sont enregistrés. ESP peut être utilisés pour acquérir facilement des données pour un consommateur qui a besoin de données proviennent de plusieurs capteurs, où chaque capteur a été enregistré avec différents SP. Les ESP communiquent avec plusieurs SP concernant l'acquisition de données pour le compte du consommateur. SP et ESP constituent ensemble l'équivalent logique, dans une certaine mesure, de la couche Cloud de l'IoT conventionnel.

1.6.4 *Couche des Consommateurs de données de capteurs*

Tous les consommateurs ont besoin pour s'enregistrer et obtenir un certificat numérique valide d'une autorité afin de consommer les données d’un capteur. Les consommateurs peuvent être des gouvernements, des organisations commerciales, des institutions universitaires, des communautés

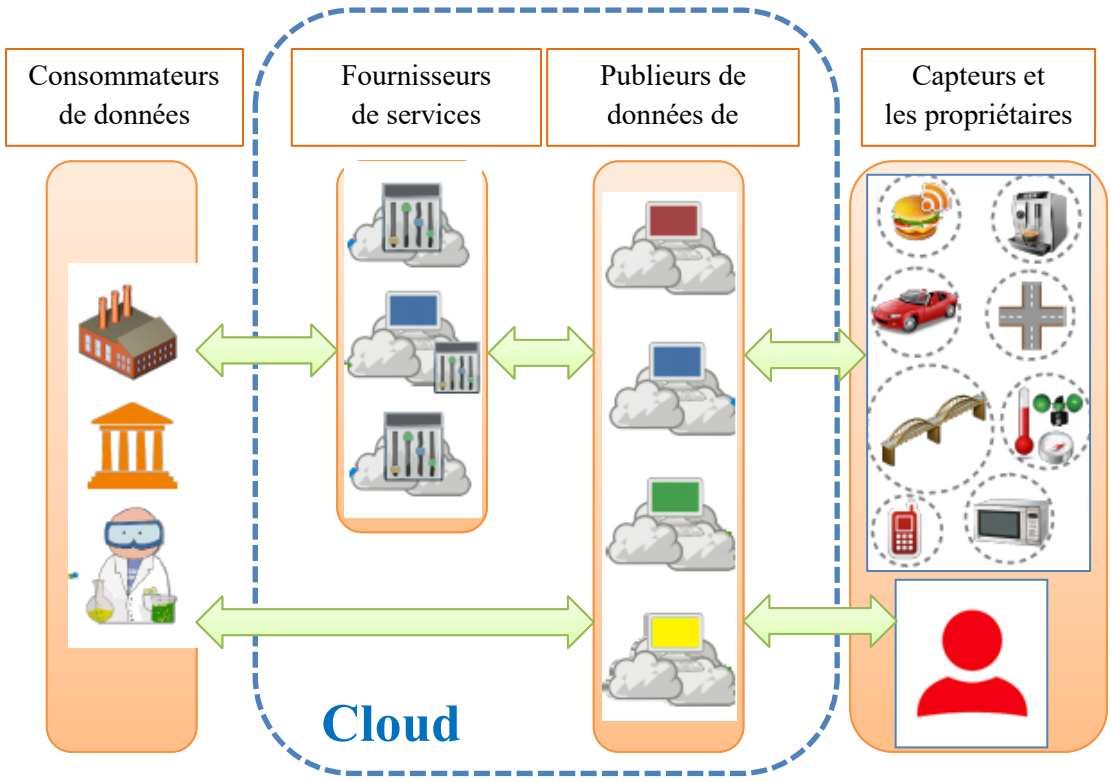


Figure 1.5- Le modèle de la perception en tant que service (S^2aaS) (Inspiré de [Perera2013]).

de recherche scientifique, etc. Les consommateurs ne communiquent pas directement avec les capteurs ou les propriétaires de capteurs. Toutes les communications et transactions doivent être effectuées via des SP ou des ESP. Notes que, le consommateur peut acquérir directement des données de SP si seulement il a la capacité technique requise. Les consommateurs ont moins de capacités techniques et d'expertise peuvent acquérir les données de capteurs requises via des ESPs où la plupart des tâches difficiles telles que la combinaison des données de plusieurs SPs et la sélection de capteurs appropriés en fonction des exigences des consommateurs sont traitées. Les consommateurs de données de capteurs est équivalent au domaine utilisateur d'IoT conventionnel.

1.7 Les défis de la sécurité IoT

L'IoT, en particulier le modèle S^2aaS, présente des caractéristiques et des contraintes uniques lorsqu'il s'agit de concevoir des mécanismes de défense efficaces contre les menaces de cyber sécurité qui peuvent se résumer par les éléments suivants [Ghorbani 2017], [Perera2013]:

- **Interopérabilité des services**: dans le modèle S^2aaS, le service public peut être constitué de différents fournisseurs de services et de différentes entités (SPs et ESPs). L'accès à ces services publics se fait dans un environnement non sécurisé avec la plupart des canaux de transmission incertains, pour cela ce modèle devient plus vulnérable aux cyberattaques différentes. De plus, ces attaques peuvent être de différents niveaux (par exemple, stockage et transmission).

- **Technologies multiples**: L'IoT combine plusieurs technologies telles que RFID, WSNs, l'informatique en nuage (*Cloud Computing)*, la virtualisation, et la création des services. Chacune de ces technologies a ses propres vulnérabilités. Le problème avec le paradigme IoT est qu'il faut sécuriser la chaîne de toutes ces technologies car la sécurité d'une application IoT sera jugée en fonction de son point le plus faible qui est généralement désigné par le talon d'Achille.
- **Applications multiples:** Le paradigme IoT aura de nombreuses applications (également appelées verticales) qui couvrent la e-santé, l'industrie, les gadgets pour la maison intelligente, les villes intelligentes, etc. Les exigences de sécurité de chaque verticale sont assez différentes des autres verticales.
- **Évolutivité (*Scalability*):** Le nombre énorme des dispositifs connecté a l'internet fait de l'évolutivité un défi important lorsqu'il s'agit de développer des mécanismes défensifs efficaces. Aucun des cadres défensifs centralisés proposés précédemment ne peut plus fonctionner avec le paradigme IoT, où l'accent doit être mis sur la recherche de mécanismes de sécurité défensive décentralisés pratiques. Une solution IoT doit évoluer de manière rentable, potentiellement jusqu'à des centaines de kilomètres, voire des millions de terminaux.
- **Big Data :** non seulement le nombre d'objets intelligents sera énorme, mais les données générées par chaque objet seront également énormes, car chaque objet intelligent devrait être fourni par de nombreux capteurs, chaque capteur générant d'énormes flux de données au fil du temps. Il est donc essentiel de mettre en place des mécanismes défensifs efficaces capables de sécuriser ces grands flux de données.
- **Disponibilité:** La disponibilité fait référence à la caractéristique d'un système ou d'un sous-système qui est continuellement opérationnel pendant une période de temps souhaitable. Il est généralement mesuré par rapport à « 100 % opérationnel » ou « jamais en panne ». Une norme de disponibilité largement répandue mais difficile à atteindre pour un système ou un produit est connue sous le nom de « cinq 9 » (disponible 99,999% du temps au cours d'une année donnée). La sécurité joue un rôle majeur dans la haute disponibilité, car les administrateurs réseau hésitent souvent à utiliser les fonctions technologiques de réponse aux menaces nécessaires de peur que ces fonctions ne mettent hors service des systèmes critiques. Dans certains cas, les administrateurs réseau préféreraient ne pas avoir de protection de cyber-sécurité plutôt que de risquer une panne en raison d'un faux positif. Cela les rend aveugles aux menaces au sein de leurs réseaux de contrôle. Les entreprises ajoutent souvent de la redondance à leurs systèmes afin que la défaillance d'un composant n'affecte pas l'ensemble du système.

- **Limitations des ressources :** C'est peut-être le plus grand défi de l'IoT. Parce que la majorité des terminaux IoT ont des capacités de ressources limitées telles que le processeur, la mémoire, le stockage, la batterie et la plage de transmission. Cela fait de ces dispositifs un fruit à portée de main pour les attaques par déni de service (*Denial-of-Service*- DoS) où l'attaquant peut facilement dépasser les capacités de ressources limitées de ces dispositifs, provoquant une interruption de service. En plus de cela, les limitations de ressources de ces dispositifs soulèvent de nouveaux défis lorsqu'il s'agit de développer des protocoles de sécurité en particulier avec le fait que les techniques de cryptographie traditionnelles et matures sont connues pour être coûteuses en calcul.
- **Emplacements distants :** dans de nombreuse applications de l'IoT (par exemple, les réseaux intelligents, les chemins de fer, les bords de routes), les dispositifs IoT, généralement des capteurs, seront installés dans des emplacements sans personnel difficiles à atteindre. Les attaquants peuvent interférer avec ces dispositifs sans être vus. Les systèmes de surveillance de la cyber-sécurité et de la sécurité physique doivent être installés dans un endroit protégé, fonctionner dans des conditions environnementales extrêmes, s'adapter à de petits espaces et fonctionner à distance pour les mises à jour et la maintenance de routine, évitant ainsi les visites retardées et coûteuses des techniciens du réseau.
- **Mobilité :** on s'attend à ce que les objets intelligents changent souvent leurs emplacements dans le paradigme IoT. Cela ajoute des difficultés supplémentaires lors du développement de mécanismes défensifs efficaces dans de tels environnements dynamiques.
- **sensibilité au délai :** la majorité des applications IoT sont censées être sensibles au délai, et il convient donc de protéger les différents composants IoT de toute attaque susceptible de dégrader leur temps de service ou de provoquer une interruption de service.

Pour relever ces défis, Les solutions de sécurité pour le modèle IoT doivent répondre aux exigences suivantes,

- **Protocoles légers:** les solutions de sécurité doivent être adoptées des protocoles légers. Il existe des protocoles de communication conçus spécifiquement pour les réseaux sans fil à faible puissance qui traitent les limitations des dispositifs IoT, telles que le faible débit, la faible puissance, la perte, et la transmission sur de courtes distances (voire Section 1.4). Ces protocoles démontrent leur efficacité dans des applications dans de vastes zones telles que les villes intelligentes [Khorov2015].
- **Petit vecteur de caractéristiques:** qu'il s'agisse d'une authentification biométrique, ou autre, les messages transmis sur le système IOT doivent être courts et de faible longueur. Par

conséquent, les méthodes et les algorithmes d'extraction des caractéristiques à partir d'objets, en particulier la biométrie humaine, devraient donner un vecteur de caractéristiques de petite longueur adapté au système IoT.

- **Algorithmes légers:** Quel que soit l'algorithme utilisé, soit des algorithmes de chiffrement/déchiffrement, de transformation des vecteurs, du masquage des données, de partage des clés, d'extraction de caractéristiques biométriques, des cryptosystèmes biométriques, voire des codes correcteurs doivent respecter les contraintes de l'IoT. Les algorithmes ne doivent inclure que des opérations simples avec de petites clés; si l'algorithme basé sur les clés. Cependant, ils doivent maintenir une sécurité élevée.
- **Surveillance et contrôle à distance:** comme les objets IoT sont distribués dans des lieux publics ou inhabités tels que les routes, les voies ferrées, et les fermes, ils sont facilement accessibles aux attaquants, en particulier par contact physique. Pour éviter ce risque, l'hôte de la sécurité doit être placé dans un endroit sûr et éloigné des mains des attaquants, ou dans un endroit avec des conditions de sécurité particulières. ces hôtes mettent en œuvre les tâches de surveillance, de contrôle, de mise à jour des protocoles et de mise à jour des mécanismes de sécurité, la législation sur la politique de sécurité, et tout ce qui concerne le bon comportement du système [Kim2016].
- **Défense collaborative (défense inter-domaines):** une solution collaborative où les différents domaines (Cloud, Brouillard, Perception, et Utilisateur) interagissent les uns avec les autres sera beaucoup plus efficaces que d'appliquer des contre-mesures à chaque domaine séparément, où les différents domaines peuvent interagir et collaborer pour arrêter toute activité malveillante en cours.

1.8 Objectifs de sécurité IoT

Dans cette section, les objectifs de sécurité de l'IoT sont résumés. Ces objectifs comprennent [Stallings2014]:

- **Authentification mutuelle:** elle garantit que toute entité impliquée dans une opération est bien celle qu'elle prétend être. Nombreuses cyberattaques ciblent cette exigence lorsqu'une entité prétend être une autre identité.
- **Confidentialité:** Elle garantit que les messages échangés ne peuvent être compris que par les entités visées.
- **Intégrité:** elle garantit que les messages échangés n'ont pas été altérés/falsifiés par un tiers.

- **Autorisation:** appelée également contrôle d'accès, elle garantit que les entités disposent des autorisations de contrôle requises pour effectuer l'opération qu'elles demandent d'effectuer.
- **Non-répudiation:** Elle garantit qu'une entité ne peut pas nier une action qu'elle a effectuée. cette répudiation se fait généralement dans la cryptographie symétrique où tout participant possède la même clé.
- **Disponibilité:** Il garantit que le service n'est pas interrompu. Les attaques DoS ciblent cette exigence car elles provoquent une interruption de service.
- **Récence de système (*Freshness*):** garantit que les données sont à jour. Les cybers attaques ciblent cette exigence lors du renvoi d'un ancien message afin de restaurer une entité à son ancien état.
- **Confidentialité persistante:** Elle garantit que lorsqu'un objet quitte le réseau, il ne comprendra pas les communications qui sont échangées après son départ, cette exigence dans ce cas appelée « *Forward Secrecy* ». La confidentialité persistante garantit également que tout nouvel objet qui rejoint le réseau ne sera pas en mesure de comprendre les communications qui ont été échangées avant de rejoindre le réseau, cette exigence dans ce cas appelée «*Backward Secrecy*» [Dabbagh2016].

1.9 Attaqués et contre-mesures

Dans cette section, nous aborderons les attaques connues dans les quatre domaines suivants: Cloud, Brouillard, Perception, et Utilisateur.

1.9.1 *Domaine Cloud*

Dans le modèle S^2aaS, Le domaine de cloud est présenté par le fournisseur de service. Le fournisseur de services contient les SP et les ESP. Comme mentionné précédemment, SPs et ESPs contiennent les applications IoT qui effectuent différentes opérations sur les données collectées par les objets IoT. Chaque application IoT est dédiée à une ou plusieurs machines virtuelles (*Virtual Machines* - VMs) où chaque VM est attribuée à l'un des serveurs du centre de données cloud et se voit attribuer une certaine montante de ressources CPU et mémoire afin d'effectuer certaines tâches. Cloud est composé de milliers de serveurs où chaque serveur a certaines capacités de CPU, de mémoire, et de stockage et donc chaque serveur a une limite sur le nombre de machines virtuelles qu'il peut accueillir. Plusieurs VMs à attribuer au même serveur tant que le serveur a une capacité de ressources suffisante pour prendre en charge les besoins en ressources de chaque hébergé. [Huang2017], [Barrett2010].

Les applications IoT exécutées dans le domaine de cloud sont sujettes à de nombreuses attaques de sécurité. Nous résumons ensuite les plus populaires [Dabbagh2016]:

A) **Attaques par canaux cachés**

Bien qu'il existe une séparation logique entre les VMs exécutées sur le même serveur, certains composants matériels sont toujours partagés, tels que la mémoire cache. Cela ouvre des opportunités de fuite de données sur les VMs qui résident sur le même serveur. Cette attaque s'appuie sur un canal caché contre une VM dans le cloud pour localiser la VM cible. Il place ensuite une VM malveillante sur le même serveur que la VM cible.

✓ ***Contre-mesures***

Différentes contre-mesures peuvent être prises pour empêcher les attaques par canaux cachés notamment :

- Ne pas autoriser les VMs hébergées sur le cloud à envoyer des paquets d'analyse tels que des paquets *Traceroute*
- Séparer le cache dédié à chaque VM via le matériel ou le logiciel.
- Attribuer une seule VM par un serveur.
- Laisser chaque client spécifier une liste d'utilisateurs de confiance appelée liste blanche.
- Vider le cache partagé à chaque fois que l'allocation est commutée d'une VM à une autre (*Cache Flushing*).
- Ajouter du bruit aléatoire au temps nécessaire pour extraire les données, ce qui rend difficile de dire si les données ont été extraites du cache ou de la mémoire.

B) **Attaques par la migration de VM**

Chaque serveur dispose d'un module de migration qui est chargé d'envoyer le contenu de la VM du serveur source ou de recevoir le contenu de la VM depuis autres serveurs. La migration des VMs soulève de nouvelles menaces de sécurité telles que:

- ***Attaques du plan de contrôle:*** En exploitant un bug dans le logiciel du module de migration, l'attaquant peut pirater le serveur et prendre le contrôle total du module de migration. Cela donne à l'attaquant la possibilité de lancer des activités malveillantes, notamment les suivantes: Inondations des migrations (*Migration Flooding*), Publicité de fausses ressources (*False Resource Advertising*).
- ***Attaques de plan de données (Data Plane Attacks) :*** La technologie de virtualisation permet le déplacement d'une machine virtuelle d'un serveur à un autre. Ensuite, l'application en cours d'exécution sur la machine virtuelle peut être interrompue pendant un temps très court en raison

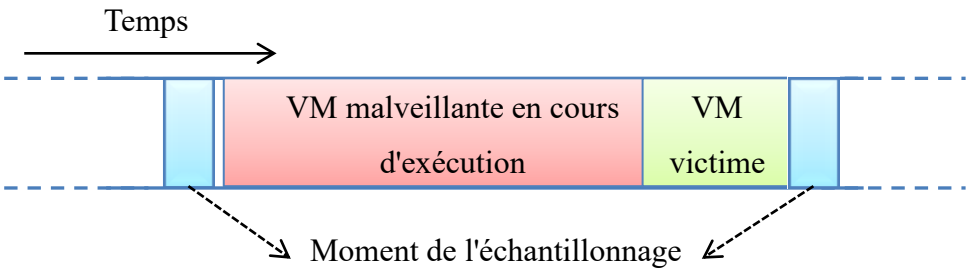


Figure 1.6- Illustration de l'attaque par vol de service (Inspiré de [Dabbagh2016]).

de ce déplacement où la turbulence est aussi faible que des centaines de millisecondes. Malheureusement, les attaquants peuvent profiter de cette courte période.. Voici des exemples d'attaques de plan de données: Attaque par reniflement (*Sniffing Attack*), Attaque de l'homme du milieu (*Man-In-The-Middle Attack*- MITM).

✓ ***Contre-mesures***

Afin de sécuriser la migration de la VM, une authentification mutuelle doit être effectuée entre le serveur initiant la migration et le serveur qui hébergera la VM migrée. Les messages de contrôle qui sont échangés entre les serveurs pour gérer la migration doivent également être chiffrés et signés par l'entité qui génère ces messages afin d'éviter de les modifier et afin d'empêcher d'autres entités de fabriquer de faux contrôles. Les numéros de séquence ou les horodatages (*Timestamps*) devraient également être inclus dans les messages échangés afin d'empêcher une entité malveillante de rejouer un ancien message qui a été envoyé plus tôt (*Replay Attack*). De plus, les paquets de réponse ARP qui mettent à jour l'adresse physique de la VM ne doivent être acceptés qu'après authentification [Perez2011].

C) **Attaque par vol de service (*Theft-of-Service Attack*)**

Dans cette attaque, une VM malveillante se comporte mal de manière à ce que l'hyperviseur (Gestionnaire de machines virtuelles) lui attribue plus de ressources que le partage qu'elle est censée obtenir. Cette allocation supplémentaire de ressources pour la VM malveillante se fait au détriment des autres VMs qui partagent le même serveur que la VM malveillante, où ces VMs victimes se voient allouer moins de ressources que ce qu'elles devraient réellement obtenir, ce qui à son tour dégrade leurs performances. Comme illustré à la Figure 1.6, la VM malveillante peut céder le cœur acquis à une autre VM peu de temps avant le moment d'échantillonnage. L'hyperviseur suppose alors que l'autre VM qui a cédé le cœur a utilisé le cœur pendant toute la durée. La VM malveillante n'est pas enregistrée comme utilisant le cœur et conserve donc une priorité élevée pour utiliser les cœurs à l'avenir.

✓ ***Contre-mesures***

Deux contre-mesures ont été proposées pour faire face à cette attaque. La première contre-mesure consiste à enregistrer plus précisément l'heure de début et de fin lorsque chaque VM utilisait les cœurs à l'aide d'horloges précises. Une autre solution consiste à randomiser les temps d'échantillonnage.

D) **Attaques d'initiés (*Insider Attacks*)**

Dans toutes les attaques décrites précédemment, les administrateurs du centre de données cloud ont été traité comme des entités de confiance. Cependant, certaines applications sensibles peuvent avoir de sérieuses inquiétudes quant à l'hébergement de leurs informations collectées sur cloud, car les administrateurs du cloud auront dans ce cas la possibilité d'accéder aux données collectées et de les modifier.

✓ ***Contre-mesures***

Différentes techniques ont été proposées pour protéger les données de ces attaques internes. Le cryptage homomorphe (*Homomorphic Encryption*) [Zeroual2021] est une forme de cryptage qui peut être utilisée pour empêcher de telles attaques. Parce qu'il permet aux serveurs cloud d'effectuer des traitements sur des données d'entrée chiffrées pour générer un résultat chiffré. Ce résultat chiffré une fois déchiffré correspond au résultat de traitement sur les données d'entrée non chiffrées.

Une autre forme de protection contre les attaques internes consiste à découper les données collectées par l'objet intelligent en plusieurs morceaux, puis à utiliser une clé secrète pour effectuer certaines permutations sur ces morceaux avant d'envoyer les données aux serveurs cloud. Cela permet de stocker les données sur les serveurs cloud sous une forme non interprétable pour les administrateurs cloud. Seules les entités autorisées qui possèdent la clé secrète peuvent renvoyer les données stockées sous une forme interprétable en effectuant les permutations correctes.

1.9.2 *Domaine Brouillard*

Les menaces de sécurité spécifiques au domaine du brouillard sont les suivantes [Dabbagh2016]:

A) **Problèmes d'authentification et de confiance:**

Contrairement aux Clouds proposés par des entreprises bien connues, les dispositifs de brouillard devraient appartenir à plusieurs entités moins connues. Un problème important d'authentification du propriétaire du dispositif de brouillard qui doit alors être pris en compte lors de l'attribution d'un objet à un dispositif de brouillard. L'objet doit également décider si le propriétaire du dispositif de brouillard est digne de confiance. La confiance est un aspect important car un objet sera attribué à

différents dispositifs de brouillard appartenant à différentes entités car leur emplacement peut changer au fil du temps. Des systèmes de réputation tels que ceux qui ont été proposés dans les réseaux pair-à-pair ou le cloud qui peuvent être utilisés pour sélectionner un dispositif de brouillard digne de confiance parmi ceux disponibles dans la zone entourant chaque objet.

B) **Risques de sécurité de migration plus élevés**

Alors que les VMs migrées dans le domaine cloud sont transportées sur le réseau interne ou sur un réseau privé virtuel sécurisé (VPN), les migrations d'un périphérique de brouillard vers un autre sont effectuées via Internet. Ainsi, il existe une probabilité plus élevée que les VMs migrées soient exposées à des liens réseau ou à des routeurs réseau compromis. Il est donc essentiel de chiffrer la VM migrée et d'authentifier les messages de migration de VM qui sont échangés entre les dispositifs de brouillard.

C) **Vulnérabilité plus élevée aux attaques DoS**

Étant donné que les dispositifs de brouillard ont des capacités de calcul inférieures, cela en fait un fruit à portée de main pour les attaques par déni de service (DoS) où les attaquants peuvent facilement submerger les dispositifs de brouillard par rapport au cloud, où un grand nombre de serveurs à haute capacité de calcul sont disponibles.

D) **Problèmes d'Intimité**

Le dispositif de brouillard peut déduire l'emplacement de tous les objets connectés. Cela permet au dispositif de brouillard de suivre les utilisateurs ou de connaître leurs habitudes de déplacement, ce qui peut porter atteinte à la vie privée des personnes qui utilisent ces objets.

E) **Capture physique de nœud de brouillard (*Node Capture Attack*),**

Puisque des nœuds brouillard peuvent être utilisés dans un lieu loin de voir les gens, en particulier les propriétaires des nœuds brouillard, un adversaire peut y accéder physiquement, volant ainsi des informations stockées (*Stolen Verifier*).

✓ ***Contre-mesures***

un dispositif appelé un "*obfuscator*" a été proposé dans [Yu2015] qui empêche les fuites d'informations en émettant des signaux qui rendent difficile pour un récepteur non autorisé de déduire l'amplitude, la fréquence , et le décalage temporel des signaux originaux échangés.

1.9.3 *Domaine Perception*

Le domaine de perception est sensible à de multiples attaques. Certaines des plus connues seront résumées ci-après.

A) **Attaques de brouillage (*Jamming Attack*)**

L'attaque provoque une interruption de service en envoyant des signaux de brouillage. Il existe différentes stratégies de brouillage qu'un brouilleur peut suivre pour lancer une attaque de brouillage. Les plus connus sont résumés ci-dessous:

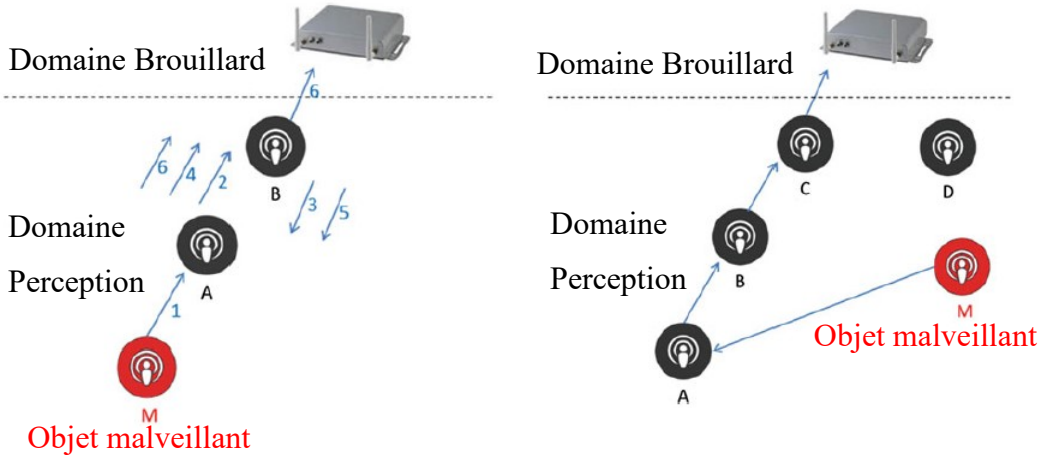
- **Brouillage constant:** L'attaquant transmet en permanence un signal de brouillage aléatoire.
- **Brouillage trompeur:** Ceci est similaire au brouillage constant à l'exception du fait que le brouilleur dissimule son comportement malveillant en transmettant des paquets légitimes qui suivent la structure du protocole MAC plutôt que d'envoyer des bits aléatoires.
- **Brouillage réactif:** Le brouilleur dans ce cas écoute le support et transmet un signal de brouillage seulement après avoir détecté qu'un signal légitime est transmis dans le support.
- **Brouillage aléatoire:** Le brouilleur alterne entre l'envoi d'un signal de brouillage et le fait de rester inactif pendant des périodes aléatoires afin de masquer l'activité malveillante.
- **Brouillage d'accusé de réception:** qui consiste à brouiller uniquement les paquets d'accusé de réception que les nœuds échangent plutôt que de brouiller l'ensemble des paquets de données transmis.

✓ ***Contre-mesures***

Différentes techniques de prévention et de détection ont été proposées pour faire face aux attaques de brouillage. Les plus populaires sont: Sauter la fréquence, Étaler le spectre, Utiliser des antennes directionnelles, Détecter le brouillage. Dans ce dernier, Le récepteur peut détecter qu'il est victime d'une attaque de brouillage en collectant des caractéristiques telles que la force du signal reçu (*Received Signal Strength* - RSS) [Xi2016] et le taux de paquets reçus corrompus. Une technique avancée d'apprentissage automatique peut ensuite être utilisée pour différencier les attaques de brouillage de la dégradation causée par la mauvaise qualité du canal due aux changements normaux de la liaison sans fil.

B) **Attaque de vampire**

Cette attaque exploite le fait que la majorité des objets IoT ont une autonomie de batterie limitée. Le malveillant se comporte mal d'une manière qui oblige les objets à consommer plus d'énergie afin que la batterie s'épuise plus tôt. Quatre types d'attaques de vampires seront identifiés ci-dessous en fonction de la stratégie utilisée pour drainer l'énergie:



(a) (b)

Figure 1.7 - Illustration de l'attaque carrousel et du attaque par inondation, (a) attaque carrousel où les flèches numérotées indiquent le chemin spécifié par l'objet malveillant, (b) attaque par étirement où l'objet malveillant génère des paquets avec un long chemin (Source [Dabbagh2016]).

- **Déni de veille (*Denial of Sleep*):** Un adversaire peut désormais lancer une attaque par déni de veille qui empêche les objets de passer en mode veille en envoyant simplement des signaux de contrôle qui modifient leurs cycles d'utilisation en les maintenant actifs plus longtemps.
- **Attaque par inondation (*Flooding Attack*):** l'adversaire peut inonder les nœuds voisins de paquets fictifs et leur demander de livrer ces paquets au dispositif de brouillard, où les dispositifs gaspillent de l'énergie à recevoir et à transmettre ces paquets fictifs.
- **Attaque carrousel (*Carrousel Attack*):** L'adversaire dans cette attaque spécifie des chemins de routage qui incluent des boucles où le même paquet est routé dans les deux sens parmi les autres objets gaspillant leur puissance. Figure 1.7 (a) illustre cette attaque.
- **Attaque par étirement (*Stretch Attack*):** un objet malveillant dans cette attaque peut envoyer les paquets qu'il est censé rapporter au dispositif de brouillard via de très longs chemins plutôt que des chemins directs et courts comme illustré sur la Figure 1.7 (b).

✓ ***Contre-mesures***

Les attaques par déni de veille peuvent être atténuées en cryptant le message de contrôle qui organise les horaires de la veille tout en incluant un horodatage ou un numéro de séquence. le message de contrôle crypté empêche l'adversaire de rejouer un ancien message de contrôle. En vérifiant l'horodatage crypté ou le numéro de séquence crypté, les objets peuvent reconnaître que le message de contrôle rejoué n'est pas un nouveau message mais un ancien que certains ont relu pour provoquer une perturbation. Les attaques par inondation peuvent être atténuées en limitant le taux

de paquets que chaque objet peut générer. Les attaques carrousel et les attaques par étirement peuvent être atténuées en désactivant le routage source ou en faisant en sorte que chaque objet à qui l'on demande de transférer un paquet, en fonction d'un itinéraire spécifié par la source, vérifie le chemin spécifié. Les paquets avec des boucles à l'intérieur ou suivent de longs chemins sont abandonnés car ils proviennent très probablement d'objets malveillants.

C) **Attaque par transfert sélectif (*Selective-Forwarding Attack*)**

Cette attaque a lieu dans le cas où l'objet ne peut pas envoyer ses paquets générés directement au dispositif de brouillard mais doit s'appuyer sur d'autres objets qui se trouvent le long du chemin pour livrer ces paquets. Un objet malveillant dans cette attaque ne transmet pas une partie des paquets qu'il reçoit des objets voisins. Un cas particulier de cette attaque est l'attaque de trou noir (*Blackhole Attack*) où l'attaquant abandonne complètement tous les paquets qu'il reçoit des objets voisins.

✓ ***Contre-mesures***

Le meilleur moyen d'empêcher les pertes de paquets est d'augmenter la capacité de transmission des objets afin qu'ils puissent atteindre le dispositif de brouillard directement sans avoir besoin de l'aide d'objets intermédiaires. Malheureusement, tous les objets IoT ne sont pas censés avoir cette capacité.

Différentes solutions ont été proposées pour atténuer le nombre de paquets abandonnés. La redondance des chemins est l'une de ces solutions, où plusieurs copies du même paquet sont livrées au dispositif de brouillard via différents chemins. La principale limitation de cette technique d'atténuation est qu'elle a une surcharge énergétique élevée car elle augmente considérablement le trafic.

Les systèmes de détection et prévention d'intrusion peuvent être efficaces dans la détection des objets malveillants qui abandonnent les paquets envoyés afin que les paquets puissent être acheminés via différents chemins qui évitent ces objets [Park2012], [Xiao2007].

D) **Attaque Sinkhole**

Un objet malveillant prétend avoir le chemin le plus court vers le dispositif de brouillard qui attire tous les objets à proximité ; qui n'ont pas la capacité de transmission pour atteindre le dispositif de brouillard, pour transmettre leurs paquets à cet objet malveillant et s'appuyer sur cet objet pour livrer leurs paquets [Ahmad2013]. Désormais, tous les paquets des nœuds voisins passent par ce nœud malveillant.

✓ ***Contre-mesures***

Des techniques de détection et d'isolement d'objets malveillants ont été proposées et reposent sur l'idée de collecter des informations auprès de différents objets où chaque objet signale des objets voisins ainsi que la distance pour atteindre ces objets [Cervantes2015].

➢ Comme dans le domaine de brouillard, l''attaque par capture physique est applicable au domaine de la perception, puisqu'un nœud de perception peut être utilisé dans un lieu publics ou inhabités, un adversaire peut y accéder physiquement, volant ainsi des informations stockées (*Stolen Verifier*).

1.9.4 *Domaine Utilisateur*

Est également appelé « consommateur de données – *Data Consumer*». Nous entendons souvent par attaques du domaine utilisateur les attaques sur les appareils de l'utilisateur (terminaux utilisateur) tel qu'un smartphone ou une carte à puce (*Smart Card*). Il existe de nombreuses attaques communes entre le domaine de la perception et le domaine de l'utilisateur. Nous nous concentrons sur la présentation d'attaques qui distinguent le domaine utilisateur des autres domaines. Parmi eux on trouve:

A) **Attaque par terminal d'utilisateur volé (*Stolen User Device Attack*)**

Si l'appareil intelligent d'un utilisateur est perdu ou volé, un attaquant peut récupérer toutes les informations sensibles stockées dans la mémoire de l'appareil intelligent volé à l'aide de l'attaque par Analyse de consommation [Kocher1999]. Ensuite, en utilisant ces informations récupérées, l'attaquant peut récupérer d'autres informations secrètes des parties communicantes. Avec une attaque physique comme le vol ; si un attaquant possède un appareil utilisateur tel qu'une carte à puce ou un smartphone, toutes les informations stockées peuvent être récupérées en même temps.

B) **Attaque d'usurpation d'identité (*Impersonation Attack*)**

Egalement appelé « attaque de mascarade (*Masquerade Attack*) », est une attaque qui utilise une fausse identité, telle qu'une identité personnelle, pour obtenir un accès non autorisé au système fermé via une identification d'accès légitime. Si un processus d'autorisation n'est pas entièrement protégé, il peut devenir extrêmement vulnérable à cette attaque. Les attaques d'usurpation d'identité peuvent être perpétrées en utilisant des mots de passe, cookies, et des connexions volés, en localisant des lacunes dans les programmes ou en trouvant un moyen de contourner le processus d'authentification. L'attaque peut être déclenchée soit par une personne au sein de l'organisation, soit par un étranger si l'organisation est connectée à un réseau public. En tant que tels, les attaquants de mascarade peuvent avoir un éventail complet d'opportunités de cybercriminalité s'ils

ont obtenu la plus haute autorité d'accès à une organisation commerciale. Les attaques personnelles, bien que moins fréquentes, peuvent aussi être nocives. On peut distinguer deux type de cette attaque, (I) Cette attaque se produit lorsqu'un utilisateur illégal prétend être une personne morale en rejouant un message d'authentification authentique intercepté à partir d'une précédente communication réussie. (II) l'adversaire essaie de présenter un nouvel identifiant faux en utilisant les messages d'authentification interceptés précédents. Ces messages sont utilisés pour contourner l'authentification du système en tant que nouveau nœud authentique.

C) **Attaque de mot de passe hors ligne (*Offline Password Attack*)**

Toute entité illégale peut acquérir des mots de passe en utilisant l'attaque par force brute " *Brute-force Attack*" pour deviner les mots de passe (*Offline Guessing Attack*). En utilisant le mode de devinette hors ligne, l'attaquant peut facilement deviner les mots de passe. pour augmenter la probabilité de récupérer les mots de passe corrects, une méthode systématique de deviner un mot de passe appelée attaque par dictionnaire (*Offline Dictionary Attack*) est utilisée. Où l'attaquant essaie de nombreux mots courants et leurs variations simples. Les attaquants utilisent des listes exhaustives des mots de passe les plus couramment utilisés, des noms d'animaux populaires, des personnages fictifs ou littéralement simplement des mots d'un dictionnaire, d'où le nom de l'attaque. L'attaquant tente également d'utiliser les informations personnelles de la victime, telles que son nom, son prénom, sa date de naissance, etc. L'attaquant peut enregistrer la liste cryptée des mots de passe et tenter de rassembler les paramètres de sécurité de la liste en cryptant n'importe quel mot de passe (utiliser la fonction de hachage comme exemple) afin de retrouver les mêmes textes cryptés enregistrés.

✓ ***Contre-mesures***

Différentes techniques ont été proposées pour protéger les données dans les terminaux des utilisateurs contre les attaques physiques telles que le vol de terminal. Le cryptage homomorphe (*Homomorphic Encryption*) est une de ces solutions [Zeroual2021]. Alors la possession du terminal entraîne la possession d'informations sensibles déjà chiffrées. La même solution fonctionne contre les attaques d'usurpation d'identité où les mots de passe, cookies et connexions écoutés.

Concernant l'attaque par mot de passe hors ligne, les mots de passe doivent être très complexes. Ils doivent nécessairement contenir mélange des chiffres, des lettres majuscules, des lettres minuscules, des symboles, etc. Ainsi que l'éloignement de tout ce qui concerne les informations personnelles ou les informations courantes. Cependant, les appareils puissants peuvent facilement calculer et trouver les mots de passe en utilisant " *Brute-force Attack*".

La meilleure solution, en particulier pour l'utilisateur humain, est la technologie biométrique. Plusieurs avantages peuvent être tirés de l'utilisation de la technologie biométrique, car ils ne peuvent pas être perdus, volés ou devinés, ils sont inoubliables, difficiles à copier, difficiles à falsifier, et difficiles à casser. Par conséquent, l'authentification des utilisateurs basée sur la biométrie est considérée comme plus sûre et fiable que les schémas d'authentification conventionnels.

1.10 Conclusion

L'IoT d'aujourd'hui est assez différent de l'IoT conventionnel connu auparavant. Le nouvel IoT permet aux entreprises d'acheter, de vendre et de payer en ligne des informations collectées par d'autres objets auprès d'autres entreprises. Les données sont analysées et agrégées via des processus intelligents, puis utilisées pour fournir un service cohérent proactif, prédictif et préventif.

Dans ce chapitre, les deux principales architectures IoT ont été présentées, l'architecture IoT basée sur trois domaines, également appelée architecture de bout en bout. La deuxième architecture est l'IoT basé sur quatre domaines, également appelé IoT basé sur le cloud ; y compris l'architecture IoT basée sur les services représentée dans le modèle S^2aaS.

Le chapitre a également analysé l'IoT du point de vue de la sécurité et de l'intimité. Le modèle S^2aaS a été pris en compte dans cette analyse où les attaques ciblaient différents domaines de l'IoT et les contre-mesures défensives ont été décrites.

Puisque l'utilisateur joue un rôle primordial dans la LoT basée sur les services ; souvent cet utilisateur est un être humain, la meilleure façon de protéger et de sécuriser ce rôle est d'utiliser la technologie biométrique. C'est ce que nous verrons dans le prochain chapitre.

Biométrie : Principes, Applications, et Sécurité

Chapitre 2: Biométrie : Principes, Applications, et Sécurité

2.1 Introduction

Aujourd'hui, avec la croissance du réseau internet et la croissance du nombre d'éléments connectés à ces réseaux, la quantité d'informations transmises entre ces éléments a aussi connu une grande augmentation. En parallèle, les opérations illégales et les cyberattaques sont de plus en plus nombreuses. Comme l'être humain fait partie de ces éléments connectés, l'incorporation des technologies biométriques présente un moyen de sécurité fiable en général, et un moyen d'authentification et d'identification en particulier.

Au cours de la dernière décennie, l'exploitation et l'utilisation de la technologie biométrique ont connu une croissance considérable soit d'une manière directe ou indirecte. Le domaine de la biométrie présente aujourd'hui un segment de marché bouillonnant qui concurrence les autres domaines [Global2021].

De ce fait, elles sont facilement accessibles et également de bon marché. De même, leur intégration dans divers appareils et leur mise en œuvre sont devenus maîtrisables. Mais faire répandre cette technologie n'est pas sans des effets négatifs, en effet, l'information biométrique émergées dans ces technologies est plus vulnérables aux imposteurs et aux attaques qui sont aussi développés. L'attaque la plus dommageable concerne les vecteurs biométriques c'est L'attaque après l'extraction car une fois l'extraction effectuée, l'information biométrique est présentée par un vecteur qui se vole facilement. Par conséquent, il est nécessaire de fournir des moyens et des mécanismes de protection du vecteur des caractéristiques.

Tous les détails concernant la technologie biométrique, le fonctionnement de ses systèmes, et les mécanismes de protection du vecteur des caractéristiques seront présentés dans ce chapitre.

2.2 Biométrie

La "Biométrie[3]" est «l'application de l'analyse statistique aux données biologiques» [Pearsall2002]. On peut dire que la biométrie c'est le processus par lequel les signaux biologiques d'une personne sont détecté, traités et enregistrés par un système électronique

[3] Selon l'*Oxford English Dictionary*, Le mot «Biométrie» comprend deux parties qui sont : «Bio» indique «Biologique» et «Métrie» indique «Métrique». La biométrie renvoie à la notion de mesurabilité des caractéristiques biologiques.

comme un moyen de confirmation de l'identité. En général, pour prouver l'identité d'une personne, au moins l'une de ces quatre questions doivent être répondues:

- Ce que nous savons? Comme le mot de passe, le numéro de code, etc.
- Ce que nous avons? Comme le badge, le jeton, la clé, etc.
- Que sommes-nous? Comme la biométrie physiologique
- Comment faisons-nous? Comme la biométrie comportementale

Les deux premières peuvent authentifier l'utilisateur mais n'indiquent pas nécessairement son existence. En conséquence, le système ne peut pas distinguer l'imposteur du l'utilisateur légitime lorsque la clé est volée ou le mot de passe est deviné. La biométrie est une solution d'authentification alternative qui a réussi à surmonter les inconvénients suscités. Sans doute, la biométrie est la plus utilisée car elle est la seule qui garantit une grande unicité et elle ne peut être oubliée, perdue ou volée. La technologie biométrique est utilisée comme un outil de sécurité pour vérifier l'identité d'une personne ou pour l'identifier. Cependant, autres exigences de la sécurité peuvent être envisageables comme nous allons le montrer dans la partie expérimentale.

Il existe de nombreux signaux biologiques qui peuvent être utilisés pour l'identification humaine, mais seuls quelques-uns sont mesurables. La technologie biométriques peuvent exploiter les différentes caractéristiques physiques ou comportementales pour le processus d'authentification/identification. Cependant, seuls quelques-uns de ces caractéristiques humaines peuvent être considérés comme biométries. Le principal obstacle est le défi lié à la récupérabilité et à la mesurabilité. Au moins actuellement, l'incapacité de mesurer bon nombre de ces signaux, en raison du manque de capteurs adéquats les exclut en tant que données biométriques acceptables. La biométrie mesurable et récupérable peut être catégorisée en plusieurs classes.

2.3 Technologies biométriques

Les technologies biométriques peuvent être classées en fonction du type de signaux sur lesquels elles s'appuient, nous trouvons les classes suivantes: physiologiques, comportementales et cognitives [Obaidat 2019].

Les caractéristiques physiologiques sont inhérentes à la physiologie humaine. Cette classe peut également être divisée en deux sous-classes : caractéristiques morphologiques comme les empreintes digitales, la géométrie de la main, la minutie des doigts, la forme du visage, la forme de l'iris, la forme de la rétine. Etc. La deuxième sous-classe contient les caractéristiques biologiques non morphologiques comme le sang, ADN, l'urine, salive .Etc.

Les caractéristiques comportementales sont des traits qui sont acquis ou appris à partir d'actions humaines. Comme exemples des caractéristiques comportementales nous trouvons la dynamique de frappe au clavier, de mouvement de la souris, des gestes, ainsi que la signature, le ton, les caractéristiques des démarches (*Gait Recognition*). Etc.

Les caractéristiques cognitives reposent sur l'état cognitif, émotionnel et conatif comme base pour reconnaître les individus. En général, ces caractéristiques cognitives sont extraites en enregistrant des bio-signaux physiologiques ou comportementaux tels qu'électroencéphalogramme (EEG), électrocardiogramme (ECG) et réponse électro-cutanée (EDR) de l'individu sous l'effet de stimuli cognitifs.

2.4 Caractéristiques et exigences

Les êtres humains possèdent des qualités très spécifiques, qui sont uniques et suffisamment stables pour être utilisées comme identifiants. Dans la mesure où nous les utilisons pour l'identification biométrique, nous les appelons des «caractéristiques biométriques».

Toute caractéristique biométrique peut être utilisée comme trait biométrique doit répondre aux exigences suivantes [Jain2004]:

- Universalité: la caractéristique doit être applicable à chaque individu.
- Unicité: deux personnes doivent être suffisamment différentes en termes de la caractéristique biométrique.
- Permanence: la caractéristique doit être suffisamment stable sur une période de temps.
- Mesurabilité: la caractéristique doit être mesurée quantitativement.

Plus les critères susmentionnés, la caractéristique biométrique en pratique peut répondre aux exigences supplémentaires suivantes :

- Performance: La précision et la vitesse du processus de la reconnaissance.
- Acceptabilité: Mesure dans laquelle les gens sont prêts à accepter l'utilisation d'un identifiant biométrique particulier dans leur vie quotidienne
- Résistance au contournement: Reflète la facilité avec laquelle le système peut être trompé en utilisant des méthodes frauduleuses telles que des contrefaçons basées sur des échantillons synthétiques et d'autres techniques d'évasion.
- Préservation de la confidentialité: protection des informations des utilisateurs privés doivent être intégrées dans les modèles biométriques et les technologies sous-ajacentes.

Chaque trait biométrique peut être fort dans certaines exigences et faibl dans autres. Une étude comparative des caractéristiques biométriques, réalisée par le groupe international de la biométrie (*International Biometric Group*) fournit des lignes directrices pour la sélection des différentes technologies biométriques en fonction du caractère acceptable par l'utilisateur, du caractère distinctif (unicité), du coût et des efforts menés par les utilisateurs comme il est indiqué dans la Figure 2.1.

Généralement, les traits biométriques sont distingués par des modalités fortes, faibles et douces [Mordini2012].

La «biométrie forte» a des caractéristiques qui peuvent répondre à la plupart des exigences mentionnées au-dessus. Aucun trait biométrique n'est parfaitement unique ni parfaitement permanent, mais certaine traits peuvent être mieux que les autres. Par exemple, chez les êtres humains, l'empreinte digitale, l'empreinte palmaire, l'iris, les veines de la main et la rétine peuvent être pratiquement considérés comme uniques et stables. Autrement dit, leur présence est considérée comme univoque liée à un seul spécifique de l'être humain. Autres traits biométriques peuvent être pratiquement considérés moins acceptables par l'utilisateur. Effectivement, l'utilisateur n'accepte pas par crainte que certaines biométries de valeur soient utilisées pour l'identification, par exemple : La rétine et l'iris. Le développement précoce des technologies biométriques était principalement basé sur la forte biométrie. D'un point de vue

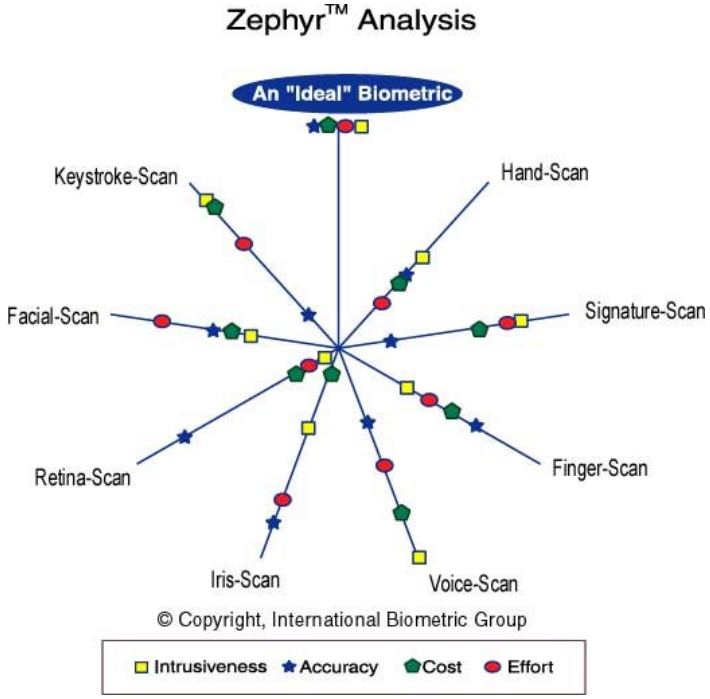


Figure 2.1- Les caractéristiques de chaque biométrie et les critères d'applicabilité (Source: [Dasgupta2017])

logique, une identification basée sur une biométrie forte n'est guère différente de l'identification basée sur des jetons artificiels (*tokens*). Ce qui change, c'est que la biométrie n'est pas un objet externe associé à la personne, mais une caractéristique corporelle interne.

La «biométrie faible» a des caractéristiques qui sont «moins uniques» ou «moins stables» que celles de la biométrie forte. Chez les humains, la biométrie faible comprend des caractéristiques telles que la forme du corps, les odeurs, le comportement, la voix, les phénomènes électro-physiologiques. Des données biométriques faibles peuvent être utilisées à des fins d'identification à condition qu'elles ne soient pas utilisées seules sans aucun autre appui ou information supplémentaire. Etant donné qu'on ne peut établir une correspondance cohérente univoque entre les biométries faible et l'individu, elles ne doivent être utilisées que dans leur contexte en considérant également les coordonnées spatiales et temporelles ou en association. Cela implique que pour les utiliser efficacement, nous devons collecter également d'autres détails tels que la localisation géo-spatiale, le moment où l'élément a été collecté, etc. Nous pouvons également associer et fusionner de manière fructueuse deux ou plusieurs données biométriques faibles, et les associer davantage à la biométrie douce (voir ci-dessous). En faisant cette association, nous pouvons atteindre une capacité élevée de discernement. Ce qui est pleinement satisfaisant pour vérifier les identités présumées (authentification) et également assez bon pour identifier les personnes dans des groupes de dimensions limitées (identification).

L'expression «biométrie douce : *soft biometric*» réfère à des caractéristiques qui sont génériques et ne peuvent pas être liées à un individu spécifique. Ils comprennent des catégories telles que le sexe, âge, race /ethnicité, poids, la taille, couleur des yeux, couleur de la peau et des cheveux .etc. Néanmoins, la biométrie douce peut être utilisée avec succès pour renforcer la biométrie forte et faible. Fondamentalement, elle permet de limiter le nombre des possibilités et par conséquent d'affiner le processus d'identification.

En général, les traits biométriques physiologiques sont souvent de la biométrie forte, alors que ceux comportementaux sont de la biométrie faible.

2.5 Applications

Le progrès des technologies de capteurs, traitement, l'intelligence computationnelle et d'autres domaines connexes ont contribué à des améliorations spectaculaires de la fiabilité et de la robustesse des technologies biométriques. Cela a contribué à augmenter le niveau de confiance et à améliorer la perception de ces technologies par les différentes parties prenantes, par exemple le public, les entrepreneurs, etc.

La biométrie est utilisée dans diverses industries, y compris le gouvernement et l'application de la loi, le commerce et la vente, les soins de santé, les voyages et l'immigration, les finances et les banques, etc. Ainsi les documents gouvernementaux qui concernent les cartes d'identité nationales, les passeports, les permis de conduire, les cartes de sécurité sociale, l'inscription des électeurs, l'enregistrement de l'aide sociale et le contrôle de leurs intégrités. Les technologies sont aussi utilisées pour renforcer ou remplacer certaines de ces documentations ou processus critiques.

L'un des domaines d'application les plus importants est les systèmes automatisés d'identification des empreintes digitales (AFIS), qui sont utilisés dans la plus part ou presque tous les domaines résumés ci-dessous :

- Contrôle d'accès physique et surveillance
- Authentification
- Médecine légale numérique
- Temps de présence
- La sécurité des frontières
- Intégrité du passeport
- Vote électronique
- Commerce électronique
- l'internet des objets.

Avec l'amélioration spectaculaire des capacités de calcul, de stockage et les progrès réalisés dans le développement de capteurs intelligents, le paysage de la technologie biométrique est également témoin d'un passage des systèmes principalement matériels à des solutions logicielles intégrant les *Smartphones* et le *Cloud*. Parmi ces technologies biométriques améliorées, l'inclusion de la technologie biométrique dans l'environnement de l'internet des objets. Cette inclusion permet de authentifier/identifier les utilisateurs afin de leur permettre de bénéficier des différents services disponibles. Ainsi, elle permet à l'utilisateur légitime de gérer ses propres objets.

2.6 Systèmes biométriques

Un système biométrique est fondamentalement un système de reconnaissance de formes qui fonctionne en acquérant des données biométriques brutes d'un individu. Les données acquises par des capteurs appropriés sont traitées en extrayant des paramètres d'identification également appelés caractéristiques biométriques. Les caractéristiques extraites forment la base de la soi-

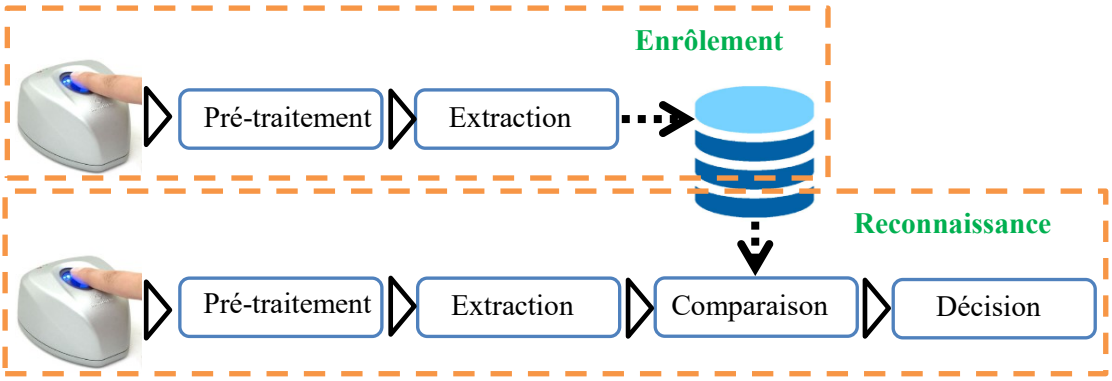


Figure 2.2- Les deux phases principales de fonctionnement d'un system biométrique.

disant signature biométrique individuelle, également appelée « gabarit biométrique » ou plus couramment « vecteur des caractéristiques ». Les vecteurs sont stockés dans une base de données et utilisés ultérieurement pour vérifier l'identité de l'utilisateur en les comparants à autres échantillons biométriques.

2.6.1 *Phases des systèmes biométriques*

En général, les systèmes biométriques impliquent deux phases de fonctionnement principales selon l'application désirée: phase d'enrôlement (Enregistrement) et phase de la reconnaissance [Jain2004]. La Figure 2.2 illustre ces deux phases d'un system biométrique.

Durant la phase d'enrôlement, le vecteur biométrique est extrait d'un nouvel utilisateur en acquissent les données d'échantillon biométrique. En construisant une représentation mathématique, le vecteur caractéristique biométrique dit vecteur modèle peut être obtenu et ensuite mis dans la base des données pour la première fois. Ce vecteur sera le point référentiel de la comparaison durant la phase de reconnaissance. Des informations supplémentaires biographiques sur l'utilisateur peuvent être ajoutées conjointement avec le vecteur biométrique dans la base de données.

Dans la phase de la reconnaissance, un autre vecteur dit vecteur requête est extrait d'un utilisateur déjà enregistré dans la base de données. L'extraction se fait de la même manière et suivant les mêmes étapes que la phase d'enrôlement. Ce vecteur est comparé avec celui de la base de données. La décision est prise selon le degré prédéfini de la similitude entre les deux vecteurs.

2.6.2 *Modes de fonctionnement*

Dans un système biométrique, durant la phase de reconnaissance, l'étape de la comparaison se diffère d'un système à l'autre selon les modes opératoires. Ces modes sont divisés en deux catégories principales :

✓ **Mode de vérification :** Aussi nommé « authentification », dans ce mode le système est simili à une réponse à la question suivante: Es-tu la personne prédite? Donc, ce mode dépend de la comparaison d'un vecteur requête avec un seul vecteur modèle (1: 1). De ce fait, le processus d'authentification est généralement cohérent avec un mot secret, un mot de passe ou tout autre moyen pour achever l'authentification avec une seule comparaison.

Il existe trois types d'approches d'authentification biométrique (Authentification statique, Authentification active, Authentification continue) [Obaidat 2019]:

- **Authentification statique :** Consiste à vérifier l'identité de l'individu une fois, généralement au moment de la connexion (*Log-in*). Bien que l'authentification statique soit cruciale pour le contrôle d'accès, elle n'est pas suffisante pour assurer la sécurité de la session. La session restante peut toujours être détournée par un pirate informatique à des fins néfastes.

- **Authentification active :** consiste à ré-authentifier l'individu; cela se produit généralement une fois après la connexion. L'authentification active a le potentiel d'attraper le détournement de la session, si une telle occurrence a eu lieu avant la réauthentification.

- **Authentification continue :** est une forme plus permanente de l'authentification active. L'identité de l'utilisateur est vérifiée à plusieurs reprises après la connexion. L'authentification continue peut se produire périodiquement, après une certaine quantité d'activité ou à l'expiration d'un intervalle de temps prédéfini. Bien qu'une approche stricte réduisant considérablement les fenêtres de vulnérabilité, elle pourrait augmenter la surcharge du système d'authentification.

✓ **Mode d'identification :** Contrairement du mode précédent, la reconnaissance est faite après plusieurs comparaisons. Souvent, un vecteur requête est comparé à tous les vecteurs modèle de la base de données (1 : N). le système est simili à une réponse à la question suivante: Qui es-tu ?

En outre, le mode d'identification biométrique peut être réalisé dans deux classes d'ensembles (Ensemble fermé et Ensemble ouvert).

- **L'identification en ensemble fermé :** Dans cet ensemble, tous les utilisateurs doivent être préalablement inscrits dans la base de données du système avant de l'utiliser. Et par conséquent, le système peut identifier l'identité de l'utilisateur exacte. La prise de décision

dépend du degré de similitude le plus proche entre le vecteur requête présenté en entrée et tous les vecteurs modèles.

- **L'identification en ensemble ouvert :** Cela signifie que toute personne que ce soit inscrite ou non peut tenter d'accéder au système. Ensuite, le système peut l'accepter ou le rejeter en fonction du score de similarité dépendant d'un seuil de sécurité prédéfini.

2.7 Biométrie multimodale

Dans le système biométrique, l'utilisateur interagit avec une interface simple. En quelques secondes, le système scanne les caractéristiques biométriques sélectionnées et décide si l'utilisateur est autorisé à y accéder ou non. Cependant, ces systèmes ne sont pas parfaits et il y a toujours des lacunes qui doivent être traités ou surmontés. Récemment, il a été découvert qu'un cours viable de la biométrie pourrait être basé sur une utilisation plus large des systèmes multimodaux à l'avenir [Ross2006]. Le multi modal consiste à combiner plus qu'une caractéristique biométrique pour l'identification/authentification ; par exemple, plusieurs échantillons (instances) de la même biométrie ou bien plusieurs traits biométries différents comme les empreintes digitales et veines de la paume, etc.

Les systèmes multi-biométriques peuvent aider dans les situations où les systèmes mono-biométrique sont considérés comme des systèmes d'exclusions, c'est-à-dire, un individu n'a pas un trait particulier ou si un trait est gravement déformé de sorte que le capteur ne peut pas l'acquérir.

2.7.1 *Scénarios de combinaison*

En fait, lorsqu'on parle de système multimodal au sens général, il existe de nombreux éléments de prouve d'identité qui posent différents scénarios possibles [Nandakumar2005]. Figure 2.3 montre ce propos.

- **Systèmes multi-échantillons (multi-instances) :** Ça veut dire plusieurs instances de la même biométrie avec un seul capteur probablement, par exemple, plusieurs images de visage d'une personne obtenues dans différentes conditions de pose / d'éclairage .etc. Il est également utilisé pour créer une image complète d'une certaine modalité biométrique (image 360° du visage).

- **Systèmes multi-capteurs :** Ça veut dire plusieurs capteurs pour acquérir la même modalité, par exemple, capteurs d'empreintes digitales optiques et thermiques. Cette technique est

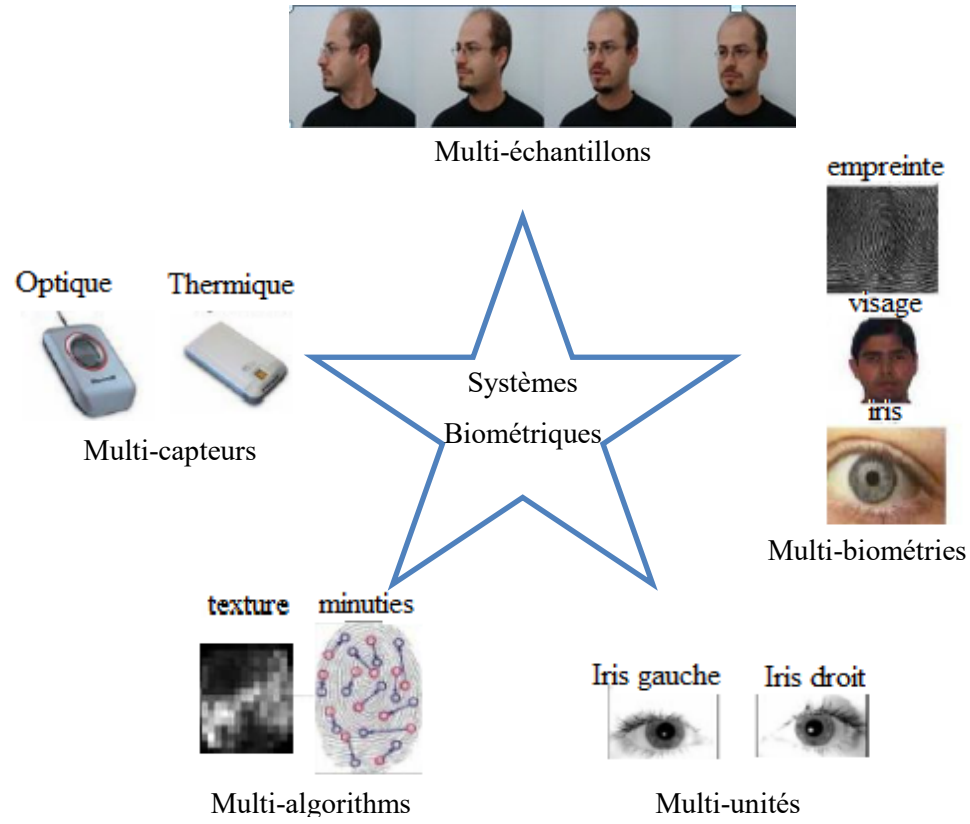


Figure 2.3- Sources de multiples éléments dans les systèmes biométriques multimodaux (Inspiré de [Nandakumar2005]).

utilisée surtout dans le cas de plusieurs emplacements avec différents type dispositifs d'identification de la même modalité tels que les aéroports.

- **Systèmes multi-algorithmes :** Ça veut dire plusieurs représentations ou plusieurs algorithmes sont utilisés pour traiter la même biométrie acquise. Ces algorithme peuvent être intervenir dans les différents niveaux d'un système biométrique tel que le niveau de l'extraction de caractéristiques et/ou le niveau de comparaison, par exemple, des algorithmes d'analyse de texture ou de minuties peuvent être utilisé pour traiter la même image d'empreinte digitale.

- **Systèmes multi-unités:** Ça veut dire plusieurs unités de la même modalité biométrique, par exemple, des images d'iris gauche et droit, deux empreintes digitales de doigts différents. Dans ce système les données nécessitent des bases de données référencées par chaque unité contrairement aux systèmes multi-échantillons qui n'a besoin qu'une seule base référentielle.

- **Systèmes multi-traits biométriques :** Ça veut dire plusieurs traits biométries différentes, par exemple, le visage, l'empreinte digitale et l'iris. Dans les quatre premiers scénarios,

plusieurs sources d'informations sont dérivées du même trait biométrique. Dans le cinquième scénario, les informations sont dérivées de différents traits biométriques. Ce scénario a particulièrement attiré beaucoup d'attention de la part des recherches car les scenarios d'un seul trait biométrique ne traitent pas le problème de la non-universalité. Ainsi que, les systèmes multi-traits biométriques traitent des caractéristiques bien dé-corrélés, contrairement d'un système mono-trait biométrique. Cela contribue à l'amélioration de la performance du système. En plus, les systèmes multi-biométriques se caractérise par sa résistance aux fraudes comme nous l'expliquerons en-dessous.

2.7.2 *Niveaux de fusion*

Dans un système biométrique, La forme d'informations biométriques change du niveau de capture jusqu'à le niveau de la décision, en particulier, dans les systèmes biométrique multimodaux où ces informations sont acquises de multi-sources. Un aspect important d'un système multi-biométrique est la fusion des informations acquises. À un certain moment de la routine de la reconnaissance, il est nécessaire de fusionner les données en une seule entité avant d'aller plus loin, ce qui pose un défi de dimensionnement dans la phase de la conception du développement d'un système multimodal. Comme le montre la Figure 2.4, le système effectue quatre opérations distinctes ; Acquisition, Extraction, Comparaison et Décision. Au niveau de chacune, la fusion peut généralement être introduite. Il convient de noter qu'au fur et à mesure que les données progressées dans le système, leur quantité est compressée en cours du chemin. Cependant, cela n'implique pas nécessairement que plus la fusion se produit tôt, meilleurs sont les résultats [Faundez2005].

La fusion biométrique peut être intrinsèquement divisée en deux sections, fusion avant la comparaison (*Matching*) et après la comparaison [Ghayoumi2015]. La raison de cette classification résulte du fait qu'après la comparaison, la quantité d'informations dont dispose le système diminue d'une marge significative qui est généralement bien plus importante que dans les autres cas [Ross2006].

A) **Fusion avant la comparaison**

Cette fusion comporte deux niveaux de fusion, fusion au niveau du capteur et fusion au niveau du vecteur des caractéristiques.

1) Fusion niveau-capteur (sensor-level)

La fusion au niveau du capteur [Nandakumar2008] consiste à joindre plusieurs sources de preuves brutes avant l'extraction de caractéristiques. Cela peut englober des images, des vidéos,

etc. La fusion à ce niveau, peut se faire uniquement dans les Systèmes multi-échantillons ou les Systèmes multi-capteurs compatibles entre eux pour le même trait biométrique. Si les sources de preuves sont incompatibles, il est difficile de les fusionner, par exemple, des images acquises de caméras ont des résolutions différentes. À ce niveau, les données obtenues contiennent le plus d'informations disponibles. Dans le traitement d'image, une méthode particulière de fusion est employée, consiste à mettre en mosaïque des images (à partir de chevauchement) pour former une seule Image composite [Ross2006].

2) Fusion niveau-vecteur (feature-level)

Dans la fusion au niveau du vecteur des caractéristiques, les sources de preuves sont consolidées une fois que les caractéristiques ont été déjà extraites de la biométrie. Suite à cela, les caractéristiques fusionnées sont ensuite transmises à un module de comparaison, et le système procède comme s'il traitait une seule source de preuves biométriques. Les ensembles de caractéristiques extraites par différents algorithmes posent un défi pour de nombreuses raisons [Ross2006]. Il peut être difficile de fusionner deux modalités choisies, si la base sur laquelle elles doivent être fusionnées n'est pas connue. Dans ces cas, il peut être difficile de produire un vecteur de caractéristiques fusionné qui satisferaient les demandes d'amélioration par rapport à un système biométrique uni-modal. Cela pourrait être exacerbé par la situation dans laquelle des ensembles de vecteurs de caractéristiques de différentes modalités ne sont pas compatibles. L'un d'eux peut varier en longueur, tandis que l'autre peut être représenté par un ensemble d'éléments de longueur fixe, ou bien un grand vecteur peut être obtenu dans le cas d'une concaténation de ces vecteurs.

La plupart des systèmes biométriques commerciaux ne donnent pas l'accès aux vecteurs des caractéristiques qu'ils utilisent dans leurs produits. Par conséquent, très peu de chercheurs ont étudié l'intégration au niveau des vecteurs caractéristiques et la plupart d'entre eux préfèrent généralement la fusion après la comparaison.

B) **Fusion après la comparaison**

Cette fusion comporte trois niveaux, fusion au niveau du score, fusion au niveau du rang et fusion au niveau de la décision.

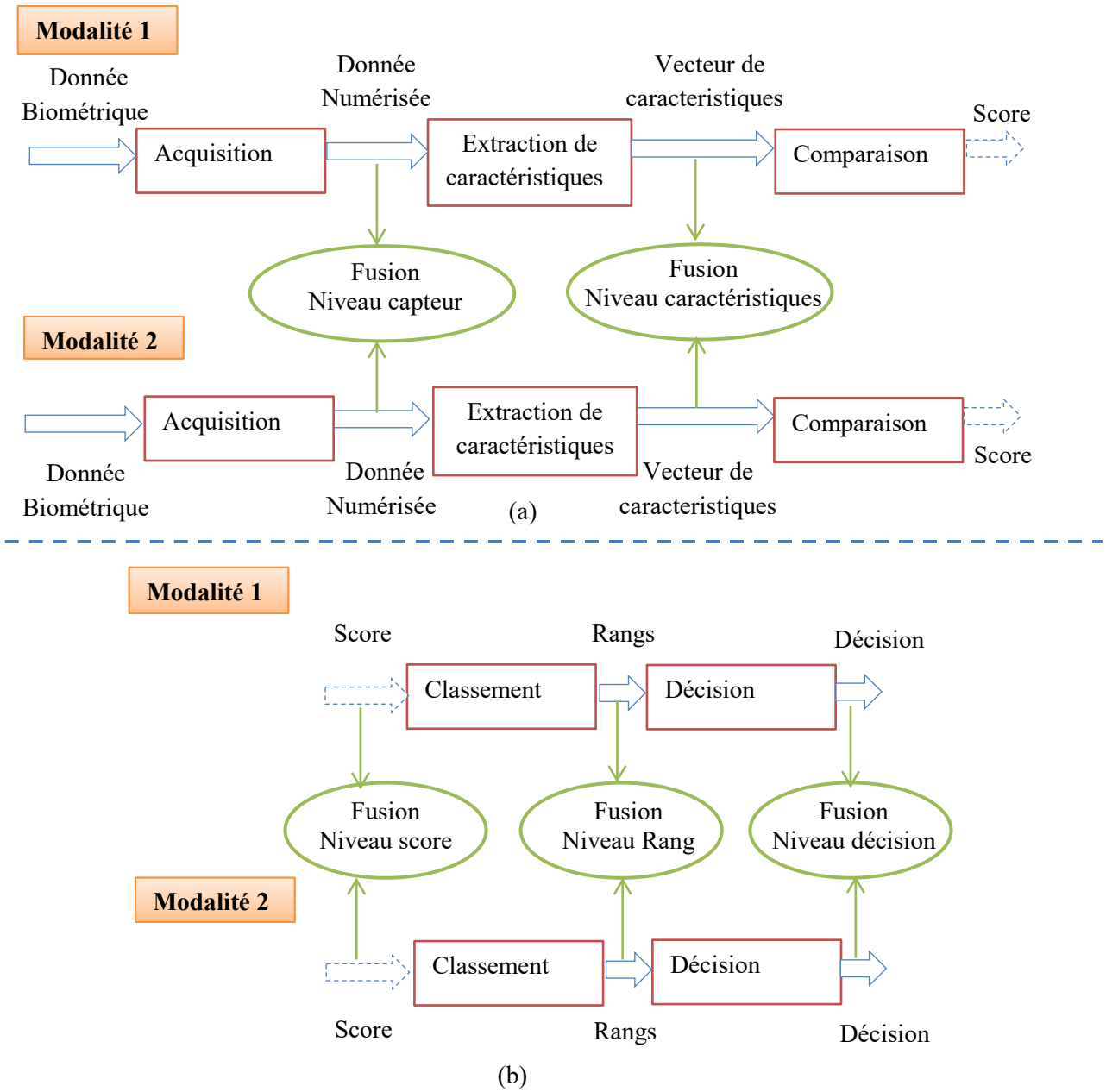


Figure 2.4- Différents niveaux de fusion, (a) fusion avant la comparaison, (b) fusion après la comparaison (Inspiré de [Nandakumar2008]).

1) Fusion niveau-score (score-level)

La fusion au niveau des scores est couramment utilisée et préférée dans les systèmes biométriques multimodaux en général, parce que les scores correspondants contiennent suffisamment d'informations permettant de distinguer les cas authentiques et les imposteurs. Ainsi qu'ils sont relativement faciles à obtenir. Étant donné un certain nombre de systèmes biométriques, des scores de comparaison pour un nombre prédéfini d'utilisateurs peuvent être fusionnés même sans que les techniques de l'extraction de caractéristiques ou les algorithmes de comparaison de chaque système soient connues.

2) Fusion niveau-rang (rank-level)

C'est toujours l'une des méthodes de fusion les plus fréquemment appliquées. Il convient de noter que la fusion au niveau du rang n'est applicable que dans les systèmes biométriques qui sont configurés pour les systèmes d'identification et non pour les systèmes de vérification (authentification) [Kumar2009]. Après l'extraction du vecteur des caractéristiques et l'acquisition du score de comparaison, l'ensemble d'identités correspondantes probables peut être trié par un ordre croissant ou décroissant, et ainsi, une liste classée d'identités des candidates peut être créée. Le but de ce niveau de fusion est de fusionner les rangs produits par les différentes modalités biométriques individuels afin d'obtenir une liste consolidée des rangs pour chaque identité.

3) Fusion niveau-décision (decision-level)

La fusion au niveau décisionnel est particulièrement utile dans les situations où les deux ou plusieurs systèmes biométriques finis sont disponibles et doivent être combinés [Nandakumar2008]. Le plus souvent, la fusion au niveau de la décision est la seule option dans ce cas.

Egalement, Il y a deux architectures de la fusion, si l'acquisition et le traitement sont faits successivement alors cette architecture est une architecture en série, mais s'ils sont faits simultanément on parle alors d'architecture en parallèle. La différence entre ces deux architectures est particulièrement évidente quand il y en a fusion au niveau score.

2.8 Sécurité des systèmes biométriques

Dans cette section, nous étudions les exigences de protection du vecteur de caractéristiques biométriques, ainsi que les propriétés nécessaires des méthodes de protection.

2.8.1 *Exigences de protection de vecteur biométriques*

Alors que les technologies biométriques peuvent s'avérer être des outils relativement robustes et efficaces pour restreindre la fraude, les systèmes biométriques eux-mêmes ne sont pas à l'abri des attaques frauduleuses et d'exploitation illégales. Ces attaques peuvent être classées en trois catégories principales: Attaques au niveau de la base de données, attaques au niveau du capteur, attaques au niveau du traitement et de la transmission.

A) Attaques au niveau de la base de données

Selon *Jain et al.* [Jain2005], Les attaques possibles sur la base de données sont, Écrasement et remplacement du vecteur (*Template Overwriting*), Attaque par le rejoue (*Replay Attack*),

Usurpation d'identité (*Template Spoofing*), et Fonction fluage (*Creep Function*), où le vecteur volé est utilisé dans un autre système à d'autres fins.

B) **Attaques au niveau du capteur**

Les attaques par usurpation d'identité au niveau capteur consistent à tenter de tromper les capteurs du système biométrique en leur faisant accepter un artefact comme un échantillon biométrique légitime, généralement à des fins d'inscription, de vérification ou d'identification erronées. Il existe d'autres attaques ciblant spécifiquement le niveau du capteur qui ne sont pas moins dangereuses que l'usurpation d'identité, tels que l'attaque de contournement (*Bypassing Attack*) et l'attaque par déni de service (*DoS*) [Lee2009].

C) **Attaques au niveau du traitement et de la transmission.**

Les attaques au niveau du traitement et de la transmission appartiennent généralement à l'une des trois types suivants: L'attaque de l'homme du milieu, l'attaque de reniflement et l'attaque par escalade [Lee2009].

- **L'attaque de l'homme du milieu (*Man-In-The-Middle Attack*):** Un adversaire pourrait modifier les algorithmes d'enrôlement ou de reconnaissance d'un système biométrique, en abaissant les seuils pour minimiser la robustesse du système et augmenter sa tolérance. Ils pourraient reprogrammer le système pour leur transmettre des copies d'échantillons légitimes ou demander au système de leur permettre un accès spécial s'il n'est pas autorisé.
- **L'attaque de reniflement (*Sniffing Attack*):** Un reniflement peut se produire si des programmes de surveillance sont mis en place pour capturer les paquets de données envoyés du capteur à la base de données.
- **L'attaque Par Escalade (*Hill-Climbing Attack*):** Les attaques par escalade consistent à présenter un échantillon biométrique de test à un algorithme biométrique pour comparaison avec un échantillon inscrit. Un score de correspondance est ensuite obtenu et étudié afin qu'un nouvel échantillon de test puisse être présenté pour une nouvelle comparaison et l'obtention d'un score de correspondance plus élevé. Ce processus est réitéré jusqu'à ce que le seuil du système biométrique soit découvert et le système soit par conséquent pénétrable.

2.8.2 *Propriétés des Méthodes de protection*

Pour contrer Les attaques ci-dessus, le système biométrique doit fournir des méthodes de protection des vecteurs biométriques plus appropriés. Ces méthodes doivent avoir certaines propriétés pour correctement et complètement accomplir sa tâche. Ses propriétés sont [Jain2008]:

- **La diversité** : le modèle stocké du même trait biométrique dans les différentes bases de données de deux systèmes différents doit être suffisamment diversifié pour ne pas permettre une correspondance (*Cross Matching*) si l'une des bases de données est compromise.
- **La révocabilité** : le modèle attaqué doit être révoqué facilement et remplacé sur la base des mêmes données biométriques obtenues auprès l'utilisateur.
- **La sécurité :** le modèle divulgué de la base de données ne doit pas divulguer les informations biométriques originales, empêchant ainsi la création d'une usurpation physique.
- **la précision:** la méthode de protection du modèle ne doit pas dégrader les performances de reconnaissance. Les performances de reconnaissance sont mesurées par le taux de faux rejet et le taux de fausse acceptation, comme nous l'avons expliqué dans le chapitre précédant.

2.9 Méthodes de sécurité des systèmes biométriques

Plusieurs méthodes ont été développées pour sécuriser les systèmes biométriques. nous pouvons les regrouper en trois sections principales qui sont : Cryptographie, Annulabilité, Stéganographie, et les crypto-systèmes biométriques. Ces méthodes peuvent être combinées pour en obtenir un niveau de sécurité plus élevé, mais cela peut entraîner des coûts plus élevés.

2.9.1 *Cryptographie*

Le processus de cryptage dans cette catégorie est achevé par l'utilisation des algorithmes mathématiques transformant les données en une forme illisible. La transformation et la récupération ultérieure des données dépendent d'un algorithme qui peut nécessiter une clé ou non. Cette catégorie est à son tour diviser en trois sous-catégories principales, La cryptographie symétrique, La cryptographie asymétrique et les fonctions de hachage cryptographique.

A) **Cryptographie symétrique**

Cryptage symétrique, également dit cryptage à une clé unique ou à une clé secrète, était le seul type de cryptage utilisé avant le développement du cryptage à clé public dans les années 1970. Cependant, il est encore assez largement utilisé. Un schéma de cryptographie symétrique comporte cinq composants (Figure 2.5):Texte-clair, Algorithme de chiffrement, Clé secrète, Texte chiffrée. Algorithme de déchiffrement.

Souvent, L'application de chiffrement symétrique dans le domaine biométrique peut être obtenue en utilisant comme une entrée en clair les caractéristiques extraites du trait biométrique.

Il existe de nombreux systèmes de cryptage symétriques robustes qui ont prouvé leur efficacité au fil du temps. Cependant, l'un des principaux défis confrontés est de savoir comment partager

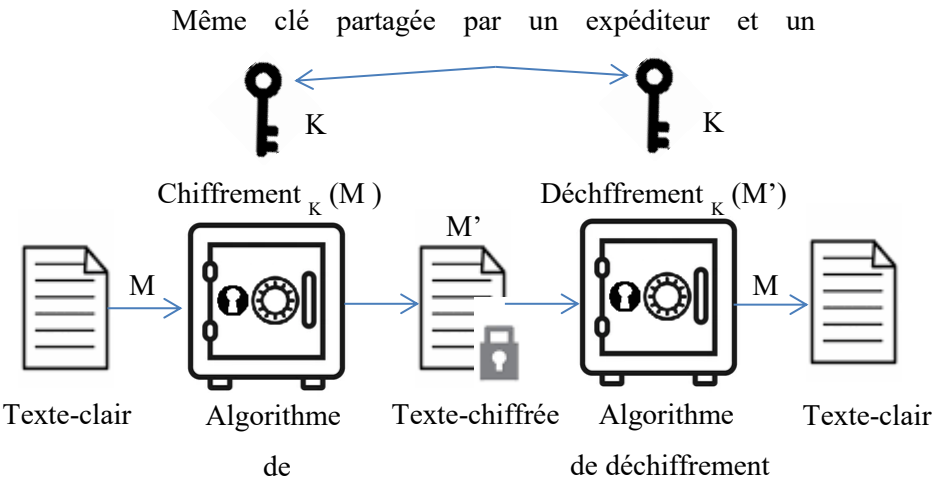


Figure 2.5- Modèle simplifié de cryptage symétrique (Inspiré de [Stallings2014]).

et transmettre en toute sécurité des clés secrètes. Pour cela, les chercheurs ont développé des systèmes de cryptage qui ne reposent pas sur le partage de la même clé pour le cryptage et le décryptage, et c'est ce qu'on appelle un système de cryptage asymétrique.

B) **Cryptographie asymétrique**

Cryptage asymétrique, également dite cryptographie à clé publique, est la plus grande et peut-être la seule véritable révolution de toute l'histoire de la cryptographie. La cryptographie à clé publique est radicalement différente de ce dont nous avons constaté dans cryptographie symétrique. D'une part, les algorithmes à clé publique sont basés sur des fonctions mathématiques plutôt que sur la substitution et la permutation. Plus important encore, la cryptographie à clé publique est asymétrique, impliquant l'utilisation de deux clés distinctes, contrairement au cryptage symétrique, qui n'utilise qu'une seule clé. L'utilisation de deux clés a des conséquences profondes dans les domaines de la confidentialité, de la distribution des clés et de l'authentification.

Un schéma de chiffrement à clé publique comporte six composants (Figure 2.6), En plus des six composants de la cryptographie symétrique qu'ils sont, texte-clair, algorithme de chiffrement, texte chiffrée et algorithme de déchiffrement, ici il s'agit d'une paire de clés qui ont été sélectionnées de sorte que si l'une est utilisée pour le chiffrement, l'autre est utilisée pour le déchiffrement.

Une fausse notion est que le chiffrement à clé publique est une technique à usage général qui a rendu le chiffrement symétrique obsolète. Au contraire, en raison de la surcharge de calcul des schémas actuels de chiffrement à clé publique, il ne semble pas probable que le chiffrement symétrique soit totalement abandonné.

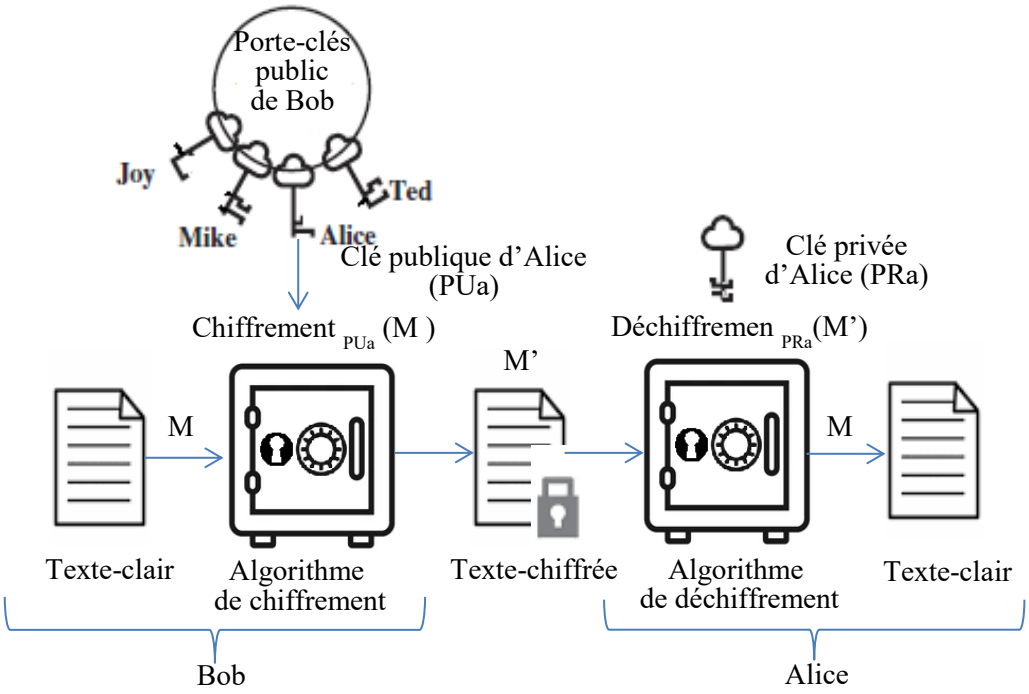


Figure 2.6- Cryptographie à clé publique(Inspiré de [Stallings2014]).

Nous pouvons dire que les deux catégories(cryptage symétrique et asymétrique) peuvent être combinées pour perfectionner le processus et achever à un cryptage complet, en partageant la clé à l'aide d'un système de cryptage asymétrique, puis en utilisant cette clé partagée pour le cryptage symétrique, où les messages et les textes sont chiffrés/déchiffrés à l'aide de ce clé afin de réduire le coût de cryptage de longs textes.

L'utilisation de la cryptographie à clé publique peut être classée en trois sous-catégories Chiffrement/Déchiffrement, Signature numérique (*Digital Signature*), et Échange de clé (*Key Exchange*). Il existe des nombreux crypto-systèmes à clé publique robustes, , dont les plus connus sont: RSA[Rivest1978], *Diffie-Hellman Key Exchange* [Diffie1976], et ECC [Stallings2014].

- ***Diffie-Hellman Key Exchange***

Le premier algorithme à clé publique publié est apparu dans l'article de *Diffie* et *Hellman* qui définissait la cryptographie à clé publique et est généralement appelé *Diffie-Hellman Key Exchange* (DH). Le but de l'algorithme est de permettre à deux utilisateurs d'échanger en toute sécurité une clé qui peut ensuite être utilisée pour le chiffrement symétrique ultérieur des messages. DH est dédié à l'échange de valeurs secrètes. Supposons qu'Alice et Bob souhaitent créer une clé à partager. Figure 2.7 montre comment ce processus déroule en utilisant l'algorithme DH.

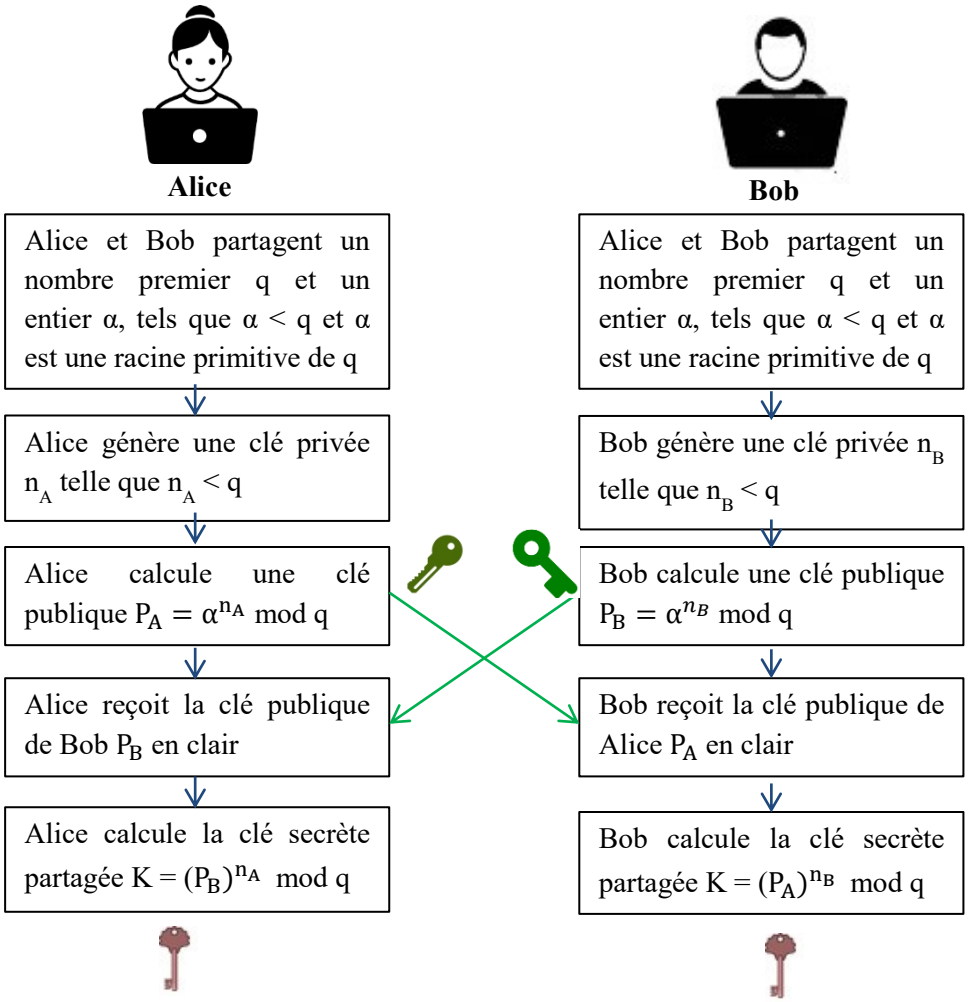


Figure 2.7- Algorithme d'échange de clé Diffie-Hellman (Inspiré de [Stallings2014]).

La sécurité de l'échange de clés par l'algorithme de Diffie-Hellman réside dans le fait que, s'il est relativement facile de calculer des exponentielles de modulo à nombre premier, Mais il est très difficile de calculer des logarithmes discrets. Pour les grands nombres premiers, cette dernière tâche est considérée comme irréalisable.

Étant donné que nous avons utilisé l'algorithme asymétrique ECC dans notre partie expérimentale, nous détaillerons cet algorithme cryptographique dans une section distincte ci-dessous.

C) **Cryptographie sur les courbes elliptiques**

Cryptographie sur les courbes elliptiques, en anglais, *Elliptic Curve Cryptography* (ECC), est l'une des techniques de cryptographie qui utilise la théorie des courbes elliptiques. ECC fait partie des algorithmes de cryptage asymétrique et il est applicable pour le chiffrement, l'établissement de clés, les signatures numériques, les générateurs pseudo-aléatoires, et autant autres tâches. L'utilisation de courbes elliptiques en cryptographie a été suggérée

indépendamment par *Victor S. Miller* en 1985 [Miller1986] et *Neal Koblitz* en 1987 [Koblitz1987]. Les algorithmes de cryptographie à courbe elliptique sont devenus largement utilisés depuis 2004. ECC est connu comme une méthode de cryptage robuste, qui offre le même niveau de sécurité que l'algorithme RSA mais en utilisant une clé plus courte [Barker2020].

Dans la prochaine section, les théories des courbes elliptiques sur les nombres réels $\mathbb{R}$ et les entiers *Zq*, ainsi que quelques applications de la cryptographie basées sur les courbes elliptiques seront expliqués.

1) Courbe elliptique sur $\mathbb{R}$

Soit *E(a, b)* le groupe (ensemble de points (*x, y*)) de la courbe elliptique: $y^2 = x^3 + ax + b$, où *a* et *b* doivent satisfaire:

$$4a^3 + 27b^2 \neq 0 \quad (1)$$

Par exemple la courbe correspondant à *E(1,1)* est $y^2 = x^3 + x + 1$. Figure 2.8 montre cette courbe.

Dans le groupe *E(a, b)*, l'addition est définie de telle sorte que trois points quelconques sur une ligne droite leur somme soit *O*, où *O* est l'identité additive. Par conséquent, les propriétés principales sont:(I) $O=-O$, (II) Pour tout point $P = (x_P,y_P)$, $P+O=P$, d'où $P-P= O$, (III) $P = (x_P,y_P)$, alors $-P = (x_P,-y_P)$, (IV) Pour deux points $P= (x_P,y_P)$ et $Q = (x_Q,y_Q)$ avec $P \neq Q$, $P + Q = N = -(P + Q)$. Où:

$$\begin{cases} x_N = \Delta^2 - x_P - x_Q \\ y_N = \Delta(x_P - x_N) - y_P \\ \Delta = (y_Q - y_P)/(x_Q - x_P) \end{cases} \quad (2)$$

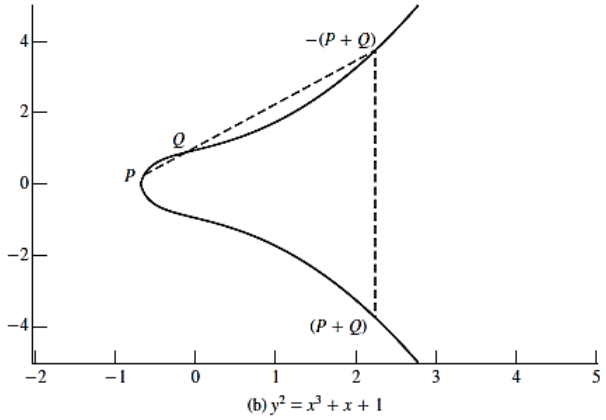


Figure 2.8- Exemple de courbe elliptique correspondant à E (1,1) (Extrait de [Stallings2014]).

Pour la multiplication, ajoutez le nombre à lui-même, par exemple $2P = P + P$. la règle d'addition change et l'équation (17) devient:

$$\begin{cases} x_N = \left(\frac{3x_P^2+a}{2y_P}\right)^2 - 2x_P, \ y_P \neq 0 \\ y_N = \left(\frac{3x_P^2+a}{2y_P}\right)(x_P - x_N) - y_P, \ y_P \neq 0 \end{cases} \quad (3)$$

2) *Courbe elliptique sur Zq*

En *Zq*, tous les entiers appartiennent à l'ensemble fini {0 à *q*-1} et toutes les opérations sont en modulo *q*, où *q* est un nombre premier. Dans ce champ, la formule de la courbe devient: y^2 mod $q=x^3+ax+b$ mod *q*, avec $(4a^3 + 27b^2)$ mod $q \neq 0$ mod *q*, et le groupe correspondant est noté $E_q(a,b)$.

L'addition dans le groupe $E_q(a,b)$ de deux points $P= (x_P,y_P)$ et $Q = (x_Q,y_Q)$ est déduite directement des équations (2) et (3) comme indiqué dans l'équation (4).

$$\begin{cases} x_N = (\Delta^2 - x_P - x_Q) mod\ q \\ y_N = (\Delta(x_P - x_N) - y_P) mod\ q \\ Si\ P \neq Q, \Delta = \left(\frac{y_Q - y_P}{x_Q - x_P}\right) mod\ q. if\ P = Q, \Delta = \left(\frac{3x_P^2+a}{2y_P}\right) mod\ q \end{cases} \quad (4)$$

Pour la multiplication, le produit peut être exprimé en additions successives [Silverman2009], alors l'équation (4) sera utilisée dans les deux cas $P \neq Q$ et $P=Q$ *(2P)*. Exemple: pour calculer *5P*, nous calculons *2 (2P) + P*.

3) *Échange de clé par ECC Diffie-Hellman (ECDH)*

Parmi les principaux algorithmes d'échange de clés basés sur ECC, on trouve le schéma *ECC Diffie-Hellman* (ECDH). ECDH est un algorithme d'établissement de clés basé sur le principe *ECC*. En supposant Alice veut partager une clé secrète dénotée $ECDH_{AB}$ avec Bob, pour ce but, ils acceptent d'utiliser le schéma ECDH. Alice et Bob doivent connaître $E_q(a,b)$ et $G=(x_G,y_G)$ où l'ordre de *G* est un grand nombre n. L'ordre de tout point *N* en $E_q(a,b)$ est défini comme le plus petit entier positif qui satisfait $nN=0$. Tout d'abord, Alice choisit un entier n_A (clé privée) inférieur à n (l'ordre de *G*) et calcule sa clé publique $P_A= n_AG$, notez que la clé publique est également en $E_q(a,b)$. Bob fait la même chose pour calculer n_B et P_B (sa pair de clés privée et publique). Deuxièmement, Alice peut calculer la clé secrète $ECDH_{AB} = n_A P_B$. De même, Bob calcule $ECDH_{AB} = n_B P_A$, notez que les deux calculs aboutissent au même point en $E_q(a,b)$. Comme propriété de $E_q(a,b)$, il est difficile de calculer n_A ou n_B à partir de n_iG même en connaissant *G*, P_A et/ou P_B.

4) *Chiffrement/déchiffrement par la courbe elliptique*

Plusieurs approches de cryptage/décryptage utilisant des courbes elliptiques ont été analysées dans les littératures. Dans cette sous-section, nous expliquons le plus simple. La première tâche

dans ce système est de coder le message en clair M à envoyer en tant que point P_m (x, y). C'est le point P_m qui sera chiffré sous forme de texte chiffré puis le déchiffré. Notez que nous ne pouvons pas simplement coder le message comme la coordonnée x ou y d'un point, car toutes ces coordonnées ne sont pas dans $E_q(a,b)$, il existe plusieurs approches de cet encodage que nous n'aborderons pas ici, mais il suffit de dire qu'il existe des techniques relativement simples qui peuvent être utilisées.

Pour chiffrer et envoyer un message M à Bob, Alice choisit un entier positif aléatoire k et produit le texte chiffré C_m constitué de la paire de points : $C_m = \{kG, P_m + kP_B\}$.

Pour déchiffrer le texte chiffré, Bob multiplie le premier point de la paire par sa clé privée et soustrait le résultat du deuxième point (Equation 5):

$$P_m + kP_B - n_B(kG) = P_m + k(n_B G) - n_B(kG) = P_m \quad (5)$$

Alice a masqué le message P_m en y ajoutant kP_B. Personne ne connaît la valeur k qu'Alice, donc même si P_B est une clé publique, personne ne peut supprimer le masque kP_B. La sécurité de l'ECC dépend de la difficulté de déterminer k étant donné kP et P. C'est ce qu'on appelle le défi du logarithme de la courbe elliptique.

Selon la comparaison NIST [Barker2020] entre RSA et ECC montrant des tailles de clés comparables en termes d'effort de calcul pour la cryptanalyse, une taille de clé considérablement plus petite peut être utilisée pour ECC par rapport à RSA. En outre, pour des longueurs de clé égales, l'effort de calcul requis pour ECC et RSA est convergent.

D) **Les fonctions de hachage cryptographique**

Une fonction de hachage H (.) accepte un bloc de données M de longueur variable en entrée et produit une valeur de hachage de taille fixe h = H (M) à la sortie. Une «bonne» fonction de hachage a la propriété que les résultats de l'application de la fonction à un grand ensemble d'entrées produiront des sorties qui sont uniformément réparties et apparemment aléatoires. En termes généraux, le principal objet d'une fonction de hachage est l'intégrité des données. Une modification de n'importe quel bit ou bits de M entraîne une modification totale du h. Les fonctions de hachage sont des fonctions à sens unique « *One Way Function* » qui sont faciles à calculer dans une direction mais très difficiles à inverser.

Il existe beaucoup de fonctions de hachage cryptographique sans clé, la plus robuste et la plus utilisée est la famille *Secure Hash Algorithm* (SHA) et la famille *Message-Digest Algorithm* (MD).

Les Méthodes de cryptage standard ne peuvent pas être utilisées pour crypter les vecteurs biométriques car, (*I*) Il y a un problème de variabilité intra-classe, c'est-à-dire la différence entre les vecteurs biométriques obtenus du même utilisateur. Chaque fois qu'un utilisateur accède au système, les informations fournies au système et les informations qui y sont stockées se diffèrent même légèrement. Les fonctions de cryptage standard est intolérante, même une différence d'un bit à l'entrée entraîne une modification de plus de 50% des bits à la sortie, donc la mise en correspondance n'est pas possible. (*II*) déchiffrer le modèle à chaque fois avant la mise en correspondance présente un risque potentiel, car cela dévoilera le modèle d'origine durant chaque session d'authentification. Donc, il est impératif de fournir des méthodes de protection tolérantes et non rigides qui correspondent aux caractéristiques des technologies biométriques.

2.9.2 ***Annulabilité***

Aussi connu sous le nom « données biométriques annulables » (*Cancellable Biometrics*). Les données biométriques annulables sont conçues pour permettre à une personne d'enregistrer et de révoquer un grand nombre des vecteurs modèles différents. Chaque image biométrique est codée avec un motif de distorsion qui varie pour chaque application. Le concept a été développé pour répondre aux problèmes de confidentialité et de sécurité selon lesquels les échantillons biométriques sont limités et doivent être utilisés pour de multiples applications, et quand ils sont volés, ils ne sont pas indemnisés [Adler2009].

Lors de l'enregistrement, le capteur recueille les caractéristiques biométriques de l'utilisateur, le système extrait ensuite le vecteur biométrique requis sous la forme d'un modèle. Ce modèle est distordu par une fonction de transformation et le modèle transformé résultant est stocké dans la base de données. La nature annulable de ce schéma est fournie par la distorsion, car les données biométriques réelles de l'utilisateur ne sont pas stockées. Lorsque l'utilisateur doit être authentifié, le vecteur requête est formé en utilisant les caractéristiques biométriques fournies, et ce vecteur est transformé en utilisant la même fonction. Le modèle transformé est ensuite sera comparé avec celui stocké dans la base de données. Dans le cas où le modèle stocké est compromis, il doit être remplacé par un autre transformé par une autre application.

Les méthodes de transformation des vecteurs biométriques sont divisées en deux catégories [Riaz2018], (*I*) Méthodes de transformation inversible, Ils sont également appelés biohachage (*Biohashing* ou *Biometric Salting*). Ils sont souvent utilisés avec une clé. (*II*) Méthodes de transformation non inversible.

2.9.3 ***Stéganographie***

La stéganographie est une technique de dissimulation des données dans laquelle un message secret est caché dans un autre message indépendant, tel qu'une image ; puis communiqué à l'autre partie [McBride2005]. Le processus de stéganographie comporte trois éléments de base: (*I*) les données à masquer, (*II*) le fichier de couverture, dans lequel les données secrètes doivent être intégrées, (*III*) le fichier résultant. Au moyen du cryptage, le contenu secret du message envoyé est visible et illisible, mais par la stéganographie, le contenu secret de l'envoi est à la fois lisible et invisible. Il existe de nombreux algorithmes pour les systèmes stéganographiques, les plus connus sont, l'algorithme du bit de poids faible (*Least Significant Bit* -LSB) [Ker2007] et les algorithmes basés sur la théorie de chaos [Britannica2021].

2.10 Crypto-systèmes Biométriques

Dans un crypto-système biométrique, le vecteur de caractéristique biométrique est essentiellement codé avec une clé secrète pour en résulter une forme sécurisée (vecteur protégé), qui ne révèle aucune information significative ou cruciale sur le vecteur de caractéristique biométrique ou sur la clé secrète associée. Plus la sécurisation du vecteur biométrique, ces systèmes peuvent être également utilisés pour assurer le partage confidentiel de la clé secrète.

Dans tels systèmes, si la clé secrète est générée de manière aléatoire quelles que soient les données biométriques, alors le crypto-système appartient aux techniques de liaison de clé (*Key Binding Techniques*). Alors que s'il est généré à partir de données biométriques (à partir du vecteur de caractéristiques), le crypto-système appartient aux techniques de génération de clé (*Key Generation Techniques*). Dans ce second cas, la clé générée ne doit aussi révéler aucune information par laquelle les données biométriques peuvent être récupérées.

2.10.1 ***Techniques de liaison des clés***

La clé de chiffrement/déchiffrement dans ces techniques est générée indépendamment du caractéristiques biométriques. Ensuite, il est joint au le vecteur biométrique pour former une seul forme sécurisée. Les algorithmes les plus connus basés sur ce type de technique sont, *Fuzzy Commitment* [Juels1999] et *Fuzzy Vault* [Juels2006].

A) **Fuzzy Commitment**

Le schéma d'engagement cryptographique conventionnel (*Commitment Scheme*) est un type de dissimulation et de liaison de clé avec un vecteur appelé le témoin (*Witness*). Depuis plusieurs années, le schéma d'engagement a été largement adopté dans les données en texte brut (texte en clair). La propriété de dissimulation de ce schéma rend les informations cruciales (la clé ou le

témoin) impossibles à deviner. La forme sécurisée issue du schéma d'engagement appelée *Commitment*, et elle ne peut pas être déchiffrée ce *Commitment* (séparer la clé du témoin) avec un autre témoin différent.

Pour expliquer techniquement le schéma d'engagement, Nous supposons: Bob est un utilisateur avec un mot de passe ou un identifiant personnel W, et un système de contrôle d'accès qui enregistre le W et une clé secrète K dans un *Commitmnet*. Ce *Commitment* est produit par un simple XOR entre W et K. Aussi le système enregistre le hach de K. Si Bob souhaite accéder le système, il doit présenter le même W pour pouvoir récupérer K par un autre XOR entre le *Commitmnet* et le W présenté. La même fonction de hachage est appliquée à la sortie K, afin de la comparer avec le hach de K enregistré. L'équation (6) explique le schéma d'engagement.

$$(W \oplus K) \oplus W = K \tag{6}$$

Jusqu'à là, W présenté et W enregistré doivent être strictement identiques et il n'y a aucune possibilité de tolérance d'erreur. Malheureusement, si W est un vecteur de caractéristiques biométriques, ce schéma n'est pas adéquat, car les deux vecteurs biométriques de caractéristiques, même s'ils sont extraits de la même personne, ne sont pas strictement identiques. Si nous voulons obtenir une certaine tolérance d'erreur dans W, c'est-à-dire que l'utilisateur peut présenter un témoin proche W', certains codes correcteurs d'erreurs doivent être employés car $W \oplus K \oplus W' = K'$ et K' est également proche de K .C'était l'idée du schéma d'engagement flou (*Fuzzy Commitment*) proposé par *Juels* et *Wattenberg* [Juels1999], nous parlons de flou car les témoins ne doivent pas être strictement identiques et une certaine proximité peut être tolérée.

Comme tous les systèmes biométriques, le crypto-système biométrique basé sur l'engagement flou se compose de deux phases principales: l'enregistrement (enrôlement) et l'authentification (contrôle d'accès).

Durant la phase d'enrôlement, le vecteur modèle W_i est extrait d'un nouvel utilisateur U_i. Puis transformé sous forme binaire à l'aide d'une étape quantification [Nandakumar2015] suivie d'une étape de codage. Le nombre des niveaux de quantification doit être soigneusement choisi car autant des niveaux impliquent une entropie élevée et donc un vecteur long, tandis qu'un petit nombre peut entraîner une perte d'informations et donc une perte de la précision (semblable au théorème d'échantillonnage de *Shonnon-Nequist*). Par la suite, une clé secrète aléatoire K est générée et codée par un code correcteur pour former un code binaire C. Le code de correction est généralement utilisé pour corriger les erreurs de transmission. Dans l'engagement flou, ce code

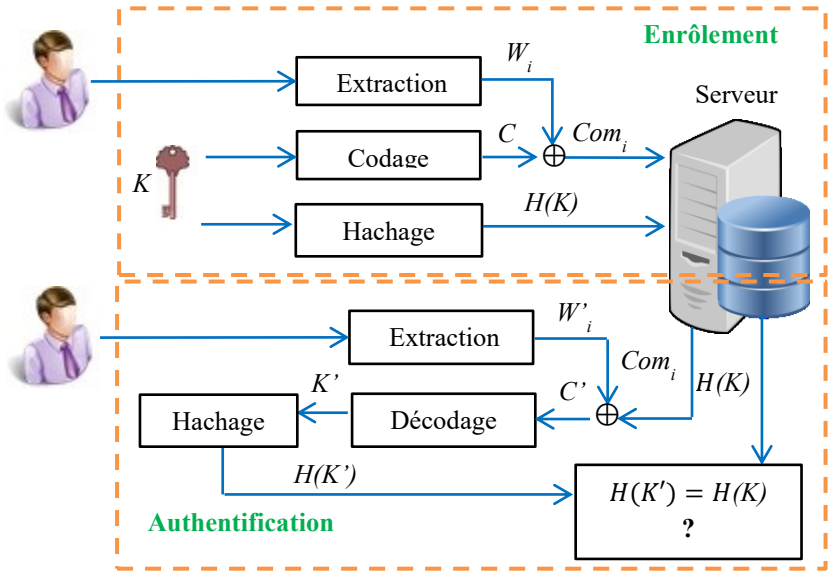


Figure 2.9- Diagramme de flux d'un crypto-système biométrique basé sur *Fuzzy Commitment* (Inspiré de [Adamovic2016]).

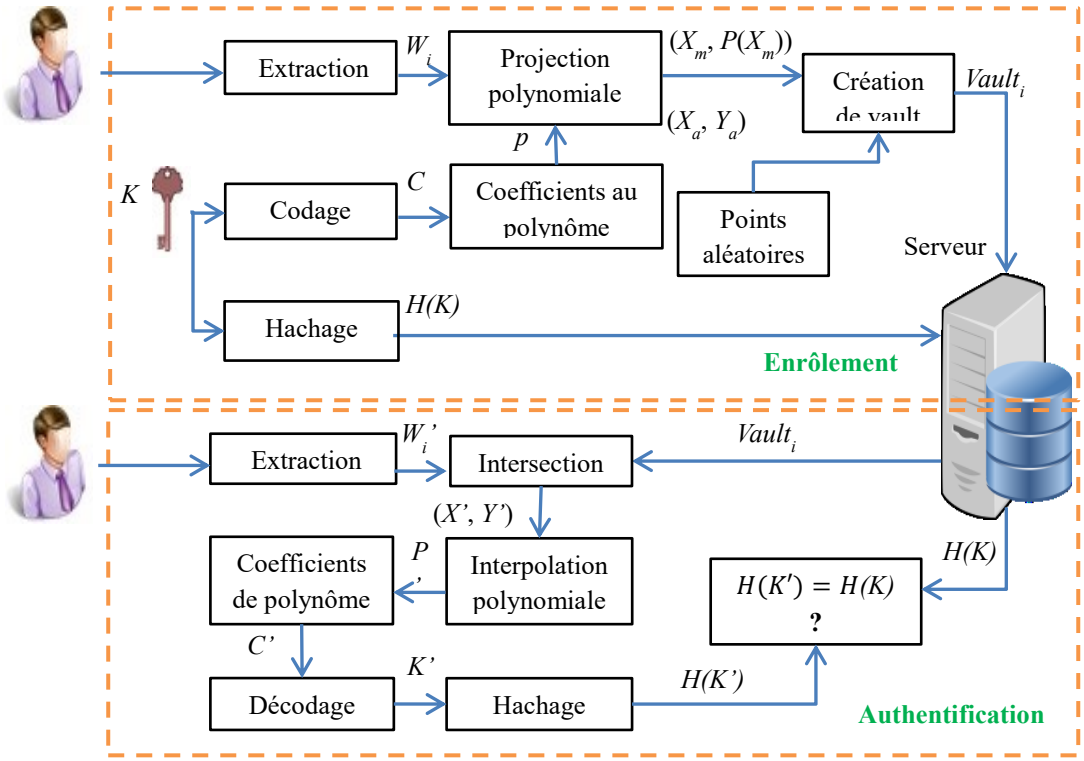


Figure 2.10- Diagramme de flux d'un crypto-système biométrique basé sur *Fuzzy Vault* (Inspiré de [Uludag2006]).

peut résoudre le problème de la divergence d'échantillon intra-classe et conférer au crypto-système une particularité tolérante aux erreurs. Ainsi, W_i et C sont liés par un XOR pour former un *Commitment* noté Com_i, ($Com_i = W_i \oplus C$), puis la fonction de hachage (fonction à sens unique) est appliquée sur la clé d'origine K pour obtenir une signature $H(K)$. Enfin, le système enregistre $H(K)$ et Com_i de chaque utilisateur dans sa base de donnés.

Durant la phase d'authentification, U_i présente à nouveau son W_i'. Le système applique d'abord une fonction XOR entre W_i' présenté et Com_i stocké. Ensuite, le résultat (qui est une chaîne binaire) C' est décodé pour obtenir K' puis il est haché pour obtenir la signature ($H(K')$). Si $H(K) = H(K')$, la clé est alors récupérée et par conséquent W_i et W_i' sont alors de la même personne, ce qui implique que U_i est authentifié avec succès. Notez qu'une grande différence entre W_i et W_i' (personne différente) conduit à un échec de décodage. Dans ce cas, la clé ne peut pas être récupérée et le candidat est supposé imposteur. Figure 2.9 explique les deux phases (l'enrôlement, l'authentification) d'un crypto-système biométrique basé sur l'engagement flou.

B) **Fuzzy Vault**

Un autre type d'algorithme populaire basé sur les techniques de liaison de clé a été proposé par *Juels* et *Sudan* sous le nom "*Fuzzy Vault*" [Juels2006]. La traduction terminologique de *Fuzzy Vault* est le coffre-fort flou.

Il lie la clé et le vecteur biométrique dans une forme sécurisée connue sous le nom de "*Vault*". Tel *Vault* contient des points extraits de la biométrie et correspondent à une projection d'un polynôme extrait de la clé. Pour le camouflage, quelques points aléatoires appelés *Chaff Points ;* qui ne sont pas calculés par le polynôme, sont ajoutés à la *Vault*. Bien entendu le vecteur est un vecteur biométrique, le système doit être tolérant aux erreurs, et ainsi certains codes correcteurs doivent être inclus dans ce schéma pour corriger les légères différences. Un nombre suffisant des points (pairs) du vecteur requête doivent être extrait pour une interpolation polynomiale efficace.

Premièrement, dans la phase d'enrôlement d'un système coffre-fort flou, W_i et la clé codée C de la clé K. sont utilisés pour former des paires de points $(X_m, P(X_m))$, où P est un polynôme dont les coefficients sont construits à partir d'éléments de C, X_m sont les points extraits du vecteur biométrique modèle et $P(X_m)$ sont les points résultants de la projection de X_m sur P. Ensuite, $Vault_i$ est créée en ajoutant un nombre suffisant de pairs des points aléatoires (X_a, Y_a) sans projection de X_a sur P (c'est-à-dire $Y_a \neq P(X_a)$). Enfin, le système enregistre $H(K)$ et $Vault_i$ de chaque utilisateur U_i dans sa base de donnés. Il est important de noter qu'un bon camouflage peut être obtenu en ajoutant un grand nombre des points aléatoires, ce qui améliore la sécurité de la sauvegarde.

Durant la phase d'authentification, *Ui* présente à nouveau son W_i'. Le système recherche l'intersection entre W_i' présenté et le $Vault_i$ stocké. Les points obtenus (X', Y'), qui peuvent inclure quelques points erronées; sont analysés par interpolation polynomiale telle que Lagrange

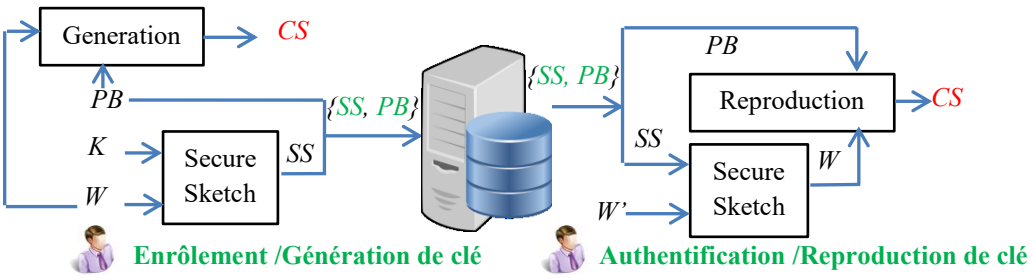


Figure 2.11- Diagramme de flux d'un crypto-système biométrique basé sur *Fuzzy Extracor*.

ou la méthode de moindre carré [Hazewinkel2002] pour trouver le polynôme P'. Il est important de noter que même avec quelques paires erronées, P' peut être correctement récupéré, mais avec certaines limitations. Ensuite, les coefficients C' de P' sont décodés pour déterminer K'. K' est haché pour obtenir la signature ($H(K')$). Si $H(K) = H(K')$, la clé est alors récupérée et par conséquent W_i et W_i' appartiennent à la même personne (U_i est authentifiée avec succès). Figure 2.10 montre les deux phases (l'enrôlement, l'authentification) d'un crypto-système biométrique basé sur le coffre-fort flou.

2.10.2 *Techniques de génération des clés*

Dans tels systèmes, une clé secrète *CS* est générée directement à partir du vecteur de caractéristiques biométriques. Le schéma le plus populaire basé sur les techniques de génération de clé est l'extracteur flou (*Fuzzy Extractor - FE*). FE a été proposé par *Dodis et al.* en 2004 [Dodis2004]. Dans leur proposition, l'une des techniques de de liaison de clé soit le *Fuzzy Vault* ou le *Fuzzy Commitment* est utilisée. Dans les deux cas, la forme sécurisé résultant est appelé "*Secure Sketch - SS*". Dans ce schéma, *CS* est reproduit à partir de *SS* à l'aide d'une chaîne aléatoire publique *PB,* pour rendre une attaque par force brute plus difficile. L'objectif principal de FE est de récupérer le vecteur modèle d'origine utilisé pour calculer *SS* stockée sur le serveur. À partir de ce vecteur exact, *CS* peut être reproduit.

FE se compose de deux fonctions : Gen (.) pour générer la clé et Rep (.) pour reproduire la clé. Gen (.) prend le vecteur biométrique *W* comme un entrée. Alors que la clé secrète *CS* et le paramètre public *PB* sont des sorties, cette fonction est abrégée: Gen (*W*) = (*CS*, *PB*). Il convient de noter que *W* est lié à une chaîne aléatoire *K* en utilisant l'une des techniques de liaison de clé (FC ou FV) pour générer une *SS*. Après la génération de la clé, *CS* et *K* doivent être supprimés. La deuxième fonction, Rep (.) prend le vecteur biométrique *W'* et *PB* comme entrées, et donne *CS*' comme sortie. Donc *W* original peut être récupéré à partir de *SS*, puis *CS'* peut être reproduite

à partir de *W*. Nous notons cette fonction *CS*'=Rep (*W', PB*). Figure 2.11 montre le schéma bloc du *Fuzzy Extractor*.

En d'autres termes, FE permet d'extraire un certain caractère aléatoire *CS* de *W*, puis de reproduire avec succès *CS* à partir de n'importe quelle chaîne *W'* proche de *W*. La reproduction se fait à l'aide de la chaîne publique *PB* produite lors de l'extraction initiale. Mathématiquement on dit, pour tout *W, W'* satisfaisant Dis (*W*, *W'*) < *t*, si *CS*, *PB* ← Gen (*W*), alors on a Rep (*W', PB*) = *CS*, où Dis est la distance (la différence) entre (*W*, *W')* et *t* est la plus grande différence possible (nous en discuterons dans la prochaine section des codes correcteurs).

Récemment, l'extracteur flou est utilisé dans des nombreuses applications. Outre l'authentification, il a prouvé son efficacité dans le partage sécurisé de la clé qui rivalise avec les techniques de cryptage traditionnelles. En plus, il est flexible et compatible avec les exigences des technologies biométriques. Pourtant la génération directe des clés à partir de la biométrie est attrayante et utile dans divers applications cryptographiques, les clés à entropie élevée peuvent toujours être difficiles à générer.

2.11 Conclusion

En général, ce chapitre fournit tout ce qui concerne la biométrie et ses systèmes. Le chapitre montre la possibilité de collecter plus d'une modalité biométrique dans un seul système pour combler les lacunes d'un système unimodal. Les systèmes de technologie biométrique en termes de sécurité et de protection ont également été abordés. L'utilité de protéger le vecteur des caractéristiques biométriques a été présentée où les différents risques informatiques et les cybers attaques sont élaborés. Quelques méthodes de protections pionnières sont détaillées en révélant les avantages et les limites de chacune. Plusieurs algorithmes répandus ont été expliqués où l'accent est mis sur les crypto-systèmes biométriques, car notre travail en dépend principalement.

Les systèmes biométriques sont inclus dans presque tous les systèmes d'identification et de sécurité, surtout dans le plus grand réseau d'internet tel que l'IoT. L'objectif de notre thèse se distingue notamment par l'implication des systèmes biométriques dans l'IoT, où l'être humain joue un rôle primordial dans tel réseau.

Le chapitre suivant présente une synthèse de l'état de l'art de l'implication de la biométrie dans l'IoT, tout en notant les lacunes qui existent. Nous suggérons également des solutions possibles pour combler ces lacunes.

Chapitre 3 : Biométrie dans l'IoT : Implications, Problèmes, et Contributions

Chapitre 3: Biométrie dans l'IoT : Implications, Problèmes, et Contributions

3.1 Introduction

Dans le premier chapitre, nous avons tout couvert sur l'IoT, son architecture et sa sécurité. Tandis que le deuxième chapitre traitait des technologies biométriques et de leur sécurité. Le présent chapitre est comme une sorte d'une combinaison de tous ces concepts.

Ce chapitre traite de l'inclusion et de l'implication de la technologie biométrique dans l'IoT. Le chapitre est divisé en deux parties principales, la première présente un état de l'art et les progrès menés dans cet axe. Tandis que la deuxième partie est un recueil de nos contributions à l'égard de l'intégration de la biométrie dans l'environnement IoT. L'état de l'art et nos contributions se font à trois niveaux, au niveau de l'architecture IoT de bout-en-bout, au niveau de l'accès au cloud, et au niveau du modèle S^2aaS. A chaque niveau, nos contributions et le système d'authentification et d'échange de clé proposé sont discutés.

3.2 Où la biométrie peut- être impliquée dans l'IoT?

Avant de parler de « Où et Comment » utiliser la technologie biométrique pour sécuriser l'IoT, nous éclaircissons où l'être humain peut s'intégrer dans le paradigme IoT. D'après ce qui précède dans le chapitre précédent, nous constatons qu'un être humain peut interagir avec l'IoT à travers trois situations (Figure 3.1). La première situation est en tant qu'un consommateur de données dans le domaine utilisateur. La deuxième en tant qu'un producteur de données dans le domaine de la perception, et la troisième en tant qu'un propriétaire de données dans le domaine des capteurs et propriétaires de données selon le modèle S^2aaS.

Dans la première situation, Le propriétaire des données est l'être humain qui possède tous les objets personnels, tels que les téléphones portables, les montres bracelets, les lunettes, les ordinateurs portables, les produits alimentaires, et les articles électro-ménagers. Etc. Ou tout simplement, n'importe quel objet personnel embarquant des capteurs, des actionneurs, des RFIDs. Les pouvoirs du propriétaire du service peuvent être résumés comme suit:

- Le propriétaire doit pouvoir décider avec qui il souhaite partager ses données et dans quelles conditions.
- Le propriétaire devrait motiver à partager ses données et recevoir des avantages en retour.

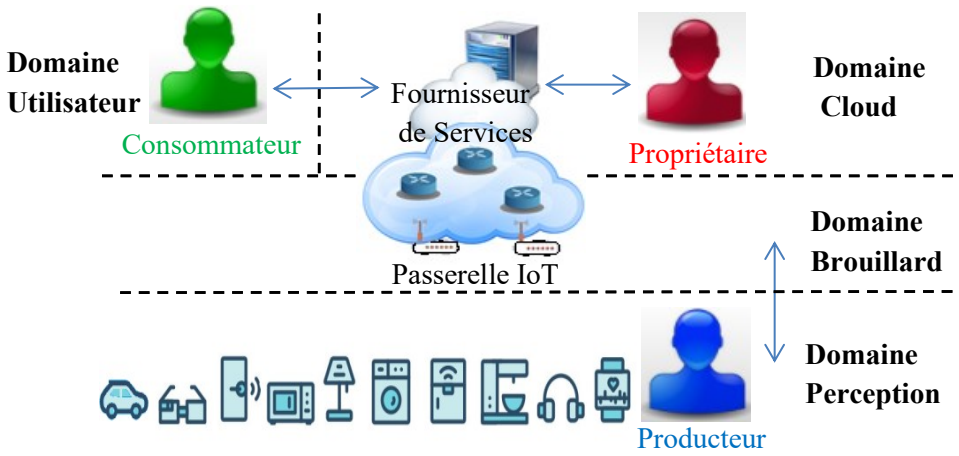


Figure 3.1- Les trois rôles que les humains peuvent jouer dans l'IoT.

- En même temps, cela devrait inciter des tiers à proposer des services d'analyse de données afin que même les propriétaires non spécialisés puissent bénéficier du partage de leurs données.

Un être humain peut être considéré comme un producteur de données si les données proviennent principalement de son corps. Dans ce cas, une personne est traitée comme n'importe quel d'objet intégré avec des capteurs, des actionneurs et/ou RFID. Le meilleur exemple de ce cas est celui des applications IoMT, ces applications incluent la surveillance du niveau de glucose, de l'électrocardiogramme (ECG), de la pression artérielle, de la température corporelle, de la saturation en oxygène, Etc.

L'être humain en tant que consommateur des données ; appelé aussi l'utilisateur, est la personne qui utilise les données produites dans le domaine de la perception. C'est le cas le plus courant dans le rôle de l'être humain dans l'IoT, où l'esprit imaginé est immédiatement le rôle de l'humain dans l'IoT en tant qu'un utilisateur.

Où l'être humain peut être intégré dans l'IoT, la technologie biométrique peut être impliquée. Cela signifie que la technologie biométrique peut être mise en œuvre dans le domaine d'utilisateur, le domaine du cloud, ou le domaine de la perception.

Les mécanismes de sécurité utilisés dans les trois situations, y compris les mécanismes biométriques, sont les mêmes. Les protocoles de sécurité IoT et les algorithmes légers peuvent être mis en œuvre pour n'importe quel rôle humain, qu'il soit producteur, consommateur ou propriétaire. Pour cela, il suffit que nous nous concentrons uniquement dans cette thèse sur le rôle du consommateur, et cela reflète les autres rôles. Cependant, nous avons abordé ce rôle à plusieurs niveaux, au niveau IoT de bout en bout, au niveau d'accès au cloud, et au niveau IoT à quatre domaines (modèle S^2aaS).

3.3 Extraction de caractéristiques biométriques

Techniquement, le premier processus d'intégration de la biométrie dans l'IoT est l'extraction de caractéristiques biométriques, et cette section est dédiée à ce point.

Les techniques d'extraction des caractéristiques peuvent utiliser des méthodes de filtrage, de transformation, d'analyse ou d'hybridation entre eux pour extraire le maximum d'informations utiles et uniques. Pour obtenir un bon taux de reconnaissance, l'extraction des caractéristiques doit garantir une petite variabilité intra-classe et une grande variabilité interclasse. En général, les techniques d'extraction peuvent être classées en deux catégories principales, techniques classiques et techniques basées sur l'apprentissage automatique « *Machine Learnnig*».

3.3.1 *Techniques classiques*

L'image contient de nombreuses caractéristiques qui peuvent être exploitées pour l'analyser. Les types des caractéristiques extraites varient en fonction de la technique utilisée. Donc, il est nécessaire de connaître le type de la scène pour choisir efficacement la technique appropriée. Par conséquent, Comme ces techniques nécessitent une expertise humaine, on les appelle des classiques. Cette catégorie est également divisée en plusieurs classes, celle basé sur les lignes, sur la texture et sur l'apparence. Ces classes incluent de nombreuses techniques, nous citons celle utilisée dans ou liée à notre travail, par exemple celles basées sur les lignes, le filtre de Gabor [Lee1999], [Meraoumia2014] et le modèle binaire local (*Local Binary Pattern - LBP*) [Mohammadi2015]. Parmi les méthodes basées sur la texture nous trouvons principalement, la transformation de Fourier rapide (*Fast Fourier Transform - FFT*) [Li2002], [Tachaphetpiboon2006], la transformation en ondelettes discrète (*Discrete Wavelet Transform - DWT*) [Huang2015] et la transformation en cosinus discret (*Discrete Cosine Transform - DCT*) [Jing2004], [Badrinath2012]. Et enfin, l'analyse en composantes principales (*Principal Component Analysis - PCA)* [Yang2002], [Chen2004] comme une méthode basées sur l'apparence. Parmi les différentes techniques d'extraction classiques, nous entamons DWT, DCT et PCA car ils font partie de nos travaux expérimentaux.

A) Transformation en ondelettes discrète (DWT)

L'une des techniques de présentation des propriétés de la dé-corrélation spatio-fréquentielle est la transformé en ondelettes. L'analyse d'une image en utilisant cette transformée est très adaptée aux applications biométriques [Akansu2010]. A l'instar de FFT et DCT, l'idée de base est de séparer entre les hautes et les basses fréquences d'un signal y compris l'image. Les basses fréquences présentent l'allure générale (version grossière) de l'image. Alors que les hautes

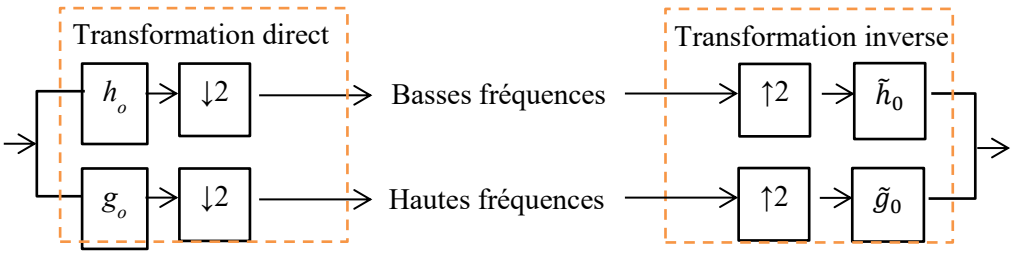


Figure 3.2- Filtres PB, PH, sous-échantillonnage et sur-échantillonnage utilisés dans DWT-1D (Inspiré de [Huang2015]).

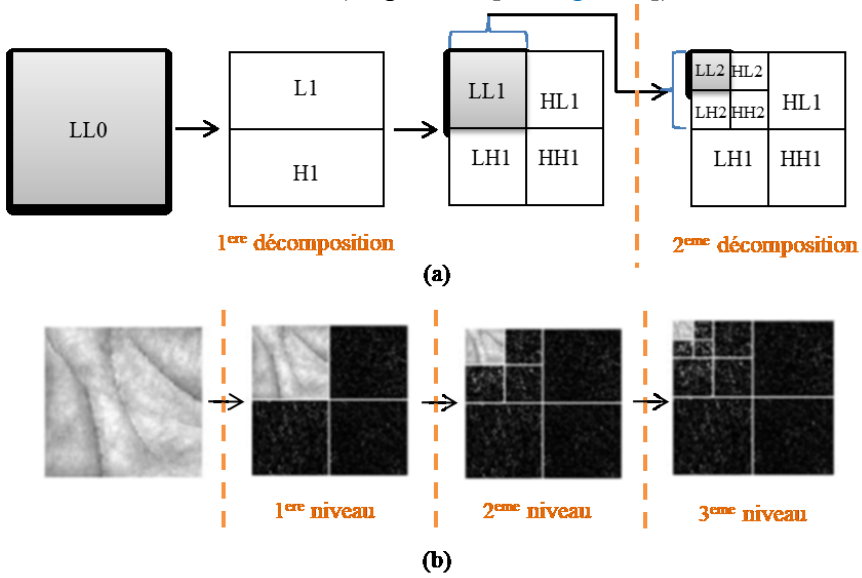


Figure 3.3- les étapes de la décomposition en ondelette d'une image, (a) décomposition en ondelette à deux niveaux, (b) exemple d'application pour une image de paume (Inspiré de [Huang2015]).

fréquences expriment les détails d'image ou les variations brusques entre les pixels voisins. De points de vue de l'énergie, l'énergie de l'image est concentrée dans les basses fréquences.

Le premier qui a développé les ondelettes est *Morlet* basé sur les études de *Haar* [Theodoridis2009]. *Moret* l'a défini comme une famille des fonctions d'énergie finie et de moyenne nulle dilatées et translatées à partir d'une fonction de base. L'équation (7) présente les dilations et les transitions d'une fonction de base $\psi(x)$ qui peuvent être générées pour définir une famille des fonctions ondelettes.

Où a et b sont respectivement les paramètres de dilatation et de translation.

La transformée en ondelette continue d'une fonction *f* à l'abscisse *b* et à l'échelle *a* est définie par le produit scalaire entre *f* et l'ondelette de l'équation (7).

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}}\psi\left(\frac{x-b}{a}\right) \quad (b,a) \in \mathbb{R}^2, a \neq 0 \qquad (7)$$

$$wf(a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(x)|a|^{-\frac{1}{2}} \psi^* \left(\frac{x-b}{a}\right) dx \quad (8)$$

Analysons l'équation (8), il s'est constaté que la transformée en ondelette est une présentation bidimensionnelle d'un signal unidimensionnel. Ce qui n'est pas le cas de la transformée de Fourier.

Le calcul des coefficients de la transformée en ondelettes continue peut être lourd en générant un volume important des données à chaque échelle possible qui est aussi continu. Donc l'application de ce mode continu n'est pas pertinente pour les signaux discrets tels que l'image.

En choisissant des sous-ensembles de (a, b); quelques échelles, nous serons face à un autre mode de la transformée en ondelette dite discrète (DWT). Pour une machine, l'implémentation de DWT est beaucoup plus facile que la transformée continu (voir impossible d'implémenter le continu). En effet, l'implémentation de DWT n'est qu'une application successive de paires de filtres passe-bas et passe-haut (h_o et g_o), suivis d'un sous-échantillonnage (↓) de facteur deux. Pour la transformée inverse, un autre filtre passe-bas et passe-haut ($\tilde{h}_0$ et $\tilde{g}_0$) suivis d'un sur-échantillonnage (↑) de facteur deux sont utilisés. Figure 3.2 montre ce propos.

La DWT 1D peut facilement être étendue à deux dimensions en appliquant les filtres successivement dans les deux directions de l'image. La décomposition en ondelettes d'une image se déroule donc selon les étapes suivantes [Rabbani2002]:

Etape 1. Décomposer chaque colonne d'image en utilisant verticalement les filtres 1D.

Etape 2. Appliquer les mêmes filtres pour les deux bandes résultantes de la première étape mais suivant les lignes.

Etape 3. Répéter le même processus pour la sous-bande basse-basse fréquence (LL) (le quart Haut-Gauche).

L'application de DWT pour des raisons de reconnaissance biométrique il faut tout d'abord recadrer (prendre une partie) l'image selon des points référentiels appelés les régions d'intérêts (*Region of Interest - ROI*). Puis, déterminer les niveaux qui peuvent être utilisés. Plus le nombre des niveaux est grand, plus la discriminabilité est élevée. Mais en détriment de nombre des blocs engendrés qui influe directement sur la longueur du vecteur des caractéristiques. C'est pour cette raison que le choix de nombre des niveaux doit être un compromis entre les deux propriétés. A titre d'exemple la Figure 3.3 (a) montre l'application de DWT à trois niveaux ce qui engendre 9 blocs. Pour chaque bloc nous calculons soit l'énergie moyenne, l'écart-type ou l'entropie soit

une combinaison de ces trois paramètres [Tewari2014]. L'équation (9) montre la formulation pour calculer ces paramètres.

$$\begin{cases} énergie\ moyenne = \frac{1}{N^2}\sum_{k=1}^{N}\left|C_k^{DWT}\right| \\ écart-type = \sqrt{\frac{1}{N^2}\sum_{k=1}^{N}\left(\left|C_k^{DWT}\right| - \bar{C}\right)^2} \\ entropie = \frac{1}{N^2}\sum_{k=1}^{N}\left|C_k^{DWT}\right|^2 \log\left|C_k^{DWT}\right|^2 \end{cases} \quad (9)$$

Où N est le nombre des coefficients dans un bloc, k est le numéro du coefficient, C_k^{DWT} le coefficient DWT et $\bar{C}$ est la moyenne des coefficients du bloc. A ce stade nous avons obtenu un vecteur contenant des valeurs réelles (coefficients DWT transformés). Pour avoir un vecteur final approprié il reste deux étapes : la quantification et le codage.

B) **Transformation en cosinus discret (DCT)**

La transformée en cosinus discret (DCT) expresse une séquence finie en terme de la somme des fonctions cosinus de différentes fréquences. DCT a été proposée la première fois par *Nasir Ahmed* en 1972 [Ahmed1974]. C'est une technique de transformation largement utilisée en traitement de signal et la compression des données. Il est connu que l'œil humain est moins sensible aux certaines fréquences de l'image.

La transformée DCT permet d'éliminer certaines fréquences que l'œil humain n'aperçoit pas. Ce qui revient à supprimer les hautes fréquences de l'image tout en gardant les données importantes représentées par les basses fréquences. DCT est transformée liée à celle de Fourier et très similaire à la transformée de Fourier discrète (DFT), mais elle utilise des nombres réels seulement. Plus exactement, les coefficients DCT sont déduits des coefficients des séries de Fourier de la séquence étendue périodiquement et symétriquement, alors que ceux de DFT sont déduits de la séquence périodiquement étendue. Les DCTs sont équivalentes aux DFTs d'environ deux fois la longueur, car la fonctionnalité est sur des données réelles avec une symétrie paire (puisque la transformée de Fourier d'une fonction est réelle et paire), alors que dans certaines variantes les données d'entrée et/ou de sortie sont décalées de moitié des échantillons.

Les DCT multidimensionnels, dénommé MD DCT, sont développés pour étendre le concept de DCT sur les signaux de média numérique. DCT se calcule comme il est indiqué par l'équation (10). Pour la généralisation en 2D l'équation (11) peut être utilisée.

$$F(u) = \frac{1}{\sqrt{2M}} C_u \sum_{i=0}^{N-1} f(j) \cos\left(\frac{(2i+1)u\pi}{2N}\right) \quad (10)$$

$$F(u,v) = \frac{1}{\sqrt{2N}} C_u C_v \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} f(j,j) \cos\left(\frac{(2i+1)u\pi}{2N}\right)\cos\left(\frac{(2j+1)v\pi}{2M}\right) \quad (11)$$

$$f(\gamma) = \begin{cases} \sqrt{\frac{1}{N}} \; Si \; \gamma = 0 \\ \sqrt{\frac{2}{N}} \; Si \; \gamma \neq 0 \end{cases}$$

Pour les systèmes de la reconnaissance dédiés aux images biométriques, le DCT le plus approprié est le DCT par bloc (BDCT). Dans cette transformée par bloc, l'image biométrique est découpée en plusieurs bloc (appelées sous-images) de taille inférieure. Ce qui induit un temps de calcul et un espace de mémoire très importants et très contraignant [Nixon2014]. La méthode la plus connue et populaire pour la transformée DCT par bloc est celle de deux dimensions (2D Block Based Discrete Cosine Transform-2D-BDCT) [Meraoumia2013].

Après calculer le DCT, les coefficients de grandes amplitudes et basses fréquence se localisent dans le coin supérieure-gauche. En parcourant la matrice en Zig-Zag nous aurons un vecteur des coefficients qui nécessite également une quantification et un codage.

C) ***Analyse en composantes principales (PCA)***

PCA est définie comme une transformation linéaire orthogonale transformant les données (séries) à un nouveau système de coordonnées de telle façon que la grande variance issue de quelques projections scalaires soit liée de la première coordonnée (appelée la première composante pincipale), la seconde grande variance soit sur la seconde coordonnée et ainsi de suite [Jolliffe2002].

Soit une matrice de données X de dimension $n \times p$ avec une moyenne empirique nulle par colonne (la moyenne d'échantillon de chaque colonne est décalée au zéro). Chaque ligne (n) de la matrice représente les différentes répétions (mesures et prises dans certains cas) de l'expérience, et chaque (p) fournie un type particulier de caractéristique (résultat d'une mesure, d'un prélèvement d'échantillon ou pour notre cas le vecteur biométrique par exemple).

Mathématiquement, la transformation PCA de X est définie par un ensemble de ℓ vecteurs de longueur p dont les poids ou les coefficients $w_{(k)} = (w_1,\ldots,w_p)$ qui organise chaque vecteur ligne $x_{(i)}$ de X à un nouveau vecteur des scores des composantes principales $t_{(i)} = (t_1,\ldots,t_\ell)$ où : $tk(i)=x(i).w(k)$ pour i=$1,..,n$ et $k=1,\ldots,\ell$.

De telle manière que les variables individuelles $t_1,\ldots,t_\ell$ hérite le maximum possible des variance du X, avec chaque vecteur de coefficients w contraint à être un vecteur unitaire. ℓ est généralement choisi inférieure à p afin de réduire la dimensionnalité).

- *La première composante* :

Afin de maximiser la variance, le premier vecteur de poids w_1 doit donc satisfaire :

$$w_1 = \arg max_{\|w\|=1} \left\{ \sum_i (t_1)^2_{(i)} \right\} = \arg max_{\|w\|=1} \left\{ \sum_i \left(x_{(i)} . w \right)^2 \right\} \quad (12)$$

De manière équivalente, cela peut s'écrire sous forme matricielle de l'équation suivante:

$$w_1 = \arg max_{\|w\|=1} \{ \|Xw\|^2 \} = \arg max_{\|w\|=1} \{ w^T X^T X w \} \quad (13)$$

Puisque w_1 a été défini comme un vecteur unitaire, donc :

$$w_1 = \arg max \left\{ \frac{w^T X^T X w}{w^T w} \right\} \quad (14)$$

Avec w_1 trouvé, la première composante principale d'un vecteur de données $x_{(i)}$ peut alors être donnée comme un score$(t_1)_{(i)} = x_{(i)} . w_1$ dans les coordonnées transformées, et encore le vecteur correspondant dans les variables d'origine, $\{x_{(i)} . w_1\} . w_1$.

- *Les autres composantes :*

La *k-i*^{ème} composante peut être trouvée en soustrayant tout simplement les premières $k - 1$ composantes principales de X :

$$\hat{X}_k = X - \sum_{s=1}^{k-1} X w_{(s)} w_{(s)}^T \quad (15)$$

Puis trouver le vecteur de poids qui extrait la variance maximale de cette nouvelle matrice de données:

$$w_k = \arg max_{\|w\|=1} \left\{ \left\| \hat{X}_k w \right\|^2 \right\} = \arg max_{\|w\|=1} \left\{ \frac{w^T \hat{X}_k^T \hat{X}_k w}{w^T w} \right\} \quad (16)$$

La décomposition complète en PCA de X peut donc être donnée par : $T = Xw$.

Où w est une matrice p-par-p de poids dont les colonnes sont les vecteurs propres de $X^T X$. La transposition de w est parfois appelée transformation de blanchiment ou de sphère. Les colonnes de w multipliées par la racine carrée des valeurs propres correspondantes, ou encore les vecteurs propres agrandis par les variances, sont appelées chargements en PCA ou en analyse factorielle.

- L'une des inconvénients des techniques traditionnelles est que les caractéristiques de l'image sont calculées selon un ensemble d'étapes préconçues et ne sont pas apprises dans le contexte auquel l'image appartient.

3.3.2 *Techniques basées sur l'apprentissage automatique*

L'apprentissage automatique peut être intégré dans l'étape d'extraction elle-même, dans l'étape de classification, ou il peut même être intégré dans toutes les étapes de la reconnaissance. Sur la

base de celles-ci, ces techniques peuvent être classées en plusieurs sous-classes, l'apprentissage automatique au niveau de la classification et l'apprentissage profond « *deep learning* ».

A) **L'apprentissage automatique au niveau de classification**

En effet, les systèmes de reconnaissance de formes incorporent généralement des classificateurs avec un comportement d'apprentissage, tels que Machine à Vecteurs de Support (*Support Vector Machine - SVM*), Forêt d'Arbres Décisionnels (*Random Decision Forests*) et Fonction de Base Radiale (*Radial Basis Function - RBF*). Ces types de système peuvent rencontrer des problèmes importants lorsque les vecteurs de caractéristiques d'apprentissage ont un taux de corrélation élevé, ce qui affecte de manière significative la précision du système pendant la phase de classification. Pour améliorer l'efficacité de la méthode d'extraction des caractéristiques, de nombreux chercheurs ont tenté d'incorporer l'idée d'apprentissage dans les méthodes d'extraction des caractéristiques classiques pour extraire des vecteurs de bonne variabilité. Bien entendu, la nécessité de dé-corréler les différents vecteurs de caractéristiques de différentes classes (interclasses) est plus importante pour obtenir des systèmes de reconnaissance efficaces, ce qui nécessite l'utilisation des techniques à un comportement d'apprentissage. À titre d'exemples de ces techniques, nous trouvons, DCTNet [Hossain2019] et PCANet [Chlaoua2018].

B) **Apprentissage profond**

Ces dernières années, l'apprentissage automatique a été inclus même dans l'étape d'extraction des caractéristiques d'une image, y compris les images biométriques. De cette manière toutes les étapes d'un système de reconnaissance des formes sont soumises à un apprentissage automatique, où l'entrée n'est qu'une image brute, et le système effectue toutes les tâches. Cela s'appelle l'apprentissage profond. Un système de reconnaissance des formes typiquement profond est illustré dans la Figure 3.4. Les deux étapes (couches) sont montrées, l'une pour l'analyse qui inclue plusieurs niveaux et l'autre pour la classification.

1) Couche d'extraction des caractéristiques

La principale différence entre l'apprentissage automatique et l'apprentissage profond réside dans la partie d'analyse. En fait, ces deux méthodes concordent dans la partie classification, qui est généralement un classificateur intelligent. Dans l'apprentissage automatique, les caractéristiques de l'image sont extraites indépendamment du classificateur, ce qui peut affecter son résultat si la méthode d'extraction de caractéristiques utilisée est inappropriée (des vecteurs des caractéristiques avec une forte corrélation inter-classe peuvent être obtenus). Au contraire, en l'apprentissage profond, les vecteurs des caractéristiques extraits sont automatiquement adaptés

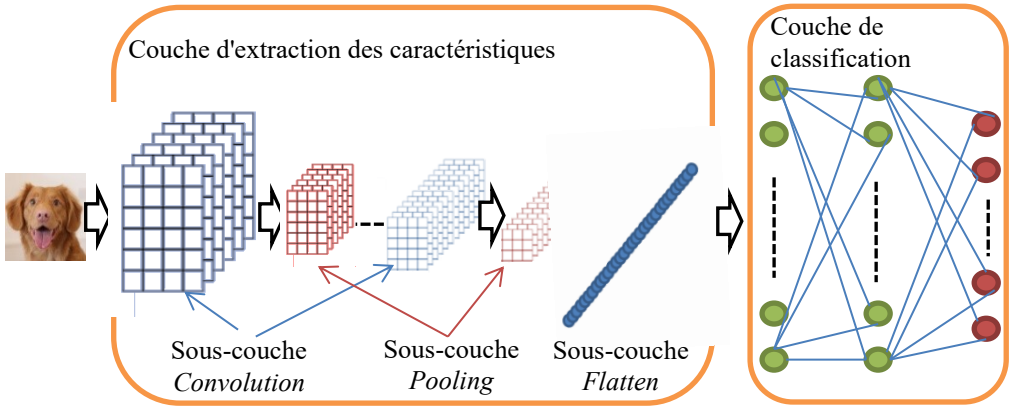


Figure 3.4- Système de reconnaissance des formes typiquement profond dépend du réseau de neurones artificiels (Réseau Neuronal Convolutif).

au classificateur. Par ailleurs, en l'apprentissage profond, l'image est analysée à plusieurs niveaux, ce qui permet de pénétrer dans la profondeur de l'image et d'extraire le maximum des caractéristiques appropriées pour une classification plus optimale.

Comme le montre la Figure 3.4, la couche de l'analyse se compose de trois sous-couche secondaires, qui sont *i)* la sous-couche *Convolution*, dans laquelle l'image est analysée à l'aide de plusieurs filtres afin de capturer uniquement les caractéristiques importantes, *ii)* La sous-couche de *Pooling*, dont l'objectif principal est de réduire le volume de données en éliminant les données moins importantes. Il est à noter que l'image peut être analysée sous plusieurs niveaux, de sorte que chaque niveau peut être composé de la sous-couche de *Convolution* suivie de la sous-couche *Pooling*. Enfin *iii)* la sous-couche *Flatten*, qui permet de normaliser le vecteur de caractéristique final pour qu'il soit approprié à l'entrée à la dernière couche entièrement connectée (*Fully Connected Neural Network*). La longueur du vecteur de caractéristique d'entité finale est principalement liée à la mise en commun du dernier niveau, qui est liée au nombre de niveaux d'analyse et au nombre de filtres convolutifs dans chaque niveau [Kim2017].

2) Couche de classification

La deuxième couche, consacrée à la classification des images, est basée sur les vecteurs de caractéristiques extraits de la première couche. Dans cette couche, nous pouvons utiliser n'importe quel classificateur dépendant de l'apprentissage automatique tel que le réseau de neurones artificiels, Méthode des k -plus proches voisins (*K - Nearest Neighbors - KNN*), SVM .Etc. Par exemple, si le système profond de reconnaissance de formes dépend du réseau de neurones artificiels, il s'appelle « Réseau Neuronal Convolutif *(convolutional neural network - CNN)*» [Zeroual2018].

3.4 Implication de la biométrie dans l'IoT de bout-en-bout

Architecture IoT de bout-en-bout (*End-to-End*) est également appelée architecture à trois domaines, domaine de la perception, domaine réseau, et domaine d'utilisateur. Le composant manquant dans cette architecture est le cloud, où les objets du domaine de la perception sont utilisés, surveillés, contrôlés et gérés directement depuis l'application utilisateur. Dans cette architecture l'humain peut communiquer directement avec ses propres objets à travers la passerelle responsable soit le *Sink node* dans WSN ou le *Gateway* dans l'IoT. Dans certaines littératures, le nœud de passerelle IoT est également appelé nœud de brouillard. Bien qu'il existe une différence entre les deux termes, le rôle du premier est plus proche de celui des routeurs de réseau avec moins de capacités de mise en réseau, tandis que le second a une plus grande capacité de calcul, de stockage et d'économiser de l'énergie.

La sécurité des schémas d'authentification à deux facteurs (cartes à puce et des mots de passe) sont vulnérables aux attaques par la devinette et soumis à des politiques de changement de mot de passe inefficaces dans IoT. Récemment, pour améliorer la sécurité de ces schémas les technologies biométriques sont combinées avec les techniques des cartes à puce et des mots de passe. Plusieurs publications traitent de la sécurité de l'architecture IoT de bout en bout à l'aide de la biométrie. Ce qui suit un résumé de quelques publications récentes.

Das [Das2015], a proposé un schéma de sécurité pour l'architecture IoT de bout en bout basé sur trois acteurs, nœud de passerelle (*Fog Node - FN*), nœud de capteur (*Sensor Node - SN*) et nœud d'utilisateur (*User Node - U*). *Das* a utilisé la carte à puce et l'algorithme Extracteur Flou (*Fuzzy extractor -FE*, référez-vous à la section 2.10.2) pour protéger le schéma. Il a suggéré qu'il devrait y avoir un canal sécurisé pour l'enregistrement des utilisateurs. En fait, tels canaux sécurisés ne sont pas toujours disponibles sur le marché en ligne ouvert. Des recherches approfondies [Park2016], [Wang2017], [Moon2017] ont également amélioré la sécurité de bout en bout en utilisant une conception d'authentification utilisateur à trois facteurs (mot de passe + carte à puce + biométrie). *Mishra et al.* [Mishra2017] a proposé un protocole d'authentification pour des communications multimédia sécurisées dans des réseaux WSNs compatibles à l'IoT, ainsi qu'une méthode de reconnaissance biométrique basée sur biohachage (faire référence à Section 2.9.2). L'enregistrement des utilisateurs est également réalisé via des canaux sécurisés. *Maurya et al.* [Maurya2017] et *Li et al.* [Li2018] ont proposé une stratégie de sécurité IoT bout en bout basée sur FE et le schéma *ECC Diffie-Hellman* (ECDH) (voir Section 2.9.1.).

Les recherches antérieures consacrées à l'IoT de bout en bout se caractérisent par les points communs suivants:

- Trois rôles sont adoptés: utilisateur, passerelle (brouillard) et capteur.
- Utilisation des cartes à puce.
- Utilisation des canaux sécurisés

Par conséquent:

- Les schémas précédents ne peuvent être appliqués que lorsque l'utilisateur communique directement avec le nœud de brouillard ou la passerelle appropriée, et cela ne semble pas être le cas pour l'IoT basé sur les services. Car le concept de service n'est pas supporté malgré qu'il soit la pierre angulaire du paradigme IoT actuel.
- l'utilisation d'une carte à puce spéciale pour chaque application IoT est une solution lourde et coûteuse, et pose beaucoup des problèmes de possession d'autant des cartes.
- Les canaux sécurisés ne sont pas toujours disponibles sur les marchés en ligne

3.5 Implication de la biométrie dans l'accès au cloud

L'accès sécurisé au cloud peut être considéré comme la première étape de la planification d'un schéma de sécurité pour l'IoT basé sur le cloud, car il en fait partie intégrante. En général, les techniques biométriques sont mises en œuvre à travers deux méthodes principales, les techniques traditionnelles et les techniques basées sur l'apprentissage automatique « *Machine Learning* ». Jusqu'à ses jours, Les techniques d'apprentissage automatique; soit l'apprentissage automatique seulement au niveau de classification ou l'apprentissage profond (faire référence à Section 3.3.2), se réalisent au niveau du cloud ou au niveau des stations de calcul à haute capacité. *Zeroual et al.* [Zeroual2019] suggèrent l'utilisation de Framework d'apprentissage profond pour mobile « *Tensorflow Lite*» pour faire la reconnaissance sur mobile sans avoir besoin du cloud ou des ressources de calcul et pour éviter que les utilisateurs envoient leur trais biométrique à chaque fois que l'authentification est demandé. Cependant, lors de la phase d'enrôlement, il faudrait utiliser le cloud pour construire le modèle d'apprentissage automatique (*Machine-Learned Model*), puis laisser les utilisateurs exécuter ces modèles sur des appareils mobiles à faible latence. En outre les données biométriques sont transmisses et enregistrées dans le cloud en clair. Dans un autre travail [Zeroual2021] *Zeroual et al*. ont essayé de résoudre ce problème. Ils ont mis au point un schéma sécurisé pour authentifier les utilisateurs dans le cloud. Ce schéma est basé sur CNN en conjonction avec un chiffrement homomorphe. Cette approche permet l'authentification sans exposer le trait biométrique aux risques d'attaques en chiffrant

simplement le vecteur de caractéristiques. Malgré l'efficacité de chiffrement homomorphe élevée et la haute précision de CNN, l'extraction de caractéristiques biométriques par des méthodes d'apprentissage automatique nécessite une capacité supplémentaire qui n'est pas suffisamment disponible dans autres entités IoT, telles que les dispositifs des utilisateurs, les dispositifs de brouillard, et les dispositifs de perception. Les techniques d'extraction biométrique traditionnelles (faire référence à Section 3.3.1) sont plus convenables aux dispositifs IoT, y compris les capteurs, les actionneurs, les nœuds de brouillard, et les terminaux utilisateurs. L'attaque la plus dommageable concerne les vecteurs biométriques après l'extraction. Par conséquent, il est nécessaire de fournir des moyens et des mécanismes de protection du vecteur des caractéristiques. Ces moyens sont discutés en détail dans le deuxième chapitre. Parmi eux, nous nous concentrons sur les travaux actuels qui traitent des techniques de Cryptosystèmes Biométriques car nos solutions proposées pour un accès sécurisé au cloud sont basées sur ces systèmes. Sans aucun doute, les algorithmes de *Fuzzy Commitment* et le *Fuzzy Vault* sont les algorithmes les plus connus dans ce domaine (faire référence à Section 2.10.1).

Dans la littérature, le schéma *Fuzzy Commitment* (FC) a été mis en œuvre dans plusieurs modalités biométriques avec différentes techniques d'extraction et différents codes correcteurs (faire référence à Annexe B). *Ao et Li* dans [Ao2009], et *Lu et al.* dans [Lu2009] ont proposé des schémas de reconnaissance de visage basés sur le FC en tant qu'un algorithme de cryptosystème et BCH en tant que code correcteur. le LBP a été utilisé comme une technique d'extraction des caractéristiques biométriques dans le schéma d'*Ao* et *Li*, alors que le PCA a été utilisé dans le schéma de *Lu et al*. Nous notons également qu'il y a une grande différence dans la longueur du code de chiffrement dans les deux schémas, où nous trouvons 707 bits dans le premier, alors que nous trouvons que 63 bits dans le deuxième. *Teoh* et *Kim* [Teoh2007], et *Nandakumar* [Nandakumar2010] ont choisi FC mais pour la reconnaissance des empreintes digitales au lieu des visages. Les techniques d'extraction et le codes correcteurs utilisés sont le filtre de Gabor et *Reed-Solomon* (RS), la transformation de Fourier et Code Convolutionnel Récursif pour le schéma de *Teoh* et *Kim,* et le schéma de *Nandakumar* respectivement. *Nandakumar* a utilisé un code de chiffrement de 2048 bits, alors que *Teoh* et *Kim* ont utilisé un autre de 375 bits. L'article de *Shukla* et *Patel* [Shukla2021] propose une FC basée sur les empreintes digitales. Ils ont incorporé BCH pour la correction d'erreurs et ils ont sécurisé le code de chiffrement par mappage de hachage SHA-256. La méthodologie proposée a une taille de clé de 1512 bits. *Shukla* et *Patel* ont incorporé un schéma de codage dépendant du nombre et

du type des minuties[4] présents à proximité du point central de l'empreinte digitale. Dans les années 2000, les minuties étaient considérées comme la caractéristique la plus discriminante et la plus fiable d'une empreinte digitale, mais malheureusement son application est uniquement dans la modalité biométrique des empreintes digitales [Lee2007]. Kausar [Kausar2021] a utilisé le FC pour lier la clé de chiffrement secrète avec le modèle d'iris annulable du patient (faire référence à Section 2.9.2). Les caractéristiques comportementales ont récemment suscité l'intérêt des chercheurs et FC en fait partie, tel que le travail de *lamiaa et al,* où un schéma de FC est appliqué avec un système biométrique basé sur une vision industrielle de démarches (*Machine Vision Gait*) [Elrefaei2019]. Les caractéristiques de la démarche sont extraites des images de la marche à l'aide de LBP, puis elles sont réduites à l'aide de PCA pour produire le vecteur des caractéristiques final. Bien que FC soit apparu il y a trois décennies, il est encore utilisé à ce jour dans de nombreuses applications biométriques pour son efficacité et sa légèreté.

Plusieurs chercheurs ont travaillé sur l'axe d'accès au cloud via *Fuzzy Vault* (FV). *Wang* et *Plataniotis* [Wang2007] et *Wu et al.* [Wu2011], ont présenté des schémas de reconnaissance de visage en utilisant l'algorithme FV en tant que système cryptographique biométrique. Les deux schémas sont très similaires, car dans tous les deux, PCA a été implémenté comme technique d'extraction des caractéristiques biométriques et la même longueur de code de chiffrement de 144 bits a été utilisée. La principale différence entre eux est que *Wang* et *Plataniotis* ont appliqué RS comme code correcteur des erreurs, alors que *Wu et al.* ont appliqué CRC. On peut dire que leurs travaux se différencient seulement par l'implication des deux codes correcteurs (RS et CRC) dans le FV. *Uludag et Jain* [Uludag2006] ont conçu un schéma de reconnaissance des empreintes digitales Similaire à celui d'*Wu et al.* [Wu2011]. Alors qu'il se distingue de lui par l'utilisation des techniques de correspondance par les minuties. *Nandakumar* et ses collègues dans [Nandakumar2007] ont appliqué le schéma d'*Uludag et Jain* en utilisant un code de chiffrement de 128 bits. Un autre travail sur les minuties en utilisant FV de *Clancy et al* [Clancy2003] a été proposé, cette fois, RS a été utilisé au lieu du CRC tout en gardant la longueur du code de chiffrement de 128 bits. Dans le travail de *Baghel et al*. [Baghel2021], une technique basée sur FV est proposée pour empêcher le vol d'identité et pour sécuriser les informations d'empreintes digitales stockées dans la base de données. Ils ont proposé une nouvelle technique pour filtrer les points authentiques de FV à partir d'une combinaison de

[4] (En anglais "Minutiaes", sont les principales caractéristiques d'une image d'empreinte digitale et sont utilisés dans la correspondance (Matching). Pour plus de détails sur les Minuties voir [Maltoni2009]).

points authentiques et de points de camouflage aléatoires utilisés. Étant donné que les points de minutie sont utilisés pour construire le FV, il est difficile d'aligner les images soit de la requête ou du modèle enregistré. Pour ce but, une technique d'alignement basée sur PCA est également proposée pour aligner les modèles et les requêtes. Dans ce travail, CRC a été utilisé comme un code correcteur. L'empreinte palmaire est une autre modalité biométrique cryptée par FV comme il est proposé par certaines recherches [Wu2008], [Kumar2009a]. RS a été incorporé dans les deux schémas. 2D-Gabor a été implémenté en tant que technique d'extraction dans le schéma de *Wu et al*. Alors que DCT a été choisi dans le schéma de *Kumar*. En ce qui concerne la longueur des codes de chiffrement, elles sont 1024, 306 bits respectivement. D'autres travaux tentent de combiner FC et FV. Par exemple *Chang et al.* [Chang2021] ont proposé un Framework multi-biométrique, qui combine FC et FV en utilisant le schéma de chiffrement préservant le format.

Il est possible de noter quelques lacunes dans les cryptosystèmes biométriques précédents, notamment:

- Les travaux antérieurs traitaient le problème de la sauvegarde des données biométriques dans le cloud avec les algorithmes FC et FV. Ils n'abordaient pas le problème de la sécurité de transmission, ce qu'il est nécessaire aujourd'hui pour l'accès à distance aux services cloud. Les deux algorithmes doivent être bien exploités pour sécuriser la transmission des données biométriques.
- Puisqu'il existe de nombreuses techniques d'extraction biométriques et de nombreux types des codes correcteurs, et il existe aussi un grand espace de clés, le schéma doit être plus efficace et plus léger en optimisant le choix.
- Les travaux précédents peuvent être élargies et complétés avec les travaux de l'implication biométrique à l'architecture IoT de bout en bout pour préparer et créer un schéma complet qui inclut tous les acteurs IoT basé sur les services.

3.6 Implication de la biométrie dans l'IoT à quatre domaines

L'architecture IoT à quatre domaines est composée de, domaine de la perception (*Sensing Domain*), domaine du brouillard (*Fog Domain*), domaine de cloud (*Cloud Domain*), et domaine d'utilisateur (*User Domain*). Dans le modèle S^2aaS, le domaine-cloud comprend tous les publieurs de données de capteurs et les fournisseurs de services étendus (voir Figure 1.5). Pour S^2aaS en tant qu'un nouveau modèle, peu d'études ont traité la sécurité de ce modèle sous tous ses aspects. Nous avons constaté que la plupart de ces études gèrent le modèle sans domaine-

utilisateur. Cela signifie qu'aucune technologie biométrique n'a été utilisée et donc aucun de ses avantages n'a été bénéficié. Et autres utilisent le cloud parce qu'il s'agit d'un nœud de confiance, Ce n'est pas le cas dans un environnement S^2aaS, et d'autres ont complètement abandonné le cloud (l'architecture IoT de bout en bout). Dans la section ci-dessous, nous essaierons de présenter certains travaux de sécurité de l'IoT basés sur le cloud.

Harbi et al. [Harbi2019] ont utilisé le chiffrement ECC pour chiffrer/déchiffrer les processus de communication et pour l'échange de clé. La stratégie de *Harbi et al.* est basée sur le rôle de la station de base, de chef de cluster, et de membre du cluster, correspondant respectivement au rôle de cloud, de nœud de brouillard, et nœud de capteur. Malgré le rôle critique de l'utilisateur dans l'architecture IoT, ce dernier n'est pas abordé par leur schéma proposé. *Roy et al*. [Roy2018] ont développé un schéma d'authentification d'utilisateur efficace pour l'IoT basé sur le cloud. Ce travail s'appuie sur le concept des services publics dans l'IoT. Cependant, le plan proposé est uniquement basé sur deux rôles (Utilisateur et Service). Le travail fait partie des très rares travaux qui ont traité le cloud comme un nœud non fiable, ce point est similaire à une partie de notre contribution aux systèmes de sécurité de l'IoT basé sur les services publiques à quatre domaines. *Gupta et al*. [Gupta2019] ont intégré des services cloud dans leur conception de sécurité. Cependant, ce cloud est supposé être un nœud fiable, ceci est incompatible avec la notion des services publiques.

- L'une des lacunes majeures des travaux susmentionnés qu'il n'existe pas un schéma de sécurité complet pour l'implication des technologies biométriques dans l'IoT avec toutes ses entités, telles que les utilisateurs, les services de cloud, les nœuds de brouillard, et les nœuds de capteurs.
- De même, la plupart de ces travaux contiennent au moins un canal sûr, en particulier lors de l'enregistrement des utilisateurs, et cela n'est pas réaliste lorsque le service est public qui doit être accessible via un site web et non via un enregistrement sur un serveur local ou via une carte à puce.

3.7 Solution biométrique proposée pour l'IoT de bout-en-bout

3.7.1 *Contributions*

Les principales contributions de notre travail [Bentahar2020] peuvent être résumées comme suites:

- Utiliser la biométrie dans l'IoT de bout-en-bout pour l'authentification et pour le partage sécuritaire de la clé.

- Appliquer de FE basée sur le principe FV combiné avec un protocole d'établissement de clé efficace pour économiser du temps et de l'énergie.
- Incorporer *Reed-Solomon* et *Extended Hamming* pour fournir une certaine tolérance aux erreurs.
- Effectuer une analyse informelle pour clarifier comment notre système résiste à des attaques de transmission sans affecter la légèreté et la précision du système.
- Effectuer des analyses d'énergie et de coût de communication.
- Comparer le travail avec des travaux connexes pour montrer qu'il est plus sûr et plus précis.

3.7.2 *Système proposé d'authentification et d'échange de clé*

Notre crypto-système gère l'authentification et l'échange de clé entre les nœuds d'utilisateur et les nœuds de brouillard.

Dans la phase d'enrôlement (Enregistrement biométrique hors ligne), le vecteur de caractéristiques W est extrait de l'utilisateur biométrique en tant qu'ensemble de points X_m, ces points présentent les modèles stockés dans les nœuds de brouillard.

Dans la phase d'authentification (Contrôle d'accès), le vecteur de caractéristiques W' est extrait de la requête biométrique de l'utilisateur sous la forme d'un autre ensemble de points X'. De plus, des chaînes aléatoires r et k sont générées par le nœud utilisateur. r est concaténée avec l'horodatage t pour obtenir le paramètre publique PB tel qu'il est exprimée dans l'équation (17), où || est l'opération de concaténation. La clé secrète CS est générée en concaténant W' et PB. Puis la fonction de hachage est appliquée sur le résultat, comme il est illustré dans l'équation (18). La fonction de hachage $H(.)$ est appliquée aussi sur k. En parallèle, k est codé à l'aide d'un code correcteur (RS ou EH) pour obtenir la clé codée C (*Code Word*). C est présenté comme des coefficients du polynôme P. Ensuite, W' est projeté sur P pour former les paires des points $(X_q, P(X_q))$. Des pairs des points aléatoires (X_a, Y_a) sont ajoutés à l'ensemble des paires authentiques (X_q, Y_q), où $Y_a \neq P(X_a)$ pour le camouflage. L'ensemble des tous les paires des points $\{(X_q, P(X_q)), (X_a, Y_a)\}$ constitue la forme sécurisée SS. SS, $H(k)$, et PB sont envoyés au nœud de brouillard. Il est à noter que r et k doivent être supprimés du terminal utilisateur après l'envoi de la requête.

$$PB = r||t. \tag{17}$$

$$CS = f(PB, W') = H(PB||W') \tag{18}$$

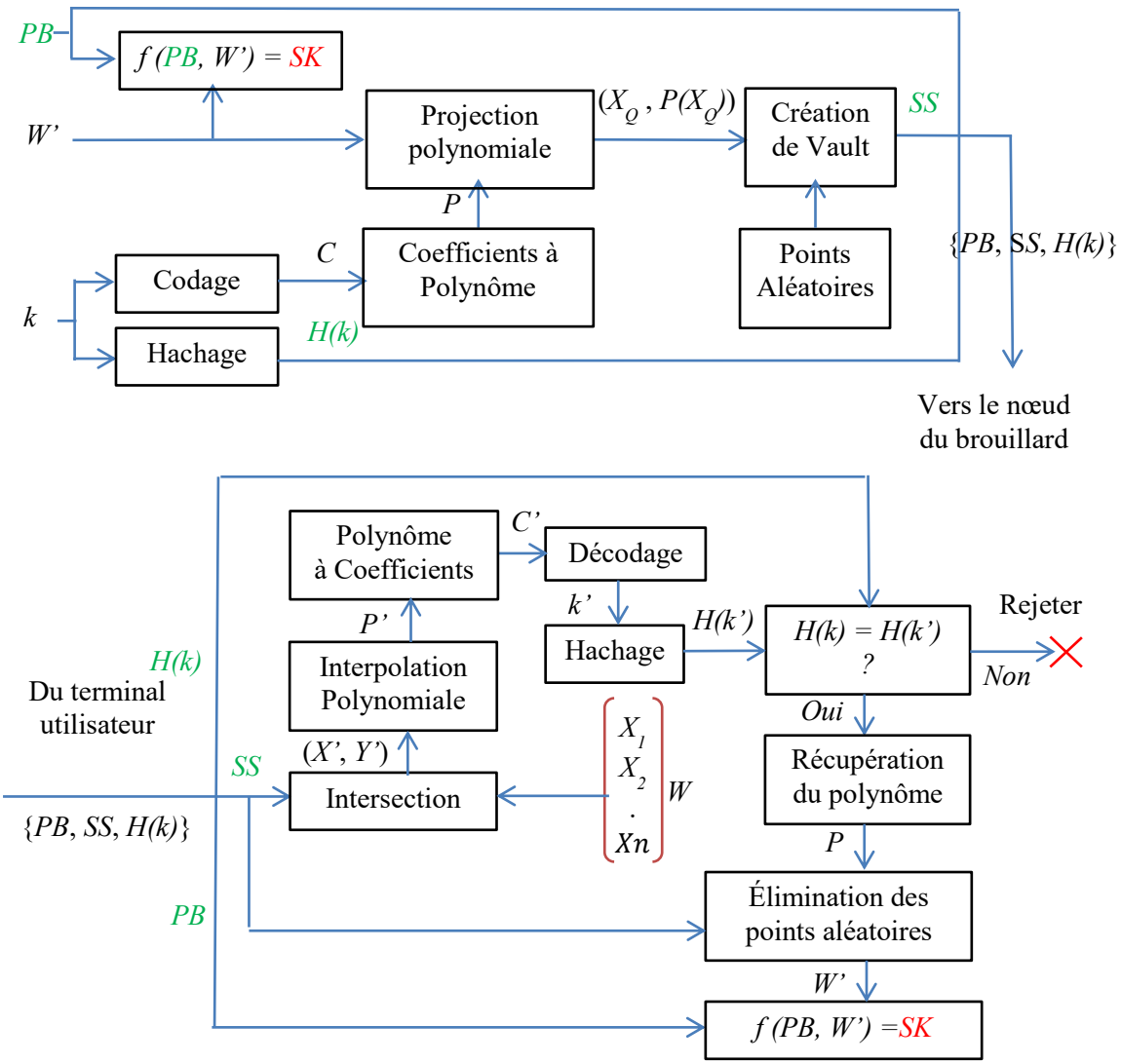


Figure 3.5- Organigramme de flux du crypto-système proposé pour l'IoT bout-en-bout (Source [Bentahar2020]).

Lorsque *SS*, *H(k)* et *PB* arrivent au nœud de brouillard, une intersection est faite entre les points *X* reçus de *SS* et les points de modèle stockés X_m. Comme il est précédemment expliqué, *X* contient des points authentiques et autres faux et aléatoires. En utilisant une technique d'interpolation polynomiale, un polynôme *P'* peut être reconstruit et converti en coefficients *C'*. *C'* peut être décodé pour récupérer *k'* en utilisant le même code correcteur utilisé dans la formation *SS*. Après le hachage *k'*, le *H(k')* résultant est comparé au *H(k)* reçu. S'ils sont identiques, cela signifie qu'ils appartiennent à la même personne et que l'authentification est réussie. Dans ce cas, *C* initial (décodé à partir de *C'*) est utilisé pour reconstruire *P* initial. Nous utilisons *P* pour faire la distinction entre les points aléatoires X_a et les points authentiques X_q du *SS* qui expriment *W'* original. Avec *PB* et *W'*, la clé secrète *CS* peut être reproduite à l'aide de la même fonction *f* de l'équation (18). La clé *CS* reproduite sera utilisée pour crypter tous les

échanges de données et les messages entre l'utilisateur et le nœud de brouillard correspondant. Figure 3.5 montre l'organigramme de l'authentification sécurisée et de l'échange de clé proposé.

3.8 Solution biométrique proposée pour l'accès au Cloud

Dans cette section nous présentons notre apport dans l'axe d'accès au cloud en précisant les différents systèmes proposés.

3.8.1 *Contributions*

Les crypto-systèmes biométriques étaient à l'origine utilisés pour le contrôle d'accès local. Ce qui signifie que les modèles biométriques sont chiffrés sur le serveur, mais les requêtes ne le sont pas (en clair). De nos jours, les serveurs locaux n'existent plus et ont été remplacés par le cloud. Le concept d'origine des crypto-systèmes biométriques présente un risque important pour l'authentification au Cloud-IoT lorsque les requêtes sont envoyées en clair sur des canaux sans fil non sécurisés. Nous avons donc inversé ce concept, de sorte que dans un tel nouveau concept inversé, les solutions proposées fournissent une authentification à distance sécurisée et en même temps fournissent un partage des clés secrètes.

Nos contributions pour sécuriser l'accès au cloud peuvent être résumées comme suites:

- Invertir le concept des crypto-systèmes biométriques pour sécuriser les transmissions et pour partager les clés plus tôt que sécuriser seulement la sauvegarde dans la base de données.
- Proposer des crypto-systèmes biométriques adaptés aux contraints IoT basés sur le DCT et le DWT en utilisant le FC [Bentahar2018], [Bentahar2018a].
- Proposer un crypto-système biométrique basé sur « *Eigen-Fingerprints*- PCA » en utilisant le FC et le FV [Bentahar2019a], [Bentahar2021a].
- Améliorer le processus d'authentification pour qu'elle soit plus légère et plus précise [Bentahar2019] en :
    - ✓ Proposant une extraction PCA appropriée par quantification moyenne des blocs non linéaires.
    - ✓ garantissant la fonction d'évolutivité (*Scalability*)
    - ✓ incorporant de nouveaux taux de reconnaissance.
    - ✓ Utilisant des nombreuses modalités biométriques et des nombreuse codes correcteurs.

3.8.2 *Les schémas proposés*

A) Schéma basé sur DCT

Dans [Bentahar2018], FC est utilisé pour sécuriser l'authentification biométrique basée sur le DCT et pour partager la clé secrète sans affecter la légèreté et la précision du système.

- **Extraction de vecteur de caractéristiques**

Après la détection du point singulier d'une image d'empreinte digitale (la région d'intérêts (*Region of interest* - ROI)) [Sen2002], l'image est recadrée à 64x64 pixels, puis écartelée à 4 sous images de 32 x32 pixels. 2D-DCT est calculé pour chaque sous-image, et les coefficients sont arrangés comme dans le DWT (faire référence à Section 3.3.1) mais juste pour les 9 blocs de fréquences hautes à moyennes (hors basses fréquences). Enfin nous calculons les paramètres caractéristiques. Cet arrangement nous donne un vecteur de 36 caractéristiques (éléments). Ensuite le code *Gray*[5] est utilisé pour le codage de chaque élément. 8 bits suffisent pour coder toute valeur trouvée en fonction de la valeur maximale des éléments du vecteur (coefficients calculés). Ainsi, le vecteur de caractéristique extrait W est de 36 x 8 bits.

- **Incorporation du FC**

Lors de l'implémentation de FC, une clé de 96 bits a été utilisée. Pour protéger cette clé des erreurs de transmission, deux codes correcteurs ont été utilisés. Le premier est un simple code de répétition: Code(3,1), où la clé codée $C \in \{0, 1\}^3$. Ce code est un correcteur par vote majoritaire à une efficacité de 1/3. Le second est le code EH (8, 4); pour savoir plus, reportez-vous à Annexe B. A noter que le EH ne fonctionne qu'en binaire (*galois field*- GF(1bit)), il est purement bit-à-bit. Et c'est le cas dans le schéma FC. Ainsi, une chaîne binaire de 48 bits ($k \in \{0, 1\}^1$) a été générée aléatoirement. Lorsque nous appliquons le code EH, nous obtenons une clé de 96 bits. Après l'application du code de répétition, la clé codée finale C obtenue a une longueur de 288 bits. Il est évident que la clé codée a la même taille que le vecteur W.

Comme le montre la Figure 3.6, Dans le terminal de l'utilisateur (U_i), la requête W' et la clé codée C sont *XORées* pour former un *Commitment SS* ($SS = W' \oplus C$), puis la fonction de hachage

[5] Gray Code : C'est un code de représentation binaire à une distance de *Hamming* d'un seul bit entre deux états adjacents (deux niveaux adjacents de quantification). Par exemple les deux valeurs 63 et 64 sont très proches dans la représentation décimale mais distantes de 7 bits dans la représentation binaire usuelle (00111111 et 01000000). Dans la représentation en *Gray Code*, Ces deux valeurs ne sont distantes que d'un bit (0010000 et 01100000), elles sont très proches.

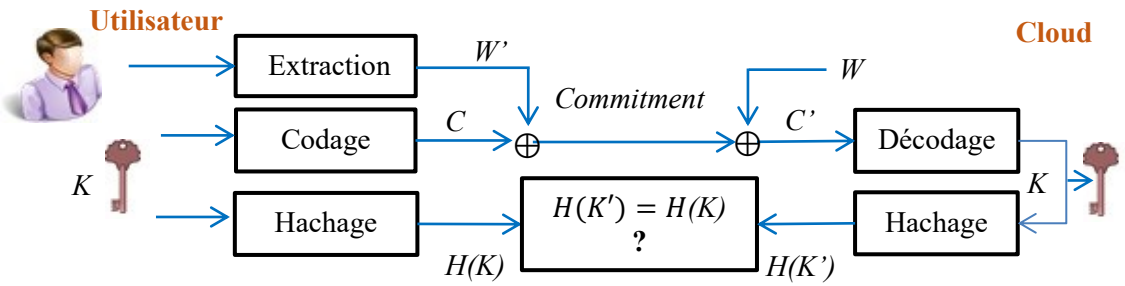


Figure 3.6- FC proposé pour l'authentification en cloud et l'échange de clé (Source [Bentahar2018]).

est appliquée à la clé d'origine *k* pour obtenir *H(k)*. Finalement, le système n'a envoyé que la paire {*SS*, *H(K)*}. Une note importante, *W'* doit être supprimé du U_i.

Lors de la réception dans le cloud, le système applique d'abord une fonction XOR entre le *SS* reçu et le *W* stocké. Ensuite, le résultat trouvé (*C'*) est décodé par le processus de répétition inverse pour produire une clé de taille de 96 bits. Pour restaurer le *k* d'origine, le vecteur résultant est décodé par le décodeur d'EH. le haché *H(k')* est comparé à *H(k)* reçu, Si *H(K) = H(K'),* la clé est alors récupérée. Et donc les *W* et *W'* sont alors de la même personne. Ce qui signifie que le la personne est authentifiée avec succès. De plus, la clé de chiffrement *k* a été parfaitement partagée en toute sécurité.

B) **Schéma basé sur DWT**

Dans [Bentahar2018a], l'image de l'empreinte digitale est recadrée avec 64 x 64 (selon ROI). DWT de 4 niveaux est appliqué sur l'image recadrée pour obtenir 12 blocs. Pour chaque bloc, l'énergie moyenne est calculée (en utilisant l'équation (8) du chapitre 1). Après la quantification, les coefficients sont codés en code *Gray* de 8 bits. Donc le vecteur obtenu est de taille de 12x8 = 96 bits. Cette longueur est adaptée aux réseaux IoT, car elle n'est pas trop longue pour déborder la faible capacité des objets de l'IoT et n'est pas trop petite pour menacer la sécurité (la clé courte est facile à trouver). Brièvement, Le travail [Bentahar2018a] est similaire au travail de [Bentahar2018], sauf que le premier est basé sur le DWT et le second sur le DCT.

C) **Schéma basé sur Eigen-Fingerprints**

Le terme « *Eign* » signifié à l'ensemble des valeurs et des vecteurs propres extraits par PCA. Par exemple "*Eigenface*" une solution basé sur PCA pour la reconnaissance faciale. "*Eigen-Fingerprints*" pour la reconnaissance des empreintes digitales. Dans [Bentahar2019a] et [Bentahar2021a], nous proposons un *Eigen-Fingerprints* où les vecteurs de caractéristiques biométriques sont extraits par PCA avec des techniques d'optimisation rendant le schéma plus

adapté aux réseaux sans fil. Les deux principes de liaison de clé: le FC et le FV sont incorporés dans ces travaux. Les codes correcteurs orientés octet et orientés bit sont utilisés pour le chiffrement des données et pour l'échange de clés.

- **Extraction de vecteur de caractéristiques**

Les données sont organisées sous forme de matrice. Les colonnes représentent les variables à analyser, tandis que les lignes représentent les observations. Dans le cas d'images de taille N×N, il faut remodeler chaque image en un vecteur colonne de taille égale à N^2 observations. Le nombre de ses vecteurs correspondant au nombre d'images. Comme dans l'apprentissage automatique, PCA utilise cette nouvelle matrice comme un ensemble d'apprentissage (EA). L'application de PCA sur la matrice de covariance de l'EA donne les valeurs propres et les vecteurs propres qui sont directement liés pour extraire des caractéristiques spécifiques de l'EA.

Si EA contient m images biométriques: Γ_i (i = 1 à m), les *Eigen-Fingerprints* peuvent être résumées dans les cinq étapes suivantes:

Étape 1: Remodeler Γ_i de dimension N×N en un vecteur colonne de longueur N^2.

Étape 2: Calculer l'écart $\Phi i = \Gamma_i - \Psi$, où Ψ est la moyenne de EA.

Étape 3: Concaténer $[\Phi_1, \Phi_2, .., \Phi_m]$ pour former une matrice A de dimension N^2x m.

Étape 4: Calculer la matrice de covariance $C = AA^T$ de dimension N^2xN^2. Dans le but de minimiser les calculs, $L = A^TA$ peut remplacer C [Yongxu2006].

Étape 5: Trouver les valeurs propres et les vecteurs propres de la matrice L, et construire la matrice de projection MP (faire référence à Section 3.3.1).

À ce stade, nous avons obtenu une matrice de projection MP qui sera utilisée pour extraire le vecteur de caractéristiques W d'une autre image Γ_q non incluse dans EA. Il suffit de projeter $\Phi_q = \Gamma_q - \Psi$ sur MP, tel que, $W = MP^T\Phi_q$. Il est à noter que la longueur de W est m.

On note W et W' le vecteur de caractéristiques de l'image modèle (*Template*) Γ_t et le vecteur de caractéristiques de l'image de requête (*Query*) Γ_q, respectivement. Si W et W' sont très proches sous une certaine échelle de distance et un seuil de sécurité prédéterminé, nous pouvons dire que les deux images appartiennent au même personne. La métrique de distance utilisée pour la correspondance des caractéristiques doit être choisie en fonction de la nature du vecteur de caractéristiques. Ainsi, la distance de *Hamming*, la distance de différence, et la distance euclidienne sont utilisées lorsque les coefficients du vecteur caractéristique sont des valeurs binaires, des valeurs quantifiées (c'est-à-dire un ensemble fini non binaire), ou des valeurs décimales (ensemble non fini), respectivement.

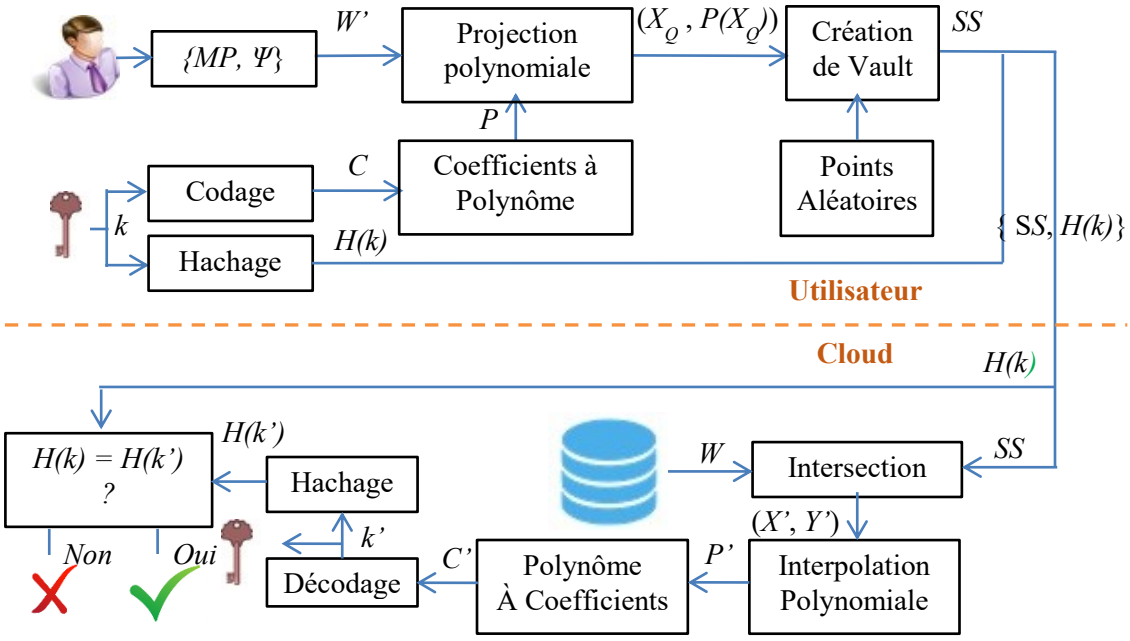


Figure 3.7- FV proposé pour l'authentification en cloud et l'échange de clé (Source [Bentahar2019]).

- **Incorporation du FC et du FV**

Incorporation du FC est basée sur notre schéma FC dans [Bentahar2018], mais un nouvel aspect de l'apprentissage est introduit. Les résultats de cet apprentissage représentés par la matrice *MP* et le vecteur *Ψ* doivent être enregistrés dans le terminal utilisateur.

Concernant l'incorporation du FV, la forme du lien entre le témoin *W* et la clé codée *C* est différente de celle du FC. Dans FC, la forme de la liaison est le XOR. Alors que dans FV, ils sont liés par projection polynomiale. Ensuite, la clé est récupérée par interpolation polynomiale. Le système d'authentification et d'échange de clé sécurisé basé sur le FV et le PCA illustré à la Figure 3.7. Nous pouvons expliquer le système comme suit:

Lors de l'implémentation du FC, deux codes correcteurs ont été utilisé: EH (8.4) de GF (1 bit), RS (10,4) de GF (8 bits). La longueur de la clé codée *C* est de 80 bits, elle correspond à la longueur du vecteur de caractéristiques. Lors de l'implémentation du FV, la clé codée par RS (10,4) se présente sous forme de coefficients à 8 bits d'un polynôme *P* de 7 degrés. La clé binaire correspondante est de 8×8 = 64 bits. L'algorithme de décodage *Berlekamp-Massey* a été utilisé. Le vecteur PCA est de 10 coefficients mais que le polynôme est de 8 termes (degré 7). Ce choix de degré est basé sur plusieurs tests qui montrent que certains vecteurs PCA peuvent avoir des répétitions dans les coefficients notamment après la quantification, et le nombre moyen de répétitions est de deux. Pour pallier l'échec de l'interpolation polynomiale sur un grand nombre

de requêtes, les termes de *P* sont fixés à 10-2 = 8. Avec cette configuration, le FV peut trouver la clé même si l'intersection entre *W'* et *SS* avait 2 points aléatoires.

D) **Améliorations suggérées pour une authentification plus légère et plus précise**

Le travail [Bentahar2019] s'agit d'une extension de nos précédents travaux sur la sécurisation des accès biométriques au cloud. Le travail consiste à renforcer la sécurité et améliorer les capacités d'authentification et d'échange de clés. Il consiste également en un ensemble de tests exhaustifs avec diverses modalités biométriques et divers codes correcteurs. Les modalités biométriques utilisées sont, les visages, les empreintes digitales, et les empreintes palmaires. Les codes correcteurs utilisés sont EH et RS. Alors que le PCA est utilisé comme une technique d'extraction. Il est important de noter que notre système est un système d'identification et pas seulement un système d'authentification.

Pour augmenter la rapidité de nos systèmes sans perdre d'informations, nous réduisons la longueur du vecteur de caractéristiques à 20 coefficients en utilisant la moyenne de bloc non linéaire. Ce vecteur réduit est obtenu en faisant la moyenne de chaque ensemble d'éléments. Où le nombre d'éléments n'est pas le même dans chaque ensemble, mais est déterminé par une suite arithmétique de raison quatre[6]. Le premier coefficient du vecteur réduit à la même valeur que le premier coefficient du vecteur d'origine. Le deuxième coefficient est la moyenne des quatre coefficients du vecteur d'origine. Le troisième est la moyenne des huit coefficients, et ainsi de suite jusqu'au dernier ($20^{ème}$) qui correspond à la moyenne des coefficients restants (de 686 à 1000, car on a 1000 images dans EA). De cette manière, les vecteurs réduits (*W* et *W'*) préservent les informations essentielles et fondamentales principalement présentes dans les premiers coefficients. De plus, l'assemblage de tous les derniers coefficients dans le $20^{ème}$ ensemble offrait la propriété d'évolutivité (*Scalability*). C'est-à-dire que de nouveaux utilisateurs peuvent rejoindre le système sans modifier la longueur du vecteur et donc la longueur de la clé. Ainsi, il maintient la stabilité de notre système.

Pour réduire la variance intra-classe, les coefficients de vecteur de caractéristiques réduits sont quantifiés. La quantification est la conversion de nombres réels en un ensemble fini d'entiers (niveaux de quantification). Grâce à de nombreux tests expérimentaux, il a été prouvé que 71 niveaux de quantification suffisent pour donner les meilleurs résultats. Deux représentations quantitatives ont été utilisées, la première étant la représentation entière (*Direct Recognition-*

[6] Une suite arithmétique de raison quatre: est une progression arithmétique avec une différence commune de quatre.

Non-Binary - DR-NB), et la seconde étant la représentation binaire (*Direct Recognition-Binary* - DR-B). Le code *Gray* a été utilisé comme codage binaire. Après la réduction, la quantification, et le codage nous obtenons un vecteur de caractéristiques binaire d'une longueur de 160 bits.

Notre travail [Bentahar2019] présente quater crypto-systèmes qui sont : FC avec code *Extended Hamming* (FC-H), FC avec code *Reed-Solomon* (FC-RS), FV avec code *Extended Hamming* (FV-H) et FV avec code *Reed-Solomon* (FV-RS). Plus techniquement, les codes correcteurs utilisés sont EH (8.4) de GF (1 bit), RS (20,8) de GF (8 bits). La longueur de la clé codée *C* pour les quatre crypto-systèmes est de 160 bits, elle correspond à la longueur du vecteur de caractéristiques. Dans FC-H, les erreurs sont découvertes et corrigées bit à bit. Dans FC-RS, le décodage nécessite de passer de GF (1 bit) à GF (8 bits), puis l'algorithme de *Berlkam-Massey* est appliqué aux 20 octets résultants. Dans FV-H et FV-RS, la clé C est convertie en une chaîne de 10 éléments (au lieu de 20) en concaténant tous les deux éléments en un seul élément (2 octets pour un élément). L'objectif de cette concaténation est de minimiser le nombre de coefficients de *P*. Grâce à une série de tests sur plusieurs modalités biométriques, nous avons trouvé que: D'une part, les vecteurs de caractéristiques peuvent avoir des redondances dans leurs éléments. D'autre part, la distance sans redondance (*Set Difference*) entre *W* et *W'* de la même personne est d'environ la moitié de la longueur du vecteur (10 éléments dans notre cas). En revanche, pour reconstruire correctement *P* par l'interpolation des moindres carrés, le nombre de paires de points (éléments) doit être supérieur ou égal au nombre de coefficients de *P*.

3.9 Solution biométrique proposé pour S^{2}aaS

3.9.1 *Contributions*

l'IoT actuels se caractérise par la fourniture des différents services publics aux différents consommateurs. Cela doit être inclus dans le système de sécurité de l'IoT. Nous avons abordé dans [Bentahar2021] ce point avec l'implication des technologies biométriques pour exploiter leurs avantages. Notre contribution dans ce contexte peut se résumer dans les points suivants:

- Suggérer un schéma d'échange de clé simple, efficace et sécurisé pour le modèle S2aaS basé sur l'IoT.
- Inclure tous les acteurs IoT possibles.
- Assurer les authentifications mutuelles, l'établissement des clés, la confidentialité des transmissions, l'intégrité des données, et la récence de système.

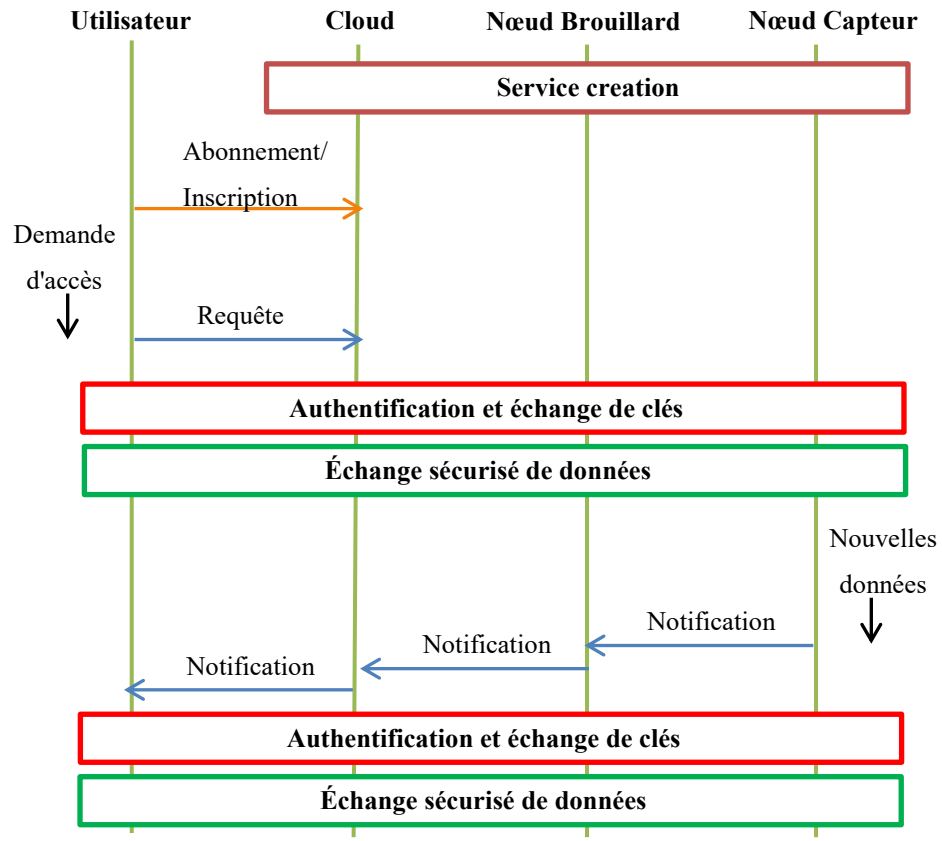


Figure 3.8- Protocole proposé basé sur le paradigme Pub/Sub.

- Attribuer des tâches de la carte à puce aux terminaux des utilisateurs (généralement des smartphones). Par conséquent, ils peuvent accéder rapidement et en toute sécurité aux services publics via un simple site Web.
- Traiter le cloud comme un nœud non fiable, et tous les canaux de transmission ne sont pas sécurisés, car S2aaS est un véritable environnement.
- Intégrer le partage de clé secrète dans le sous-système biométrique ainsi que l'authentification pour gagner du temps et de l'énergie.
- Eviter les processus lourds dans les intervalles critiques et fréquents.
- Discuter la résilience contre diverses attaques au niveau de la transmission et de la sauvegarde.
- Prouver formellement le schéma par l'outil AVISPA qui montre que tous les objectifs de sécurité requis ont été atteints.
- Comparer le schéma avec d'autres travaux récents connexes en termes de sécurité, de coûts de communication, de coûts de calcul et de certaines mesures connexes, pour montrer que notre schéma est léger et offre le plus de protection que les autres.

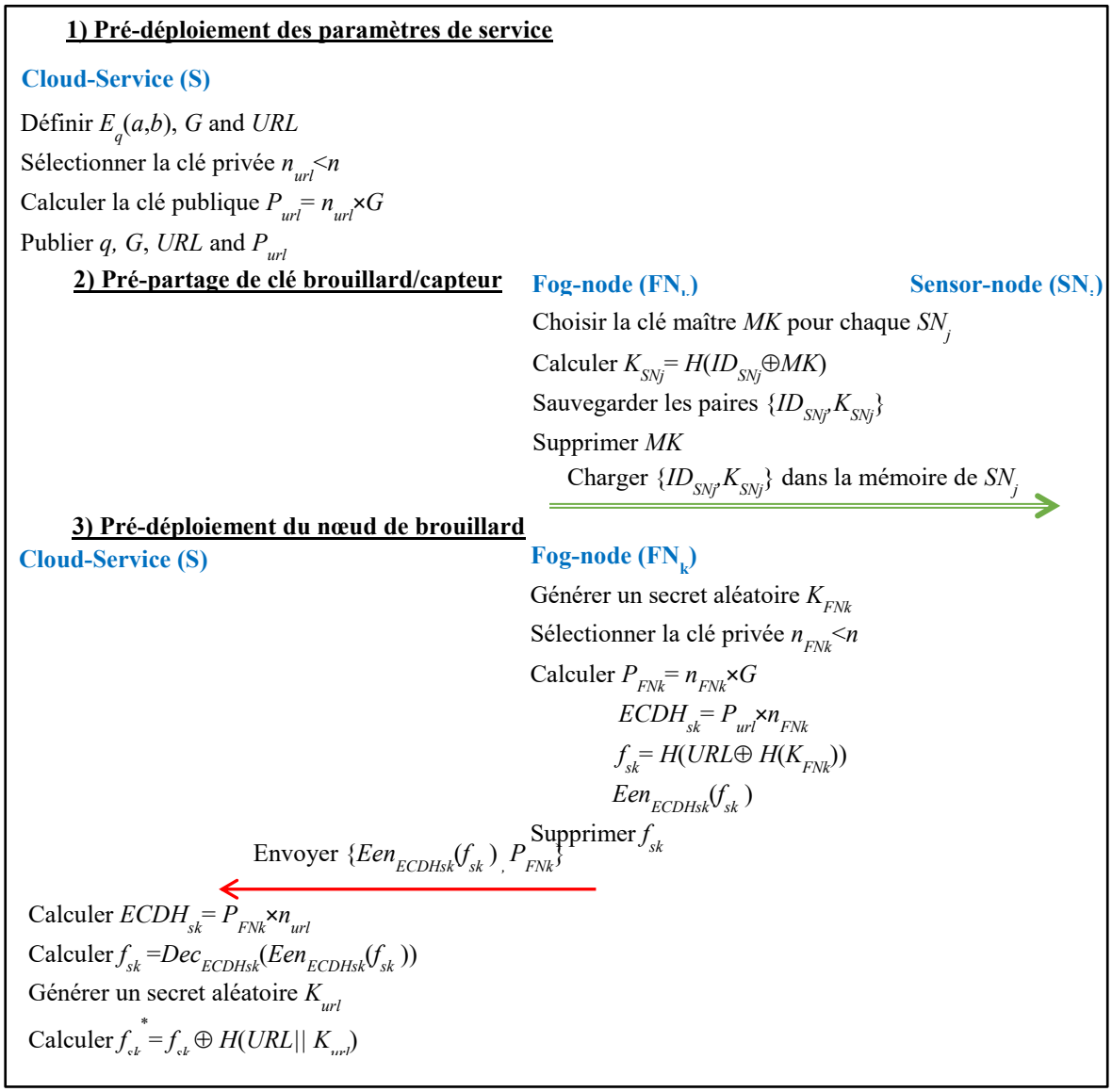


Figure 3.9- Phase de pré-déploiement du service.

3.9.2 *Schéma proposé d'authentifications mutuelles et d'échange de clé*

Cette section décrit le schéma suggéré basé sur le paradigme Pub/Sub (voir Section 1.4) comme illustré à la Figure 3.8. FE et ECDH ont été utilisés car ils sont bien adaptés à la technologie biométrique pour fournir une inscription et une connexion sécurisées. La combinaison de ces deux techniques permet de protéger la clé secrète de l'utilisateur en cas de vol ou de perte. Aussi l'utilisation d'ECDH permet la génération de clés fortes de courte longueur FE et ECDH ont été utilisés conjointement avec diverses techniques de sécurité bien connues telles que les chiffrements symétriques et les fonctions de hachage pour assurer la confidentialité du système et l'intégrité des données.

Quatre acteurs principaux sont impliqués : les utilisateurs, les services cloud, les nœuds de brouillard et les nœuds de perception (capteurs). Tout d'abord, Les utilisateurs s'abonnent/s'inscrivent au service public souhaité. Si un utilisateur (U_i) souhaite bénéficier de ce service, il se connecte (Login) et envoie une requête au Cloud-Service (S). S transférer la requête aux nœuds de brouillard (FN_k) et de là aux nœuds de capteurs (SN_j). D'autre part, si SN_j obtient des données intéressantes, il informe FN_k, S, U_i dans le bon ordre. A ce stade, U_i sera invité à se connecter et à consommer ces données.

Dans ce qui suit, le système crypto-biométrique proposé est détaillé, a noter que le message Msg chiffré par la clé symétrique KEY est noté: $Een_{KEY}(Msg)$, et le message Msg déchiffré par la même clé est noté: $Dec_{KEY}(Msg)$.

A) **Phase de pré-déploiement du service**

Cette phase se compose de trois parties principales : le pré-déploiement des paramètres de service, la pré-partage de clé brouillard/capteur, et le pré-déploiement du nœud de brouillard. Figure. 3.9 résume les parties de la phase de pré-déploiement.

- **Pré-déploiement des paramètres de service**

Lors de la création du service, les administrateurs (SP ou ESP dans le modèle S^2aaS) créent un service avec un identifiant unique (*Uniform Resource Locator* - URL). Et ils choisissent les éléments publics du ECC qui sont, $Eq(a,b)$ et G. la clé privée de service n_{url} est sélectionnée, puis la clé publique P_{url} est calculée comme suit: $n_{url} \times G$. Ensuite les paramètres: q, G, URL et P_{url} sont publiés sur le web.

- **Pré-partage de clé brouillard/capteur**

FN_k choisit aléatoirement une clé maître MK et un identifiant unique ID_{FNk}, et choisit un ID_{SNj} unique pour chaque SN_j déployé. Puis, il calcule le secret partagé : $K_{SNj} = H(ID_{SNj} \oplus MK)$. Enfin le couple $\{ID_{SNj}, K_{SNj}\}$ est chargé dans la mémoire de SN_j et de FN_k. Noter qu'il n'y a pas de menace pour la sécurité si ID_{SNj} est publiquement connu comme on le verra plus tard.

- **Pré-déploiement du nœud de brouillard**

FN_k génère aléatoirement un secret K_{FNk} et sélectionne une clé privée n_{FNk}. Puis il calcule la clé publique P_{FNk} comme suit: $P_{FNk} = n_{FNk} \times G$. La clé secrète partagée $ECDH_{sk}$ entre FN_k et S est calculée telle que: $ECDH_{sk} = P_{url} \times n_{FNk}$. La fonction f_{sk} est calculée comme suit: $f_{sk} = H(URL \oplus H(K_{FNk}))$, et chiffrée par la clé $ECDH_{sk}$. Le résultant $Een_{ECDHsk}(f_{sk})$ est envoyé à S. Il est important de supprimer f_{sk} du nœud.

Lorsque $Een_{ECDHsk}(f_{sk})$ atteint *S,* ce dernier calcule $ECDH_{sk} = P_{FNk} \times n_{url}$ et l'utilise pour déchiffrer le reçu. Ensuite, *S* génère aléatoirement un secret K_{url} et calcule $f_{sk}^* = f_{sk} \oplus H(URL || K_{url})$.

B) **Phase d'inscription de l'utilisateur**

Tout d'abord, U_i choisit et s'abonne au service souhaité via l'internet (*URL*). U_i obtient les paramètres publics du service: (q, G, P_{url}), sélectionne sa clé privée appropriée n_i, et calcule sa clé publique P_i correspondante telle que $P_i = n_i \times G$. P_i est envoyé à *S* de même que ID_i. *S* Calcule la clé qui sera partagée avec U_i comme suit : $ECDH_{si} = P_i \times n_{url}$. Ensuite, *S* calcule et envoie $Een_{ECDHsi}(f_{si})$, où $f_{si} = H(ID_i \oplus H(K_{url}))$ à U_i.

Lorsque $Een_{ECDHsi}(f_{si})$ arrive à U_i , ce dernier calcule $ECDH_{si} = P_{url} \times n_l$, et décrypte $Een_{ECDHsi}(f_{si})$ reçu. En même temps, il choisit un mot de passe PW_i, génère un secret aléatoire K_i, et calcule $RPW_i = H(ID_i || K_i || PW_i)$. Le vecteur de caractéristiques W_i est extrait de la biométrie de *Ui*. La chaîne CS_i est générée à partir de W_i en utilisant le paramètre public *PB* (voire FE dans la Section 2.10.2). Utilisant CS_i, U_i calcule ce qui suit : $f_{si}^* = f_{si} \oplus H(ID_i || CS_i || K_i)$, $e_i = H(ID_i || RPW_i || CS_i)$, et $r_i = H(ID_i || CS_i) \oplus K_i$. $ECDH_{si}$ est cachée dans le paramètre bu_i comme suit: $bu_i = H(ID_i || CS_i) \oplus ECDH_{si}$. Enfin, $\{PB, e_i, r_i, bu_i, f_{si}^*\}$ sont sauvegardés dans le nœud utilisateur et tous les autres paramètres sont supprimés. La Figure 3.10 résume la phase d'inscription de l'utilisateur.

C) **Phase de connexion (Login)**

Lorsque U_i veut accéder à un service; où il y est déjà abonné et inscrit, il saisit son PW_i et ID_i, fournit son W_i', reproduit CS_i" à partir de W_i' et *PB*, et calcule les paramètres suivants: $K_i' = PB \oplus H(ID_i || CS_i')$, $RPW_i' = H(ID_i || K_i' || PW_i)$, et $e_i' = H(ID_i || RPW_i' || CS_i')$. Si $e_i' = e_i$ alors PW_i et W_i' sont authentiques, donc U_i peut envoyer la requête à *S*. La phase de connexion est illustrée à la Figure 3.11.

D) **Phase d'authentification de l'utilisateur et d'échange de clé**

Une fois la requête parvenue à *S*, ce dernier vérifie la validité de ID_i, et répond avec un défi aléatoire R_1. À ce stade, *Ui* récupère $ECDHsi = H(ID_i || CS'_i) \oplus bu_i$, et envoie $Een_{ECDHsi}(R_1, T_1, URL)$ à *S.* Noter qu'à partir d'ici, nous indiquons T_i $(i=1,2,3,4,5,6)$ pour l'horodatage de l'événement instantané. Après déchiffrement du message reçu en utilisant $ECDH_{si}$. *S* vérifie si R_1 reçu est le même que R_1 envoyé, si c'est le cas, *Ui* sera authentifiée avec succès. *S* vérifie également la récence de l'horodatage sous la condition $|T_1 - T_1^*| < \Delta$, où T_1^* c'est l'heure à laquelle le message est reçu, et Δ c'est le seuil de différence de temps qui a une valeur

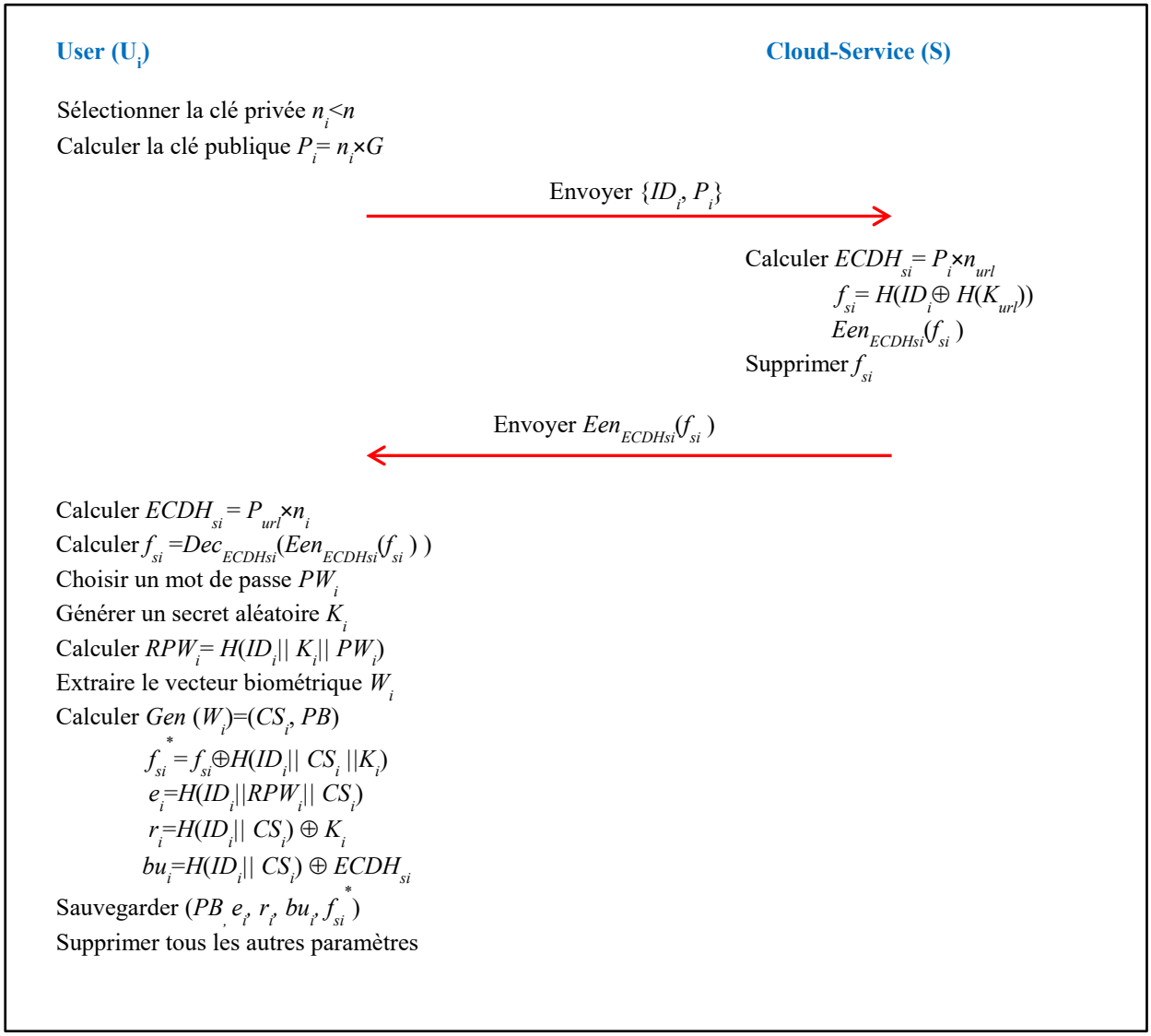


Figure 3.10- Phase d'inscription de l'utilisateur.

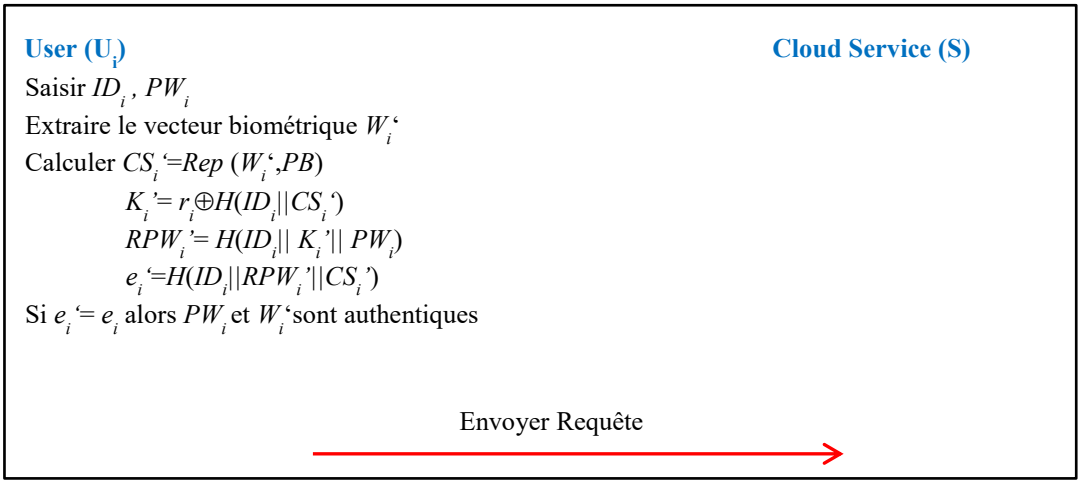


Figure 3.11- Phase Login.

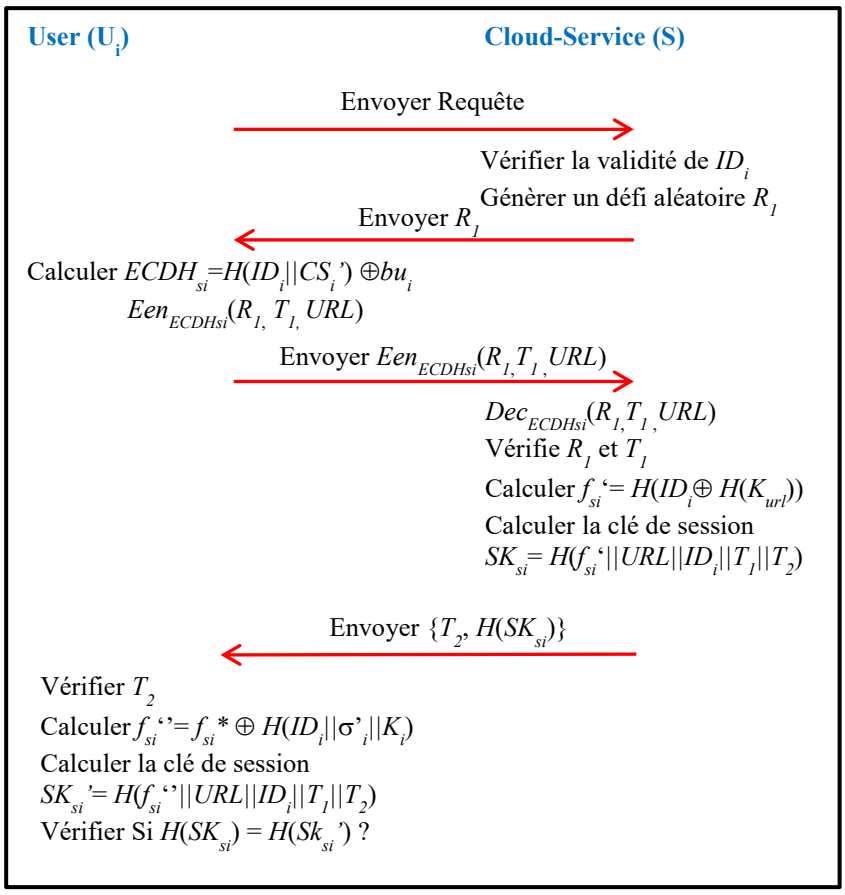


Figure 3.12- Processus d'authentification mutuelle et d'échange de clé entre l'utilisateur et le Cloud-Service.

suffisamment petite. Puis, il calcule $f_{si}'=H(ID_i \oplus H(K_{url}))$, ainsi la clé de session $SK_{si} = H(f_{si}'||URL||ID_i||T_1||T_2)$. La clé hachée $H(SK_{si})$ et T_2 sont envoyées à *Ui*.

Ui calcule $f_{si}'' = f_{si}^* \oplus H(ID_i||CS_i'||K_i)$, afin que la clé de session SK_{si}' puisse être récupérée en calculant $H(f_{si}''||URL||ID_i||T_1||T_2)$. La clé hachée calculée $H(Sk_{si}')$ est ensuite comparée à la clé hachée reçue $H(SK_{si})$. S'ils sont identiques, le processus d'établissement et d'échange de clés entre l'utilisateur et le Cloud-Service est réussi et l'authentification mutuelle est effectuée. U_i vérifie également la récence du message reçu par la valeur de T_2. Figure 3.12 montre le flux de messages pour le processus d'authentification et d'échange de clé entre l'utilisateur et le Cloud-Service.

E) **Phase d'authentification mutuelle et d'échange de clé entre Cloud-Service/Brouillard/Capteur**

En revanche, lorsque S reçoit la requête de *Ui*, ce dernier envoie *URL* au FN_k correspondant. FN_k génère un défi aléatoire R_2 et l'envoie à *S*. *S* sélectionne les capteurs SN_j appropriés pour FN_k requis et envoie $Een_{ECDHsk}(R_2, T_3, ID_{SNj})$ à FN_k. Le message reçu est déchiffré à l'aide de

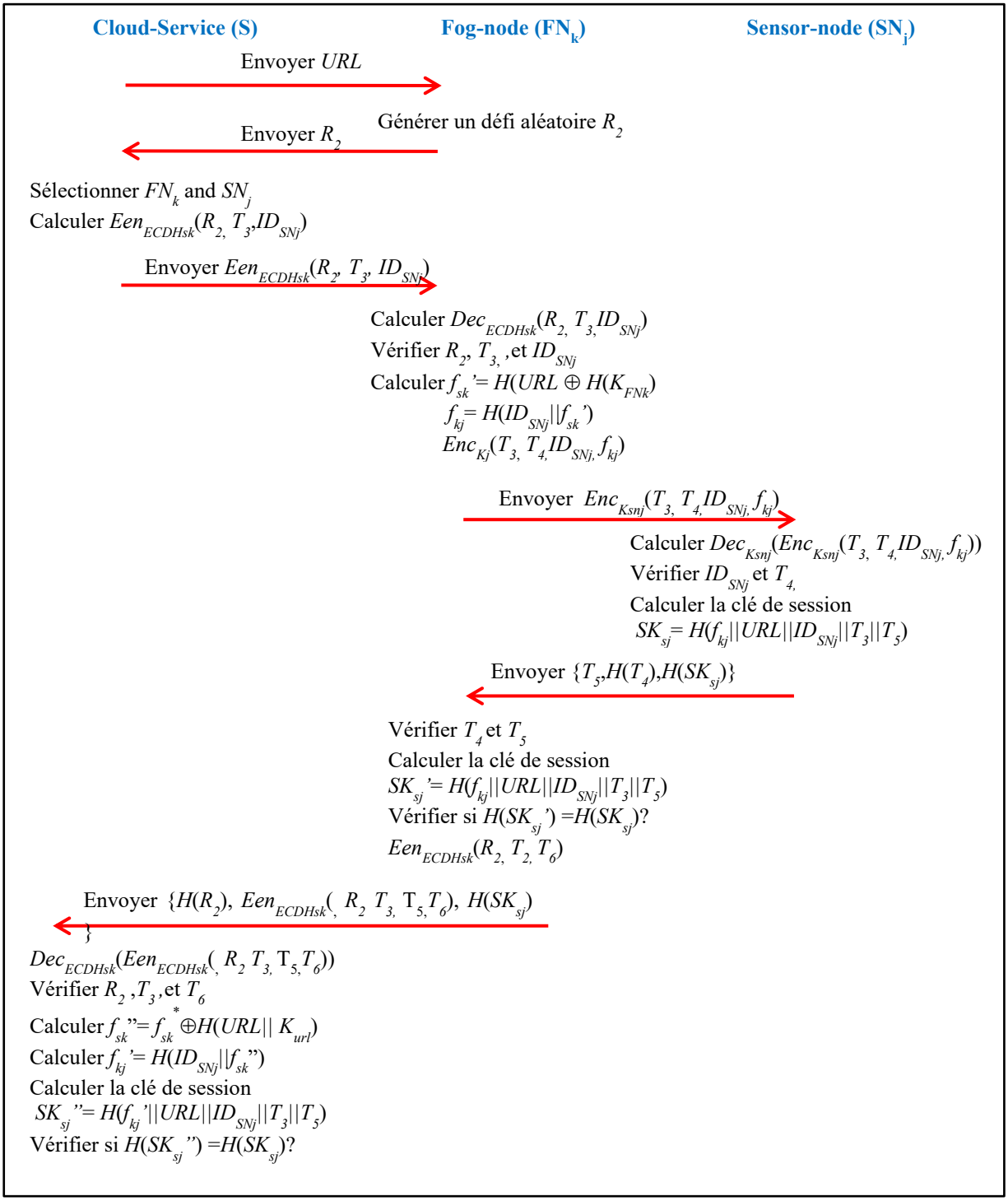


Figure 3.13- Processus d'authentification mutuelle et d'échange de clé entre Cloud-Service/ Brouillard/Capteur.

$ECDH_{sk}$. R_2, et T_3 sont vérifiés pour assurer respectivement l'authentification S et la récence du message. Puis, FN_k calcule $f_{sk}'=H(URL \oplus H(K_{FNk})$ et $f_{kj}=H(ID_{SNj}||f_{sk}')$, chiffre La chaîne ($T_3$, T_4,ID_{SNj},f_{kj}) par la clé K_{SNj} appropriée et pré-partagée et l'envoie à SNj. Noter que, ce processus est effectué avec tous les capteurs impliqués.

Lorsque La chaîne chiffrée est arrivée à *SNj*, elle est déchiffrée, et les paramètres ID_{SNj} et T_4 sont vérifiés. Par conséquent, le nœud de brouillard est authentifié auprès du nœud de capteur. A ce point, SN_j calcule la clé de session $SK_{sj}= H(f_{kj}||URL||ID_{SNj}||T_3||T_5)$. $H(SK_{sj})$, $H(T_4)$, et envoie T_5 à FN_k. Ce dernier vérifie T_4 et T_5, en conséquence, le nœud de capteur est authentifié auprès du nœud de brouillard. Ensuite, il récupère la clé de session $SK_{sj}'= H(f_{kj}||URL||ID_{SNj}||T_3||T_5)$. Le hach de la clé de session calculée est comparée au hach de la clé reçue. Si la même valeur est trouvée, le processus d'échange de clé et l'authentification mutuelle entre le nœud de brouillard et le nœud de capteur est réussi.

FN_k envoie $H(R_2)$, $H(SK_{sj})$, et $Een_{ECDHsk}(, R_2\ T_3,\ T_5, T_6)$ à *S.* Pour que le Cloud-Service authentifie le nœud de brouillard, *S* déchiffré la chaîne reçue et vérifie R_2, T_3, et T_6. La clé de session SK_{sj}'' est récupérée en calculant $f_{sk}''= f_{sk}^{*} \oplus H(URL||K_{url})$, $f_{kj}'= H(ID_{SNj}||f_{sk}'')$, et $H(f_{kj}'||URL||ID_{SNj}||T_3||T_5)$. Si la clé récupérée est égale à celle reçue, le processus d'échange de clé et l'authentification mutuelle entre le Cloud-Service et le nœud de brouillard est réussi. Enfin, le nœud capteur est authentifie avec succès par le Cloud-Service, si le T_3 reçu à la même valeur que le T_3 envoyé. La figure 3.13 montre le flux de messages de mutuelle authentification et d'échange de clé du Cloud-Service, du nœud de brouillard, et du nœud de capteur.

3.10 Conclusion

L'implication des technologies biométriques dans l'IoT est le point le plus important de notre thèse. Notre contribution dans cet axe est décrite dans ce chapitre. Nous avons d'abord décrit « Où et Comment la biométrie peut être impliquée dans l'IoT. Pour répondre à la question « Où », nous avons clarifié les trois situations dans lesquelles une personne peut s'intégrer à l'IoT. La première est en tant que consommateur de données, la deuxième en tant que producteur de données, et la troisième en tant que propriétaire de données. Pour répondre à la question « Comment », nous avons expliqué ici les techniques permettant l'intégration des technologies biométriques dans ces situations. Étant donné que toutes les technologies biométriques ont la même manière de s'intégrer dans toutes les situations, Nous avons choisi d'axer notre travail sur l'intégration de l'humain en tant qu'utilisateur, car c'est plus courant.

Au début de ce chapitre, une partie est consacrée à l'état de l'art concernant l'implication des technologies biométriques dans l'IoT. Dans cette partie, nous avons abordé le rôle de l'utilisateur à trois niveaux, au niveau de l'architecture IoT de bout-en-bout, au niveau de l'accès au cloud, et au niveau du modèle S^2aaS. Selon cette classification, nous avons fourni une solution pour chaque niveau Dans la dernière partie.

Chapitre 4 : Résultats expérimentaux : Evaluation, Discussion, et Commentaires

Chapitre 4 : Résultats expérimentaux : Evaluation, Discussion, et Commentaires

4.1 Introduction

Ce dernier chapitre est consacré aux résultats expérimentaux obtenus. Pour chacune de nos solutions, nous avons abordé les points suivants, la sélection des paramètres expérimentaux et de l'environnement, la discussion des résultats, la comparaison avec d'autres travaux récents pertinents, et un bref résumé des résultats. Enfin, une conclusion complète résume tous les résultats finaux.

4.2 Sélection des paramètres expérimentaux et de l'environnement

Les paramètres expérimentaux et l'environnement de mise en œuvre de chaque solution seront précisés dans cette section. Noter que Tous les processus ont été exécutés sous *Matlab R2017a* sous un système d'exploitation *Windows 10* et un PC embarqué avec un processeur *Intel Pentium i5,* 2 cœurs 2,5 GHz avec 6 Go de RAM.

4.2.1 *L'environnement expérimental de l'IoT de bout-en-bout*

Dans la partie sous-système biométrique du schéma proposé, nous pouvons utiliser n'importe quelle méthode d'extraction de caractéristiques à partir d'une image biométrique. Par conséquent, des techniques allant des méthodes classiques aux méthodes modernes peuvent être utilisées. Notre système étant principalement dédié au cryptage biométrique, nous nous concentrerons davantage sur l'efficacité du schéma du point de vue de la sécurité de l'information. Pour ce faire, PCA a été choisie car il s'agit d'une méthode largement utilisée, efficace et rapide.

Pour rendre la phase de validation plus crédible et plus équitable, le crypto-système proposé a été testé avec différentes modalités biométriques, qui sont le visage, l'empreinte digitale et l'empreinte palmaire. Les bases de données d'empreintes digitales et d'empreintes palmaires de l'université *Polytechnique de Hong Kong* (POLY) [Fingerprint-Bases], [Palmprint-Bases], tandis que la base de données *FEI Face* [Face-Bases] est utilisée pour valider l'efficacité du schéma par la modalité de visage. Chaque base de données contient 100 personnes, chaque personne à 12 instances, 10 pour l'apprentissage, une image pour la phase d'enrôlement, et une pour la requête. Le RS a été utilisé comme un code correcteur orienté octet et EH comme un code orienté bit. L'algorithme de décodage implémenté dans le RS code est *Berlekamp-Massey* [Berlekamp-Massey].

4.2.2 *L'environnement expérimental de l'accès au Cloud*

Dans la solution proposé pour de l'accès au Cloud [Bentahar2018] et [Bentahar2018a], la base de données de FVC2000 [FVC2000-Base] a été utilisée comme base de données interne, et la base de données du NIST [NIST-Base] a été utilisée comme base de données externe. La base de données interne contient 40 personnes en tant qu'utilisateurs authentiques, et la base de données externe contient 40 personnes en tant qu'imposteurs. Chaque personne dispose 8 images d'empreintes digitales (instances), dont 75 % pour l'enrôlement et 25 % pour l'interrogation (test). La base de données externe a été utilisée pour déterminer le FAR pour les imposteurs. Il existe un autre type de FAR dans la base de données interne, lorsqu'un utilisateur authentique est accepté comme un autre utilisateur authentique. Nous avons donc:

$$\text{FAR} = \frac{\text{FAR de base de données interne} + \text{FAR de base de données externe}}{2} \quad (19)$$

Les méthodes suggérées dans [Bentahar2019a] et [Bentahar2021a] ont été testées et validées sur la base de données [Fingerprint-Bases] de l'Université polytechnique de Hong Kong (PolyU)[7]. Afin d'accueillir deux bases de données, cette base a été divisée en deux groupes: (*I*) Interne inclut les utilisateurs authentiques, et (*II*) Externe inclut les utilisateurs imposteurs. La base de données interne contient 40 personnes. Chaque personne soumet 12 images d'empreintes digitales, dix pour le EA, une en tant que modèle et une en tant que requête. La base de données externe contient 40 autres empreintes digitales de requête qui ne sont pas incluses dans EA. Toutes les images d'empreintes digitales ont été recadrées à 64 x 64 pixels en fonction du point singulier (ROI), ce qui a donné une matrice de *MP* de 4096 × 10 pixels.

Les bases de données d'empreintes digitales [Fingerprint-Bases], et d'empreintes palmaires [Palmprint-Bases] de l'Université Polytechnique de Hong Kong ont été utilisées pour notre travail dans [Bentahar2019]. Les images de ces bases de données ont été collectées auprès de volontaires parmi les étudiants et les employés de cette université des deux sexes d'âges différents (entre 20 et 60 ans). La base de données d'empreintes digitales comprenait 336 personnes lors de la première session et 160 lors de la deuxième session. Les personnes impliquées dans les deux sessions disposent 12 images sans contact et 12 images avec contact avec une variance intra-classe significative. La base de données d'empreintes digitales a été obtenue auprès de 250 volontaires, contenant chacune 12 images par paume, soit 24 images par

[7] La base de données de l'Université polytechnique de Hong Kong (PolyU) consiste à convertir une empreinte sans contact en une empreinte avec contact, et nous avons choisi cette base de données pour un test plus fiable car l'image générée par cette conversion est généralement de qualité moyenne par rapport à l'image source avec contact.

personne avec différentes positions de main et différentes conditions d'éclairage. La base de données *FEI Face* [Face-Bases] a été utilisée pour tester les performances des crypto-systèmes biométriques basés sur le visage. Cette base de données contient des images des visages de 200 personnes, chaque personne a 14 images. Des variances intra-classes importantes ont également été introduites. Cet ensemble de données contient des personnes de sexes différents et d'âges différents (entre 19 et 40 ans).

Les crypto-systèmes décrits dans [Bentahar2019] sont conçus pour fonctionner avec 150 personnes (pour chaque modalité biométrique). Donc, nous avons divisé les bases de données expérimentales en deux sous-ensembles. Le premier contient 100 personnes qui est une base de données interne. Ce sous-ensemble a été également divisé en trois galeries, où ≈ 83% (100 images, 10 images par personne) est sélectionné pour la phase d'apprentissage, ≈8,5% (100 images, 1 image par personne) pour la phase d'enrôlement, et les autres, ≈8,5%, (100 images, 1 image par personne) sont réservés aux différents tests (Identification). Le deuxième sous-ensemble contient 50 personnes qui est une base de données externe. Ce sous-ensemble consiste à simuler l'intrusion des imposteurs utilisant une image de requête pour chaque imposteur. Il convient de noter que chaque image biométrique, représentée par sa région d'intérêt (ROI) [Sen2002], [David2003], [Hjelmås2001] a été recadrée et redimensionnée à 64 x 64 pixels. Après l'application de la technique PCA sur EA, la matrice de projection *MP* a été obtenue. Ainsi, les deux images, Γ_t pour l'enrôlement et Γ_q pour l'authentification ont été projetées sur la matrice *MP* afin d'obtenir les vecteurs de caractéristiques *W* et *W'*.

L'évaluation des taux de reconnaissance de notre système a été faite à l'aide des critères conventionnels à savoir, le taux de reconnaissance au rang un (*Rank One Recognition*-ROR), le taux de faux rejet (*False Rejection Rate* - FRR), le taux de fausse acceptation (*False Acceptance Rate* - FAR), et le taux d'acceptation des utilisateurs autorisés (*Genuine Acceptance Rate* - GAR), ainsi que de nouveaux critères sont introduits à savoir, le taux d'erreur d'identification (*Identification Error Rate* - IER), le taux de réussite du système (*Success Rate* - SR), et le taux de client exclu (*Excluded Client Rate* - ECR). Pour plus de détails (faire référence à Annexe A).

Tous les travaux [Bentahar2019], [Bentahar2018], [Bentahar2018a], [Bentahar2019a], et [Bentahar2021a] ont été testés et validés dans deux états principaux, qui sont, système non sécurisé (non crypté) et système sécurisé (crypté). Les résultats ont été analysés d'abord entre les deux états, puis dans le même état. Noter que seules les requêtes authentiques proviennent de la base de données interne, tandis que pour la base de données externe, toutes les requêtes proviennent d'imposteurs.

4.2.3 *L'environnement expérimental du modèle S²aaS*

Le schéma proposé pour le S²aaS a été formellement validé par un outil largement utilisé sous le nom « (AVISPA) » (Automated Validation of Internet Security Protocols and Applications) [AVISPA]. Dans l'AVISPA, le schéma ou le protocole doit être implémenté en langage HLPSL (*High-Level Protocol Specification Language*) [Viganò2006]. Notre implémentation HLPSL est simulée à l'aide du simulateur SPAN (*Security Protocol ANimator for AVISPA*) [SPAN].

A) **Outil AVISPA**

AVISPA offre quatre *back-ends* qui sont: (I) «*On-the-fly Model-Checker*» (OFMC), (II) «*Constraint Logic-based Attack Searcher*» (CL-AtSe), (III) «*SAT-based Model-Checker*» (SATMC), et (IV) «*Tree Automata based on automatic approximations for the analysis of security protocols*» (TA4SP). Les détails des *back-ends* et leurs fonctions sont expliqués dans [Armando2006]. Dans l'AVISPA, le schéma ou le protocole doit être implémenté en langage HLPSL. HLPSL est un langage qui décrit les rôles, les sessions, l'environnement, et les objectifs. Les rôles de base sont appelés « Agents », les rôles composés sont appelés « Session», et l'Environnement décrit la composition de multi sessions. Le résultat révèle si les objectifs de sécurité ont été atteints ou non.

En HLPSL, les commandes: « SND » et « RCV » signifient respectivement les canaux d'envoi et de réception, « hash_func » signifie la fonction de hachage. Le chiffrement et le déchiffrement du message à l'aide d'une clé symétrique sont décrits comme suit : « {message}_key ». La spécification des objectifs consistent en des protocoles de de secret et des protocoles d'authentification. Les protocoles de secret sont exprimés dans la spécification des objectifs sous la forme : « *secrecy_of Sec*», où *Sec* est le nom du protocole, et exprimé dans la spécification de rôle sous la forme: « *secret(Key',Sec,Agent)* ». Cette dernière expression est similaire à la question: la clé « *Key* » est-elle sécurisée dans le « Agent » par le protocole « *Sec* » ? Les protocoles de d'authentification sont exprimés dans la spécification des objectifs sous la forme: « *authentication_on auth*», où *auth* est le nom de protocole d'authentification, et exprimé dans la spécification de rôle sous deux formes: (i) « *witness(node1,node2,auth,P)*» où *node1* et *node2* sont deux Agents, *P* est le paramètre utilisé pour assurer l'authentification. Cette expression signifie que *node1* demande à *node2* de l'authentifier à l'aide de *P* (ii) « *request(node1,node2,auth,P)*» indique si la réponse à la demande d'authentification est assuré ou non?

```
role session(Ui,S,FNk,SNj:agent,IDi,URL,IDsnj:text,H:hash_func)
def=
    local
        SND1,RCV1,SND2,RCV2,SND3,RCV3,SND4,RCV4:channel(dy)
    composition
            user (Ui,S,FNk,SNj,IDi,URL,H,SND1,RCV1)
        /\ servise_cloud (Ui,S,FNk,SNj,IDi,URL,IDsnj,H,SND2,RCV2)
        /\ fog (Ui,S,FNk,SNj,URL,IDsnj,H,SND3,RCV3)
        /\ sensor (Ui,S,FNk,SNj,URL,IDsnj,H,SND4,RCV4)
end role
role environment()
def=
    const
ui,s,fnk,snj:agent,
h,gen,rep:hash_func,
idi,url,idsnj,t1,t3,t2,t4,t5,t6:text,
sec_1,sec_2,sec_3,sec_4,
u_auth_s,s_auth_u,s_auth_f,f_auth_s,f_auth_sn,sn_auth_f, s_auth_sn:protocol_id
    intruder_knowledge = {idi,url,idsnj,h,gen,rep,t2,t5}
    composition
                session(ui,s,fnk,snj,idi,url,idsnj,h)
            /\ session(ui,s,fnk,snj,idi,url,idsnj,h)
end role
goal
secrecy_of sec_1
secrecy_of sec_2
secrecy_of sec_3
secrecy_of sec_4
authentication_on u_auth_s
authentication_on s_auth_u
authentication_on s_auth_f
authentication_on f_auth_s
authentication_on f_auth_sn
authentication_on sn_auth_f
authentication_on s_auth_sn
end goal
environment()
```

Figure 4.1- Spécification du rôle de la session, du goal, et de l'environnement en HLPSL (Source [Bentahar2021])

B) **Spécification du schéma proposé en HLPSL**

Les quatre agents impliqués dans notre implémentation HLPSL sont: User (*Ui*), Cloud- Service (*S*), Fog-Node (*FNk*), et Sensor-Node (*SNj*). La spécification de la session, de l'environnement, et des objectifs sont décrits dans la Figure 4.1. Afin de tester correctement la résistance du schéma à diverses attaques; surtout les attaques qui utilisent des messages hérités, et afin de vérifier également les performances dans un environnement multisession, le schéma proposé doit être stimulé pendant au moins deux sessions. Cet environnement multisession est décrit dans la figure 4.1. La spécification des objectifs de notre schéma se compose de quatre protocoles de secret et sept protocoles d'authentification. Les protocoles de secret pour les clés *Ki, Kurl, Kfnk* and *Ksnj* sont exprimés dans le rôle de « *goal* » respectivement comme suit: *secrecy_of sec_1*,

secrecy_of sec_2, secrecy_of sec_3, et *secrecy_of sec_4*. Les protocoles d'authentification sont: *authentication_on u_auth_s* lorsque l'utilisateur authentifie le service, *authentication_on s_auth_u* lorsque le service authentifie l'utilisateur, *authentication_on s_auth_f* lorsque le service authentifie le brouillard, *authentication_on f_auth_s* lorsque le nœud de brouillard authentifie le service, *authentication_on f_auth_sn* lorsque le nœud de brouillard authentifie le capteur, *authentication_on sn_authons_f* lorsque le capteur authentifie le brouillard, et enfin *authentication_on s_auth_sn* lorsque le service authentifie le capteur.

C) **Simulateur SPAN**

Notre implémentation HLPSL est simulée à l'aide du simulateur SPAN [SPAN] sous un système d'exploitation *Ubuntu* par *VirtualBox* (processeur Intel Pentium i5, 2 cœurs de 2,5 GHz et DRAM de 2Go). Le schéma est vérifié à l'aide du *back-end* connu par OFMC.

4.3 Mesure de performance

Dans cette section nous présentons les performances de nos systèmes de sécurité pour: IoT de bout-en-bout, Accès au Cloud, et Modèle S^2aaS.

4.3.1 *IoT de bout-en-bout*

Dans la solution biométrique de l'IoT de bout-en-bout, les performances ont été évaluées en termes de : Longueur appropriée de la clé, Taux de reconnaissance et coûts de calcul, Analyse d'énergie, et Analyse de sécurité.

A) **Longueur appropriée de la clé**

La longueur du vecteur de caractéristiques est de 20 éléments. Par conséquent, le nombre de coefficients de polynôme doit être égal à leur demi-longueur (10), de sorte qu'une interpolation peut être obtenue en utilisant l'approximation des moindres carrés par chaque coefficient avec la moindre erreur et une longueur minimale (en bits). La longueur de clé optimale utilisée est obtenue en analysant les taux de reconnaissance par rapport à la longueur de clé. Comme le montre la Figure 4.2, FRR est directement proportionnel à la longueur de la clé et le FAR est inversement proportionnel. Avec une augmentation de la longueur de la clé, le niveau de sécurité augmente également, mais la précision diminue. La Figure 4.2 montre que la longueur de clé de 160 bits est le meilleur compromis entre sécurité (Zéro FAR) et précision (Minimum FRR). Ce dernier correspond à 5 et 6, respectivement pour les codes correcteurs RS et EH.

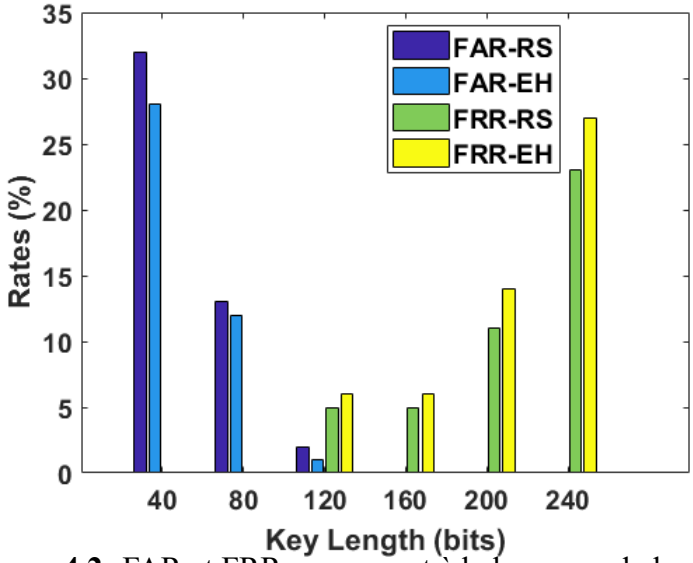


Figure 4.2- FAR et FRR par rapport à la longueur de la clé (la reconnaissance de visage extrait de [Bentahar2020]).

B) **Taux de reconnaissance et coûts de calcul**

Tableau 4.1 montre le GAR, le FRR, le FAR et le temps de calcul en utilisant FE basé sur le EH et le RS pour les trois modalités biométriques. Le système crypté se caractérise par l'absence de seuil. Les résultats sont basés sur la récupération de la clé, si la clé est récupérée, la requête est acceptée, sinon, elle est rejetée.

Tableau 4.1- Taux de reconnaissance et temps de calcul du système proposé (Source [Bentahar2020]).

Modalité Biométrique	Critère d'évaluation	FE-EH	FE-RS
Visage	GAR (%)	94	95
	FRR (%)	6	5
	FAR (%)	0	0
	Time (s)	1.63	12.55
Empreinte Digitale	GAR (%)	88	89
	FRR (%)	12	11
	FAR (%)	0	0
	Time (s)	1.40	12.35
Empreinte Palmaire	GAR (%)	86	90
	FRR (%)	14	10
	FAR (%)	0	0
	Time (s)	1.44	12.41

D'après le Tableau 4.1, Nous constatons que tous les FARs sont nules et que les FRRs sont satisfaisantes par rapport au haut niveau de sécurité offert. En termes de code correcteur, le RS est légèrement meilleur que le EH (1% à 3%), mais en termes de temps de calcul, il est environ 12 fois plus lent. Étant donné que l'IoT nécessite une réponse rapide et une faible latence, il est

préférable d'utiliser un code correcteur rapide. Dans notre cas, EH répond à ces exigences, notamment, car il n'est pas très différent du code RS en termes de taux de reconnaissance.

C) **Analyse d'énergie**

Comme nous constatons dans la Figure 3.5, L'authentification et l'échange de la clé sont effectués dans un seul message plutôt que plusieurs échanges séparément pour chaque processus. Ce massage se compose de *SS, PB* et *H (k). SS* contient au maximum de 20 paires des points (X_Q, *P* (X_Q)), et 20 autres faux points (X_a, Y_a). Alors la longueur de *SS* est de 640 bits (80 x 8 bits). *PB* est une concaténation de *t* et *r* (voir équation (31)). *t* n'est que de 26 bits comme suit: 6, 6, 5, 5, et 4 bits correspondant respectivement aux secondes, minutes, heures, jours et mois. *r* est une chaîne aléatoire de 230 bits. La longueur de *PB* donc de 256 bits. La fonction de hachage SHA-256 a été utilisée dans notre expérimentation car il est largement utilisé dans la transmission sécurisée dans les réseaux sans fil. Et alors, la longueur *H (k)* est de 256 bits. La longueur finale du massage envoyé n'est que de 1152 bits ≈1Kbits afin d'effectuer l'authentification et le partage de la clé en même temps.

Au niveau de la sauvegarde, le nœud d'utilisateur n'enregistre aucune information et le nœud du brouillard ne sauvegarde que 160 bits du vecteur de caractéristiques pour chaque utilisateur.

Tableau 4.2- Comparaison entre le schéma proposé et d'autres travaux connexes (Source [Bentahar2020]).

Modalité Biométrique	Auteurs	FRR	FAR	Code Correcteur	Longueur du code (bits)	Méthode d'extraction	Base de données	Nombre d'individus
Visage	[Wang2007]	0.5	7.38	RS	144	PCA	ORL	40
	[Wu2011]	8.5	0	CRC	144	PCA	ORL FRAV2D	40 100
	The proposed FE	6/5	0	EH/RS	160	PCA	FEI Face	100
Empreinte Digitale	[Uludag2006]	27	0	CRC	144	Minutiae Points	FVC 2002	100
	[Nandakumar2007]	4	0.04	CRC	128	Minutiae Points	FVC 2002 MSU	100 160
	The proposed FE	12/11	0	EH/RS	160	PCA	PolyU	100
Empreinte Palmaire	[Kumar2009a]	1	0.3	RS	306 to 309	DCT	Live Data	85
	The proposed FE	14/10	0	EH/RS	160	PCA	PolyU	100
Iris et Empreinte Digitale	[Kaur2017]	12.2 10.8	0.26 0.5	BCH BCH	150 150	Minutiae Points	CASIA Live Data	18 18

D) **Analyse de sécurité**

Diverses attaques de transmission peuvent cibler notre système. Dans cette section, ces attaques sont analysées et les exigences de sécurité suivantes sont confirmées.

- **Authentification:** selon les taux de reconnaissance indiqués ci-dessus, l'authentification est assurée.

- **Confidentialité:** Nous supposons qu'un intercepteur peut renifler *SS* et *PB*. Dans ce cas, d'une part il ne peut pas obtenir les données biométriques sensibles, car *SS* contient un mélange des points authentiques et des faux points qui ne peuvent être distingués sans connaissance de *P*. d'autre part, La possession de *PB* sans connaître *W* n'a aucun effet. Le crypto-système proposé est donc résistant aux attaques par reniflement et aux attaques par usurpation d'identité. De plus, ne pas obtenir les informations biométriques signifie ne pas obtenir la clé secrète. alors la communication cryptée utilisant cette clé reste confidentielle. De ce fait, le système assure la confidentialité.
- **Intégrité:** L'attaque bien connue « l'Homme du Milieu (MITM) » cible cette exigence. Si un MITM modifie *PB* ou *SS* pendant la transmission, à la fois *H(k)* et *CS* seront complètement changés, de sorte que l'intégrité des données est garantie.
- **Fraîcheur:** En utilisant l'horodatage, l'ancienne demande sera rejetée.
- **Confidentialité persistante:** En générant une clé aléatoire pour chaque session, le piratage d'une session ne révèle pas les autres sessions.
- **Résistance contre « *Replay Attack* »:** Si l'adversaire essaie de renvoyer *PB*, le nœud de brouillard ne l'acceptera pas car le *PB* aléatoire itère à nouveau.

E) **Analyse comparative**

Le temps de calcul n'est pas pris en compte dans cette comparaison, car il dépend des matériels informatiques utilisés, qui varie d'un travail à l'autre.

Tableau 4.2 montre que le nombre d'échantillons de base de données dans notre implémentation est parmi les plus grands. Et notre clé a une longueur appropriée (160 bits) qui n'est pas assez longue pour menacer la précision ou assez courte pour menacer la sécurité. De plus, notre FAR est toujours nul, même dans les cas où d'autres FRR sont meilleurs que les nôtres, leurs FAR ne le sont pas.

F) **Résumé des résultats**

Notre solution pour sécuriser l'IoT de bout-en-bout se caractérise par un niveau de sécurité élevé et un taux de reconnaissance raisonnable. La solution respecte les contraintes de l'IoT lors des deux phases, authentification et échange de clé. De plus, la clé secrète partagée est utilisée pour sécuriser l'échange de données et de messages entre les différents acteurs IoT de bout-en-bout. La comparaison avec des travaux connexes montre que notre système est plus sûr et plus précis.

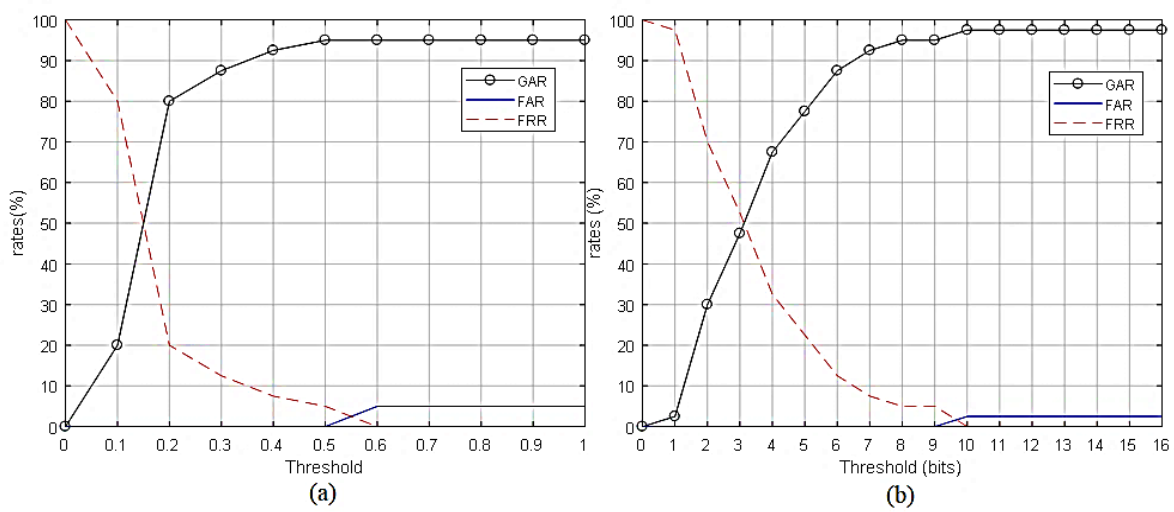


Figure 4.3- GAR, FAR, et FRR pour (a) DR-NB, (b) DR-B (source [Bentahar2021a])

4.3.2 *Accès au Cloud*

Dans les systèmes non cryptés, le seuil est déterminé par ERR. Et inversement dans les systèmes cryptés où il n'y a pas de seuil à à déterminer. FAR/FRR dans tels systèmes sont trouvés grâce à une récupération de clé réussie, si la clé est récupérée la requête est acceptée. Le temps de calcul pour chaque système crypté est le temps moyen de traitement d'une requête, composé du temps de la recherche de modèle, le calcul de clé, la correction des erreurs, et la prise de décision.

A) **Performance des systèmes : [Bentahar2018a], [Bentahar2018] [Bentahar2019a], et [Bentahar2021a]**

On note d'après le Tableau 4.3 [Bentahar2018a] et [Bentahar2018] que les systèmes cryptés proposés sont efficaces pour toutes les bases de données et pour les deux méthodes (DWT et DCT) selon le FAR nul. FAR minimum est l'indicateur le plus important en termes de sécurité car rejeter un utilisateur authentique est plus sûr que d'accepter un imposteur. Les valeurs FRR sont respectivement de 2,5 et 5 pour DR-DCT et DR-DWT, et de 7,5 et 10 pour FC-DCT et FC-DWT. Cela donne la préférence au DCT.

Le FRR de DR_B est plus petit que FC-H, cela a du sens car le système est devenu plus sécurisé, et donc plus sévère. Cependant FRR reste faible et acceptable (petite différence - 5%).

Comme nous le savons, le seuil optimal doit être proche de l'EER et pas trop petit pour avoir un compromis entre le minimum FRR/FAR et le maximum GAR. D'après la Figure 4.3 extraite de [Bentahar2021a], le seuil optimal est respectivement de 0,5 et 9 bits pour DR-NB et DR-B, ce qui donne un FRR de 5% et un FAR de 0%. Ces seuils optimaux seront utilisés comme exigence

stricte de la valeur de distance minimale acceptable pour empêcher toute requête d'imposteurs. Le FAR obtenu selon ces seuils est nul pour toutes les requêtes externes.

Tableau 4.3- Taux de reconnaissance des systèmes proposés dans [Bentahar2018a] et [Bentahar2018].

Bases de données	**Rate %**	**Méthodes basé sur DWT** [Bentahar2018a]		**Méthodes basé sur DCT** [Bentahar2018]	
		DR-B	***FC-H***	***DR-B***	***FC-H***
Base de données interne	***FRR %***	5%	10%	2.5%	7.5%
	FAR %	5%	0%	2.5%	0%
	ERR %	5%	5%	2.5%	5%
	GAR %	95%	90%	97.5%	92.5%
Base de données externe	***FAR %***	5%	0%	2.5%	0%
Toutes les bases de données	***FRR %***	5%	10%	2.5%	7.5%
	FAR %	5%	0%	2.5%	0%
	EER %	5%	5%	2.5%	3.75%
	GAR %	95%	90%	97.5%	92.5%

Tableau 4.4- Taux de reconnaissance et le cout de calcule des systèmes proposés dans [Bentahar2019a] et [Bentahar2021a].

	Système non crypté		**Système crypté**		
	DR-NB	***DR-B***	***FC-H***	***FC-RS***	***FV-RS***
FAR (%)	0	0	0	0	0
FRR (%)	5	5	27.5	22.5	12.5
Temps (sec)	9.82×10^{-5}	2.16×10^{-4}	0.96	14.56	9.71

D'après les résultats du Tableau 4.4 [Bentahar2019a] et [Bentahar2021a] pour les systèmes cryptés, FV-RS montre de meilleurs résultats que les FCs schémas avec un FRR de 12,5 % par rapport au code orienté bit FC-H (22,5 %) et par rapport au code orienté octé FC-SR (27,5 %). Selon le temps de calcul, les systèmes DR-NB et DR-B sont les plus rapides ($9,82 \times 10^{-5}$ et $2,16 \times 10^{-4}$, respectivement). Lorsqu'il est limité aux systèmes cryptés, FC-H est le plus rapide, le plus lent est FC-RS, FV-RS est un compromis.

Étant donné que FC est naturellement orienté bit et code RS utilisé est orienté octet, une conversion de présentation binaire (GF(1 bit)) à une présentation entière GF (8 bits) et un codage/décodage, puis une inversion en binaire est nécessaire. Cela explique que FC-RS est prend plus de temps à s'exécuter que FC-H (14,56 contre 0,96 s). Cela montre également la différence des FRRs, parce que les codes orientés octet ont une plus grande capacité de correction que les codes orientés bit.

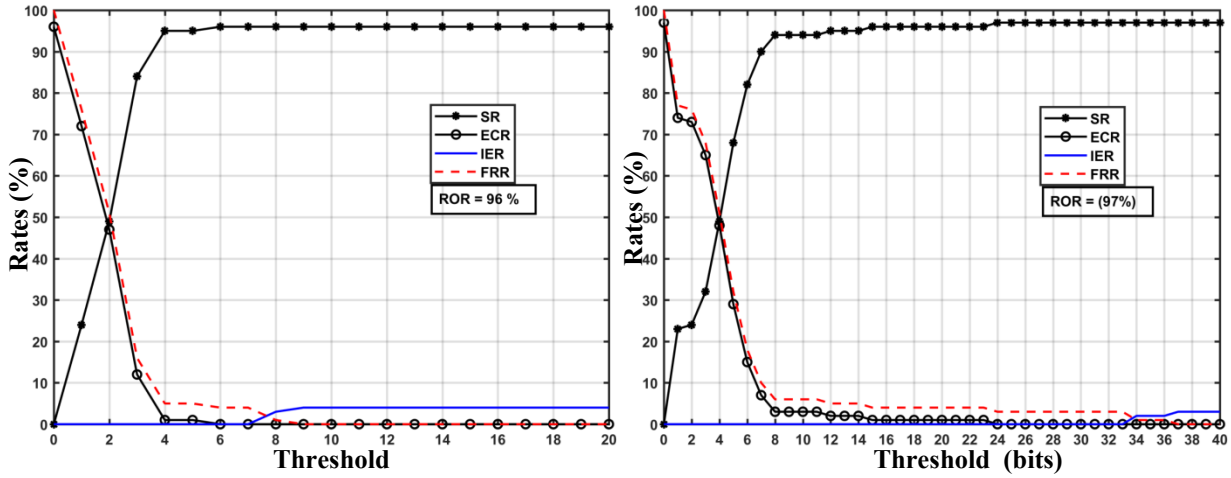
B) **Performance des systèmes non cryptés [Bentahar2019]**

Figures 4.4, 4.5, et 4.6 [Bentahar2019] montrent les résultats SR, ECR, IER, FRR, et ROR pour les visages, les empreintes digitales et les empreintes palmaires, respectivement. Diverses courbes du DR-NB sont capturées pour toute la plage (seuils 0-20 éléments), tandis que les courbes du DR-B ne sont capturées que pour les 40 premiers bits (seuil 0-40 bits au lieu de 0-160 bits). Lorsque le seuil de 40 bits est dépassé, toutes les courbes sont stabilisées.

En général, SR et IER sont proportionnels au seuil, ECR et FRR sont inversement proportionnels. Cela a du sens car lorsque le seuil est 0, aucune requête n'est acceptée, qu'elle soit authentique (SR) ou non(IER).

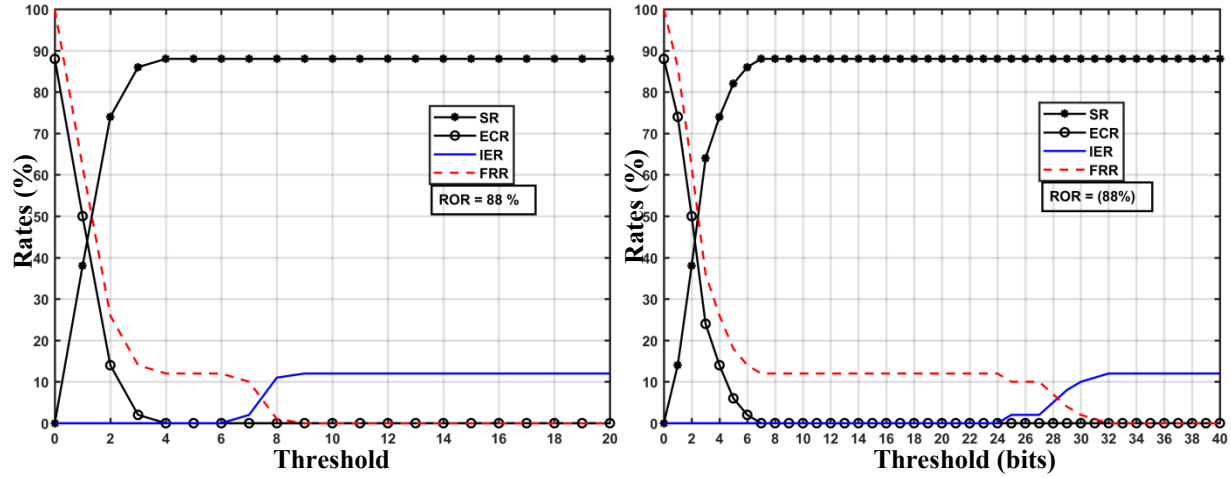
A ce stade, le FRR est de 100% car toutes les requêtes ont été rejetées. Noter que l'ECR est également à son taux maximum à ce stade, mais n'atteint pas 100% en raison de certains utilisateurs authentiques étaient auparavant éliminés selon les critères de distance minimale. ECR au seuil 0 est identique à la valeur ROR fixe. A l'autre extrémité (seuil maximum), ECR et FRR sont nuls car toutes les requêtes sont acceptées (sous forme d'un SR ou IER).

ECR est toujours inférieur au FRR pour les seuils bas et nul pour les seuils hauts. Cette infériorité s'explique par le fait que le FRR n'est obtenu qu'en utilisant le seuil alors que la distance minimale et le seuil sont utilisés ensemble pour l'ECR (deux étapes d'élimination). La plage des seuils bas est connue pour être plus sévère que la plage de seuils hauts qui est plus tolérante. Les deux plages sont séparées par le point EER. EER est défini comme la valeur à laquelle la différence entre ECR, IER et FRR est minimale. Ce point correspond toujours à l'intersection entre IER et FRR dans nos courbes. Étant donné que les abscisses sont quantitatifs (discrètes), le seuil optimal n'a pas besoin d'être exactement sur le point EER mais aussi proche que possible pour assurer un SR élevé et un ECR, IER, et FRR bas. Le seuil optimal est choisi pour offrir un compromis entre la sécurité et la précision. Dans la RD-NB, ce seuil optimal est de 8, 7 et 8, respectivement, pour les visages, les empreintes digitales et les empreintes palmaires (Figure. 4.4 (a), 4.5 (a) et 4.6 (a)). Est de 34, 28 et 34 bits, respectivement, dans le DR-B (voir Figure. 4.4 (b), 4.5 (b) et 4.6 (b)). Passé le seuil 9 dans DR-NB et de 32 à 38 bits (selon la modalité biométrique) dans DR-B, l'IER atteint une petite valeur constante. Le FAR dans le système d'identification ouvert est jusqu'à 100 % (Figure. 4.7). Alors que la valeur maximale de l'IER en DR-NB est de 4 %, 12 % et 12 %, et en DR-B est de 3 %, 12 % et 9 %, respectivement pour, le visage, les empreintes digitales, et les empreintes palmaires. Concernant SR, les courbes se sont stabilisées à 96%, 88%, et 88 % dans DR-NB et 97%, 88%, et 91 % dans DR-B



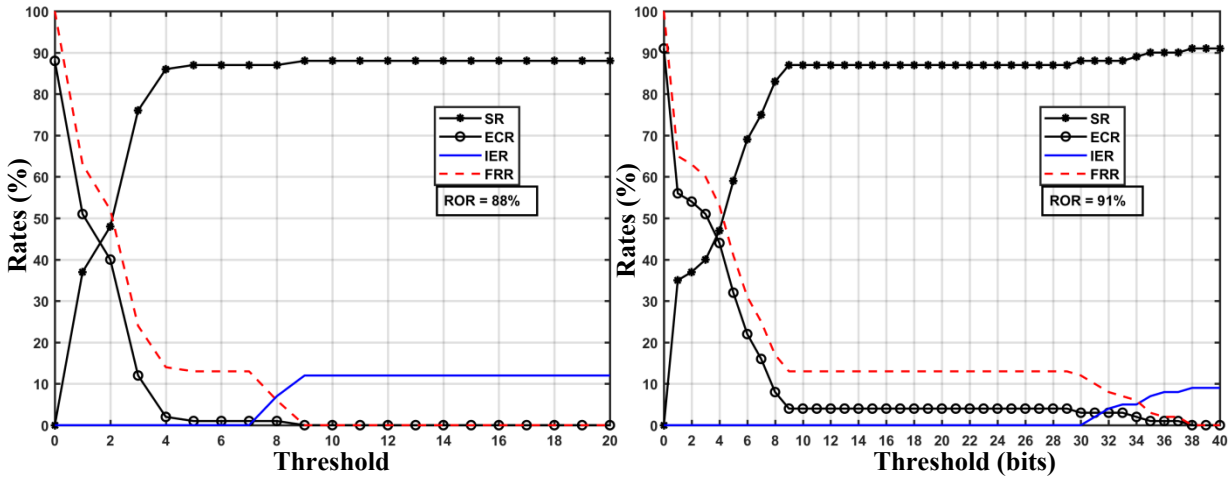
(a) (b)

Figure 4.4- Taux de la reconnaissance de visages dans DBs internes, (a) Représentation non-binaire, (b) Représentation binaire. (Source [Bentahar2019]).



(a) (b)

Figure 4.5- Taux de la reconnaissance d'empreintes digitales dans DBs internes, (a) Représentation non binaire, (b) Représentation binaire. (Source [Bentahar2019]).



(a) (b)

Figure 4.6- Taux de la reconnaissance d'empreintes palmaires dans DBs internes, (a) Représentation non binaire, (b) Représentation binaire. (Source [Bentahar2019]).

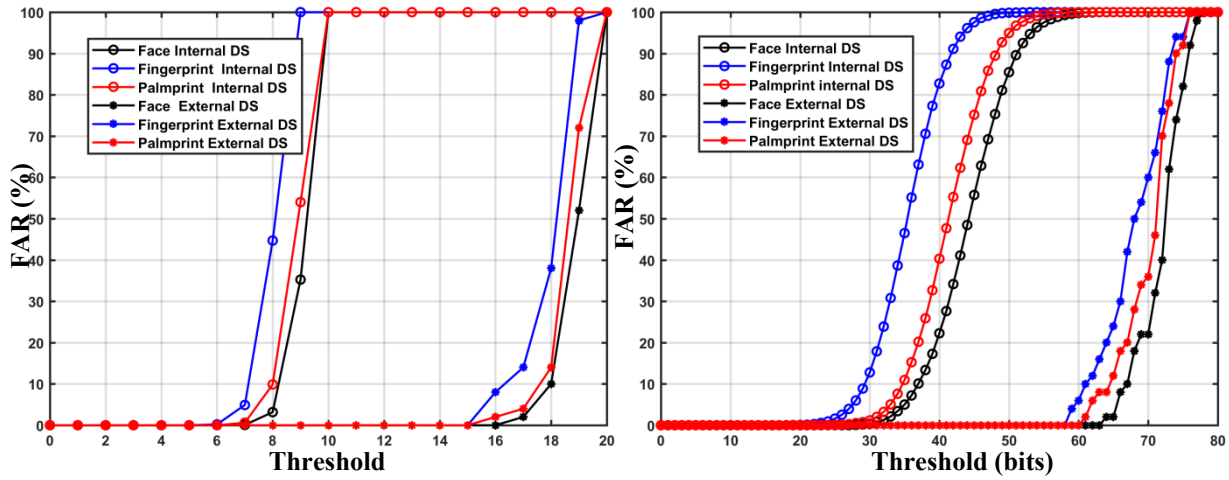


Figure 4.7- FAR des DBs internes et des DBs externes, (a) Représentation non binaire, (b) Représentation binaire. (Source [Bentahar2019]).

Tableau 4.5- Taux de reconnaissance et les couts de calcule des systèmes proposés (Source [Bentahar2019].

		Systems Non-Cryptés		Systems Cryptés			
Modalité biométrique	**Critère d'évaluation**	**DR-NB**	**DR-B**	**FC-H**	**FC-RS**	**FV-H**	**FV-RS**
Visage	SR (%)	96	97	90	91	94	95
	GAR (%)	99	99	90	91	94	95
	ROR (%)	96	97	90	91	94	95
	ECR (%)	0	0	0	0	0	0
	IER (%)	3	2	0	0	0	0
	FRR (%)	1	1	10	9	6	5
	FAR (%)	3.17	3.40	0	0	0	0
	Temps (s)	1.06×10^{-4}	5.62×10^{-4}	5.05	63	6.53	72.38
Empreinte digitale	SR (%)	88	88	87	88	88	89
	GAR (%)	90	93	87	88	88	89
	ROR (%)	88	88	87	88	88	89
	ECR (%)	0	0	0	0	0	0
	IER (%)	2	5	0	0	0	0
	FRR (%)	10	7	13	12	12	11
	FAR (%)	4.89	5.99	0	0	0	0
	Temps (s)	1.34×10^{-4}	6.81×10^{-4}	4.82	61.14	6.34	72.29
Empreinte Palmaire	SR (%)	87	89	82	84	86	90
	GAR (%)	94	94	82	84	86	90
	ROR (%)	88	91	82	84	86	90
	ECR (%)	1	2	0	0	0	0
	IER (%)	7	5	0	0	0	0
	FRR (%)	6	6	18	16	14	10
	FAR (%)	9.87	7.63	0	0	0	0
	Temps (s)	1.22×10^{-4}	6.06×10^{-4}	4.98	62.11	6.47	72.33

respectivement pour le visage, l'empreinte digitale, et l'empreinte palmaire. Ces valeurs obtenues sont satisfaisantes, et surtout, le système reste sûr et précis même avec un seuil augmenté.

Pour comparer les taux d'erreur de reconnaissance entre les bases de données internes et les bases de données externes, les courbes de la Figure 4.7 (a) et de la Figure 4.7 (b) ne montrent que le

FAR conventionnel car le FRR ne peut pas être défini pour la base de données externes (Imposteurs). Selon la Figure 4.7, la différence entre l'apparence de FAR dans la base de données interne et la base de données externe est importante. Dans DR-NB, les seuils correspondants sont respectivement de 7, 6, 7, et de 17, 16, 16 pour les bases de données internes et les bases de données externes. Dans DR-B, les seuils correspondants sont respectivement de 24, 18, 20, et de 64, 59, 61 bits pour les bases de données internes et les bases de données externes. Ainsi, le seuil optimal choisi garantit non seulement un grand SR et un petit ECR, IER, et FRR, mais garantit également zéro FAR pour tous les imposteurs qui n'ont pas été impliqués dans le processus d'apprentissage.

Selon le Tableau 4.5, il n'y a pas de différence significative dans le taux de reconnaissance entre les implémentations non binaires et binaires. En effet, cette différence n'est que de 2% pour les trois modalités biométriques. Mais en termes de temps de traitement, l'implémentation non binaire est environ cinq fois plus rapide que l'implémentation binaire car la correspondance se fait octet par octet au lieu de bit par bit.

C) **Performance des systèmes cryptés [Bentahar2019]**

Dans les systèmes cryptés, FV est plus efficace que FC en termes de taux de reconnaissance (grand SR, GAR, ROR et petit FRR). Car la correspondance basée sur *Set Difference* est plus proche du vecteur de caractéristiques d'origine que la correspondance basée sur la distance de *Hamming*. Autrement dit, dans FC, l'utilisation du code *Gray* comme conversion supplémentaire peut entraîner une perte de précision. En termes de temps de traitement, FC est plus rapide que FV car le système FC est basé sur une simple opération XOR contrairement au système FV, qui nécessite la construction et l'interpolation polynomiale.

Selon un autre point de vue, le choix du code correcteur a un impact significatif sur les performances des systèmes cryptés. En effet, les taux obtenus par le code RS sont supérieurs aux taux du code EH. Parce que les codes orientés octet (RS) ont une capacité de correction énorme par rapport aux codes orientés bit (EH). Mais en termes de temps de traitement, EH est environ 12 et 11 fois plus rapide que le code RS pour le FC et le FV respectivement. Cet écart s'explique par le fait que les codes orientés octet sont utilisés dans un grand GF plutôt que des codes binaires.

D) **Analyse comparative**

Tableau 4.6 résume les résultats des FC-RS et FV-RS proposés par rapport à d'autres travaux. Le tableau présente les taux obtenus (FAR, FRR) avec, les différents codes correcteurs utilisés, la

Tableau 4.6- Comparaison des cryptosystèmes proposés avec d'autres travaux connexes (Source [Bentahar2019].

Cryptosystème	Modalité	travaux connexes	FRR	FAR	Code correcteur	Longueur du code (bits)	Méthode d'extraction	Base de données	Nombre d'individus
Fuzzy commitment	Visage	[Ao2009]	7.99	0.11	BCH	707	MB-LBP	FRGC-ver2.0	222
		[Lu2009]	30	0	BCH	63	PCA	CMU PIE	68
		FC-RS proposé	9	0	Reed-Solomon	160	PCA	FEI Face	150
	Empreinte digitale	[Teoh2007]	0.9	0	Reed-Solomon	375	Gabor	FVC 2002	100
		[Nandakumar2010]	12.6	0	Recursive convolutional	2048	Fourier transform	FVC 2002 DB1/ DB2	200
		FC-RS proposé	12	0	Reed-Solomon	160	PCA	PolyU	150
	Empreinte palmaire	FC-RS proposé	16	0	Reed-Solomon	160	PCA	PolyU	150
Fuzzy Vault	Visage	[Wang2007]	0.5	7.38	Reed-Solomon	144	PCA	ORL	40
		[Wu2011]	8.5	0	CRC	144	PCA	ORL FRAV2D	40 100
		FV-RS proposé	5	0	Reed-Solomon	160	PCA	FEI Face	150
	Empreinte digitale	[Clancy2003]	20	0	Reed-Solomon	128	Minutiae Points	VeriFinger_ Sample_DB	51
		[Uludag2006]	27	0	CRC	144	Minutiae Points	FVC 2002	100
		[Nandakumar2007]	4	0.04	CRC	128	Minutiae Points	FVC 2002 MSU	100 160
		FV-RS proposé	11	0	Reed Solomon	160	PCA	PolyU	150
	Empreinte palmaire	[Wu2008]	0.93	0	Reed Solomon	1024	2D Gabor	PolyU	193
		[Kumar2009a]	1	0.3	Reed Solomon	306 to 309	DCT	unconstrained peg-free	85
		FV-RS proposé	10	0	Reed Solomon	160	PCA	PolyU	150

longueur de clé, la méthode d'extraction, et la base de données. Les schémas du code RS ont été choisis car ils ont de meilleurs taux par rapport aux schémas du code EH. FAR et FRR ont été sélectionnés car il s'agit des taux conventionnels communs à tous les travaux connexes.

Premier mot, nous avons utilisé des bases de données de 150 personnes pour chaque modalité, et cela fait partie des grands nombres du tableau de comparaison. Deuxième mot, la clé dans nos systèmes est la longueur appropriée, ni trop petite pour menacer la sécurité ni trop longue pour augmenter la complexité de calcul qui affecte la légèreté. Troisième mot, selon les FRR obtenus, les résultats ont été très satisfaisants. Notre FAR est toujours nul. Alors que dans d'autres travaux où leur FRR est légèrement meilleur que le nôtre, leur FAR est non nul dans la plupart des cas.

E) **Résumé des résultats**

Les résultats obtenus montrent que le système non crypté basé sur la représentation binaire est plus lent que celui basé sur la représentation non binaire, mais il est légèrement plus précis. Les systèmes cryptés par le FC avec un code orienté bit est le plus léger et que les systèmes cryptés par FV avec un code orienté octet est la plus précise. Compromis entre légèreté et précision, le système crypté par FV avec un code orienté bit est la solution la plus convenable. Cependant, pour l'IoT, il est clair que la meilleure solution est la plus rapide donc la plus légère, qui est FC-H, d'autant plus que la différence de taux de reconnaissance n'est que de 4% au pire des cas avec le reste des systèmes cryptés.

4.3.3 *Modèle S^2aaS*

Dans cette section, les performances de la solution biométrique du modèle S^2aaS sont évaluées et discutées.

A) **Analyse informelle**

Cette section décrit la résistance du système à diverses attaques largement connues et analyse dans quelle mesure les objectifs de sécurité décrits ci-dessous ont été atteints.

- **Attaque par terminal d'utilisateur volé:** Supposons que l'adversaire A réussisse à obtenir toutes les données d'un terminal volé qui contient*:* PB_i, e_i, r_i, bu_i, et f_{si}^*. Sans une acquisition biométrique authentique et une saisie correcte du PW_i, A ne peut pas calculer d'autres paramètres tels que $ECDH_{si}$, K_i, et/ou RPW_i. Par conséquent, A ne peut pas s'authentifier ou générer des données cryptées avec le service souhaité.
- **Authentification mutuelle:** l'utilisateur s'authentifie auprès de son terminal à l'aide de ses données biométriques et d'un mot de passe correspondant. Par conséquent, $ECDH_{si}$ est récupéré pour une utilisation future. U_i est authentifié par S grâce au défi R_1, et S est authentifié par U_i

grâce à T_1. Via R_2, S et FN_k sont mutuellement authentifiés, et Via T_4, FN_k et SN_J sont mutuellement authentifiés. Finalement, SN_J est authentifié par S grâce à T_3. De plus, R_1, R_2, T_1, T_3 et T_4 ne sont vérifiés qu'après le processus de chiffrement et de déchiffrement. Sur la base de ce qui précède, il est clair que notre système assure l'authentification mutuelle à tous les niveaux.

- **Attaque par reniflement (*Sniffing Attack*):** Toutes les données sensibles sont cryptées à l'aide des clés *ECDH* ou de la clé K_{SNj} , et sont renforcées par la fonction de hachage. Ces mécanismes assurent une confidentialité/intégrité des données et empêchent A de révéler les données sensibles.
- **Attaque par rejeu (*Replay Attack*):** Étant donné que tous les messages d'authentification sont datés avec des horodatages, lorsque A renvoie l'un de ces messages, le récepteur le considère comme un *Replay Attack* en raison de la répétition du même nombre aléatoire ou de la violation de l'horodatage.
- **La récence du système (*Freshness*):** Avec un horodatage exact pour chaque message, le système rejette l'ancien message, de sorte que chaque entité ne traite que le nouveau message, afin que A ne puisse pas abuser de l'ancien message**.**
- **Attaque de l'homme du milieu (MITM):** Supposons que A renifle le message $Een_{ECDHsi}(R_1, T_1, URL)$, même si R_1 et URL sont connus, il est difficile pour A le modifier sans connaître $ECDHsi$ et T_1. Dans le côté Service/Brouillard, si A renifle $Een_{ECDHsk}(R_2, T_3, ID_{SNj})$, il ne peut faire aucune modification sans connaître $ECDHsk$ and T_3. Dans le côté Brouillard/Capteur, si A renifle $Enc_{Kj}(T_3, T_4, ID_{SNj}, f_{kj})$**,** le message ne peut pas être modifié sans connaître K_{SNj}, T_3 et T_4. De plus A ne peut modifier ni calculer la clé de session $SK_{si}= H(f_{si}'||URL||ID_i||T_1||T_2)$ ou $SK_{sj}= H(f_{kj}||URL||ID_{SNj}||T_3||T_5)$ sans connaître f_{si}' , T_1, f_{kj} et T_3. Comme nous le notons, A ne peut en aucun cas modifier ou calculer des informations sensibles. Par conséquent, le système assure la résilience aux attaques MITM.
- **Attaque d'usurpation d'identité (*Impersonation Attack*):** Supposons que A a obtenu K_{SNj} en capturant physiquement un nœud capteur, et il a intercepté le message d'authentification $Enc_{Ksnj}(T_3, T_4, ID_{SNj}, f_{kj})$. S'il peut imiter un vrai identifiant ID_{SNj}', il ne peut reconstruire $SK_{sj}= h(f_{kj}||URL||ID_{SNj}||T_3||T_5)$ ni reconstruire un véritable K_{SNj}. Car K_{SNj} est le résultat hachée de l'opération XoR entre le véritable ID_{SNj} et la clé maître MK, et ce dernier a été supprimé de tous les dispositifs. Pour cette raison, les nœuds de brouillard et les nœuds de capteur ne peuvent pas être usurpés. Sans oublier que, A ne peut fournir un vecteur biologique W_i' suffisamment proche d'un W_i natif sauvegardé, A ne connaît pas le mot de passe correct PW_i, Par conséquent, A ne peut pas s'authentifier à cause du mal calculé: $CS_i'=Rep(W_i', PB)$, $RPW_i'= H(ID_i|| K_i'|| PW_i)$ et

e_i'=$H(ID_i||RPW_i'||CS_i')$. Egalement, A ne peut pas récupérer la clé secrète de l'utilisateur $K_i' = r_i \oplus H(ID_i||CS_i')$, ni de récupérer $ECDHsi= H(ID_i||CS'_i) \oplus bu_i$. Il est clair que *A* est impuissant d'imiter l'utilisateur ou tout autre nœud.

- **Attaque par capture physique:** En supposant que *A* a physiquement atteint le nœud capteur, toutes les données sur ce capteur seront compromises y compris ID_{SNj} et K_{SNj}. Étant donné que différents horodatages T_3 et T_4 sont à nouveau générés pour chaque session, cet attribut rend la session unique. Par conséquent, *A* ne peut rien faire avec les données capturées sans recevoir T_3 et T_4 du nœud de brouillard. Le même mécanisme protège le système contre l'attaque physique du nœud de brouillard. L'utilisation des horodatages T_2, T_3, T_5, T_6 et du défi R_2 pour chaque session empêche *A* de créer la session en tant que véritable nœud de brouillard. Donc, notre système offre une résistance aux attaques de capture physique des nœuds IoT.
- **Attaques d'initiés au Cloud (*Insider Attack*):** Après des authentifications mutuelles réussies entre les acteurs de l'IoT, le Cloud ne traite plus que des acteurs authentiques. Supposons que *A* a piraté le Cloud et récupère sa clé secrète K_{url}, *A* ne peut pas établir une session en tant que véritable U_i car il ne connaît pas le f_{si}, ni en tant que véritable FN_k car il ne connaît pas le f_{sk}.
- **Attaque par déni de service (DoS):** Le nœud capteur ne gère que le message d'authentification qui a été déchiffré par K_{SNj}, et rejette tous les messages non chiffrés avec cette clé. De plus, après déchiffrement du message, le nœud capteur vérifie ses identifiants ID_{SNj} et T_4 reçu avant d'agir. En d'autres termes, le nœud de capteur rejette tous les messages inutiles pour éviter de gaspiller de l'énergie et du temps. Cela signifie que le nœud de capteur ne conserve de l'énergie que pour les messages authentiques et rend le système résistant aux attaques DoS.
- **Attaque de mot de passe hors ligne (*Offline Password Attack*):** Tout d'abord la devinette d'un vecteur biométrique authentique W_i est presque impossible. Sans oublier que W_i, ID_i, K_i, et PW_i sont concaténés et hachés a plusieurs niveaux pour produire e_i comme indiqué dans ce qui suit : $Gen\ (W_i)=(CS_i, PB)$, $RPW_i= H(ID_i|| K_i|| PW_i)$,et $e_i=H(ID_i||RPW_i||CS_i)$. De plus, il est très difficile de deviner le mot de passe PW_i. Tous ces facteurs rendent l'intrusion très compliqué surtout que le mot de passe et la biométrie sont requis ensemble.

B) **Vérification formelle de sécurité en utilisant l'AVISPA**

Comme le montre la Figure 4.8, après l'implémentation de notre protocole en utilisant l'AVISPA, et après la simulation en utilisant le SPAN, Notre système s'est avéré efficace contre

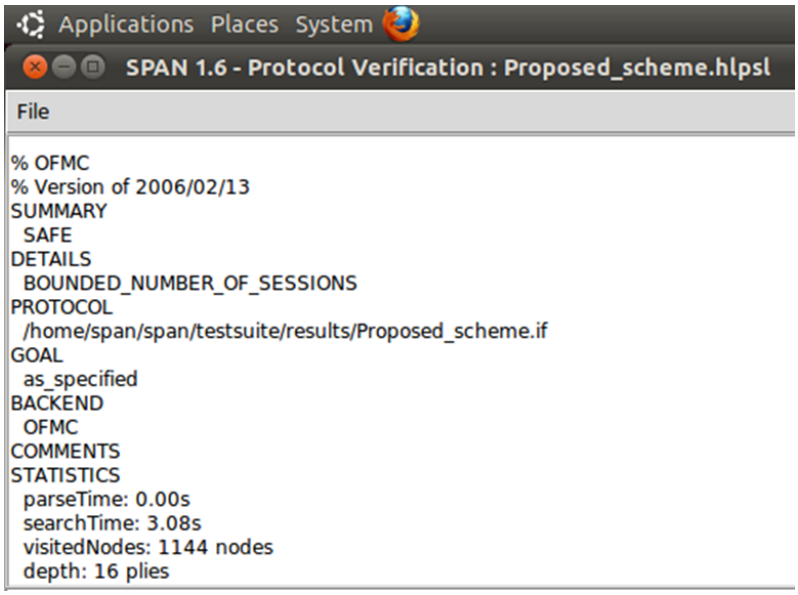


Figure 4.8- Résultats de la vérification formelle en utilisant *back-end* OFMC sous SPAN (Source [Bentahar2021])

une variété d'attaques car tous les objectifs de sécurité susmentionnés dans le rôle "Goal" sont atteints. La figure montre une capture d'écran des résultats à l'aide du *back-end* OFMC.

Les résultats de *Fuzzy Extractor* en termes de taux de reconnaissance et de coûts de calcul des phases d'inscription/de connexion des utilisateurs ont été discutés en détail dans la Section 4.3.1.

C) **Analyse comparative**

Cette section est dédiée à la comparaison du schéma proposé avec des schémas récents en termes de fonctionnalités de sécurité, de performances, et de certaines métriques connexes. Tableau 4.7 présente les éléments de sécurité et certaines des caractéristiques qui distinguent le schéma proposé d'autres schémas connexes. Noter que, F1: Résilience contre l'attaque par dispositif d'utilisateur volé (terminal utilisateur ou carte à puce), F2: Authentification mutuelle, F3: Résilience contre *Sniffing Attack*, F4: Résilience contre *Replay Attack*, F5: *Freshness*, F6: Résilience contre l'attaque d'usurpation d'identité, F7: Résilience contre *Offline Password Attack*, F8: Résilience contre l'attaque MITM, F9: Résilience contre l'attaque par capture physique (*Stolen Verifier*), F10: Résilience contre l'attaque DoS, F11: Vérification formelle du schéma, F12: Implication biométrique, F13: Evaluation de système biométrique F14: Inclusion du Cloud-Service, F15: Cloud est considéré comme un nœud non fiable, F16: Nombre de rôles.

Tout d'abord, les travaux passés en revue dans le tableau 4.7 ne traitent pas ou mentionnent certaines attaques bien connues. Sur la base de notre analyse de ces schémas, des vulnérabilités de sécurité ont été trouvées. Les fonctions de sécurité F1 à F7 sont satisfaites dans la plupart des

schémas du tableau. Cependant, la résilience de F8, F9 et F10 n'a pas été atteinte dans près de la moitié ces schémas. F13 à F16 sont les éléments très intéressants du Tableau 4.7. Effectivement, l'inclusion des services de cloud dans le système de sécurité IoT n'a été appliqué que dans [Roy2018] et [Gupta2019]. Comme indiqué par F16, dans [Roy2018], seuls deux rôles ont été incorporés (Utilisateur et Service), et trois rôles dans tous les schémas restants, alors que dans notre schéma, le nombre de rôles est de quatre. Le cloud est inclus en tant que nœud de confiance dans [Gupta2019], alors que dans le nôtre, il est traité comme un nœud non fiable. Les services publics sont déployés dans ce cloud non fiable, qui est l'un des nouveaux défis de sécurité IoT. Le schéma proposé a été démontré son efficacité à relever ce défi, comme nous l'avons décrit dans l'analyse informelle et illustré à la figure 4.8. Dans notre schéma, un sous-système biométrique est impliqué (F12) et évalué avec différents taux de reconnaissance et temps de calcul (F13).

La comparaison des coûts de calcul et de communication avec les travaux récents est présentée dans le tableau 4.8. On note les coûts de temps suivants, t_{fe} pour les opérations (*Gen(.)*, *Rep(.)*) de FE, $t_{e/d}$ pour le chiffrement/déchiffrement symétrique, t_{pm} pour la multiplication de points de courbe elliptique, t_{ch} pour l'opération de *Chebyshev Map*[8], et en fin t_h pour la fonction de hachage (*H(.)*). N/A c.-à-d. non applicable (car le cloud n'est pas inclus dans le schéma). *Comp* c'est le coût de calcul et *Comm* c'est le coût des messages de communication en bits. Les coûts du XoR et des opérations de concaténation sont négligeables. *Comp* et *Comm* sont facturés pendant les phases de connexion, d'authentification et d'échange de clé. Pour une comparaison crédible, les travaux, [Das2015], [Park2016], [Wang2017], [Moon2017], [Maurya2017], [Mishra2017], [Li2018], et [Roy2018] ont été choisis car les techniques biométriques y sont incluses, ce qui affecte le coût de calcul et de communication.

On sait que les opérations de la multiplication de points de courbe elliptique sont des opérations coûteuses et lourdes, et qu'elles ont donc été évités dans les périodes critiques et fréquentes telles que la connexion et l'authentification/l'échange de clés. Cela a été pratiquement démontré dans [Kilinc2014], où t_h= *0.0023 ms,* $t_{e/d}$= *0.0046 ms,* et t_{pm}=*2.226 ms,* il est clair que t_{pm} coûte environ 1000 fois plus que t_h et 500 fois plus que $t_{e/d}$.

Notre requête est ID_i pour U_i , et *URL* pour. Une longueur de 160 bits a été utilisée pour chacune de ces requêtes, car c'est largement suffisant, et il est également équivalent à d'autres travaux connexes. Les longueurs de ID_{SNj}, *H(.),* R_i, et T_i sont respectivement de 160, 160, 32, et 32 bits.

[8] *Chebyshev Map* est une carte chaotique avec sa fonction de densité invariante.

La taille de sortie Een_{KEY} *(.)* ou Dec_{KEY} *(.)* est égale à la taille des données d'entrée. Ainsi, les coûts de communication du connexion, de l'authentification, et de l'échange de clé pour Utilisateur/Service sont obtenus comme suit: requête = 160 bits, R_1= 32 bits, $Een_{ECDHsi}(R_1, T_1, URL)$=224 bits, et $\{T_2, H(SK_{si})\}$=192 bits, alors total: 608 bits. Pour Service/Brouillard/Capteur, les résultats sont: transfert de requête = 160 bits, R_2= 32 bits, $Een_{ECDHsk}(R_2, T_3, ID_{SNj})$=224 bits, $Enc_{Ksnj}(T_3, T_4, ID_{SNj}, f_{kj})$= 384 bits, $\{T_5, H(T_4), H(SK_{sj})\}$=352 bits, et $\{H(R_2), Een_{ECDHsk}(, R_2\ T_3, T_5, T_6), H(SK_{sj})\}$=448 bits, alors total : 1600 bits. Globalement: 608+1600=2208 bits.

Les coûts de calcul et de communication obtenus sont satisfaisants comme indiqué dans Tableau 4.8, malgré le schéma proposé offre, (i) Sécurité plus sûre que les autres (selon le Tableau 4.7), (ii) Quatre rôles, (iii) Résistance aux attaques ciblant le cloud, (iv) Tous les canaux de transmission ne sont pas sécurisés, (v) Plus de tâches aux technologies biométriques que l'authentification (masque et récupère les clés secrètes K_i and $ECDH_{si}$).

De plus, dans notre schéma, l'incorporation de quatre rôles nécessite sept processus d'authentification pour réaliser l'authentification mutuelle entre tous les acteurs IoT. Étant donné que les schémas connexes incluent jusqu'à un maximum de trois rôles, seules quatre processus d'authentifications sont requis, authentification mutuelle entre l'utilisateur et le nœud de brouillard, et authentification mutuelle entre le nœud de brouillard et le nœud de capteur. Idéalement, le coût des fonctions d'encodeur/déchiffrement et de hachage devrait être calculé comme suit: $t_i = t_i \times 4/7$, où i= *h* ou *e/d*, donc le coût de calcul sera de 12 t_h + 4.57 $t_{e/d}$ au lieu de *21* t_h + *8* $t_{e/d}$. Le coût de FE n'est pas pris en compte car il n'est effectué que dans le nœud d'utilisateur. Le coût de communication diminue également à mesure que le nombre de messages diminue, de sorte que la longueur des messages de communication devient ≈ 1 262 bits. Sur la base de ces résultats, nous pouvons voir que le schéma proposé est efficace et rentable du point de vue des coûts de calcul et des coûts de communication.

D) **Résumé des résultats**

Le système crypto-biométrique proposé répond aux exigences de l'IoT, il offre une légèreté en termes de coûts de calcul et de communication. Le système répond aux exigences de sécurité en garantissant la confidentialité de la transmission, l'intégrité des données, la disponibilité du service, la récence du système, et l'authentification mutuelle entre tous les acteurs IoT. Les résultats ont été comparés à d'autres dans des travaux récents connexes. Cette comparaison a montré que notre schéma est léger et offre plus de sécurité que les autres. En incluant tous les acteurs possibles de l'IoT; notamment la notion de service public, nous pouvons dire que les résultats indiquent le schéma proposé est adapté au modèle S^2aaS basé sur l'IoT.

4.4 Conclusion

Les résultats empiriques obtenus pour chacune de nos solutions s'avèrent efficaces en termes de sécurité, de légèreté, et d'évolutivité, compte tenu des limites de l'Internet des Objets. Les solutions suggérées à tous ces niveaux incluent le partage sécurisé de clés secrètes à utiliser pour assurer les objectifs de sécurité requis.

Tableau 4.7- Comparaison des fonctions de sécurité et d'autres caractéristiques. (Source [Bentahar2021].

Caractéristique	[Jiang 2014]	[Choi 2016]	[Das 2015]	[Park 2016]	[Wang 2017]	[Moon 2017]	[Farash 2016]	[Wu 2016]	[Maurya 2017]	[Mishra 2017]	[Li 2018]	[Harbi 2019]	[Roy 2018]	[Gupta2 019]	la notre
F1	✓	×	✓	✓	×	✓	✓	×	✓	✓	✓	×	✓	×	✓
F2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
F3	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
F4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
F5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
F6	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓
F7	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	✓
F8	✓	×	✓	✓	×	✓	✓	×	✓	×	×	×	✓	✓	✓
F9	✓	✓	✓	✓	×	✓	✓	✓	✓	×	×	×	×	✓	✓
F10	×	✓	✓	×	×	✓	✓	×	×	×	×	✓	✓	×	✓
F11	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
F12	×	✓	✓	✓	✓	✓	×	×	✓	✓	✓	×	✓	×	✓
F13	×	×	×	×	×	×	×	×	×	×	×	×	×	×	✓
F14	×	×	×	×	×	×	×	×	×	×	×	×	✓	✓	✓
F15	×	×	×	×	×	×	×	×	×	×	×	×	✓	×	✓
F16	3	3	3	3	3	3	3	3	3	3	3	3	2	3	4

Tableau 4.8- Comparaison des coûts de calcul et de communication. (Source [Bentahar2021]).

Entité	[Das2015]	[Park2016]	[Wang2017]	[Moon2017]	[Maurya 2017]	[Mishra 2017]	[Li2018]	[Roy2018]	la notre
Comp Ui	$t_{fe}+ t_{e/d} +6\ t_h$	$2t_{fe} + 2t_{pm}+10t_h$	Non évalué	$t_{fe}+3t_{pm} +6\ t_h$	$3\ t_h$	$8\ t_h$	$3t_{pm} +8t_h$	$t_{fe}+ 2t_{ch}+9\ t_h$	$t_{fe}+ t_{e/d} +7\ t_h$
Comp S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	$t_{ch}+5\ t_h$	$3t_{e/d} +7\ t_h$
Comp FNk	$2\ t_{e/d} +3t_h$	$16t_h$	Non évalué	$t_{e/d}+t_{pm} +6\ t_h$	$t_{pm}+3t_h$	$10\ t_h$	$t_{pm}+7t_h$	N/A	$3t_{e/d} +4\ t_h$
Comp SNj	$t_{e/d} + 2t_h$	$2t_{pm}+4t_h$	Non évalué	$t_{e/d} +2t_{pm} +4\ t_h$	$t_{e/d}$	$5\ t_h$	$2t_{pm} +4t_h$	N/A	$t_{e/d} +3\ t_h$
Total *Comp*	$t_{fe}+4t_{e/d} +11t_h$	$2t_{fe} + 4t_{pm}+30t_h$	$4t_{pm}+18t_h$	$t_{fe}+2t_{e/d} +6t_{pm}+16\ t_h$	$t_{e/d} +t_{pm}+6t_h$	$23\ t_h$	$6t_{pm} +19t_h$	$t_{fe}+ 3t_{ch}+14\ t_h$	$t_{fe}+ 8t_{e/d} +21\ t_h$
Comm (bits)	832	Non évalué	3968	Non évalué	713	16384	2720	992	2208

Générale

Conclusion Générale

L'internet des objets semble devenir plus répandu, notamment avec le développement des transactions financières en ligne. Aujourd'hui, l'IoT se concentre sur les services publics plutôt que privés où il n'y a qu'une seule partie qui peut y accéder (*Sensing as a Service*). Il est maintenant possible de combiner ces services hétérogènes en un seul service unifié à large exploitation. Mais cette tendance a aussi rendu l'Internet des objets plus vulnérable aux cyberattaques car il n'y a pas de propriétaire unique des services. Dès lors, les efforts sont redoublés pour apporter des solutions de sécurité à ces nouveaux défis. D'un autre point de vue, comme l'être humain est partie de ces objets connectés, l'incorporation des technologies biométriques est devenue un moyen incontournable.

1. Rappel des objectifs de la thèse

La sécurité de l'IoT utilisant les technologies biométriques est abordée dans cette thèse. Afin d'atteindre le niveau de sécurité requis, les objectifs de sécurité à savoir, authentification mutuelle, confidentialité, intégrité, autorisation, non-répudiation, disponibilité, fraîcheur, confidentialité persistante, ont été atteints. Les solutions proposées respectent les contraintes IoT et surmonte ses limitations. Les solutions définissent clairement « Où et Comment » la technologie biométrique doit être inclue. Nous avons expliqué à travers les différentes sections que être humain peut interagir avec S^2aaS-IoTen tant qu'un consommateur, producteur et propriétaire de données.

Il s'est montré que la technologie biométrique peut être mise en œuvre dans tous les niveaux IoT. le niveau d'utilisateur, niveau du cloud et niveau de la perception. Les mécanismes de sécurité utilisés dans les trois niveaux, y compris les mécanismes biométriques, sont les mêmes. Les protocoles de sécurité IoT et les algorithmes légers sont mis en œuvre pour n'importe quel rôle humain. Les solutions proposées traitent les éléments suivants, IoT de bout-en-bout, l'accès au cloud et S^2aaS-IoT. Pour chaque élément, nous avons proposé une solution de sécurité qui assure à la fois l'authentification mutuelle et l'échange de clés.

Les lacunes communes extraites des travaux connexes se résument en, tous les acteurs ne sont pas abordés, l'utilisation des cartes à puces, les solutions sont lourdes et coûteuses, le Cloud et les canaux sont considérés sécuritaires et les solutions sont conçues soit pour la sauvegarde soit pour la transmission. Afin de combler ces lacunes nous avons proposé, une solution globale qui inclut tous les acteurs IoT possibles, une nouvelle application crypto-systèmes pour sécuriser la

transmission et la sauvegarde ensemble, des crypto-systèmes associés à un protocole d'échange de clé efficace pour économiser du temps et de l'énergie, une amélioration pour une authentification plus légère et plus précise en utilisant une quantification moyenne des blocs non linéaires. Notons que les processus lourds sont évités dans les intervalles critiques et fréquents. Nous avons également attribué les tâches de la carte à puce aux terminaux des utilisateurs qui peuvent grâce à nos solutions accéder rapidement et en toute sécurité aux services publics via un simple site Web. Pour donner plus de crédibilité à nos solutions, le Cloud est traité comme un nœud non fiable, et tous les canaux de transmission ne sont pas sécurisés, pour assimiler un environnement S^2aaS réel.

2. Bilan de la thèse

Nos solutions reposent conjointement sur les techniques d'extraction biométrique DCT, DWT et PCA. Sur le plan de cryptographie, nous avons utilisé la fonction de hachage, les techniques de cryptage symétrique/asymétrique en particulier ECC/ ECDH, et Les crypto-systèmes biométriques *Fuzzy Commitment*, *Fuzzy Vault*, *Fussy Extractor*. Il est à noter que les codes correcteurs adoptés dans nos solutions sont *Extended Hamming* et *Reed-Solomon* pour plus de tolérance convenable à la biométrie.

Les solutions sont à faible coût de calcule en tenant en compte le niveau de sécurité requis. Cela grâce à la légèreté des protocoles et des algorithmes, au petit vecteur de caractéristiques biométrique, et à la défense collaborative.

Pour valider les solutions proposées, nous avons fait une analyse informelle pour montrer comment nos systèmes peuvent résister aux diverses attaques de transmission et du sauvegarde sans affecter l'agilité et la précision, une analyse de coût d'énergie et de coût communication, une analyse formelle en utilisant l'outil AVISPA pour évaluer le modèle S^2aaS-IoT, une validation des sous-systèmes biométriques en termes de taux de reconnaissance conventionnels ainsi que des nouveaux taux et en termes du temps de calcul, et finalement une comparaison avec autres travaux récents en termes de la sécurité, des coûts de communication, des coûts de calcul, et autres métriques pertinentes.

Les solutions proposées répondent aux exigences de l'IoT en offrant des schémas simples et efficaces. Elles répondent également aux exigences de la sécurité en garantissant le partage sécuritaire des clés, la confidentialité de la transmission, l'intégrité des données, la disponibilité du service, la récence du système, et l'authentification mutuelle entre tous les acteurs IoT. Les sous-systèmes biométriques se caractérisent par un taux de reconnaissance acceptable et un haut

niveau de sécurité. La comparaison avec autres travaux connexes indique que nos systèmes sont légers, plus précis et offre plus de protection.

Des résultats supplémentaires peuvent être aussi mentionnés, nous avons déduit que le système non crypté basé sur la représentation binaire est plus lent que celui basé sur la représentation non binaire, mais il est légèrement plus précis. *Fuzzy Commitment* avec code orienté bit est le plus léger en termes du temps de calcule tandis que le *Fuzzy Vault* avec code orienté octet est la plus précise en termes de taux de reconnaissance. Un compromis entre la légèreté et la précision peut être achevé avec le *Fuzzy Vault* de code orienté bit. Dernier point est que le *Fuzzy Commitment* avec un code orienté bit est plus adapté à l'IoT grâce à sa rapidité qui n'influe pas sur la précision, effectivement la différence des taux de reconnaissance ne dépasse pas 4% par rapport aux autres crypto-systèmes analysés.

3. Perspectives de la recherche

Selon un communiqué de presse publié en janvier 2018 par ABI Research, la maturité des technologies de sécurité IoT est à la hausse dans les environnements industriels, les transports et l'automobile, les gouvernements et les services publics. Si vous suivez comment au cours des dernières années, il y a eu des cyberattaques vraiment massives utilisant des dispositifs compatibles IoT et découvrir à chaque fois de nouvelles cybermenaces dans l'espace IoT dans l'actualité. Il est très facile de s'attendre à ce que cela ne soit pas différent dans les années à venir. C'est probablement une évidence que nous pouvons nous attendre à plus de failles de sécurité et que l'industrie proposera plus d'initiatives de sécurité. Il existe un marché émergeant pour l'IoT de bout-en-bout ou la gestion de la sécurité IoT basée sur le cloud. En parallèle, il existe de nombreux défis qui s'appliquent certainement dans le contexte de l'IoT basé sur les services publics, l'IoT industriel, et l'industrie 4.0, où la sécurité fait partie intégrante de l'architecture. Des défis supplémentaires sous l'angle de la sécurité de l'IoT concernent les réglementations relatives à la protection des données personnelles. Et lorsque nous parlons des données personnelles, des données biométriques en font certainement partie.

Pour relever ce défi particulier, l'utilisation de l'intelligence artificielle pour la surveillance et le contrôle en temps réel doit être accrue, en s'appuyant sur l'apprentissage automatique pour la reconnaissance biométrique, mais toute en respectant les contraintes et limitations de l'IoT.

Autre défi présent dans l'amélioration des processus de reconnaissance et de sécurité qui nécessitent l'automatisation de la Qualité de Service (QoS). Par exemple, le changement automatique de la longueur de la clé et du seuil de reconnaissance peut être défini en fonction de la sécurité et de la précision souhaitées.

D'autre part, avec l'apparition de la *Blockchain* (technologie de registre distribué) dans la sécurité de l'IoT et une intégration continue de l'IoT et de la *Blockchain*, les technologies biométriques doivent être incluses de manière appropriée dans ce domaine, parce que c'est la technologie la plus efficace pour l'authentification humaine. L'amélioration dans le domaine de la sécurité peut également être obtenue en réduisant le coût de communication et le coût de calcul entre les différentes entités IoT, autrement dit, un schéma de sécurité moins coûteux et une authentification plus précise, mais plus sécurisée. Autre défi à rencontrer, qui est l'élargissement des technologies biométriques et IoT pour inclure toutes les applications y compris les moins importantes.

L'IoT n'est pas un exemple d'histoire de science-fiction futuriste. Des milliards d'objets et des infrastructures logicielles (*Frameworks*) faisant partie de l'IoT sont déployés aujourd'hui. C'était inimaginable, mais c'est juste arrivé. De même, des choses que nous considérons comme lointaines et impossibles peuvent se produire plus tôt que prévu. Certaines personnes seront heureuses du monde à venir et des choses avancées. D'autres regretteront le bon vieux temps où une table n'était en effet qu'une table.

Liste des Publications et Communications

- Bentahar, A., Meraoumia, A., Bradji, L., & Bendjenna, H. (2021). Sensing as a Service in Internet of Things: Efficient Authentication and key Agreement Scheme. In Journal of King Saud University - Computer and Information Sciences. Elsevier BV. https://doi.org/10.1016/j.jksuci.2021.06.007.
- Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2019). Biometric Cryptosystems: Towards a Light and Precise Remote Authentication. In Recent Advances in Computer Science and Communications (Vol. 13). Bentham Science Publishers Ltd. https://doi.org/10.2174/2666255813666191223115223.
- Bentahar, A., Meraoumia, A., Bendjenna, H., & Zeroual, A. (2018). IoT Securing System using Fuzzy Commitment for DCT-based Fingerprint Recognition. In 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS). IEEE. https://doi.org/10.1109/pais.2018.8598511.
- Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2018). Biometric Cryptosystem Scheme for Internet of Things using Fuzzy Commitment principle. In 2018 International Conference on Signal, Image, Vision and their Applications (SIVA). IEEE. https://doi.org/10.1109/siva.2018.8660993.
- Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2019).Securing Remote Authentication Using Fuzzy Commitment and Fuzzy Vault International Conference on Pattern Analysis and Recognition 2019-10 (ICPAR 2019). https://icpar2019.sciencesconf.org/.
- Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2020). Fuzzy Extractor-Based Key Agreement for Internet of Things. In 020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP). IEEE. https://doi.org/10.1109/ccssp49278.2020.9151574.
- Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2021). Eigen-Fingerprints-Based Remote Authentication Cryptosystem. In 2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI). IEEE. https://doi.org/10.1109/icrami52622.2021.9585979.
- Zeroual, A., Amroune, M., Derdour, M., & Bentahar, A. (2021). Lightweight deep learning model to secure authentication in Mobile Cloud Computing. In Journal of King Saud University - Computer and Information Sciences. Elsevier BV. https://doi.org/10.1016/j.jksuci.2021.09.016.
- Zeroual, A., Amroune, M., Derdour, M., Meraoumia, A., & Bentahar, A. (2018). Deep authentication model in Mobile Cloud Computing. In 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS). https://doi.org/10.1109/pais.2018.8598508.
- Zeroual, A., Derdour, M., Amroune, M., & Bentahar, A. (2019). Using a Fine-Tuning Method for a Deep Authentication in Mobile Cloud Computing Based on Tensorflow Lite Framework. In 2019 International Conference on Networking and Advanced Systems (ICNAS). https://doi.org/10.1109/icnas.2019.8807440.

Bibliographie

[Adamovic2016] Adamovic, S., Milosavljevic, M., Veinovic, M., Sarac, M., & Jevremovic, A. (2016). Fuzzy commitment scheme for generation of cryptographic keys based on iris biometrics. In IET Biometrics (Vol. 6, Issue 2, pp. 89–96). https://doi.org/10.1049/iet-bmt.2016.0061.

[Adler2009] Adler, A. (2009). Cancelable Biometrics. In Encyclopedia of Biometrics (pp. 175–178). Springer US. https://doi.org/10.1007/978-0-387-73003-5_66.

[Ahmad2013] Ahmad Salehi, S., Razzaque, M. A., Naraei, P., & Farrokhtala, A. (2013). Detection of sinkhole attack in wireless sensor networks. In 2013 IEEE International Conference on Space Science and Communication (IconSpace). https://doi.org/10.1109/iconspace.2013.6599496.

[Ahmed1974] Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete Cosine Transform. In IEEE Transactions on Computers: Vol. C–23 (Issue 1, pp. 90–93). https://doi.org/10.1109/t-c.1974.223784.

[Akansu2010] Akansu, A. N., Serdijn, W. A., & Selesnick, I. W. (2010). Emerging applications of wavelets: A review. In Physical Communication (Vol. 3, Issue 1, pp. 1–18). Elsevier BV. https://doi.org/10.1016/j.phycom.2009.07.001.

[Al-Turjman2020] Trends in Cloud-based IoT. (2020). In F. Al-Turjman (Ed.), EAI/Springer Innovations in Communication and Computing. Springer International Publishing. https://doi.org/10.1007/978-3-030-40037-8.

[Ao2009] Ao, M., & Li, S. Z. (2009). Near Infrared Face Based Biometric Key Binding. In Advances in Biometrics (pp. 376–385). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-01793-3_39.

[Armando2006] Armando, A., Basin, D., Boichut, Y., Chevalier, Y., Compagna, L., Cuellar, J., Drielsma, P. H., Heám, P. C., Kouchnarenko, O., Mantovani, J., Mödersheim, S., von Oheimb, D., Rusinowitch, M., Santiago, J., Turuani, M., Viganò, L., & Vigneron, L. (2005). The AVISPA Tool for the Automated Validation of Internet Security Protocols and Applications. In Computer Aided Verification (pp. 281–285). Springer Berlin Heidelberg. https://doi.org/10.1007/11513988_27.

[AVISPA] AVISPA Automated Validation of Internet Security Protocols and Applications, AVISPA Project.(2006). (Online). Accessed, december 2019. Available, http://www.avispa-project.org/.

[Badrinath2012] Badrinath, G. S., Tiwari, K., & Gupta, P. (2012). An Efficient Palmprint Based Recognition System Using 1D-DCT Features. In Lecture Notes in Computer Science (pp. 594–601). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31588-6_76.

[Baghel2021] Baghel, V. S., Prakash, S., & Agrawal, I. (2021). An enhanced fuzzy vault to secure the fingerprint templates. In Multimedia Tools and Applications (Vol. 80, Issues 21–23, pp. 33055–33073). Springer Science and Business Media LLC. https://doi.org/10.1007/s11042-021-11325-w.

[Barbuddhe2020] Barbuddhe, V., Zanjat, S. N., & Karmore, B. S. (2020). Information theory, coding and cryptography. LAP Lambert Academic Publishing.

[Barker2017] Barker, E., & Mouha, N. (2017). Recommendation for the Triple Data Encryption Algorithm (TDEA) block cipher. National Institute of Standards and Technology. https://doi.org/10.6028/nist.sp.800-67r2.

[Barker2020] Barker, E. (2020). Recommendation for key management: National Institute of Standards and Technology. https://doi.org/10.6028/nist.sp.800-57pt1r5.

[Barrett2010] Barrett, D., & Kipper, G. (2010). Visions of the Future. In Virtualization and Forensics (pp. 211–220). Elsevier. https://doi.org/10.1016/b978-1-59749-557-8.00011-4.

[Beaver2016] Beaver, K. (2016). Hacking For Dummies (5th ed.). John Wiley & Sons. ISBN 978-1-119-15468-6.

[Bentahar2018] Bentahar, A., Meraoumia, A., Bendjenna, H., & Zeroual, A. (2018). IoT Securing System using Fuzzy Commitment for DCT-based Fingerprint Recognition. In 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS). IEEE. https://doi.org/10.1109/pais.2018.8598511.

[Bentahar2018a] Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2018). Biometric Cryptosystem Scheme for Internet of Things using Fuzzy Commitment principle. In 2018 International Conference on Signal, Image, Vision and their Applications (SIVA). IEEE. https://doi.org/10.1109/siva.2018.8660993.

[Bentahar2019] Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2019). Biometric Cryptosystems: Towards a Light and Precise Remote Authentication. In Recent Advances in Computer Science and Communications (Vol. 13). Bentham Science Publishers Ltd. https://doi.org/10.2174/2666255813666191223115223.

[Bentahar2019a] Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2019).Securing Remote Authentication Using Fuzzy Commitment and Fuzzy Vault International Conference on Pattern Analysis and Recognition 2019-10 (ICPAR 2019). https://icpar2019.sciencesconf.org/.

[Bentahar2020] Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2020). Fuzzy Extractor-Based Key Agreement for Internet of Things. In 020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP). IEEE. https://doi.org/10.1109/ccssp49278.2020.9151574.

[Bentahar2021] Bentahar, A., Meraoumia, A., Bradji, L., & Bendjenna, H. (2021). Sensing as a Service in Internet of Things: Efficient Authentication and key Agreement Scheme. In Journal of King Saud University - Computer and Information Sciences. Elsevier BV. https://doi.org/10.1016/j.jksuci.2021.06.007.

[Bentahar2021a] Bentahar, A., Meraoumia, A., Bendjenna, H., Chitroub, S., & Zeroual, A. (2021). Eigen-Fingerprints-Based Remote Authentication Cryptosystem. International Conference on Recent Advances in Mathematics and Informatics (ICRAMI 2021). IEEE. Tebessa, Algeria, September 21-22, 2021 http://icrami.rf.gd/?i=1.

[Berlekamp-Massey] Berlekamp-Massey algorithm. Hazewinkel, M. (n.d.). Encyclopaedia of Mathematics. CWI. URL: http://encyclopediaofmath.org/index.php?title=Berlekamp-Massey_algorithm&oldid=50141.

[Britannica2021] Britannica, T. Editors of Encyclopaedia (2021, April 25). Chaos theory. Encyclopedia Britannica. https://www.britannica.com/science/chaos-theory.

[Cerf1974] Cerf, V., & Kahn, R. (1974). A Protocol for Packet Network Intercommunication. In IEEE Transactions on Communications (Vol. 22, Issue 5, pp. 637–648). https://doi.org/10.1109/tcom.1974.1092259.

[Cervantes2015] Cervantes, C., Poplade, D., Nogueira, M., & Santos, A. (2015). Detection of sinkhole attacks for supporting secure routing on 6LoWPAN for Internet of Things. In 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM). https://doi.org/10.1109/inm.2015.7140344.

[Chang2021] Chang, D., Garg, S., Ghosh, M., & Hasan, M. (2021). BIOFUSE: A framework for multi-biometric fusion on biocryptosystem level. In Information Sciences (Vol. 546, pp. 481–511). Elsevier BV. https://doi.org/10.1016/j.ins.2020.08.065.

[Chen2004] Chen, S., & Zhu, Y. (2004). Subpattern-based principle component analysis. In Pattern Recognition (Vol. 37, Issue 5, pp. 1081–1083). Elsevier BV. https://doi.org/10.1016/j.patcog.2003.09.004.

[Chlaoua2018] Chlaoua, R., Meraoumia, A., Aiadi, K. E., & Korichi, M. (2018). Deep learning for finger-knuckle-print identification system based on PCANet and SVM classifier. In Evolving Systems (Vol. 10, Issue 2, pp. 261–272). Springer Science and Business Media LLC. https://doi.org/10.1007/s12530-018-9227-y.

[Choi2016] Choi, Y., Lee, Y., & Won, D. (2016). Security Improvement on Biometric Based Authentication Scheme for Wireless Sensor Networks Using Fuzzy Extraction. In International Journal of Distributed Sensor Networks (Vol. 12, Issue 1, p. 8572410). SAGE Publications. https://doi.org/10.1155/2016/8572410.

[Clancy2003] Clancy, T. C., Kiyavash, N., & Lin, D. J. (2003). Secure smartcard-based fingerprint authentication. In Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications - WBMA '03. https://doi.org/10.1145/982507.982516.

[Dabbagh2016] Dabbagh, M., & Rayes, A. (2016). Internet of Things Security and Privacy. In Internet of Things From Hype to Reality (pp. 195–223). Springer International Publishing. https://doi.org/10.1007/978-3-319-44860-2_8.

[Das2015] Das, A. K. (2015). A secure and effective biometric-based user authentication scheme for wireless sensor networks using smart card and fuzzy extractor. In International Journal of Communication Systems (Vol. 30, Issue 1, p. e2933). Wiley. https://doi.org/10.1002/dac.2933.

[Dasgupta2017] Dasgupta, D., Roy, A., & Nag, A. (2017). Biometric Authentication. In Infosys Science Foundation Series (pp. 37–84). Springer International Publishing. https://doi.org/10.1007/978-3-319-58808-7_2.

[David2003] David Zhang, Wai-Kin Kong, Jane You, & Michael Wong. (2003). Online palmprint identification. In IEEE Transactions on Pattern Analysis and Machine Intelligence (Vol. 25, Issue 9, pp. 1041–1050). https://doi.org/10.1109/tpami.2003.1227981.

[Diffie1976] Diffie, W., & Hellman, M. E. (1976). Multiuser cryptographic techniques. In Proceedings of the June 7-10, 1976, national computer conference and exposition on - AFIPS '76. https://doi.org/10.1145/1499799.1499815.

[Dodis2004] Dodis, Y., Reyzin, L., & Smith, A. (2004). Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. In Advances in Cryptology - EUROCRYPT 2004 (pp. 523–540). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-24676-3_31.

[Elrefaei2019] Elrefaei, L. A., & Al-Mohammadi, A. M. (2019). Machine vision gait-based biometric cryptosystem using a fuzzy commitment scheme. In Journal of King Saud University - Computer and Information Sciences. Elsevier BV. https://doi.org/10.1016/j.jksuci.2019.10.011.

[Face-Bases] FEI Face Database. Accessed Jun. 2018, From https://fei.edu.br/~cet/facedatabase.html.

[Farash2016] Farash, M. S., Turkanović, M., Kumari, S., & Hölbl, M. (2016). An efficient user authentication and key agreement scheme for heterogeneous wireless sensor network tailored for the Internet of Things environment. In Ad Hoc Networks (Vol. 36, pp. 152–176). Elsevier BV. https://doi.org/10.1016/j.adhoc.2015.05.014.

[Faundez2005] Faundez-Zanuy, M. (2005). Data fusion in biometrics. In IEEE Aerospace and Electronic Systems Magazine (Vol. 20, Issue 1, pp. 34–38). https://doi.org/10.1109/maes.2005.1396793.

[Fingerprint-Bases] The Hong Kong polytechnic contactless 2D to contact-based 2D fingerprint images database version 1.0. (2017). Edu.Hk. Accessed Nov. 2017, from http://www4.comp.polyu.edu.hk/~csajaykr/fingerprint.htm.

[FVC2000-Base] Fingerprint verification competition (FVC2000). (2000). Unibo.It. Accessed Jan, 2018, from http://bias.csr.unibo.it/fvc2000/download.asp.

[Ghayoumi2015] Ghayoumi, M. (2015). A review of multimodal biometric systems: Fusion methods and their applications. In 2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS). https://doi.org/10.1109/icis.2015.7166582.

[Ghorbani 2017] Ghorbani, H. R., & Ahmadzadegan, M. H. (2017). Security challenges in internet of things: survey. In 2017 IEEE Conference on Wireless Sensors (ICWiSe). https://doi.org/10.1109/icwise.2017.8267153.

[Global2021] Global Biometric Systems Market (2021-2026) By Technology, Functionality, Component, Authentication, End Users, Geography and the Impact of Covid-19 with Ansoff Analysis. (2021). Infogence Global Research. Report April 2021. 170 pages. ID: 5317229. URL://Global Biometric Systems Market.

[Gupta2019] Gupta, A., Tripathi, M., Shaikh, T. J., & Sharma, A. (2019). A lightweight anonymous user authentication and key establishment scheme for wearable devices. In Computer Networks (Vol. 149, pp. 29–42). Elsevier BV. https://doi.org/10.1016/j.comnet.2018.11.021.

[Harbi2019] Harbi, Y., Aliouat, Z., Refoufi, A., Harous, S., & Bentaleb, A. (2019). Enhanced authentication and key management scheme for securing data transmission in the internet of things. In Ad Hoc Networks (Vol. 94, p. 101948). Elsevier BV. https://doi.org/10.1016/j.adhoc.2019.101948.

[Hazewinkel2002] Hazewinkel, M. (Ed.). (2002). Encyclopaedia of Mathematics. Springer. ISBN 1402006098. http://encyclopediaofmath/Lagrange_interpolation_formula.

[Hjelmås2001] Hjelmås, E., & Low, B. K. (2001). Face Detection: A Survey. In Computer Vision and Image Understanding (Vol. 83, Issue 3, pp. 236–274). Elsevier BV. https://doi.org/10.1006/cviu.2001.0921.

[Hossain2019] Hossain, M. T., Teng, S. W., Zhang, D., Lim, S., & Lu, G. (2019). Distortion Robust Image Classification Using Deep Convolutional Neural Network with Discrete Cosine Transform. In 2019 IEEE International Conference on Image Processing (ICIP). https://doi.org/10.1109/icip.2019.8803787.

[Huang2015] Huang, Z.-H., Li, W.-J., Shang, J., Wang, J., & Zhang, T. (2015). Non-uniform patch based face recognition via 2D-DWT. In Image and Vision Computing (Vol. 37, pp. 12–19). Elsevier BV. https://doi.org/10.1016/j.imavis.2014.12.005.

[Huang2017] Huang, B., & Elsevier. (2017). Comprehensive GIS : Geographic Information System. Elsevier. ISBN 978-0-12-804793-4.

[Jain2004] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An Introduction to Biometric Recognition. In IEEE Transactions on Circuits and Systems for Video Technology (Vol. 14, Issue 1, pp. 4–20). https://doi.org/10.1109/tcsvt.2003.818349.

[Jain2005] Jain, A. K., Ross, A., & Uludag, U. (2005). Biometric template security: Challenges and solutions. 2005 13th European Signal Processing Conference, (pp. 1-4). https://ieeexplore-ieee-org.sndl1.arn.dz/document/7078369.

[Jain2008] Jain, A. K., Nandakumar, K., & Nagar, A. (2008). Biometric Template Security. In EURASIP Journal on Advances in Signal Processing (Vol. 2008, Issue 1, p. 579416). Springer Science and Business Media LLC. https://doi.org/10.1155/2008/579416.

[Jiang2014] Jiang, Q., Ma, J., Lu, X., & Tian, Y. (2014). An efficient two-factor user authentication scheme with unlinkability for wireless sensor networks. In Peer-to-Peer Networking and Applications (Vol. 8, Issue 6, pp. 1070–1081). Springer Science and Business Media LLC. https://doi.org/10.1007/s12083-014-0285-z.

[Jing2004] Jing, X.-Y., & Zhang, D. (2004). A Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction. In IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics) (Vol. 34, Issue 6, pp. 2405–2415). https://doi.org/10.1109/tsmcb.2004.837586.

[Juels1999] Juels, A., & Wattenberg, M. (1999). A fuzzy commitment scheme. In Proceedings of the 6th ACM conference on Computer and communications security - CCS '99. https://doi.org/10.1145/319709.319714.

[Juels2006] Juels, A., & Sudan, M. (2006). A Fuzzy Vault Scheme. In Designs, Codes and Cryptography (Vol. 38, Issue 2, pp. 237–257). Springer Science and Business Media LLC. https://doi.org/10.1007/s10623-005-6343-z.

[Kaur2017] Kaur, T., & Kaur, M. (2017). Cryptographic key generation from multimodal template using fuzzy extractor. In 2017 Tenth International Conference on Contemporary Computing (IC3). IEEE. https://doi.org/10.1109/ic3.2017.8284321.

[Kausar2021] Kausar, F. (2021). Iris based cancelable biometric cryptosystem for secure healthcare smart card. In Egyptian Informatics Journal. Elsevier BV. https://doi.org/10.1016/j.eij.2021.01.004.

[Ker2007] Ker, A. D. (2007). Steganalysis of Embedding in Two Least-Significant Bits. In IEEE Transactions on Information Forensics and Security (Vol. 2, Issue 1, pp. 46–54). https://doi.org/10.1109/tifs.2006.890519.

[Khorov2015] Khorov, E., Lyakhov, A., Krotov, A., & Guschin, A. (2015). A survey on IEEE 802.11ah: An enabling networking technology for smart cities. In Computer Communications (Vol. 58, pp. 53–69). Elsevier BV. https://doi.org/10.1016/j.comcom.2014.08.008.

[Kilinc2014] Kilinc, H. H., & Yanik, T. (2014). A Survey of SIP Authentication and Key Agreement Schemes. In IEEE Communications Surveys & Tutorials (Vol. 16, Issue 2, pp. 1005–1023). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/surv.2013.091513.00050.

[Kim2016] Kim, H.-J., Chang, H.-S., Suh, J.-J., & Shon, T. (2016). A Study on Device Security in IoT Convergence. In 2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA). IEEE. https://doi.org/10.1109/icimsa.2016.7503989.

[Kim2017] Kim, P. (2017). Convolutional Neural Network. In MATLAB Deep Learning (pp. 121–147). Apress. https://doi.org/10.1007/978-1-4842-2845-6_6.

[Koblitz1987] Koblitz, N. (1987). Elliptic curve cryptosystems. In Mathematics of Computation (Vol. 48, Issue 177, pp. 203–203). American Mathematical Society (AMS). https://doi.org/10.1090/s0025-5718-1987-0866109-5.

[Kocher1999] Kocher, P., Jaffe, J., & Jun, B. (1999). Differential Power Analysis. In Advances in Cryptology — CRYPTO' 99 (pp. 388–397). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-48405-1_25.

[Kumar2009] Kumar, A. (2009). Fusion, Rank-Level. In Encyclopedia of Biometrics (pp. 607–611). Springer US. https://doi.org/10.1007/978-0-387-73003-5_159.

[Kumar2009a] Kumar, A., & Kumar, A. (2009). Development of a new cryptographic construct using palmprint-based fuzzy vault. EURASIP Journal on Advances in Signal Processing, 2009(1). https://doi.org/10.1155/2009/967046.

[Kumar2017] Kumar, N. (2017). IoT architecture and system design for healthcare systems. In 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon). IEEE. https://doi.org/10.1109/smarttechcon.2017.8358543.

[Lee1999] Lee, C.-J., & Wang, S.-D. (1999). Fingerprint feature extraction using Gabor filters. In Electronics Letters (Vol. 35, Issue 4, p. 288). Institution of Engineering and Technology (IET). https://doi.org/10.1049/el:19990213.

[Lee2007] Lee, S.-W., & Li, S. Z. (2007). Advances in Biometrics. Springer-Verlag Berlin Heidelberg. ISBN 978-3-540-74549-5.

[Lee2009] Lee, V. M. (2009). Fraud Reduction, Overview. In Encyclopedia of Biometrics (pp. 584–592). Springer US. https://doi.org/10.1007/978-0-387-73003-5_25.

[Lee2014] Lee, J., Kao, H.-A., & Yang, S. (2014). Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. In Procedia CIRP (Vol. 16, pp. 3–8). Elsevier BV. https://doi.org/10.1016/j.procir.2014.02.001.

[Leiner2009] Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., & Wolff, S. (2009). A brief history of the internet. In ACM SIGCOMM Computer Communication Review (Vol. 39, Issue 5, pp. 22–31). Association for Computing Machinery (ACM). https://doi.org/10.1145/1629607.1629613.

[Li2002] Li, W., Zhang, D., & Xu, Z. (2002). Palmprint Identification by Fourier Transform. In International Journal of Pattern Recognition and Artificial Intelligence (Vol. 16, Issue 04, pp. 417–432). World Scientific Pub Co Pte Lt. https://doi.org/10.1142/s0218001402001757.

[Li2018] Li, X., Niu, J., Bhuiyan, M. Z. A., Wu, F., Karuppiah, M., & Kumari, S. (2018). A Robust ECC-Based Provable Secure Authentication Protocol With Privacy Preserving for Industrial Internet of Things. In IEEE Transactions on Industrial Informatics (Vol. 14, Issue 8, pp. 3599–3609). https://doi.org/10.1109/tii.2017.2773666.

[Lu2009] Lu, H., Martin, K., Bui, F., Plataniotis, K. N., & Hatzinakos, D. (2009). Face recognition with biometric encryption for privacy-enhancing self-exclusion. In 2009 16th International Conference on Digital Signal Processing (DSP). https://doi.org/10.1109/icdsp.2009.5201257.

[Maltoni2009] Maltoni, D., Maio, D., Jain, A. K., & Prabhakar, S. (2009). Handbook of fingerprint recognition (2nd ed.). Springer. https://doi.org/10.1007/978-1-84882-254-2.

[Maurya2017] Maurya, A., & Sastry, V. N. (2017). Fuzzy Extractor and Elliptic Curve Based Efficient User Authentication Protocol for Wireless Sensor Networks and Internet of Things. In Information (Vol. 8, Issue 4, p. 136). MDPI AG. https://doi.org/10.3390/info8040136.

[McBride2005] McBride, B. T., Peterson, G. L., & Gustafson, S. C. (2005). A new blind method for detecting novel steganography. In Digital Investigation (Vol. 2, Issue 1, pp. 50–70). Elsevier BV. https://doi.org/10.1016/j.diin.2005.01.003.

[Meraoumia2013] Meraoumia, A., Chitroub, S., & Bouridane, A. (2013). An Efficient Hand-Based Biometric Recognition System Using Finger- Knuckle-Print Data. In Recent Patents on Telecommunication (Vol. 1, Issue 2, pp. 151–162). Bentham Science Publishers Ltd. https://doi.org/10.2174/2211740711201020007.

[Meraoumia2014] Meraoumia, A., Chitroub, S., & Bouridane, A. (2014). An Efficient 2D and 3D Palmprint Identification System by Jointly Using Gabor Filter Response, Wavelet Transform and Radial Basis Function. In Recent Patents on Signal Processing (Vol. 4, Issue 1, pp. 18–31). Bentham Science Publishers Ltd. https://doi.org/10.2174/2210686304666140324184046.

[Miller1986] Miller, V. S. (1986). Use of Elliptic Curves in Cryptography. In Lecture Notes in Computer Science (pp. 417–426). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-39799-x_31.

[Mishra2017] Mishra, D., Vijayakumar, P., Sureshkumar, V., Amin, R., Islam, S. H., & Gope, P. (2017). Efficient authentication protocol for secure multimedia communications in IoT-enabled wireless sensor networks. In Multimedia Tools and Applications (Vol. 77, Issue 14, pp. 18295–18325). Springer Science and Business Media LLC. https://doi.org/10.1007/s11042-017-5376-4.

[Mohammadi2015] Mohammadi, S., & Hariri, M. (2015). New approaches to fingerprint authentication using software methods based on fingerprint texture. In 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI). IEEE. https://doi.org/10.1109/kbei.2015.7436198.

[Moon2017] Moon, J., Lee, D., Lee, Y., & Won, D. (2017). Improving Biometric-Based Authentication Schemes with Smart Card Revocation/Reissue for Wireless Sensor Networks. In Sensors (Vol. 17, Issue 5, p. 940). MDPI AG. https://doi.org/10.3390/s17050940.

[Mordini2012] Mordini, E., Tzovaras, D., & Ashton, H. (2012). Introduction. In The International Library of Ethics, Law and Technology (pp. 1–19). Springer Netherlands. https://doi.org/10.1007/978-94-007-3892-8_1.

[Nandakumar2005] Nandakumar, K. (2005). Integration of Multiple Cues in Biometric Systems. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.84.1779.

[Nandakumar2007] Nandakumar, K., Jain, A. K., & Pankanti, S. (2007). Fingerprint-Based Fuzzy Vault: Implementation and Performance. In IEEE Transactions on Information Forensics and Security (Vol. 2, Issue 4, pp. 744–757). https://doi.org/10.1109/tifs.2007.908165.

[Nandakumar2008] Nandakumar, K. (2008). Multibiometric systems: fusion strategies and template security. PhD. Dissertation. Michigan State University, USA. Advisor(s) Anil K. Jain. Order Number: AAI3312725. https://dl.acm.org/doi/10.5555/1467970.

[Nandakumar2010] Nandakumar, K. (2010). A fingerprint cryptosystem based on minutiae phase spectrum. In 2010 IEEE International Workshop on Information Forensics and Security (WIFS). https://doi.org/10.1109/wifs.2010.5711456.

[Nandakumar2015] Nandakumar, K., & Jain, A. K. (2015). Biometric Template Protection: Bridging the performance gap between theory and practice. In IEEE Signal Processing Magazine (Vol. 32, Issue 5, pp. 88–100). https://doi.org/10.1109/msp.2015.2427849.

[Nesse2013] Nesse, P. J., Svaet, S. W., Strasunskas, D., & Gaivoronski, A. A. (2013). Assessment and optimisation of business opportunities for telecom operators in the cloud value network. In Transactions on Emerging Telecommunications Technologies (Vol. 24, Issue 5, pp. 503–516). Wiley. https://doi.org/10.1002/ett.2666.

[NIST-Base] Biometric Special Databases and Software. (2010). Nist.Gov. Accessed Jan, 2018, from https://www.nist.gov/itl/iad/image-group/resources/biometric-special-databases-and-software.

[Nixon2014] Nixon, M., & Aguado, A. S. (2014). Feature Extraction and Image Processing. Newnes. https://doi.org/10.1016/c2009-0-25049-5.

[Obaidat 2019] Obaidat, M. S., Traore, I., & Woungang, I. (Eds.). (2019). Biometric-Based Physical and Cybersecurity Systems. Springer International Publishing. https://doi.org/10.1007/978-3-319-98734-7.

[Palmprint-Bases] The Hong Kong Polytechnic Multispectral Palmprint Database. Accessed 2011, From https://www4.comp.polyu.edu.hk/~biometrics/MultispectralPalmprint/MSP.htm.

[Park2012] Park, J., Seong, D., Yeo, M., Lee, B., & Yoo, J. (2012). An Energy-Efficient Selective Forwarding Attack Detection Scheme Using Lazy Detection in Wireless Sensor Networks. In Lecture Notes in Electrical Engineering (pp. 157–164). Springer Netherlands. https://doi.org/10.1007/978-94-007-5857-5_17.

[Park2016] Park, Y., & Park, Y. (2016). Three-Factor User Authentication and Key Agreement Using Elliptic Curve Cryptosystem in Wireless Sensor Networks. In Sensors (Vol. 16, Issue 12, p. 2123). MDPI AG. https://doi.org/10.3390/s16122123.

[Pearsall2002] Pearsall, J., & Trumble, B. (Eds.). (2002). Oxford English reference dictionary (2nd ed.). Oxford University Press.

[Perera2013] Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2013). Sensing as a service model for smart cities supported by Internet of Things. In Transactions on Emerging Telecommunications Technologies (Vol. 25, Issue 1, pp. 81–93). Wiley. https://doi.org/10.1002/ett.2704.

[Perera2019] Perera, C., Bouguettaya, A., Kanhere, S., & Liu, C. H. (2019). Guest Editorial: Introduction to the Special Section on Sensor Data Computing as a Service in Internet of Things. In IEEE Transactions on Emerging Topics in Computing (Vol. 7, Issue 2, pp. 311–313). https://doi.org/10.1109/tetc.2019.2905089.

[Perez2011] Perez-Botero, D. (2011). A Brief Tutorial on Live Virtual Machine Migration From a Security Perspective.

[Rabbani2002] Rabbani, M., & Joshi, R. (2002). An overview of the JPEG 2000 still image compression standard. In Signal Processing: Image Communication (Vol. 17, Issue 1, pp. 3–48). Elsevier BV. https://doi.org/10.1016/s0923-5965(01)00024-8.

[Rayes2016] Rayes, A., & Salam, S. (2016). The Things in IoT: Sensors and Actuators. In Internet of Things From Hype to Reality (pp. 57–77). Springer International Publishing. https://doi.org/10.1007/978-3-319-44860-2_3.

[Rayes2016a] Rayes, A., & Salam, S. (2016). IoT Protocol Stack: A Layered View. In Internet of Things From Hype to Reality (pp. 93–138). Springer International Publishing. https://doi.org/10.1007/978-3-319-44860-2_5.

[Rayes2016b] Rayes, A., & Salam, S. (2016). IoT Services Platform: Functions and Requirements. In Internet of Things From Hype to Reality (pp. 165–194). Springer International Publishing. https://doi.org/10.1007/978-3-319-44860-2_7.

[Rayes2016c] Rayes, A., & Salam, S. (2016). Fog Computing. In Internet of Things From Hype to Reality (pp. 139–164). Springer International Publishing. https://doi.org/10.1007/978-3-319-44860-2_6

[Riaz2018] Riaz, N., Riaz, A., & Khan, S. A. (2018). Biometric template security: an overview. In Sensor Review (Vol. 38, Issue 1, pp. 120–127). Emerald. https://doi.org/10.1108/sr-07-2017-0131.

[Rivest1978] Rivest, R. L., Shamir, A., & Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. In Communications of the ACM (Vol. 21, Issue 2, pp. 120–126). Association for Computing Machinery (ACM). https://doi.org/10.1145/359340.359342.

[Ross2006] Ross, A. A., Nandakumar, K., & Jain, A. K. (2006). Handbook of Multibiometrics. Springer. ISBN 978-038-7331-232.

[Roy2018] Roy, S., Chatterjee, S., Das, A. K., Chattopadhyay, S., Kumari, S., & Jo, M. (2018). Chaotic Map-Based Anonymous User Authentication Scheme With User Biometrics and Fuzzy Extractor for Crowdsourcing Internet of Things. In IEEE Internet of Things Journal (Vol. 5, Issue 4, pp. 2884–2895). https://doi.org/10.1109/jiot.2017.2714179.

[Savio2011] Savio, J. (1 April 2011). Browsing history: A heritage site has been set up in Boelter Hall 3420, the room the first Internet message originated in. Daily Bruin. UCLA. Retrieved 6 June 2020. https://dailybruin.com/2011/04/01/browsing_history.

[Sen2002] Sen Wang, Wei Wei Zhang, & Yang Sheng Wang. (2002). Fingerprint classification by directional fields. In Fourth IEEE International Conference on Multimodal Interfaces. https://doi.org/10.1109/icmi.2002.1167027.

[Shukla2021] Shukla, S., & Patel, S. J. (2021). Securing fingerprint templates by enhanced minutiae-based encoding scheme in Fuzzy Commitment. In IET Information Security (Vol. 15, Issue 3, pp. 256–266). Institution of Engineering and Technology (IET). https://doi.org/10.1049/ise2.12024.

[Silverman2009] Silverman, J. H. (2009). The arithmetic of elliptic curves (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-09494-6.

[SPAN] SPAN - security protocol animator for AVISPA. (September 2017). Irisa.Fr. Accessed december 2019, from http://people.irisa.fr/Thomas.Genet/span/.

[Stallings2014] Stallings, W. (2014). Cryptography and network security: Principles and practice, international edition. Pearson. ISBN 9780133354690.

[Statista2020] Statista Research Department, IoT: number of connected devices worldwide 2012–2025. Technology & Telecommunications- Consumer Electronics Statista, [Online]. Available: https://www.statista.com/statistics/471264/ iot-number-of-connected-devices-worldwide/. [Accessed September 2020].

[Tachaphetpiboon2006] Tachaphetpiboon, S., & Amornraksa, T. (2006). Applying FFT Features for Fingerprint Matching. In 2006 1st International Symposium on Wireless Pervasive Computing. 2006 1st International Symposium on Wireless Pervasive Computing. IEEE. https://doi.org/10.1109/iswpc.2006.1613625.

[Teicher2018] Teicher, J. (February 7 2018).The little-known story of the first IoT device In the '80s. https://www.ibm.com/blogs/industries/little-known-story-first-iot-device/.

[Teoh2007] Teoh, A. B. J., & Kim, J. (2007). Secure biometric template protection in fuzzy commitment scheme. In IEICE Electronics Express (Vol. 4, Issue 23, pp. 724–730). Institute of Electronics, Information and Communications Engineers (IEICE). https://doi.org/10.1587/elex.4.724.

[Tewari2014] Tewari, K., & Kalakoti, R. L. (2014). Fingerprint Recognition and feature extraction using transform domain techniques. In 2014 International Conference on Advances in Communication and Computing Technologies (ICACACT 2014). https://doi.org/10.1109/eic.2015.7230719.

[Theodoridis2009] Theodoridis, S., & Koutroumbas, K. (2009). Feature Generation I: Data Transformation and Dimensionality Reduction. In Pattern Recognition (pp. 323–409). Elsevier. https://doi.org/10.1016/b978-1-59749-272-0.50008-6.

[Uludag2006] Uludag, U., & Anil Jain. (2006). Securing Fingerprint Template: Fuzzy Vault with Helper Data. In 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06). IEEE. https://doi.org/10.1109/cvprw.2006.185.

[Vans2011] Vans, D. E. (2011). The Internet of things: how the next evolution of the internet is changing everything, Cisco Internet Business Solutions Group (IBSG).

[Viganò2006] Viganò, L. (2006). Automated Security Protocol Analysis With the AVISPA Tool. In Electronic Notes in Theoretical Computer Science (Vol. 155, pp. 61–86). Elsevier BV. https://doi.org/10.1016/j.entcs.2005.11.052.

[Wang2007] Wang, Y., & Plataniotis, K. N. (2007). Fuzzy Vault for Face Based Cryptographic Key Generation. In 2007 Biometrics Symposium. 2007 Biometrics Symposium. IEEE. https://doi.org/10.1109/bcc.2007.4430549.

[Wang2017] Wang, C., Xu, G., & Sun, J. (2017). An Enhanced Three-Factor User Authentication Scheme Using Elliptic Curve Cryptosystem for Wireless Sensor Networks. In Sensors (Vol. 17, Issue 12, p. 2946). MDPI AG. https://doi.org/10.3390/s17122946.

[Weiser1991] Weiser, M. (1991). The Computer for the 21 st Century. Scientific American, 265(3), 94–105. http://www.jstor.org/stable/24938718.

[Wu2008] Wu, X., Wang, K., & Zhang, D. (2008). A cryptosystem based on palmprint feature. In 2008 19th International Conference on Pattern Recognition. (ICPR). https://doi.org/10.1109/icpr.2008.4761117.

[Wu2011] Wu, L., Xiao, P., Yuan, S., Jiang, S., & Chen, C. W. (2011). A Fuzzy Vault Scheme for Ordered Biometrics. In Journal of Communications (Vol. 6, Issue 9). Engineering and Technology Publishing. https://doi.org/10.4304/jcm.6.9.682-690.

[Wu2016] Wu, F., Xu, L., Kumari, S., & Li, X. (2016). A privacy-preserving and provable user authentication scheme for wireless sensor networks based on Internet of Things security. In Journal of Ambient Intelligence and Humanized Computing (Vol. 8, Issue 1, pp. 101–116). Springer Science and Business Media LLC. https://doi.org/10.1007/s12652-016-0345-8.

[Xi2016] Xi, W., Han, J., Li, K., Jiang, Z., & Ding, H. (2016).Location Inferring in Internet of Things and Big Data. In Big Data (pp. 309–335).Elsevier.https://doi.org/10.1016/b978-0-12-805394-2.00013-1.

[Xiao2007] Xiao, B., Yu, B., & Gao, C. (2007). CHEMAS: Identify suspect nodes in selective forwarding attacks. In Journal of Parallel and Distributed Computing (Vol. 67, Issue 11, pp. 1218–1230). Elsevier BV. https://doi.org/10.1016/j.jpdc.2007.04.014.

[Yang2002] Yang, J., & Yang, J. (2002). From image vector to matrix: a straightforward image projection technique—IMPCA vs. PCA. In Pattern Recognition (Vol. 35, Issue 9, pp. 1997–1999). Elsevier BV. https://doi.org/10.1016/s0031-3203(02)00040-7.

[Yongxu2006] Yongxu, W., Xinyu, A., Yuanfeng, D., & Li Yongping. (2006). A Fingerprint Recognition Algorithm Based on Principal Component Analysis. In TENCON 2006 - 2006 IEEE Region 10 Conference. https://doi.org/10.1109/tencon.2006.344032.

[Yu2015] Yu, T., Sekar, V., Seshan, S., Agarwal, Y., & Xu, C. (2015). Handling a trillion (unfixable) flaws on a billion devices. In Proceedings of the 14th ACM Workshop on Hot Topics in Networks. HotNets-XIV. https://doi.org/10.1145/2834050.2834095.

[Zeroual2018] Zeroual, A., Amroune, M., Derdour, M., Meraoumia, A., & Bentahar, A. (2018). Deep authentication model in Mobile Cloud Computing. In 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS). https://doi.org/10.1109/pais.2018.8598508.

[Zeroual2019] Zeroual, A., Derdour, M., Amroune, M., & Bentahar, A. (2019). Using a Fine-Tuning Method for a Deep Authentication in Mobile Cloud Computing Based on Tensorflow Lite Framework. In 2019 International Conference on Networking and Advanced Systems (ICNAS). https://doi.org/10.1109/icnas.2019.8807440.

[Zeroual2021] Zeroual, A., Amroune, M., Derdour, M., & Bentahar, A. (2021). Lightweight deep learning model to secure authentication in Mobile Cloud Computing. In Journal of King Saud University - Computer and Information Sciences. Elsevier BV. https://doi.org/10.1016/j.jksuci.2021.09.016.

[Zhang2020] Zhang, W. E., Sheng, Q. Z., Mahmood, A., Tran, D. H., Zaib, M., Hamad, S. A., Aljubairy, A., Alhazmi, A. A. F., Sagar, S., & Ma, C. (2020). The 10 Research Topics in the Internet of Things. In 2020 IEEE 6th International Conference on Collaboration and Internet Computing (CIC). https://doi.org/10.1109/cic50333.2020.00015.

Annexes

Annexe A : Performance des systèmes biométriques

A.1. Introduction

Le système biométrique est généralement évalué en termes de précision, sécurité, coût de calcul, le coût de communication. Etc. En termes de précision/sécurité, les critères d'évaluation se diffèrent de l'ensemble fermé à l'ensemble ouvert.

A.2. Identification en ensemble fermé

Les critères d'évaluation utilisés dans cet ensemble sont : le taux de reconnaissance au rang un (*Rank One Recognition* - ROR) et le rang de reconnaissance parfaite (*Rank of Perfect Recognition* -RPR). Ces taux sont basés sur le fait que le vecteur authentique est choisi en fonction du score maximum de la similitude.

ROR représente le rapport entre le nombre des vecteurs correctement authentifiés et le nombre total des vecteurs de la base de données. Equation (A1) permet de calculer *ROR* où N_{ca} représente le nombre d'utilisateurs acceptés par le système et *N* est le nombre total des personnes dans la base de données.

RPR est calculé dans le cas où le système identifie l'utilisateur en utilisant des nombreux tests. Si la similitude avec la personne prédite est classée au premier ordre, dans ce cas on l'appelle le premier rang (*Rank 1*), et si elle est classée dans les deux premiers ordres du degré de la similitude, elle s'appelle le deuxième rang (*Rank 2*) et ainsi de suite. Donc, *RPR* calcule la probabilité de trouver la personne authentique dans un rang donné, par conséquence, *ROR* est obtenu comme un résultat du premier range.

$$ROR = \frac{N_{ca}}{N} \tag{A1}$$

A.3. Identification en ensemble ouvert

Comme il est précédemment mentionné, dans l'ensemble ouvert, la décision est prise en fonction d'un seuil prédéfini. La détermination du seuil optimal est critique et spécifique, et est une condition stricte qui permet de prendre une décision en fonction de l'échelle de distance. Ainsi, leur impact est perceptible, notamment en cas de biométrie de mauvaise qualité ou d'intrusion depuis une base de données externe. Parce que sans seuil prédéfini, toute requête doit trouver un modèle selon la distance minimale.

Les critères d'évaluation adoptés sont le taux de faux rejet (*False Rejection Rate - FRR*) et le taux de fausse acceptation (*False Acceptance Rate - FAR*). Le FRR représente le taux de rejet des utilisateurs qui sont normalement autorisés. Tandis que le *FAR* est le taux d'acceptation des

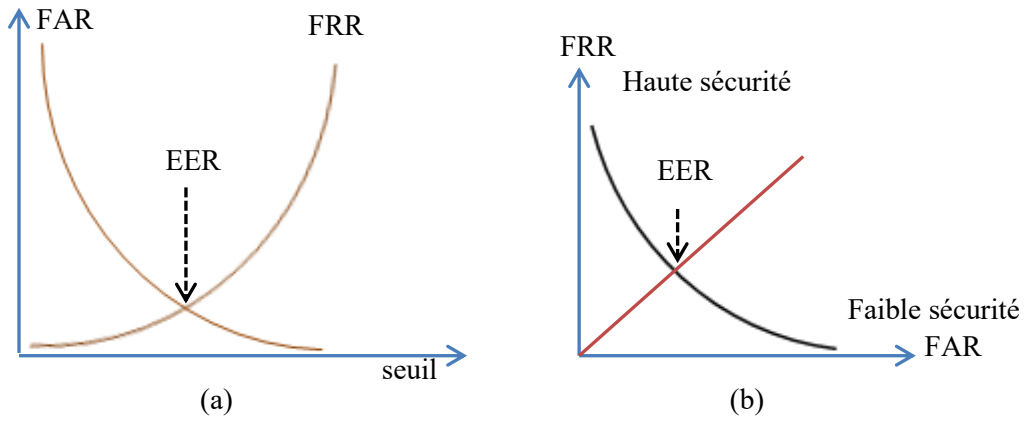


Figure A.1- La présentation de la relativité entre FAR, FRR et EER.

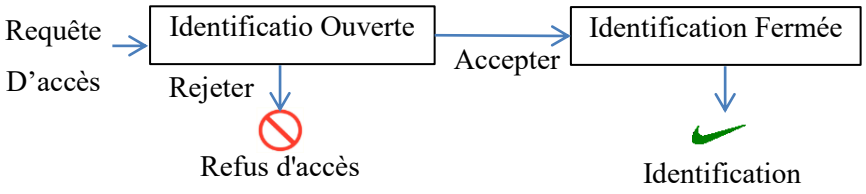


Figure A.2- Mécanisme d'identification en combinant l'identification ouverte et fermée (Source: [Bentahar2019]).

utilisateurs supposés non autorisés. En outre, un autre taux qui représente le taux d'acceptation des utilisateurs autorisés (*Genuine Acceptance Rate - GAR*), peut être obtenu en utilisant le *FRR* comme il est indiqué par l'équation (A2).

$$GAR = 1 - FRR \quad \text{(A2)}$$

FAR et *FRR* sont inversement proportionnels comme le montre la Figure A.1 (a). La Figure A.1 (b) montre que la sécurité est plus assurée avec un petit *FAR*. Effectivement, en termes de sécurité, le refus d'un utilisateur légitime mieux qu'accepter un imposteur.

Autre important taux d'évaluation de cet ensemble qui peut être tenu en compte, c'est le taux d'erreurs égales (*Equal Error Rate - EER*). Ce taux présente le point d'égalité entre les deux taux précédents et par conséquent le point d'équilibre entre la sécurité et la précision. *EER* est défini comme la valeur obtenue à un certain niveau de seuil d'un système biométrique où *FAR* et *FRR* ont la même valeur (voir Figure A.1). En général, plus *EER* est petit, la précision et la sécurité de ce système biométrique sont plus élevées [Dasgupta2017].

A.4. Combinaison d'identification en ensemble ouverte et fermée

La combinaison de ces deux modes d'identification pour en bénéficier les avantages nécessite l'utilisation d'autres critères d'évaluation [Bentahar2019]. Si le système exécute d'abord une tâche d'identification ouverte puis celle d'identification fermée (voir Figure A.2) alors les

acceptations se limitent uniquement à la personne authentique, autrement dit, l'utilisateur qui a demandé l'accès sera bien distingué du reste des utilisateurs. Durant l'identification en mode ouvert, les scores sont calculés et comparés en fonction d'un seuil par lesquels les *FRR* et *FAR* sont évidemment obtenus. En introduisant le score de la similarité maximale comme un autre critère les fausses acceptations seront limitées à une seule erreur d'identification (*Identification Error - IE*). Par conséquent, le taux d'erreur d'identification (*Identification Error Rate - IER*) est un *FAR* avec une définition restreinte. Pour calculer le taux de réussite du système (*Success Rate - SR*), l'équation (A3) est utilisée.

En revanche, l'absence de critère de seuil dans le mode d'identification fermé peut permettre aux imposteurs d'avoir accès au système avec n'importe quel score. En introduisant le critère de seuil, les requêtes imposteurs seront rejetées. Cependant, certaines demandes authentiques peuvent également être rejetées car leurs scores n'ont pas atteint le seuil exigé. Ce type d'erreur correspond aux utilisateurs authentiques exclus est appelé taux de client exclu (*Excluded Client Rate - ECR*). En basant sur ces nouvelles notions, le calcule de *SR* dans ce cas peut se faire par l'équation (A4).

$$SR = 1 - (FRR + IER) \tag{A3}$$

$$SR = ROR - ECR \tag{A4}$$

À partir des équations (A2), (A3) et (A4), nous avons:

$$ROR = (1 - FRR) + ECR - IER = GAR + ECR - IER \tag{A5}$$

A.5. Conclusion

Les critères conventionnels d'évaluation de connaissance les plus importants sont présentés dans cette annexe, aussi bien en identification en ensemble fermé qu'en ensemble ouvert. De plus, de nouveaux critères ont été introduits qui combinent les deux concepts. En plus des critères conventionnels, les systèmes proposés ont été testés par ces nouveaux critères pour combiner les avantages de l'ensemble fermé et de l'ensemble ouvert.

Annexe B : Codes Correcteurs d'Erreurs

B.1. Introduction

Dans cette annexe, nous allons présenter les fondements de base et le principe de fonctionnement des codes correcteurs d'erreurs. Ces codes sont principalement conçus pour détecter et corriger les erreurs de transmission numérique, mais certaines applications dans autres domaines tel que la reconnaissance et la protection par la biométrie sont également envisageables. Des nombreux codes et algorithmes ont été proposés [Barbuddhe2020], nous allons fournir une généralité suffisante pour bien comprendre le but de ces codes, et une explication détaillée des codes utilisés dans notre travail pour bien éclaircir les mécanismes de correction.

B.2. Généralité

Lors d'une transmission ou d'un sauvegarde numérique, certaine éléments (le bit généralement) peuvent être altérés. Pour protéger l'information de ces altérations les codes détecteurs/correcteurs sont impérativement utilisés. Ils sont omniprésents dans tous les systèmes de communication numériques. Il s'agit d'ajouter quelques éléments supplémentaires à l'information pour achever cette protection, c'est le codage du canal.

Le codage du canal peut être seulement détecteur, c.-à-d. il détecte l'erreur mais il ne peut pas localiser sa position exacte, donc il ne peut pas la corriger. Parmi ces codes nous trouvons la parité unidimensionnelle. Ou il peut être correcteur, c.-à-d. il localise l'erreur et la corrige, exemple la parité croisée (bidimensionnelle). Cependant, la capacité de correction des codes n'est pas à l'infinie, elle possède des limitations. Dans ce cas, le terme « code détecteur/correcteur » est plus pointu et plus significatif. Exemple : le code *Hamming* qui peut détecter deux erreurs et en corriger une.

Les codes peuvent se classifier en plusieurs types. Ci-dessous quelques définitions des types répandus.

- Codes linéaires : ce sont des codes dont le calcul des bits de contrôle se fait en fonction de tous les éléments du mot (information) à coder.
- Codes non-linéaires : ce sont des codes dont le calcul des bits de contrôle ne dépend pas de tous les éléments du mot.
- Codes cycliques : Si nous faisons une permutation entre le premier et le dernier bit (ou vis-versa) dans n'importe quel mot-code (information codée) nous obtenons un mot-code existant dans la table du codage.
- Codes en bloc : chaque mot se code indépendamment d'autres.

$$G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} , \; H^T = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix}$$

Figure B.1- la matrice génératrice et de contrôle d'un code Hamming C(7,4)

$$G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Hamming C(7,4) — Hamming étendu C(8,4)

$$H^T = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix}$$

Hamming C(7,4) — Hamming étendu C(8,4)

Figure B.2- La matrice génératrice et de contrôle d'un code Hamming étendu

- Codes convolutifs : les mots se codent selon les codes des blocs précédents

B.3. Code Hamming

C'est code en bloc linéaire, dans ce codage les colonnes de la matrice de contrôle H^T sont les nombres en valeurs binaires ordonnées de 1 à n qui indiqueront la position d'erreur. La longueur du mot-code n'est alors plus quelconque mais elle est $n=2^k - 1$ (le zéro est exclu comme valeur de la colonne). De ce fait la matrice génératrice n'est pas aussi quelconque mais elle doit être définie par la dimension C (n,m) = ($2^k - 1$, $2^k - 1 - k$). Où n est la longueur du mot-code, m est celle du mot et k est un entier positif. La distance minimale d'un code de *Hamming* est de 3, donc : Il détecte 2 erreurs et il en corrige une.

Exemple : le code *Hamming* C(7,4) ; Les longueurs correspondant sont : n=7, m=4 et r=3, a comme une matrice génératrice G et celle de contrôle H^T de dimension 4x7 et 7x3 respectivement. La figue B.1 montre ce propos.

Donc, pour coder un mot il suffit multiplier ce dernier par la matrice G, et pour le décoder nous multiplions le mot-code par H^T. Le vecteur obtenu par la multiplication du mot-code par H^T est appelé « syndrome » et il indique directement la position du bit erroné s'il n'est pas nul.

Soit le mot « 1011 », le mot-code est donc « 0110011 ». Après la transmission ou le transfert, si le mot-code n'était altéré nous obtenons « 000 » comme syndrome. Si une erreur est survenue sur un bit ; à titre d'exemple le deuxième bit, nous aurons « 0010011 » x H^T = « 010 » qui veut dire « 2 » en décimal. L'erreur est détectée et le deuxième bit est corrigé.

A noter que ce code corrige seulement une erreur parmi les 7 bits envoyés. Si deux ou plusieurs bits sont altérés, le décodage échoue en détectant autres positions non-pertinentes. Exemple : prenons deux erreurs dans le même mot-code précédent, soit « 0010111 » c.-à-d. dans le deuxième et le cinquième bit. « 0010111 » x H^T = « 111 » qui veut dire la septième position !. Le système en suite change le septième bit et il nous donne « 0010110 » comme mot-code corrigé alors que les erreurs étaient sur le deuxième et le cinquième bit. A noter que le mot-code « 0010110 » n'est pas arbitraire et il existe dans la table des mot-codes. Pour remédier à ce problème (discerner entre deux et une erreur), le bit de parité peut être ajouté et par conséquent la huitième position sera indiquée par le vecteur « 000 ». C'est le code *Hamming* étendu (*Extended Hamming* - EH(8,4)).

B.4. Code Hamming étendu

À partir du code de Hamming C ($2^k - 1$, $2^k - 1 - k$) de distance 3, on peut construire un autre code plus optimal de C (2^k , $2^k - 1 - k$) et de distance 4 dit code de Hamming étendu Le codage consiste à ajouter au mot-code ordinaire un bit de parité qui vaut « 0 » pour un nombre pair des « 1 » et « 1 » pour l'impair.

La matrice de contrôle H^T est obtenue par l'adjonction d'une ligne ne contenant que des « 0 » pour la huitième position et l'adjonction d'une colonne ne contenant que des « 1 » pour la parité. Figure B.2 montre les matrices d'un code Hamming étendu EH(8,4).

Le codage et décodage se font de même manière que le code ordinaire, c.-à-d. le codage est mot x G et le décodage est mot-code x H^T. Le syndrome obtenu est de longueur (r+1), r pour la position et un bit pour la parité. Selon ce syndrome, le décodage est par l'algorithme suivant:

Début

Si tous les (r+1) bits sont des zéros,

Alors la transmission est faite sans erreurs.

Sinon (Au moins un « 1 » est présent),

Si le bit de parité et « 1 »,

Alors il y a une erreur dans la position indiquée par les autres bits r (le tous-zéro de r est la dernière position).

Sinon (Bit de parité est « 0 »),

Alors il y a deux erreurs et le codage dans ce cas est seulement détecteur.

Fin.

Pour trois erreurs ou plus le code étendu revient au problème du celui ordinaire.

Soit le même mot précédent « 1011 ». Son mot-code est « 01100110 ». Nous prenons les trois cas ci-dessus et nous décodons les mot-code reçu :

- Sans erreur : « 01100110 » x H^T= « 0000 ».
- Une erreur dans « 2 » : « 00100110 » x H^T= « 0101 ».

Deux erreurs dans « 2 et 5 » : « 00101110 » x H^T= « 1110 ».

B.5. Codes BCH et Reed-Solomon

Le nom BCH vient de ses inventeurs *Bose, Chaudhuri, Hocquenghem*. Il s'agit d'un code CRC particulier avec une capacité de correction plus grande. Dans les codes cycliques CRC ordinaires la distance minimale est calculée par n(n-1)/2 combinaisons alors que par les BCH cette distance est directement déduite du polynôme générateur g(x). Il peut corriger t erreurs selon le nombre des polynômes minimaux concaténés.

Pour les codes cycliques, la factorisation de (X^n+1) est faisable si n est pair. La question qui se pose est : quels sont les codes cycliques de longueur n si n est impair ?

La factorisation de $X^n + 1$ est $(X + 1)(X^{n-1} + X^{n-2} + \dots + X + 1)$. Si $X^{n-1} + \dots + X + 1$ a des racines, la situation est simple ; mais s'il n'a pas de racine nous devons imaginer de telles racines pour que $X^n + 1$ soit factorisable. Ces racines sont appelées racines primitives et ce sont l'origine des codes BCH.

Nous appelons une racine primitive n^e de l'unité un nombre imaginaire α de telle façon que toutes les puissances de ce nombre n'égale pas à 1 sauf la dernière puissance n c.-à-d. : $\alpha \neq 1, \alpha^2 \neq 1, \alpha^3 \neq 1, \dots, \alpha^n = 1$. Alors ce nouveau nombre n'est 0 ni 1.

Si α est une racine imaginée, Alors $\alpha^2, \alpha^3, \dots, \alpha^n$ le sont aussi. Par conséquent la factorisation de (X^n+1) est faisable telle que: $(X^n + 1) = (X + \alpha)(X + \alpha^2)(X + \alpha^3) \dots (X + \alpha^n) = (X + 1)(X + \alpha)(X + \alpha^2)(X + \alpha^3) \dots (X + \alpha^{n-1})$.

A ce stade, nous avons finalement des racines de (X^n+1), mais le problème c'est qu'elles ne sont pas dans F_2 c.-à-d. elles ne sont pas binaires. Le nouveau problème est comme regrouper ces $(X+\alpha^i)$ pour que le polynôme générateur soit dans F_2. Nous ne voulons pas entrer dans plus de détails et nous focalisons directement sur l'essentiel de notre implémentation, les polynômes ci-dessous sont les résultats d'une telle factorisation.

➢ Exemple

- $X^7+1= (X+1)(X^3+X^2+1)(X^3+X+1)$ et donc les codes cycliques de longueur 7 sont (1,7), (3,7), (4,7) et (6,7).
- $X^9+1= (X+1)(X^2+X+1)(X^6+X^3+1)$ et donc les codes cycliques de longueur 9 sont (1,9), (2,9), (3,9), (6,9), (7,9) et (8,9).
- $X^{11}+1= (X+1)(X^{10}+X^9+ \ldots. +X+1)$ et donc les codes cycliques de longueur 11 sont triviaux.

Le codage est la multiplication du mot par le polynôme générateur engendré. Et le décodage se fait par l'algorithme de *Peterson*, *Gorenstein* et *Zierler* suivant. Notons P(X) la représentation polynômial du mot reçu.

Étape 1: On calcule $S_1 = P(\alpha 2) \ldots S_{2t} = P(\alpha 2^{2t})$.

Étape 2: Calcul du nombre d'erreurs, on note *v* le rang de la matrice $\begin{pmatrix} S_1 & \ldots. & S_t \\ S_2 & \ldots. & S_{t+1} \\ \vdots & \vdots & \vdots \\ S_t & \ldots. & S_{2t+1} \end{pmatrix}$.

Étape 3: On résout le système $\begin{pmatrix} S_1 & \ldots. & S_t \\ S_2 & \ldots. & S_{t+1} \\ \vdots & \vdots & \vdots \\ S_t & \ldots. & S_{2t+1} \end{pmatrix} \times \begin{pmatrix} s_v \\ s_{v-1} \\ \vdots \\ s_1 \end{pmatrix} = \begin{pmatrix} S_{v+1} \\ S_{t+1v+2} \\ \vdots \\ S_{2v} \end{pmatrix}$.

Étape 4: On détermine les *v* racines $\alpha^{-i1},\ldots, \alpha^{-iv}$ du polynôme $s_vX^v+\ldots+ s_1X+1$.

Étape 5: Les erreurs ont eu lieu en position $i_1,\ldots, i_v$, on inverse ces bits.

Prenons deux cas simples. Premier cas: t=1 (c'est-à-dire que le code considéré est un code de *Hamming*). Dans ce cas l'algorithme est très dégéné:

Début

On calcule $S = P(\alpha)$.

Si S=0 alors pas d'erreur.

Sinon on peut écrire $S= \alpha^i$, L'erreur a eu lieu en position i, on inverse ce bit.

Fin.

Second cas : t=2. Dans ce cas l'algorithme est le suivant:

Début

On calcule $S_1=P(\alpha)$, $S_2=P(\alpha^2)$, $S_3=P(\alpha^3)$, $S_4=P(\alpha^4)$.

On calcule du nombre d'erreurs.

Si $S_1=0$ Alors pas d'erreur.

Sinon

Début

Si $S_1 \times S_3 = S_2^2$, une erreur, on peut noter $S_1 = \alpha^i$, on inverse le bit en position i.

Sinon

Début

On résout le système $\begin{cases} S_1 \times s_2 + S_2 \times s_1 = S_3 \\ S_2 \times s_2 + S_3 \times s_1 = S_4 \end{cases}$.

On détermine les 2racines α^{-i} et α^{-j} du polynôme $s_2X_2 + s_1X + 1$.

Les erreurs ont eu lieu en position i et j, on inverse ces bits.

Fin sinon

Fin sinon

Fin si.

Fin.

Les codes Reed-Solomon sont une classe élargie des codes BCH. Nous travaillons sur F_n au lieu F_2, donc tous les coefficients des polynômes sont modulo *n* et pas *2*. La méthode de codage et l'algorithme de décodage reste les mêmes. Les codes Reed-Solomon sont orientés octet ce qui les rends ayant plus de capacité de correction avec un minimum de redondance par rapport au BCH.

B.6. Conclusion

Les codes correcteurs d'erreurs les plus importants sont présentés dans cette annexe, en particulier, ceux utilisés dans notre travail. Dans tous les cas, on peut dire que plus la précision de la correction d'erreur est élevée, plus la complexité des calculs est grande et, par conséquent, plus le temps est long et plus l'implementation est est lourde.