Titre: Title:	Conception d'une architecture basée sur la blockchain pour des services de commerce électronique
Auteur: Author:	Olson Italis
Date:	2023
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Italis, O. (2023). Conception d'une architecture basée sur la blockchain pour des services de commerce électronique [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/54852/

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Conception d'une architecture basée sur la blockchain pour des services de commerce électronique

OLSON ITALIS

Département de génie informatique et génie logiciel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*

Génie informatique

Août 2023

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

Conception d'une architecture basée sur la blockchain pour des services de commerce électronique

présentée par **Olson ITALIS**

en vue de l'obtention du diplôme de *Philosophiæ Doctor*

a été dûment acceptée par le jury d'examen constitué de :

Martine BELLAÏCHE, présidente
Samuel PIERRE, membre et directeur de recherche
Alejandro QUINTERO, membre et codirecteur de recherche
Tarek OULD-BACHIR, membre
Abderrezak RACHEDI, membre externe

DÉDICACE

À mes parents, Marie-Édrise et Dominique-Vital,

À mes frères et sœurs : Rachel, Noslo, Morino, Sagine, Mackel, Shedna, Freud, Edny et Lentz,

À ma tante Rose-Vilna, partie trop tôt,

À mes oncles et cousine, Joseph-Wilson, Pierre-Paul et Frantze, qui ont joué le rôle de parents quelques fois dans ma vie.

REMERCIEMENTS

Tout d'abord, je tiens à exprimer ma sincère reconnaissance à mon directeur de recherche, le professeur Samuel Pierre, pour son encadrement professionnel, son amabilité et son soutien académique et financier. Je vous remercie sincèrement pour la confiance que vous avez eue en mon travail et pour m'avoir donné cette opportunité inestimable d'acquérir de nouvelles compétences scientifiques dans le Laboratoire de recherche en Réseautique et en Informatique Mobile (LARIM), d'élargir mes perspectives et d'avoir une meilleure compréhension de mon travail. Vos réalisations, votre attitude, votre humanité et votre modestie sont exemplaires dans ma vie.

Ma sincère gratitude va également à mon codirecteur de recherche, le professeur Alejandro Quintero, pour son guide, sa disponibilité et sa manière de nous pousser à l'excellence. J'ai compris qu'il encaisse les faiblesses des étudiants, nous défend devant les institutions qui sponsorisent la recherche en nous poussant à faire mieux.

J'aimerais également remercier mes collègues des projets ATISCOM et APSEC. D'abord, j'exprime ma reconnaissance à la société Flex Group, en la personne de M. Mostafa Chafi, pour le support financier accordé à cette recherche. Je remercie aussi M. Sepehr Keykhaie et M. Gaston Blanchard du groupe Flex, qui ont orienté les recherches vers des aspects plus pratiques.

Je tiens également à remercier Mme Martine Bellaïche, M. Abderrezak Rachedi et M. Tarek Ould-Bachir pour l'intérêt qu'ils ont porté à ma recherche et pour avoir accepté de participer au jury.

Je ne saurais oublier mes collègues et ami.e.s du LARIM, pour avoir fait du laboratoire un lieu propice à la concentration et au travail collaboratif dans une atmosphère chaleureuse et amicale. Merci à Dre. Franjieh El Khoury, la superviseure du projet ATISCOM, pour son soutien inconditionnel et pour ses relectures qui ont grandement amélioré ce travail; elle a également su faire du LARIM une véritable famille. Je me rends compte que la famille est grande: certains ont laissé, d'autres arrivent et apportent toutes et tous leur énergie positive à l'atmosphère. Ainsi, je remercie Dieu pour la connexion avec Marième, Éric, Lamia, Amir, Farnoush, Nasrin, Claudy, Loic, Sanaz, Paran, Jean, Nick, Marc, Gaelle, Hadger et Somayyeh. Merci à vous pour votre collaboration, et vos commentaires constructifs sur mes recherches. Entre autres, je voudrais adresser ma gratitude envers un autre groupe de cette famille, constitué de mes ami.e.s de

l’ISTEAH qui sont comme moi attaché.e.s ou souvent de passage au LARIM: Magdala, Fania, Marie Carria, Rose-Michelle, Jean-Marie, Ary, Marc-Donald et Amélie. Certains d’entre eux me poussent à terminer la thèse rapidement tout en sachant apporter des temps de pause plein de chaleur. Merci à vous toutes et à vous tous.

Merci à mon ami Jocelyn Dupenor (qui est comme un frère), à sa femme Gerline, ses enfants Marc-Olivier et Ann-Ashlee qui sont désormais mes neveu et nièce. Je me dois de remercier également, ceux qui ont joué le rôle de parents à des périodes différentes dans ma vie : mes oncles Joseph-Wilson et Pierre-Paul, puis ma cousine germaine Frantze.

Je suis à jamais redevable à mes parents, Marie-Édrise et Dominique-Vital, pour leur soutien continu, leur amour et la confiance qu'ils ont placée en moi. Les mots ne suffisent pas pour les remercier. Leurs prières ont toujours été avec moi. Merci à mes frères et sœurs d'avoir été de bons amis et de m'avoir poussé à aller de l'avant dans les moments difficiles, ils sont parfois des conseillers aussi.

RÉSUMÉ

L'utilisation généralisée des téléphones mobiles et le développement des services de commerce qui en découlent ont permis aux utilisateurs de s'habituer à la commodité des nouveaux modes de paiement et d'accéder presqu'en permanence aux services bancaires. Dans les régions à faible taux de bancarisation mais avec un fort taux de pénétration du téléphone mobile, développer ces services sur des plateformes technologiques mobiles devient un moyen d'augmenter l'inclusion financière.

Cependant, offrir ces services implique de lier le moyen de paiement (le téléphone) à l'identité de l'utilisateur, ce qui entraîne la collecte et le transit d'informations personnelles sur les réseaux entre les différents participants de la plateforme technologique sous-jacente; ce qui peut compromettre la privacité. En outre, des menaces telles que les logiciels malveillants peuvent entraîner des vols de données confidentielles, et la perte de l'équipement mobile est également un risque.

Les plateformes actuelles utilisent des technologies qui présentent des vulnérabilités, ce qui compromet leur sécurité. Par exemple, dans le cas des systèmes s'appuyant sur *Bluetooth Low Energy* (BLE), les clés de chiffrement des échanges sont à risque, car pouvant être copiées par un attaquant qui peut aussi saturer le canal et provoquer conséquemment un déni de service ou *Denial of Service* (DoS). Pour les plateformes utilisant le *Quick Response Code* (QRC), aucune mesure ne permet de se protéger contre un code QR compromis qui redirige vers un site malicieux. Les systèmes avec communication en champ proche ou *Near Field Communication* (NFC) permettent de résoudre le problème de stockage de clé via l'élément sécurisé ou *Secure Element* (SE) qui protège les données confidentielles contre un accès non autorisé. Toutefois, l'élément sécurisé est fourni par un manufacturier ou un opérateur de téléphonie mobile, entravant l'interopérabilité de tels systèmes. L'architecture par *Host Card Emulation* (HCE) augmente l'indépendance du SE par rapport aux fabricants ou vendeurs de cartes, certes; mais le canal entre le SE dans le cloud et le module du client doit être sécurisé.

Pour le commerce mobile à distance, les plateformes technologiques offrent les services bancaires et de paiement partout et en tout temps. Toutefois, les informations confidentielles sont souvent partagées avec plusieurs participants, impliquant un problème de privacité. Aussi, plusieurs systèmes de paiement mobile ont été déployés dans les pays en voie de développement utilisant le *Short Message Service* (SMS) et l'*Unstructured Supplementary Service Data* (USSD)

avec les opérateurs de téléphonie mobile ou *Mobile Network Operators* (MNOs). Cependant, il n'y a pas une technologie standard établissant l'interopérabilité entre ces derniers systèmes et d'autres principalement en exploitation dans les pays développés. D'où la nécessité de l'interopérabilité à ce niveau sans compromettre la sécurité des systèmes les plus fiables.

D'un autre côté, ces systèmes s'appuient sur un dorsal qui supporte les échanges entre ce que nous pouvons appeler des institutions de paiement (ex. des banques, des opérateurs de téléphonie mobile, des fournisseurs de services de paiement indépendants entre autres). Ces institutions forment des autorités de confiance qui valident les transactions et font que les plateformes fonctionnent selon un modèle centralisé. Beaucoup de faiblesses sont observées dans ce modèle : la réversibilité des transactions, un point de défaillance unique, le manque de respect de la vie privée entre autres.

La blockchain a des caractéristiques qui la rendent attractives pour la résolution éventuelle des problèmes susmentionnés. D'abord, elle est décentralisée et tend vers la transparence; ce qui pourrait être un atout pour l'interopérabilité. En plus, le recouvrement en cas de panne est rapide. Et finalement, elle est résistante à la falsification et met bien à profit l'utilisation de la cryptographie asymétrique. Cependant, le secteur des paiements hésite à utiliser la blockchain pour de nombreuses raisons : problème avec la gestion des clés, manque de respect de la vie privée, performances médiocres de certains protocoles de consensus, manque de responsabilité et de mécanisme de contrôle, évolutivité et manque d'adéquation avec les règles d'affaires établies.

Cette thèse aborde plusieurs défis liés à la blockchain. Tout d'abord, elle cherche à établir une connexion transparente entre l'identité de l'utilisateur sur la blockchain et ses caractéristiques dans la réalité, tout en garantissant des propriétés de sécurité telles que la confidentialité des données, leur intégrité et l'authentification. Ensuite, elle vise à préserver la vie privée des utilisateurs. Enfin, elle aborde la question de l'interopérabilité entre les plateformes internes des institutions financières à travers une nouvelle architecture. Ces problèmes sont traités en quatre grandes étapes.

Dans la première étape, nous avons effectué une revue de littérature systématique sur les systèmes de paiement mobile et leur intégration avec la Blockchain. Ce qui a permis de mettre en lumière une taxonomie des architectures des plateformes de paiement mobile, les enjeux de sécurité

associés et les limitations de la blockchain dans la perspective d'offrir des services de paiement électronique.

Dans la deuxième étape de la thèse, nous générons une clé cryptographique à partir des données biométriques de l'utilisateur, de manière à pouvoir la reproduire ultérieurement. Pour y parvenir, il faut surmonter des défis liés à la variabilité intra-utilisateur qui va à l'encontre de la reproductibilité de la clé. Pour résoudre ce problème, nous avons conçu une fonction qui transforme les données biométriques de l'utilisateur en un code spécifique. Nous avons ajouté un mécanisme d'apprentissage à ce processus de codification, ce qui permet de capturer une esquisse minimale des données biométriques afin d'aider à la reproduction du code. En utilisant cette codification, des algorithmes intelligents ont été développés pour créer un système d'appariement biométrique simple et efficace pour l'authentification. Une autre contribution de cette partie de la thèse est une fonction de dérivation de clé basée sur les caractéristiques biométriques. Cette fonction a été construite en respectant les exigences d'une fonction de dérivation de clé basée sur HMAC (*HMAC-based Key Derivation Function*), selon les normes de l'*Internet Engineering Task Force* (IETF). Elle utilise le code biométrique précédemment généré et un identifiant de l'utilisateur comme données d'entrée. Cela nous permet d'obtenir une clé cryptographique révocable et irréversible, avec une chaîne aléatoire. Aucun des générateurs de clé basés sur les données biométriques trouvés dans la littérature ne permet de bénéficier simultanément de ces trois propriétés.

Dans la troisième étape de la thèse, nous avons abordé le problème du manque de respect de la vie privée observé dans la technologie de la blockchain. En effet, cette technologie a été conçue de manière à rendre les transactions disponibles sur tous les nœuds validateurs. Même si un pseudonyme est utilisé pour identifier l'expéditeur ou le destinataire des fonds dans le système, la traçabilité des transactions permet de découvrir leur véritable identité. Différents systèmes ont été proposés pour résoudre ce problème en utilisant des protocoles cryptographiques. Dans cette thèse, nous avons proposé un protocole de paiement qui ne nécessite pas d'opérations cryptographiques supplémentaires par rapport à celles déjà nécessaires au fonctionnement de la blockchain. Le mécanisme utilisé pour garantir la vie privée n'entraîne pas de surcoût en termes de mémoire ou de communication dans le processus de paiement.

La dernière étape de la thèse se concentre sur la conception globale de l'architecture en tenant compte des enjeux de sécurité, de performance et d'autres limitations liées au protocole de

consensus sous-jacent qui affectent l'utilisabilité. Tout d'abord, elle propose une taxonomie des différents protocoles de consensus existants, en identifiant différentes classes de protocoles en fonction des critères adaptés à des applications spécifiques. Ensuite, elle présente un cadre définissant des architectures basées sur la blockchain avec un couplage faible entre les plateformes des institutions financières; ce qui facilite l'interopérabilité. En suivant ce cadre, une architecture est conçue en intégrant les composants développés au cours des deux premières étapes. Cette conception illustre l'intégration de la fonction de dérivation de clé, du protocole de paiement décentralisé, de la blockchain Ethereum et de la connexion avec la plateforme de paiement électronique traditionnelle pour former un système décentralisé offrant des services de commerce électronique. Ce modèle permet de conserver les avantages de la blockchain, tels que la décentralisation, tout en respectant les principes établis de l'écosystème financier. Cette étape comprend également la visualisation d'une extension du protocole de consensus stellaire ou *Stellar Consensus Protocol* (SCP) pour améliorer sa sécurité. Ce protocole étant adapté aux services nécessitant une faible latence de transaction et un haut débit, il est candidat pour une architecture offrant de meilleures performances. L'extension du SCP envisage une intégration intelligente de la notion d'importance et de réputation de chaque nœud du réseau, permettant à un participant de choisir dynamiquement les nœuds de confiance.

Les résultats de l'évaluation mettent en évidence que l'architecture proposée pour les services de commerce électronique est sécurisée, respecte la vie privée des utilisateurs et facilite l'interopérabilité entre différentes plateformes de paiement électronique. De plus, cette architecture se distingue par sa flexibilité, lui permettant de s'adapter aux réglementations en vigueur.

En conclusion, la thèse offre une approche complète pour résoudre les problèmes liés à la sécurité, à la vie privée et à l'interopérabilité dans les services de commerce électronique en utilisant la blockchain comme une solution prometteuse.

ABSTRACT

The widespread use of cell phones and the development of related commerce services have enabled users to become accustomed to the convenience of new payment methods and to access banking services almost around the clock. In regions with low banking penetration but high cell phone penetration, developing these services on mobile technology platforms is becoming a means of increasing financial inclusion.

However, offering these services involves linking the means of payment (the phone) to the user's identity, resulting in the collection and transit of personal information over networks between different participants; which can compromise privacy. In addition, threats such as malware can lead to theft of confidential data, and loss of mobile equipment is also a risk.

The technologies currently used in these platforms also present vulnerabilities that affect their security. For example, in the case of systems based on Bluetooth Low Energy (BLE), the keys used to encrypt exchanges are at risk, as they can be copied by an attacker who can also saturate the channel and consequently cause a Denial of Service (DoS). For platforms using Quick Response Code (QRC), there are no measures to protect against a compromised QR code redirecting to a malicious site. Systems with NFC solve the problem of key storage via the Secure Element (SE), which protects confidential data from unauthorized access. However, the secure element is provided by a manufacturer or a Mobile Network Operator (MNO), hindering the interoperability of such systems. The Host Card Emulation (HCE) architecture increases the independence of the SE from card manufacturers or vendors, but the channel between the SE in the cloud and the client module must be secured.

For remote mobile commerce, these platforms offer banking and payment services anywhere, anytime. However, confidential information is often shared with several participants, implying a problem of privacy. Also, several mobile payment systems have been deployed in developing countries using Short Message Service (SMS) and Unstructured Supplementary Service Data (USSD) with MNOs. However, there is no standard technology establishing interoperability between these latter systems and others mainly in operation in developed countries. Hence the need for interoperability at this level without compromising the security of the most reliable systems.

On the other hand, these systems are based on a backbone that supports exchanges between what we might call payment institutions (e.g., banks, cell phone operators, independent payment service providers, etc.). These institutions act as trusted authorities, validating transactions and ensuring that the platforms operate according to a centralized model. Many weaknesses are observed in this model: reversibility of transactions, a single point of failure, lack of respect for privacy among others.

Blockchain has a number of features that make it an attractive solution to the problems described above. Firstly, it is decentralized and tends towards transparency, which can be an asset for interoperability. Additionally, the recovery process in case of failure is quick and efficient. Finally, it is resistant to forgery and makes good use of asymmetric cryptography. However, the payments industry is reluctant to use blockchain for many reasons: key management issues, lack of confidentiality, poor performance of some consensus protocols, lack of accountability and control mechanism, scalability and lack of fit with established business rules.

This thesis addresses several challenges associated with blockchain. Firstly, it seeks to establish a transparent connection between the user's identity on the blockchain and its characteristics in reality, while guaranteeing security properties such as data confidentiality, integrity and authentication. Secondly, it aims to preserve user privacy. Finally, it discusses the ease of use of blockchain-based architecture and interoperability between financial institutions' internal platforms via this architecture. These issues are addressed in four main stages.

During the first phase, we conducted a comprehensive review of existing literature on mobile payment systems and their integration with Blockchain. This process revealed key insights, including a classification of mobile payment platform architectures, a thorough examination of the security challenges involved, and an assessment of the limitations of blockchain in relation to providing electronic payment services.

In the second stage of the thesis, we generate a cryptographic key from the user's biometric data, so that it can be reproduced at a later date. To achieve this, we need to overcome challenges related to intra-user variability, which runs counter to the reproducibility of the key. To solve this problem, we have designed a function that transforms the user's biometric data into a specific code. We have added a learning mechanism to this encoding process, enabling a minimal sketch of the biometric data to be captured to aid code reproduction. Using this coding, intelligent algorithms

have been developed to create a simple and effective biometric matching system for authentication. Another contribution of this part of the thesis is a key derivation function based on biometric features. This function has been built in compliance with the requirements of an HMAC-based Key Derivation Function, according to Internet Engineering Task Force (IETF) standards. It uses the previously generated biometric code and a user identifier as input data. This enables us to obtain a revocable and irreversible cryptographic key, with a random string. None of the key generators based on biometric data found in the literature can simultaneously benefit from these three properties.

In the third stage of the thesis, we addressed the problem of the lack of privacy observed in blockchain technology. Indeed, this technology was designed to make transactions available on all validating nodes. Even if a pseudonym is used to identify the sender or recipient of funds in the system, the traceability of transactions may enable their true identity to be discovered. Various systems have been proposed to solve this problem using cryptographic protocols. In this thesis, we have proposed a payment protocol that requires no additional cryptographic operations to those already necessary for the blockchain to function. The mechanism used to guarantee privacy does not entail any additional costs in terms of memory or communication in the payment process.

The final stage of the thesis focuses on the overall design of the architecture, considering security, performance and other limitations linked to the underlying consensus protocol that affect usability. First, it proposes a taxonomy of the various existing consensus protocols, identifying different classes of protocols based on criteria suitable for specific applications. Secondly, it presents a framework defining blockchain-based architectures with weak coupling between financial institutions' platforms; this facilitates interoperability. Following this framework, an architecture is designed by integrating the components developed in the first two steps. This design illustrates the integration of the key derivation function, the decentralized payment protocol, the Ethereum blockchain and the connection with the traditional electronic payment platform to form a decentralized system offering e-commerce services. This model retains the advantages of blockchain, such as decentralization, while respecting the established principles of the financial ecosystem. This step also includes visualizing an extension to the Stellar Consensus Protocol (SCP) to enhance its security. As this protocol is suited to services requiring low transaction latency and high throughput, it is a candidate for a higher-performance architecture. The SCP extension

envisions an intelligent integration of the notion of importance and reputation of each network node, enabling a participant to dynamically select trusted nodes.

The results of the evaluation show that the proposed architecture for e-commerce services is secure, preserves user privacy and facilitates interoperability between different platforms. In addition, the architecture stands out for its flexibility, enabling it to adapt to current regulations.

In conclusion, the thesis offers a comprehensive approach to solving security, privacy and interoperability issues in e-commerce services, using blockchain as a promising solution.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

AES	Advanced Encryption Standard
AFCEE	Association Française Pour Le Commerce Et Les Échanges Électroniques
AML	Anti-Money Laundering
BBPS	Blockchain-Based Payment System
BFT	Byzantine Fault Tolerance
BIC	Bank Identifier Code
BLE	Bluetooth Low Energy
BN	Backend-Network
BSC	Base Station Controler
BTS	Base Transciever Station
CMPS	Conventional Mobile Payment Systems
CSRK	Connection Signature Resolution Key
DoS	Denial of Service
dPoS	Delegated Proof-of-Stake
ECC	Elliptic-Curve Cryptography
eSIM	Embedded Subscriber Identity Module
FAR	False Acceptation Rate
FBA	Federated Byzantine Agreement
FE	Fuzzy Extractor

FRR	False Rejection Rate
GSM	Global System for Mobile Communication
HCE	Host Card Emulation
HMAC	Hash-Based Message Authentication Code
HSM	Hardware Security Module
IRK	Identity Resolution Key
ISM	Industrial, Scientific and Medical
JNB	Jenk Natural Break
KDF	Key Derivation Function
KYC	Know Your Customer
LTK	Long Term Key
MFA	Multi-Factor Authentication
MITM	Man In The Middle
MNO	Mobile Network Operator
MS	Mobile Station
MSC	Mobile Switching Center
NFC	Near Field Communication
NFC CLF	Near Field Communication Contactless Front-End
NIST	National Institute of Standards and Technologies

NSC	National Sort Code
OTA	Over The Air
OTP	On-Time Password
PBFT	Practical Byzantine Fault Tolerance
PIN	Personal Identification Number
PKI	Public Key Infrastructure
PoB	Proof-of-Burn
PoI	Proof-of-Importance
PoR	Proof-of-Reputation
PoS	Proof-of-Stake
PoW	Proof-of-Work
PSP	Payment Service Provider
PUF	Physically Unclonable Function
QoS	Quality of Service
QRC	Quick Response Code
RFE	Robust Fuzzy Extractor
RTGS	Real-Time Gross Settlement
SCP	Stellar Consensus Protocol
SE	Secure Element

SIM	Subscriber Identity Module
SME	Short Messaging Entity
SMPP	Short Message Peer to Peer
SMS	Short Message Service
SMSC	Short Message Service Center
SQL	Structured Query Language
SQRC	Secure Quick Response Code
SS7	Signaling System No. 7
SWIFT	Society for Worldwide Inter-Bank Financial Telecommunication
SWP	Single Wire Protocol
TCP/IP	Transfer Control Protocol / Internet Protocol
TEE	Trusted Execution Environnement
TLS/SSL	Transport Layer Security / Secure Sockets Layer
TNR	True Negative Rate
TTP	Trusted Third Party
USSD	Unstructured Supplementary Service Data
UTXO	Unspent Transaction
XSS	Cross-Site Scripting

CHAPITRE 1 INTRODUCTION

Avec l'utilisation très répandue du téléphone cellulaire et le développement des services de commerce utilisant ces équipements, les utilisateurs se sont habitués à la commodité des nouvelles formes de paiement et d'accès presqu'en tout temps aux services bancaires. Dans certaines régions à faible taux de bancarisation et fort taux de pénétration du téléphone mobile [1], le développement de ces services sur les plateformes de technologie mobile devient un moyen d'augmenter le taux d'inclusion financière.

Puisque pour offrir ces services, le moyen de paiement qui est le téléphone est lié à l'identité de l'utilisateur, des informations personnelles sont à priori enregistrées et transitent sur des réseaux entre les différents participants; ce qui entrave le respect de la vie privée. Aussi, beaucoup d'autres menaces liées à l'utilisation du téléphone, notamment des logiciels malveillants peuvent être à l'origine de vols de données confidentielles [2], sans compter la perte de l'équipement mobile.

Traditionnellement, hormis la mise en œuvre d'une couche d'authentification par mot de passe, les systèmes offrant ces services s'appuient sur la cryptographie pour garantir la sécurité des données échangées. Les clés cryptographiques peuvent être exposées dans certains cas. Dans d'autres, les algorithmes cryptographiques ont subi des processus de rétro-ingénierie. Le pire des scénarios, c'est quand des systèmes n'ont aucune mesure de chiffrement et des mécanismes d'authentification faibles ou inexistants. Même la mise en œuvre des protocoles de cryptographie est un défi important pour l'équipement mobile qui n'a pas les capacités d'un ordinateur personnel ou *Personal Computer* (PC) [3]. En effet, dans beaucoup de cas, les systèmes conçus n'étaient pas exploitables [4]. Pour les systèmes d'information qui arrivent à garantir un niveau acceptable de sécurité, la mise en œuvre et l'exploitation engendrent des coûts importants. De plus, on se retrouve dans l'environnement aujourd'hui avec un ensemble de plateformes offrant les mêmes services, mais sans interopérabilité entre elles. Il y a dans ce même ordre d'idée, un manque de standard dans cette industrie [5].

Tous les points mentionnés plus haut relatifs au transfert électronique de fonds concernent les plateformes fonctionnant selon un modèle centralisé où il y a une entité de confiance qui valide les transactions. Depuis 2008, un chercheur ou un groupe de chercheurs sous le pseudonyme de Satochi Nakamoto [6], a proposé un nouveau système de transfert électronique de fonds (le Bitcoin) dont la validation des transactions s'appuierait sur des protocoles cryptographiques. Cette

proposition semble vouloir rendre le paiement complètement anonyme, ce qui reviendrait à dissocier l'instrument de paiement de l'identité du tenant [7], propriété qui est recherchée pour le respect de la vie privée. Toutefois, cette caractéristique même peut poser un problème en cas de résolution de litiges. De plus, comme ce système crée une monnaie virtuelle, il peut conduire à la déstabilisation de l'économie en y ajoutant d'autres monnaies indépendantes des systèmes conventionnels [7].

Des avantages à considérer viennent de la technologie qui supporte le fonctionnement du Bitcoin. Elle a des caractéristiques qui la rendent attractives pour la résolution éventuelle des problèmes susmentionnés. D'abord, elle est décentralisée et tend vers la transparence; ce qui pourrait être un atout pour l'interopérabilité. En plus, le recouvrement en cas de panne est rapide. Et finalement, elle est résistante à la falsification et met bien à profit l'utilisation de la cryptographie asymétrique.

Il devient donc motivant de chercher à exploiter une telle technologie en vue de concevoir un système pour le transfert électronique des fonds. Mais, on se demande comment l'utiliser? Comment le concilier à un système conventionnel? Quels sont les défis à relever pour une exploitation réelle? Cette proposition de thèse essaiera de formuler des objectifs de recherche en vue de trouver éventuellement des réponses.

1.1 Définitions et concepts de base

Dans ce chapitre, nous précisons les sens assignés dans ce texte aux expressions utilisées se rapportant aux services du commerce électronique. Un accent particulier est mis sur les technologies qui forment les plateformes technologiques sur lesquelles ces services sont déployés.

1.1.1 Le commerce électronique

La définition adoptée dans ce texte se rapproche de celle considérée par Mostafa Hashem Sherif [7] et empruntée de l'AFCEE (*Association Française pour le Commerce et les Échanges Électroniques*). Le commerce électronique englobera toute relation d'échange de biens, de services physiques ou virtuels (logiciels, informations, etc…), utilisant des infrastructures technologiques constituées de matériels informatiques, de logiciels et de réseaux de communications [7].

Elle inclut évidemment le cas où un appareil mobile est utilisé dans le processus : le commerce électronique mobile ou simplement commerce mobile. Selon cette définition, le

commerce mobile fait référence non seulement au transfert de fonds représentant les frais de paiement pour des biens et services, mais aussi à tout transfert de fonds entre des participants sans tenir compte d'un échange de biens ou de services y relatifs. Cet aspect sera le plus souvent désigné par l'expression « paiement mobile ».

Une autre catégorie du commerce mobile vu selon la définition précédente est l'ensemble des services de la banque traditionnelle, disponible via un appareil mobile: ce sera dans ce texte, la banque mobile.

1.1.2 Les systèmes de paiement mobile

Dans cette partie du document, nous explicitons la structure typique de la plateforme de paiement mobile. Telle que présentée à la Figure 1.1, elle prend en compte et le commerce de proximité et le paiement à distance. Ensuite, nous décrivons les principales technologies de communication de ces systèmes.

1.1.2.1 Définitions et structure

Les services de commerce mobile sont conçus et déployés sur une plateforme technique supportant des services mobiles, c'est le système de paiement mobile (*SPM*) qui sera également désigné par Plateforme de Paiement Mobile (*PPM*). Il est constitué de matériels informatiques, de logiciels, interconnectés par des réseaux de communication et dont le fonctionnement est régi par des protocoles bien définis [8], [3]. La figure suivante donne une représentation possible de cette plateforme pour des scénarios du commerce mobile de proximité ainsi que le paiement à distance.

1.1.2.2 Les technologies de communication dans le commerce local

Cette sous-section est une exposition des principales technologies de communication retrouvées dans le commerce mobile de proximité.

Bluetooth Low Energy

En [9], nous trouvons une explication des fondamentaux du BLE (*Bluetooth Low Energy*), défini dans la version 4.2 des spécifications sur le Bluetooth classique; c'est un standard de communication radio sur une distance ne dépassant pas 100 m et utilisant la fréquence de 2.4 GHz de la bande ISM (*Industrial, Scientific and Medical*), pour transmettre sur des canaux de 2 MHz de largeur. La technologie est supportée même par des téléphones moyens de gamme [8], donc

assez accessible. Le BLE a été conçu surtout pour diminuer la consommation d'énergie par rapport aux premières spécifications du *Bluetooth*; certaines autres différences avec le *Bluetooth* classique incluent le nombre de canaux utilisés, certains canaux dédiés aux alertes (ce qui permet de diminuer le temps d'échanges à quelques ms) et surtout au niveau des mécanismes de sécurité.

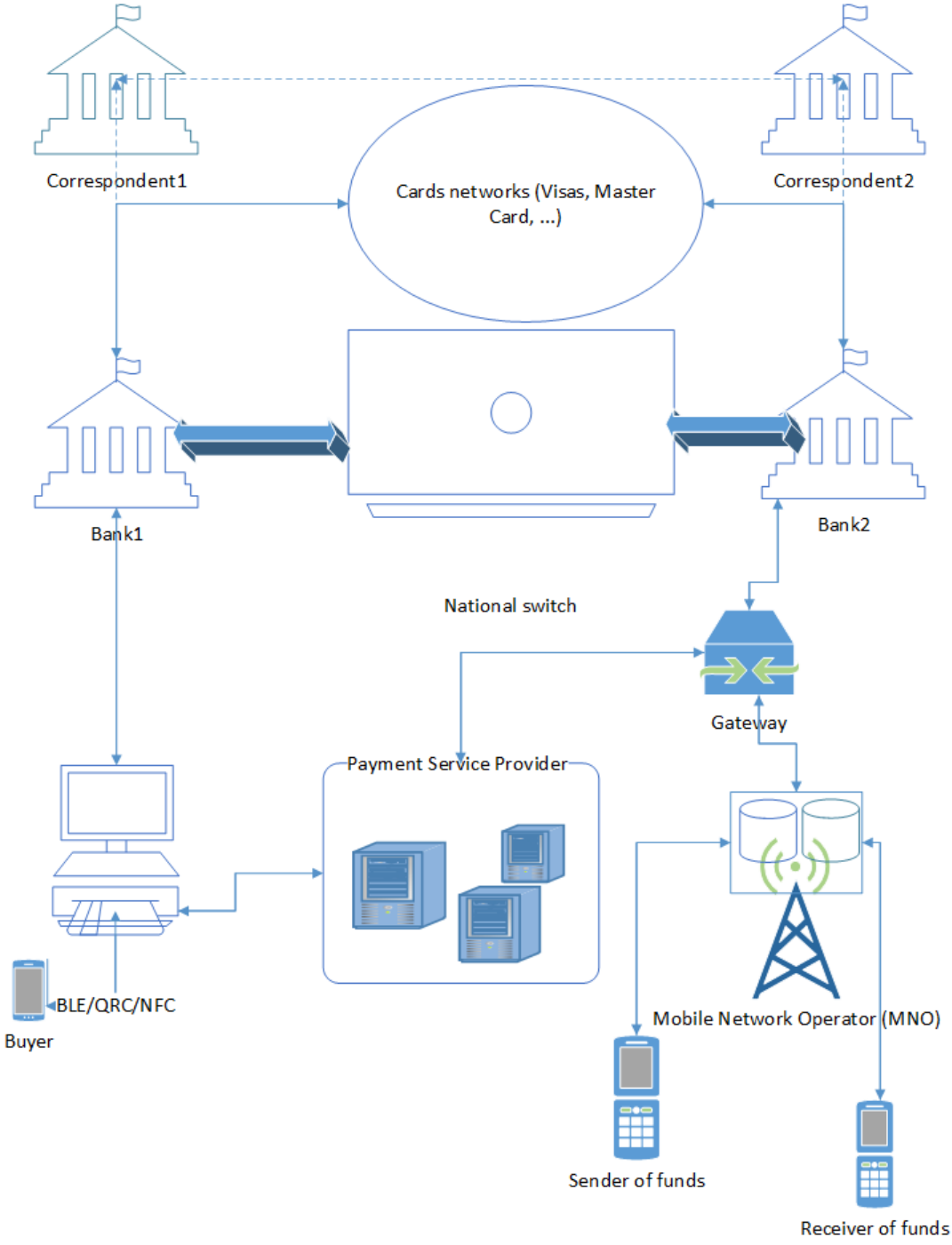


Figure 1.1 Une vue globale de la plateforme de paiement traditionnelle

Avec le BLE, des propriétés de sécurité sont garanties et mises en place au moment de l'établissement de la connexion entre les 2 équipements, le *pairing* [9]:

- La première phase de ce processus est l'échange des requis de connexion tels les modes d'authentification, les clés à s'échanger et leurs tailles, le mode de *pairing* (sans mot de passe ou *Just Works*, comparaison numérique, saisie de clé de passe, canal extérieur ou *Out of Band*).
- Au cours de la phase 2, le protocole *ECDH P-256* est utilisé pour établir une clé de chiffrement pour la session (*LTK pour Long Term Key*) : chaque équipement génère une paire de clés, la clé publique de chacun est envoyée à l'autre; ainsi, l'équipement i *Ei* peut déterminer LTK avec *P256(clé privée de Ei, clé publique reçue de l'autre)*. À l'étape d'authentification, des valeurs de confirmations sont échangées entre les équipements. Le processus s'appuie sur l'algorithme *CMAC* utilisant des clés *AES* de 128 bits.
- La phase 3 est à priori optionnelle; pendant cette phase, les clés de résolution d'identité (*IRK pour Identity Resolution Key*) et de résolution de signature de connexion (*CSRK pour Connection Signature Resolution Key*) sont distribuées entre les participants.

Avec ce processus, les propriétés de confidentialité, d'intégrité et d'authentification peuvent être garanties. Il faut pour cela que le service requiert des modes de connexion adéquats puisque, la connexion peut s'établir à des modes sans connexion (*Just Works*) ou sans signature, c'est-à-dire sans utilisation des clés IRK et CSRK pour l'authentification; Des recommandations du NIST (*National Institute of Standards and Technologies*) sont nécessaires en ce sens [10]: l'authentification des appareils et le chiffrement doivent être exigés, les clés doivent être la plus longue possible selon le standard, des mécanismes de sécurité doivent être appliqués au niveau de l'application et les utilisateurs doivent être éduqués pour contrer l'ingénierie sociale.

Quick Responce Code

Un code à barre à 2 dimensions (2D) est un ensemble de caractères alphanumériques encodés sous une forme spéciale de carrés disposés verticalement et horizontalement [3]; normalement ce code est lu par une machine. Le QRC (*Quick Response Code*) est l'une des dernières versions et des plus populaires des codes 2D. Il permet de stocker jusqu'à 4296 caractères et est utilisé dans les systèmes de paiement mobile.

Near Field Communication

La NFC (*Near Field Communication*) est une technologie basée sur le principe de couplage d'induction magnétique où un canal de communication radio (sur la fréquence de 13.56 MHz) est créé entre équipements dotés de circuits adéquats (tags, lecteurs et téléphones intelligents) qui se rapprochent d'une dizaine de centimètres environ; cette communication peut s'établir selon 3 modes d'opération [11], [12] : le mode *reader/writer* ou un équipement actif, c'est à dire disposant de sa propre énergie, peut lire/écrire des informations de/sur un autre qui est passif (un tag NFC par exemple), le mode *pair-à-pair* où les deux équipements en contact sont actifs, le mode d'émulation de carte dans lequel l'un des équipements a le logiciel et le matériel lui permettant de se comporter comme un circuit intégré passif sans contact ou un tag et l'autre peut lire ces informations.

Pour un scénario typique dans le commerce mobile, le téléphone de l'utilisateur est doté de toute la technologie lui permettant de se comporter comme une carte intelligente sécurisée dont les informations peuvent être lues par le lecteur NFC d'un marchand : l'interface de communication sans contact, le NFC CLF (*NFC Contactless Front-End*) constitué de l'antenne et du contrôleur NFC responsable de la gestion de l'émission/la réception des signaux incluant la modulation/démodulation; le contrôleur est connecté à l'élément sécurisé (*SE pour Secure Element*) à travers le SWP (*Single Wire Protocol*) et au système d'exploitation hôte via une autre interface, la NFC Controler Interface (Figure 1.2) [11]. Il faut souligner que la figure 2 présente différents cas de déploiement du SE qui est une puce offrant un espace de stockage aux données sensibles, résistant à la falsification et pouvant exécuter des fonctions cryptographiques. Il peut être conçu avec le matériel par le manufacturier, fourni par un opérateur de téléphonie mobile dans une SIM (*Subscriber Identity Module*) ou être implémenté dans le cloud sur des serveurs adéquatement configurés.

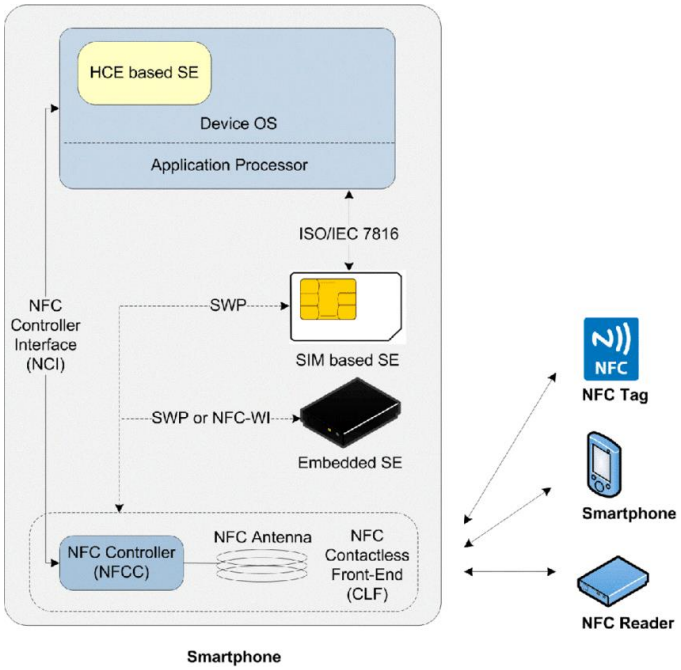


Figure 1.2 Configurations possibles du SE pour un mobile NFC [11]

1.1.2.3 Les technologies de communication dans le paiement à distance

De nos jours, les plateformes offrant le service de paiement à distance utilisent surtout les protocoles de l'internet, le service de message courts ou Short Message Service (SMS), les données de service supplémentaire non structurés ou *Unstructured Supplementary Service Data* (USSD). Nous donnons dans les lignes qui suivent une description détaillée des services SMS et USSD disponibles sur les réseaux GSM (*Global System for Mobile Communication*).

Le SMS est un service de messagerie allant jusqu'à 160 caractères alphanumériques en mode asynchrone. Une description de l'architecture de ce service est donnée dans [13] et illustrée à la Figure 3 : l'utilisateur forme le SMS et l'envoie à partir de son MS (Mobile Station) vers la BTS (*Base Transciever Station*) utilisant le protocole OTA (*Over The Air*); par le BSC (*Base Station Controler*) et le MSC (*Mobile Switching Center*), le BTS fait suivre le SMS à son SMSC (*SMS Center*) avec SS7 (*Signaling System No. 7*). Utilisant SS7, le SMSC fait suivre le SMS au BTS destinataire. Ce dernier l'envoie au MS encore appelé SME (*Short Messaging Entity*) via OTA. Le SMSC peut avoir à effectuer des taches de reformatage selon le protocole SMPP (*Short Message Peer to Peer*) et transmettre vers d'autre SMSC dans le cas où le réseau du destinataire est différent.

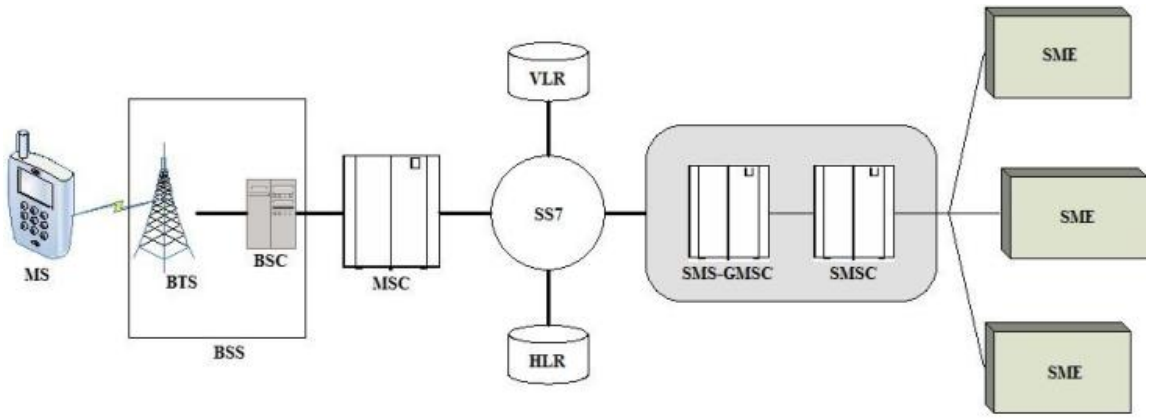


Figure 1.3 Architecture pour le service SMS [13]

L'USSD est un service de messagerie synchrone dans les réseaux GSM, allant jusqu'à 184 caractères alphanumériques; Un message USSD est une session de communication entre un MS et un serveur d'application hébergeant un service USSD disponible sur le réseau [13], [14]. L'architecture du service USSD se rapproche de celle du SMS avec des composants et protocoles du réseau utilisé pareils : une passerelle qui transforme les messages via SS7, les transfère à des applications utilisant SMPP, ces applications elles-mêmes peuvent communiquer avec d'autres serveurs de contenus externes à travers des protocoles connus tels TCP/IP. Certaines différences d'avec le service SMS peuvent être notées [13 - 14] :

- Le nombre de caractères permis par USSD pour le message est supérieur;
- Une session est établie entre le serveur et le mobile, pas de stockage à un nœud, i.e., le SMSC pour le service SMS n'est plus nécessaire pour USSD;
- Avec USSD, des menus conçus par le fournisseur de service peuvent être mis à disposition de l'utilisateur pour faciliter les entrées, ce qui n'est pas le cas avec le SMS;
- Généralement, un coût est associé au service SMS par l'opérateur du réseau ou *Mobile Network Operator* (MNO).

1.1.2.4 La Blockchain et les crypto-monnaies

Le fonctionnement de la blockchain a été initialement décrit par Satochi Nakamoto [6] qui l'a conçu pour le Bitcoin comme un réseau permettant des échanges pair-a-pair sans une entité de confiance. Chacune des transactions entre pairs est validée par un protocole de consensus utilisant à profit la cryptographie pour établir la confiance. Chaque nœud de ce réseau participe à l'exécution

des algorithmes de ce protocole et stockera les transactions par blocs liés entre eux (Figure 1.4) de manière à empêcher la modification des transactions passées [6], [15], [16]. D'où l'appellation « chaîne de blocs ». Ainsi, l'expression peut désigner cette chaîne de blocs, mais aussi et surtout la technologie qui est l'ensemble constitué du réseau pair-a-pair, du protocole de consensus, et les transactions stockées sur les nœuds du réseau. Pour une application de la Blockchain, un utilisateur est généralement défini par une paire de clés cryptographiques qui lui donne accès à ses biens sur le réseau. Son compte sur ce réseau connu de tous est tiré de la clé publique. La clé privée lui permet de disposer des biens qui lui sont destinés. Cette description nous permet de noter quelques grandes propriétés attractives de la Blockchain, telles que la décentralisation qui peut garantir un meilleur recouvrement et sur ce point une plus grande fiabilité qu'un système centralisé, la résistance à la falsification, la garantie de la vie privée dans une certaine mesure, la gestion de la confiance de manière automatique et la traçabilité [16], [17].

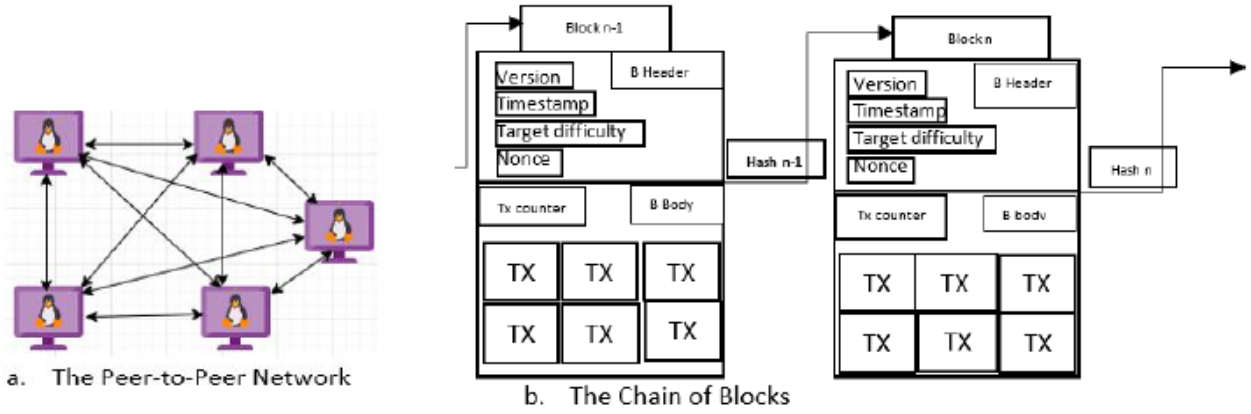


Figure 1.4 Blockchain : le modèle pair-à-pair et la chaine de blocs

Les crypto-monnaies : une application répandue de la Blockchain

Pour les crypto-monnaies (les applications les plus répandues de la Blockchain), nous pouvons d'abord considérer la définition du département du trésor américain applicable à tout type de monnaie virtuelle [18]: « un moyen d'échange pouvant être utilisé comme une monnaie dans certains environnements, mais sans les attributs d'une monnaie réelle telle le statut légal ». Les unités de la monnaie virtuelle sont produites par des algorithmes informatiques; l'une des caractéristiques spécifiques aux crypto-monnaies est l'utilisation de la cryptographie et d'un réseau

décentralisé. Elles valent aujourd'hui, plus de 230.5 milliards de dollars américains où Bitcoin représente plus de 65 % [19].

Les protocoles de consensus : un élément moteur de toute Blockchain

Le protocole de consensus permet de se mettre d'accord sur la validation des transactions soumises. L'un des premiers est la preuve de travail ou *Proof of Work* (PoW) [6]. Ce protocole permet de choisir un nœud pour valider les transactions sur la base d'un travail fourni par ce dernier. D'autres protocoles considèrent des propriétés différentes pour fournir ce privilège [20], [21], [22] : preuve d'enjeux ou *Proof of Stake* (PoS), preuve d'importance ou *Proof of Importance* (PoI), preuve de réputation ou *Proof of Reputation* (PoR), pour n'en citer que ceux-là. Plus tard, la résolution du problème des généraux byzantins [23] a donné naissance à des protocoles supportant des systèmes tolérants aux fautes byzantines ou *Byzantine Fault Tolerance* (BFT), comme le *Practical Byzantine Fault Tolerance* (PBFT) [24].

1.2 Éléments de problématique

Les plateformes de commerce mobile sont diverses par leurs composants, les technologies de communication utilisées et les protocoles d'échange. Mais, ils ont tous pour objectif de faciliter les transferts de fonds entre les comptes de participants et l'accès aux services bancaires. Dans les pays développés, ces systèmes offrent la commodité en permettant à l'utilisateur de se servir de son équipement mobile pour payer ses achats [5]. Dans d'autres régions du monde, ils facilitent l'inclusion financière [1]. La protection des informations confidentielles (cartes de paiement et clé de chiffrement) pendant les échanges et le stockage devient une nécessité; de plus, la facilité de mise en œuvre et d'exploitation de tels systèmes, leurs performances en termes de durée de réalisation de la transaction et l'interopérabilité avec d'autres sont à considérer. Relativement à ces requis, plusieurs enjeux ont été relevés après analyse des systèmes existants.

Dans le commerce de proximité, ces systèmes s'appuient majoritairement sur le BLE, le QRC et la NFC pour la communication entre le mobile de l'acheteur et le système du marchand; la connexion avec des institutions financières et des réseaux de cartes s'établit généralement par les protocoles de l'internet. Les échanges sont normalement sécurisés par la cryptographie. Une infrastructure à clé publique est souvent utilisée et les mécanismes de sécurité bien connus peuvent être adoptés entre le marchand et le système de traitement de transactions. Toutefois, pour le mobile

de l'acheteur, d'autres mécanismes doivent être mis en œuvre. Dans le cas des systèmes s'appuyant sur BLE, les clés de chiffrement des échanges est à risque, car pouvant être copiée par un attaquant qui peut aussi saturer le canal et provoquer conséquemment un déni de service. Il faut donc trouver un moyen de protéger ces clés.

De son côté, le QRC permet la communication entre le mobile du client et le marchand en offrant une grande facilité d'utilisation et une grande accessibilité. Toutefois, aucune mesure ne permet de se protéger contre un code QR compromis qui redirige vers un site malicieux. Ainsi, il faut l'utiliser convenablement et avec d'autres technologies et protocoles adéquats en vue d'avoir un système robuste et commode.

Les systèmes avec NFC permettent de résoudre le problème de stockage de clé via l'élément sécurisé qui protège les données confidentielles contre un accès non autorisé. Toutefois, l'élément sécurisé est fourni par un manufacturier ou un MNO [7], entravant l'interopérabilité de tels systèmes. L'architecture par *Host Card Emulation* (HCE) augmente l'indépendance du SE des fabricants ou vendeurs de cartes, certes; mais reste vulnérable aux attaques par relais et la sécurité du canal entre le SE dans le cloud et le module client doit être renforcée.

Pour le commerce mobile à distance, ces plateformes offrent les services bancaires et de paiement partout et en tout temps. Mais, les informations confidentielles sont souvent partagées avec les marchands. De plus, même au sein des institutions bancaires, des cas de vol de données sont connus récemment [25]. Il devient motivant de chercher à authentifier un utilisateur avec son institution financière sans pour autant garder les informations personnelles fournies parfois par l'autorité publique. Aussi, plusieurs systèmes de paiement mobile ont été déployés dans les pays en voie de développement utilisant le SMS et l'USSD avec les MNOs. Toutefois, il n'y a pas une technologie standard établissant l'interopérabilité entre ces derniers systèmes et d'autres principalement en exploitation dans les pays développés. D'où la nécessité de l'interopérabilité à ce niveau.

Aujourd'hui, la Blockchain supporte des plateformes de services liés au commerce mobile [26]. Cette technologie peut être mise en œuvre de façon complètement décentralisée (blockchain publique), complètement centralisée (blockchain privée) ou partiellement décentralisée sous la direction d'un consortium [16]. Il est préférable d'aller vers la décentralisation pour bénéficier des avantages offerts par une telle structure. Dans ce cas, la gestion des clés pose un problème, puisque

laissée à l'utilisateur et ce dernier n'est généralement lié à sa clé nulle part ailleurs. De plus, les pertes de clé privée engendrent des pertes de biens sur le réseau. Il faut donc trouver le moyen de lier une entité réelle à sa clé, et aussi de pouvoir récupérer cette clé en cas de perte, sans pour autant mettre en péril la garantie de la vie privée.

Un autre point à considérer relatif à la blockchain vient du fait que les grands succès de cette technologie sont autour de la création de crypto-monnaies. Si on envisage de créer un réseau de blockchains qui ne vise pas une nouvelle crypto-monnaie, mais le transfert des monnaies du monde réel, il faut une passerelle entre ce réseau et d'autres systèmes de comptes de paiement tels ceux des institutions financières, ou des réseaux de cartes.

Par ailleurs, plusieurs protocoles de consensus sur lesquels se basent les blockchains sont inadaptés pour des offres de services commercialisables à cause du problème d'évolutivité et de manque de contrôle entre autres. Il est difficile de concevoir des services qui impliqueraient les banques et d'autres institutions financières notables sur un réseau complètement ouvert, sans aucune capacité de restreindre certains accès en cas de constats de fraudes. D'un autre côté, les protocoles des blockchains à permission tels le PBFT et d'autres variantes du BFT [27] ont une tendance à la centralisation en exigeant la connaissance de tous les nœuds validateurs. Ainsi, il y a nécessité d'avoir un protocole de consensus offrant cette possibilité de restrictions d'accès, mais aussi assez de flexibilité pour garantir la décentralisation. Le protocole PoI semble permettre cette flexibilité; mais il n'y a pas encore une évaluation du niveau de sécurité garantie. Il devient donc nécessaire de concevoir pour la blockchain un protocole de consensus visant simultanément les propriétés de sécurité, le respect de la vie privée et une bonne qualité de service ou *Quality of Service* (QoS).

Les éléments de problématique dans les systèmes de paiement mobile tels que notés précédemment, nous conduisent aux questions de recherche suivantes :

- Peut-on parvenir à lier sans équivoque l'identité d'une entité réelle – particulièrement un utilisateur humain d'un système basé sur la Blockchain – à sa clé privée de manière à pouvoir régénérer cette clé en cas de perte, tout en évitant de la stocker chez une tierce partie?

- Est-il possible de définir un protocole d'échange entre les composants d'une plateforme de paiement, tel que, aucun participant n'ait accès à plus d'informations sur l'utilisateur que celles, qu'il a lui-même soumis dans le processus?
- Quel protocole de paiement envisager pour un système de paiement basé sur la blockchain, compte tenu du fait qu'il faut avoir une bonne performance en garantissant du même coup la sécurité du système?
- Est-il possible de permettre le transfert de fonds de manière transparente entre des utilisateurs de systèmes de paiement mobile différents, interconnectés à travers un réseau de blockchain?

Ces nombreuses questions relatives à la conception d'une architecture basée sur la Blockchain et supportant des services liés au commerce mobile nous portent à définir des objectifs de recherche.

1.3 Objectifs de recherche

L'objectif principal de cette thèse est de concevoir une architecture pour des services de commerce électronique, basée sur la blockchain, garantissant les propriétés de sécurité et du respect de la vie privée et assez flexible pour permettre l'interopérabilité entre des systèmes de paiement différents.

De manière plus spécifique, nous visons à :

- Concevoir un mécanisme de sécurité fiable, destiné à être implémenté dans un type de portefeuille électronique permettant de lier une identité sur la blockchain à une entité physique, et ainsi permettre l'authentification de cette entité, en prenant en compte les informations fournies par l'autorité publique;
- Définir un protocole de paiement sécurisé, s'appuyant sur le mécanisme d'authentification précédent, pouvant être implémenté sur un équipement léger permettant le transfert de fonds entre des comptes du système en question, tout en garantissant le respect de la vie privée;
- Proposer un cadre de conception d'architecture basée sur la blockchain pour des services de commerce électronique;
- Concevoir un modèle concret de système basé sur la blockchain pour des services de commerce électronique selon le cadre proposé;

- Implémenter et évaluer un sous-ensemble de l'architecture proposée pour offrir des services de paiement mobile.

1.4 Principales contributions de la thèse

Plusieurs enjeux étaient à considérer pour concevoir l'architecture de services liés au commerce électronique respectant les contraintes de l'écosystème financier tout en bénéficiant des avantages de la blockchain. D'abord, il y a la nécessité de connecter de manière transparente l'identité de l'utilisateur sur la blockchain avec ses caractéristiques dans la réalité. Ensuite, il fallait respecter la vie privée des utilisateurs du nouveau système. Finalement, se posait le besoin d'interopérabilité entre les plateformes internes à des institutions financières et la nouvelle architecture. La thèse aborde les problèmes susmentionnés en proposant différentes contributions pour les résoudre :

1. Une nouvelle méthode de codification, un appariеur biométrique et une fonction de dérivation de clé (KDF):

Le premier défi auquel la thèse s'attaque est relatif à la gestion des clés de l'utilisateur. En effet, dans d'autres infrastructures de sécurité basées sur la cryptographie, il est courant de s'appuyer sur une entité de confiance pour le recouvrement de ses données de connexion, notamment la clé privée. Avec le paradigme de la blockchain, l'utilisateur est rendu plus responsable de la gestion des clés. Aujourd'hui, la clé est souvent générée par un processus aléatoire rendant impossible la récupération des comptes en cas de perte. Une façon élégante de résoudre ce problème est de générer la clé à partir des données biométriques de l'utilisateur de façon à pouvoir la recréer ultérieurement. Cette voie comporte des défis de taille liés à la variabilité intra-utilisateur qui va à l'encontre de la reproductibilité de la clé. Pour surmonter ce défi, nous avons conçu une fonction permettant de faire correspondre l'ensemble des valeurs possibles d'une caractéristique biométrique à une valeur définie, formant ainsi un code pour l'utilisateur à partir du vecteur biométrique. La thèse ajoute à ce procédé de codification un mécanisme d'apprentissage qui permet de capter une esquisse minimale des données biométriques pour aider à la reproduction du code. Se basant sur cette codification, des algorithmes intelligents forment un appariеur biométrique simple et efficace pour l'authentification.

Une autre contribution de cette partie de la thèse est une fonction de dérivation de clé à partir des caractéristiques biométriques. Nous avons conçu cette fonction de dérivation de clé basée sur HMAC ou *HMAC-based Key Derivation Function* (HKDF) en intégrant comme données

d'entrée le code biométrique précédent et un identifiant de l'utilisateur. Ce qui nous permet d'obtenir une clé cryptographique révocable, irréversible avec une chaine aléatoire. Aucun des générateurs de clé à partir des données biométriques trouves dans la littérature ne permet de bénéficier de ces trois propriétés simultanément.

2. **Un protocole de paiement basé sur la blockchain garantissant le respect de la vie privée** :

L'un des défis majeurs auquel fait face l'offre des services de commerce électronique sur une blockchain est le manque de respect de la vie privée de l'utilisateur. En effet. La blockchain a été conçue avec l'idée de rendre disponible les transactions sur tous les nœuds validateurs de transactions. Malgré c'est un pseudocode, qui identifie l'envoyeur ou le receveur de fond sur le système, la traçabilité de ses transactions permet de découvrir son identité réelle. Les systèmes qui s'attaquent à ce problème s'appuient sur des protocoles cryptographiques. Dans cette thèse, nous avons proposé un protocole de paiement qui n'ajoute pas d'opérations cryptographiques supplémentaires à celles nécessaires au fonctionnement de la blockchain. Le mécanisme pour garantir la vie privée n'augmente pas le coût en mémoire ou en communication du processus de paiement.

3. **Un cadre et un modèle pour des architectures de services de commerce électronique basées sur la blockchain** :

   La conception d'une architecture basée sur la blockchain fait évidemment face aux défis de cette technologie, notamment celui de concevoir un protocole de consensus qui offre un bon niveau de sécurité, une performance acceptable, sans des limitations entravant l'utilisabilité. La considération de ces enjeux a permis une taxonomie des différents protocoles de consensus existant. L'originalité de cette taxonomie tient dans la mise en lumière des classes de protocoles eu égard aux modes de fonctionnement.

   De plus, l'interopérabilité entre les systèmes internes aux différentes institutions de paiement demeure un problème connu dans l'écosystème financier. La blockchain a certes le potentiel pour faciliter cette interopérabilité, mais les architectures proposées sont plutôt autonomes, séparées des plateformes traditionnelles et ne permettent pas toujours le respect des règles légales en vigueur. C'est pourquoi cette thèse propose **un cadre** qui définit des

architectures basées sur la blockchain avec un couplage faible entre les plateformes des institutions financières.

En suivant le cadre proposé, nous avons dans cette thèse, conçu une architecture avec les composants précédents pour offrir le commerce mobile. Cette conception montre l'intégration de la fonction de dérivation de clé, du protocole de paiement décentralisé, de la blockchain Ethereum et sa connexion avec la plateforme de paiement électronique traditionnelle pour former un système décentralisé offrant des services de commerce électronique. Ainsi, **ce modèle** permet de garder les avantages de la blockchain, notamment la décentralisation, tout en respectant des principes établis de l'écosystème financier.

4. **La visualisation d'un protocole de consensus adéquat pour des paiements interbancaires** :

Cette contribution est une spécification des modifications à introduire dans le protocole de consensus stellaire ou Stellar Consensus Protocol (SCP) pour améliorer sa sécurité. SCP est un protocole de type FBA qui permettrait d'avoir une meilleure performance que PoW ou d'autres. Les spécifications pour une extension du SCP trouvent sa motivation dans l'ensemble des limitations du protocole, dont la définition statique des nœuds de confiance. Nous avons visualisé dans cette thèse, l'intégration intelligente de la notion d'importance et de réputation de chaque nœud du réseau pour permettre à un participant de choisir sur une base dynamique les autres participants de confiance.

1.5 Plan de la thèse

Dans ce chapitre, nous avons défini les différents concepts liés au commerce électronique, analysé les principaux systèmes offrant les services relatifs et leurs rapports avec la blockchain, puis tiré de cette analyse des éléments de problématique. Ce qui nous a permis de définir des objectifs de recherche dont l'atteinte a donné naissance à des contributions de cette thèse, énoncées également dans ce chapitre. Nous allons organiser le reste de ce document de la manière suivante. Dans le chapitre 2, nous exposerons une revue de la littérature autour du sujet. Le chapitre 3 est une présentation de la méthodologie de l'ensemble de la recherche: il s'agit surtout de mettre en relation les objectifs de recherche énoncés à la section 1.3 et les articles scientifiques résultant de cette thèse.

Le chapitre 4 présente le texte intégral d'un article intitulé « *Mobile Payment Platforms : Taxonomy, Architecture, Security, and Integration with Blockchain* », publié dans la revue « *Academia Letters* ». Cet article propose une analyse des principaux systèmes de paiements mobiles permettant de mettre en lumière les points suivants. D'abord, une classification entre des plateformes de paiement électronique se basant sur la proximité entre l'expéditeur et le récepteur de fonds. Cela a permis de classer les plateformes dans le commerce local et ceux qui sont dans le paiement à distance. Ensuite, l'analyse des différents composants de ces deux types de plateformes et des technologies de communication utilisées, a permis de mettre en lumière les traits caractéristiques de chaque type d'architecture, des enjeux de sécurité et de respect de la vie privée liés à ces plateformes. Nous avons également analysé la blockchain et mis en lumière certaines limitations de cette technologie en ce qui a trait à l'offre des paiements mobiles. Finalement, nous avons indiqué certaines directions de recherche en vue de trouver des solutions aux problèmes mentionnés.

Le chapitre 5 présente le texte intégral d'un article intitulé « *Privacy-Preserving Model for Biometric-based Authentication and Key Derivation Function* », soumis pour publication dans la revue « *Journal of Information Security and Applications* ». Cet article propose un mécanisme d'authentification biométrique et une fonction de dérivation de clé ou *Key Derivation Function* (KDF). Cette dernière extrait une clé cryptographique à partir de données biométriques. Notre nouvelle approche, (basée sur des algorithmes de regroupement pour créer une méthode) détecte des caractéristiques cohérentes et distinctives à partir de celles qui sont extraites pour créer un code pour l'utilisateur. Pour reproduire le code à l'avenir, ce qui sert également à l'authentification, notre fonction de dérivation de clé stocke les données auxiliaires en garantissant le respect de la vie privée. Nous suivons les normes de l'IETF pour concevoir une HKDF, en vue de générer une clé cryptographique à partir de ce code.

Le chapitre 6 est le texte intégral de l'article intitulé « *Blockchain and Mobile Payment : Assessment on Privacy and Usability and a Scheme for Enhancement* », publié dans la conférence, « *2022 Fourth International Conference on Blockchain Computing and Applications (BCCA)* ». Cet article aborde deux points principaux. Tout d'abord, il présente une analyse des architectures basées sur la blockchain pour les paiements mobiles, en évaluant les problèmes de respect de la vie privée et de performance qui y sont associés. Ensuite, il propose un protocole de paiement pour le commerce local qui garantit le respect de la vie privée à la fois de l'acheteur et du commerçant. Ce

protocole assure qu'un nœud participant n'a accès qu'à une partie des données significatives d'une transaction, préservant ainsi la confidentialité des informations sensibles. Le document discute également des astuces pour améliorer les performances d'un tel schéma en utilisant la blockchain Ethereum. De plus, la faisabilité de la mise en œuvre de ce schéma, tant sur des blockchains publiques que sur des blockchains de consortium, a été analysée.

Le chapitre 7 est le texte intégral de l'article intitulé « *Framework and Architecture for Blockchain-based e-Commerce Services* », soumis pour publication dans la revue, « *IEEE Access* ». Cet article établit un cadre de conception pour une architecture de paiement électronique basée sur la blockchain qui tienne compte des limites de cette technologie. Il analyse d'abord la structure du système de paiement électronique traditionnel d'une part, et d'autre part les caractéristiques de la blockchain. Ensuite, il effectue une comparaison des principaux protocoles de consensus afin de proposer une taxonomie de ces protocoles. Dans cette conception d'architecture basée sur la blockchain, quatre éléments essentiels sont intégrés : (1) une fonction de dérivation de clé (KDF) qui est un mécanisme léger de récupération de clé pour les systèmes basés sur la blockchain, (2) un appariteur biométrique, (3) un protocole de paiement décentralisé pour garantir la privacité de l'utilisateur, et (4) la blockchain. L'article examine la faisabilité de mettre en œuvre cette architecture en utilisant la blockchain Ethereum. Il propose finalement des spécifications pour une extension du protocole de consensus stellaire (SCP) en vue d'améliorer la performance avec cette architecture.

Au chapitre 8, nous discutons de la méthodologie de cette recherche et des résultats obtenus. Finalement, le chapitre 9 conclut la thèse en recapitulant sur les contributions de la recherche, en exposant les limitations des travaux et en indiquant des directions de recherche futures.

CHAPITRE 2 REVUE DE LITTÉRATURE

Le transfert électronique de fonds en environnement mobile peut être représenté selon 2 grandes tendances [4] : Une tendance vers des systèmes dédiés au commerce mobile de proximité et un autre pour les services de commerce mobile à distance. Les plateformes de la première tendance peuvent être illustrées selon l'architecture de la Figure 2.1. Le canal de communication entre l'acheteur et le marchand est une technologie de communication à faible portée telle le BLE, le QRC, la NFC pour ne citer que les plus répandues. Les plateformes de la seconde tendance ont deux sous-embranchements : un premier avec des opérateurs de téléphonie mobile ou *Mobile Network Operator* (MNO*)* – Figure 2.2 – et rencontré dans les pays en voie de développement, le second concerne les transferts de fonds via l'infrastructure bancaire et à l'international : principalement des pays développés vers des pays moins avancés. Nous allons exposer dans les lignes qui suivent, les caractéristiques et défis de ces différents types de plateformes de paiement mobile.

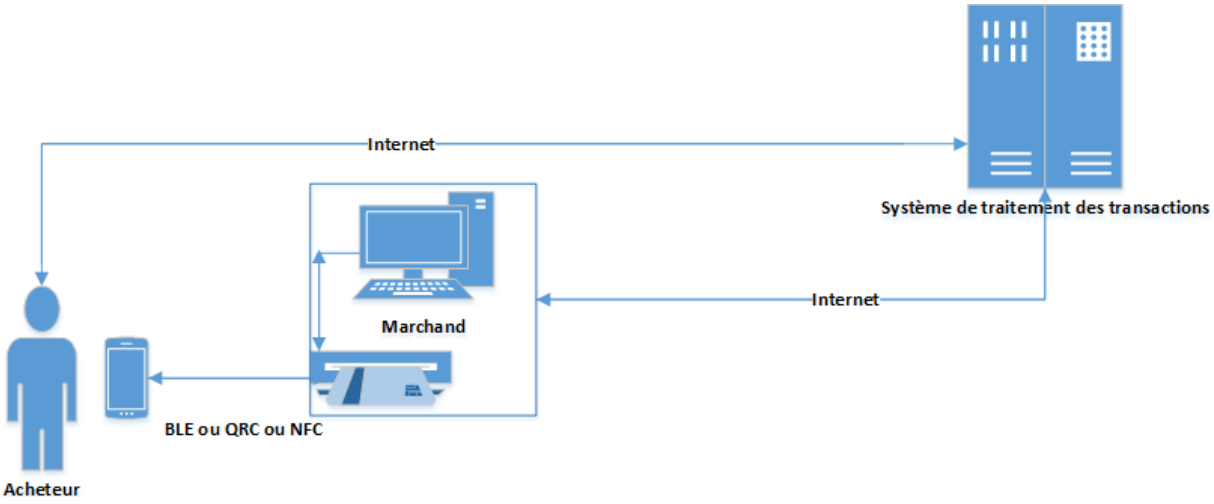


Figure 2.1 Vue globale des systèmes du commerce mobile de proximité [4]

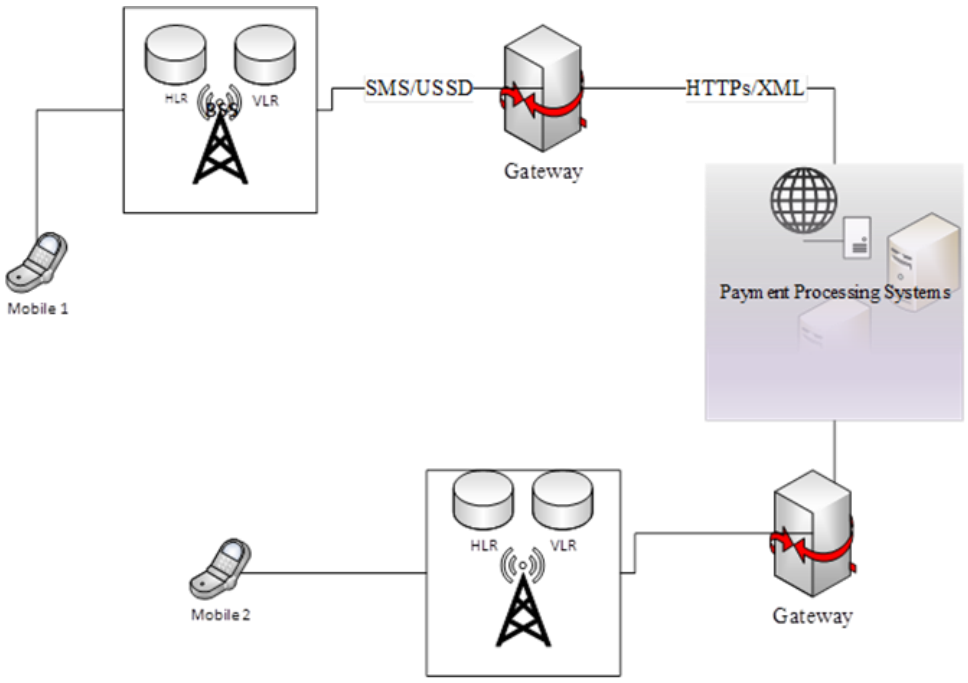


Figure 2.2 Des systèmes de transfert à distance avec des MNOs [4]

2.1 Analyse sommaire du problème

Une plateforme de paiement mobile a pour objectif ultime de permettre la réalisation convenable de l'échange de biens et d'affectation des comptes de la façon la plus sécuritaire et la plus commode possible. La technologie de paiement mobile quoique déjà très répandue dans l'industrie [5], fait face à des menaces et des risques qui sont documentés dans la littérature scientifique [5], [2]:

- Les cyberattaques telles que le reniflage, le pourriel, la mystification, l'hameçonnage, le dévoiement et les programmes malveillants;
- Des cas de fraude résultant d'une mauvaise utilisation de l'équipement ou de vol de cet équipement;
- Des vulnérabilités inhérentes aux applications et au système d'exploitation.
- D'autres attaques connues comme l'homme du milieu ou *Man In The Middle* (MITM), le déni de service ou *Denial of Service* (DoS) entre autres.

Ainsi, un premier défi des systèmes de paiement mobile est la protection contre ces différentes attaques, en garantissant les propriétés de sécurité : authentification des participants, confidentialité des échanges, intégrité des données, non répudiation d'une transaction, traçabilité, fiabilité du système et on peut même y inclure la performance dans l'offre de services.

En plus, avec le grand nombre de participants dans la plupart de ces systèmes (banques, marchands, opérateurs de téléphonie mobile, réseaux de cartes, fournisseurs de services indépendants, manufacturiers) [2], la garantie du respect de la vie privée devient un autre défi important pour les PPM.

La commodité quant à elle, fait référence ici à la facilité d'utilisation. Elle inclut la facilité de mise en œuvre, l'installation du système, la capacité du système à être exploité avec des équipements répandus et peu couteux, le degré de dépendance du fournisseur de service par rapport à d'autres entités. Ce dernier aspect ouvre sur l'interopérabilité qui est la capacité de coexistence et de coopération entre des systèmes basés sur des technologies différentes sans être intrinsèquement liés [28]; elle est la possibilité de règlementer des transactions entre deux systèmes différents.

Plusieurs solutions à ces divers défis ont été proposées, elles se différencient principalement par les composants du système, les technologies de communication et les protocoles utilisés. Les différentes tendances, classifiées selon les technologies de communication sont présentées dans les sections à venir.

2.2 Systèmes et technologies dans le commerce de proximité

Dans la plupart des systèmes du commerce mobile de proximité, le module de l'acheteur entre en contact avec celui du vendeur via la NFC ou le QRC ou le BLE. Acheteurs et vendeurs communiquent avec le système de traitement des transactions via l'internet. Ce dernier sous-système représenté comme un bloc dans la Figure 2.1 est typiquement une interconnexion entre les systèmes des banques émettrices et négociatrices de cartes de paiement, des réseaux de cartes (Visa, MasterCard entre autres) et éventuellement des fournisseurs de services indépendants (Apple Pay, Samsung Pay, Google Wallet). Les composants du bloc de traitement des transactions sont généralement interconnectés suivant les protocoles de l'Internet (TCP/IP). Les propositions les plus récentes de tels systèmes sont présentées dans les prochaines sous-sections.

2.2.1 Systèmes de paiement mobile basés sur le BLE

La forme la plus basique d'un système de paiement mobile à proximité avec Bluetooth, a la topologie suivante dont un exemple tiré de [29] est montré à la Figure 2.3:

- Le client dispose d'un téléphone intelligent sur lequel est installé une application mobile pouvant communiquer avec le point-de-vente via BLE et aux systèmes validant son compte via internet.
- Le marchand identifié par le point-de-vente pouvant se connecter au client et à la passerelle de paiement, mais aussi peut disposer d'un système de traitement des transactions.
- Le tiers de confiance ou *Trusted Third Party* (TTP) garantit l'authentification des participants. Ces derniers sont identifiés par des clés gérées par le TTP qui est souvent une autorité de certification.
- Le Serveur de traitement des paiements ou *Backend Payment Processing Server* peut représenter le système de traitement constitué d'institutions financières et de réseaux de cartes. Les institutions financières détiennent les comptes des participants et assurent leurs mises à jour selon le montant échangé.

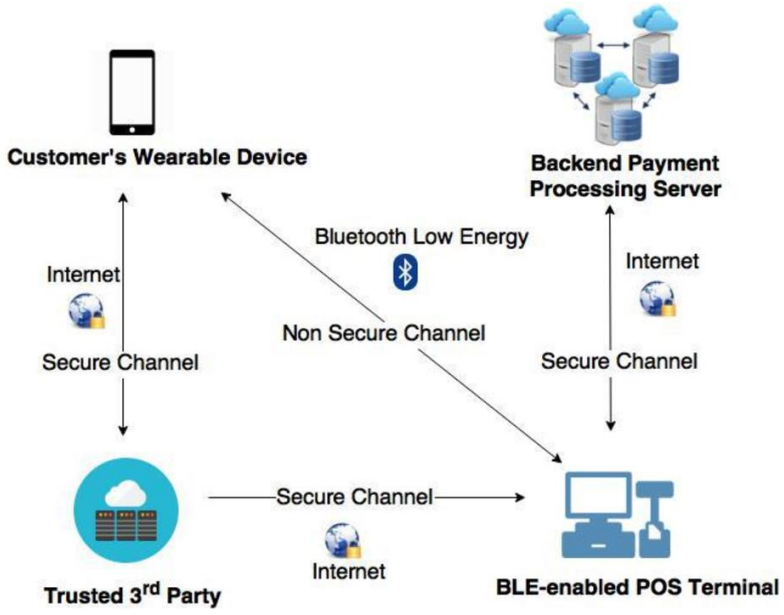


Figure 2.3 Architecture physique d'un système de paiement basé sur BLE [29]

Partant de ce schéma de base, plusieurs systèmes ont été proposés en vue de résoudre les problèmes de sécurité découverts avec l'utilisation du Bluetooth. En général, les systèmes s'évertuent à authentifier l'utilisateur et à sécuriser le canal de communication par une infrastructure à clé publique. Le traitement de la transaction par le système central est typique de tout système de paiement par cartes bancaires.

Le système proposé en [29], un peu différent, utilise un service de localisation intérieure avec des signaux Bluetooth (*Beacons*) pour garantir la sécurité des échanges. Le protocole permet l'authentification de l'utilisateur sans une infrastructure à clé publique. Les Beacons permettent de localiser un client qui, une fois dans la zone de paiement peut entamer le processus de transfert de fonds. Pour l'authentification, le TTP enverra au client et au marchand un secret (construit à partir de la localisation, de l'heure exacte et d'un identifiant du mobile); les deux génèreront une clé de partage à partir de ce secret qui sera utilisée pour l'authentification et le chiffrement des données de la transaction. Le serveur de traitement des transactions utilise la *tokenisation* adoptant ainsi des mécanismes de paiement bien connus.

Il faut noter la nécessité d'avoir un canal sécurisé entre l'entité de confiance et le client. De plus, d'autres problèmes de sécurité liés au BLE, pouvant exposer la clé de session ou amener au déni de service ou *Denial of Service* (DoS) sont bien connus [7], [30]:

- Le *Bluejacking* où l'attaquant peut juste envoyer des messages non sollicités au téléphone de la victime, ce qui peut mener à la saturation du canal de communication;
- Le *Bluesnarfing* où l'attaquant peut accéder au téléphone de la victime sans autorisation; il peut alors copier des informations confidentielles. Ce qui met en péril la propriété de garantie de la vie privée.

2.2.2 Les systèmes de paiement basés sur le QRC

Dans le commerce mobile, le QRC peut contenir une description du produit, un lien pointant vers le site du marchand entre autres [7]. Cette technologie est utilisée pour permettre la communication entre l'acheteur et le vendeur où il suffit de scanner le code sur l'écran d'un appareil avec la caméra de l'autre; plusieurs scenarios sont possibles, typiquement, un portefeuille mobile peut lire le code à un point de vente et l'application peut déduire du compte le montant de la transaction [7]. Plusieurs systèmes de paiement mobiles très répandus ont utilisé le QRC dont *Starbucks* en Amérique du Nord et *Alipay* en Chine.

Des protocoles de sécurité sont conçus (avec comme base la cryptographie) pour sécuriser les systèmes se basant sur QRC comme technologie de communication. Le principe est de chiffrer les informations sensibles, les ajouter à celles qui sont publiques avant d'encoder; ce qui est connu comme le SQRC (*Secure QRC*). Parmi ces protocoles, on retrouve une proposition de Ariana T. P.

où une autorité de confiance garde des informations d'identification des acheteurs et des vendeurs et génèrent une paire de clés temporaires pour un acheteur au moment de la transaction. La clé publique est envoyée au marchand et la clé privée à l'acheteur [31]. Ce système tel que décrit permet la garantie de la confidentialité, l'intégrité et l'authentification si le canal entre l'acheteur et l'autorité de confiance est sécurisé. Ce qui entraine éventuellement l'utilisation d'autres paires de clés. De plus, avec RSA, la performance (résultats d'expériences absents) n'est pas garantie.

Par ailleurs, plusieurs vulnérabilités du QRC ont été documentées, notamment [32]: la falsification du code pouvant détenir un lien pointant vers un site malicieux, un logiciel malveillant qui va s'installer sur le mobile de l'acheteur, des scripts croisés ou *Cross-Site Scripting* (XSS), des commandes d'injections via le langage de requête structuré ou *Structured Query Language* (SQL). Les mesures palliatives pour ces problèmes se basent surtout sur des bonnes pratiques de développement et d'exploitation telles, la validation d'entrées, la vigilance de l'utilisateur pour contrer l'ingénierie sociale.

D'autres aspects notables par rapport au QRC est sa grande utilisabilité. Pour la plupart des systèmes, il suffit à l'un des utilisateurs de scanner le code QR sur l'écran de l'autre. Les téléphones pouvant lire les codes QR sont très répandus impliquant une grande accessibilité aux services. Cependant, on ne sait pas encore comment empêcher la redirection vers un site malicieux lorsque le code QR est compromis [7].

2.2.3 Les systèmes de paiement basés sur la NFC

L'utilisation de la NFC est courante dans les services de paiement mobile. Parmi les systèmes les plus connus dans l'industrie s'appuyant sur cette technologie, on retrouve *Apple Pay*, *Samsung Pay* et *Google Wallet*. Ces systèmes offrent généralement une grande facilité d'utilisation une fois déployés, il suffit d'approcher le mobile à moins de 10 cm du point-de-vente du marchand et la transaction est réalisée. D'où l'appellation commune de paiement sans contact. Le module de l'acheteur inclut le SE pour la protection des données confidentielles. Les architectures proposées dans la littérature se différencient surtout sur la manière d'implémenter le SE.

Les systèmes avec le SE intégré requiert en général une bonne collaboration avec le manufacturier et le fournisseur de services de paiement; la possibilité qu'un opérateur de téléphonie mobile fournisse cet élément sécurisé intégré dans le module d'identité de l'abonné ou *Subscriber*

Identity Module (SIM) existe aussi [7]. Dans un cas comme dans l'autre, on est conduit à la dépendance soit du manufacturier, soit du MNO.

L'un des plus récents systèmes de ce type est la proposition faite en [33]. Ce système se base sur une infrastructure à clé publique ou *Public Key Infrastructure* (PKI) pour garantir les propriétés de sécurité utilisant les algorithmes de la cryptographie par courbes elliptiques ou Elliptic-Curve Cryptography (ECC). Les banques respectives du client et du marchand enregistrent les comptes associés et leur assignent un identifiant. Les échanges (toujours liés aux temps) sont chiffrés avec des clés asymétriques, les utilisateurs authentifiés par des certificats émis par l'autorité. Les auteurs démontrent par des analyses que le système est sécuritaire et performant. Rien n'est dit cependant sur l'implémentation du SE. Ainsi, le problème de dépendance des manufacturiers/MNOs ou de coût de mise en œuvre n'est pas considéré dans ce travail. Le système suppose implicitement que les tiers de confiance (banque et autorité de certification) ne sont pas compromis et un canal sécurisé existe entre ces entités.

D'autres architectures utilisent un SE dans les nuages (*communément désignées par HCE pour Host Card Emulation*) visant à diminuer sa dépendance d'un manufacturier ou d'un MNO [34], [35]; le scénario typique considère l'enregistrement des informations sensibles dans le cloud et envoient un jeton correspondant au terminal lorsque le paiement doit se faire. Turk et *al.* [36] vont jusqu'à clamer une indépendance complète dans une proposition où l'autorité de certification et les fournisseurs de cartes de paiement contactent le fournisseur du terminal ou le MNO pour valider les informations venant du SE. Ce jeton envoyé par le marchand à sa banque sera soumis au réseau de carte de paiement pour vérification et traitement avec la banque émettrice de la carte. Des problèmes avec cette approche : il faut savoir authentifier le mobile connecté pour la transaction et la plateforme est vulnérable aux attaques par rejeu [29], [30].

Par ailleurs, un autre aspect lié au NFC est que les téléphones supportant cette technologie sont les moins accessibles dans certaines régions du monde. Pour cet aspect, BLE et QRC offrent de meilleures perspectives pour la mise en exploitation.

2.2.4 Analyse des plateformes dans le commerce de proximité

Les plateformes supportant le service du commerce de proximité s'appuient sur des entités de confiance (comme les banques), fournisseurs de services financiers; parfois, une autorité de certification indépendante peut participer pour garantir l'authentification. Le réseau d'arrière-plan

ou Backend-Network (BN) utilise l'internet pour la communication entre les nœuds distants; le protocole de sécurité de la couche transport / couche de sockets sécurisés ou Transport Layer Security / Secure Sockets Layer (TLS/SSL) permet de protéger les échanges. Entre les marchands et l'acheteur, les technologies de communication de proximité (QRC, NFC et BLE) sont utilisées. Aussi, pour l'authentification de l'utilisateur un PIN peut être utilisé ainsi qu'un mot de passe de session ou On-Time Password (OTP). Les plus sécurisés utilisent des caractéristiques biométriques également. Spécifiquement à chaque technologie de communication de proximité vue précédemment, les points suivants doivent être mentionnés :

- Le QRC offre une grande utilisabilité et peut être couplé avec n'importe quels autres réseaux de d'arrière-plan pour faciliter l'exploitation par l'utilisateur final. D'autres mesures de sécurité liées à l'utilisation qu'on fait du QRC et aux autres technologies utilisées dans le système sont nécessaires.
- Les systèmes basés sur le BLE, offre une facilité de mise en œuvre (il suffit que le fournisseur de service financier soit reconnu par la passerelle de paiement pour accéder aux systèmes) et offre une accessibilité moyenne à l'utilisateur en termes de coût de l'équipement. Toutefois, la sécurité des clés stockées sur l'appareil reste un souci majeur. Aussi, il faut s'assurer d'une implémentation adéquate au niveau de la couche d'application pour prendre en compte les requis de sécurité.
- La NFC permet de mettre en place des systèmes robustes d'un point de vue de sécurité et aussi d'une grande utilisabilité (le paiement sans contact), mais ces derniers nécessitent de grands efforts pour la mise en œuvre et l'exploitation; ils peuvent être implémentés avec le SE intégré dans le mobile, auquel cas, l'appareil est lié au manufacturier ou avec une SIM géré par un MNO. Dans ces deux cas, la gestion du SE et la mise à jour des données est une tache coûteuse. Si le fournisseur de services est dans l'industrie de développement des cartes sécurisées, il peut lui être avantageux de mettre en place une telle plateforme. Dans le cas contraire, la tendance est vers l'architecture HCE. Il faut donc pour cette dernière architecture, authentifier le mobile et sécuriser le canal de partage des informations sensibles avec le serveur dans le cloud. Aussi, les équipements supportant la NFC sont les plus couteux pour l'utilisateur.

2.3 Systèmes et technologies dans le paiement à distance

Dans cette section, nous traitons de SPM permettant le transfert de fonds entre deux utilisateurs distants. Ils sont classés en 3 catégories : la première est caractérisée par l'utilisation des services SMS ou USSD des réseaux GSM, la 2ème est cet ensemble de système dont tous les échanges se font par les protocoles de l'Internet et la dernière rassemble des systèmes basés sur la Blockchain.

2.3.1 Les systèmes basés sur le SMS et l'USSD

Des systèmes de services dans le commerce mobile ont été conçus pour permettre les échanges via le SMS et/ou l'USSD [37]. Une première sous-catégorie n'utilise que le SMS pour la communication entre les différents modules du système (terminal de l'utilisateur, institutions financières, opérateurs de téléphonie mobile ou autres fournisseurs de service). Dans ce cas, le terminal mobile inclut une application offrant l'interface pour la saisie des données et les envoient au destinataire sous forme de SMS via le réseau de téléphonie mobile. Pour les services bancaires, c'est la banque de l'utilisateur qui reçoit la requête. Elle peut répondre directement (afficher la balance du compte par exemple) ou contacter une passerelle si une autre banque est impliquée dans la transaction (cas d'un transfert de fonds) selon le schéma de la Figure 2.4 inspirée de [38]. Souvent, la cryptographie est utilisée pour garantir des services de sécurité. L'application peut être installée dans la SIM par l'opérateur de téléphonie mobile ou sur le terminal (ce qui peut se faire dans ce cas par l'utilisateur). Plusieurs problèmes ont été relevés dans les premiers systèmes conçus tels que rapportés dans [4]: « Certains ont mis trop de temps pour générer les clés ou sont vulnérables aux attaques de reconnaissance de clés, d'autres exigeaient trop de puissances de calcul et enfin certains ne garantissaient pas la protection contre tous les types d'attaque ».

L'un des derniers systèmes de services bancaires mobiles proposés en 2017, utilise la cryptographie basée sur les propriétés de courbes elliptiques et implémentée avec J2ME (*Java 2 Mobile Edition*) [33]. Les auteurs ont montré la garantie des différentes propriétés de sécurité avec une performance acceptable. Toutefois, comme ses prédécesseurs, la clé privée stockée sur le mobile est à risque. De plus, il faut s'assurer d'une technologie facilitant l'entrée puisque la saisie de longues chaines de caractères n'offre pas une bonne commodité et n'est pas pratique dans des régions à faible taux d'alphabétisation. Aussi, la performance du système varie en fonction du réseau de téléphonie dans la région de service. Ce qui peut être un souci selon le contexte d'utilisation de ces systèmes.

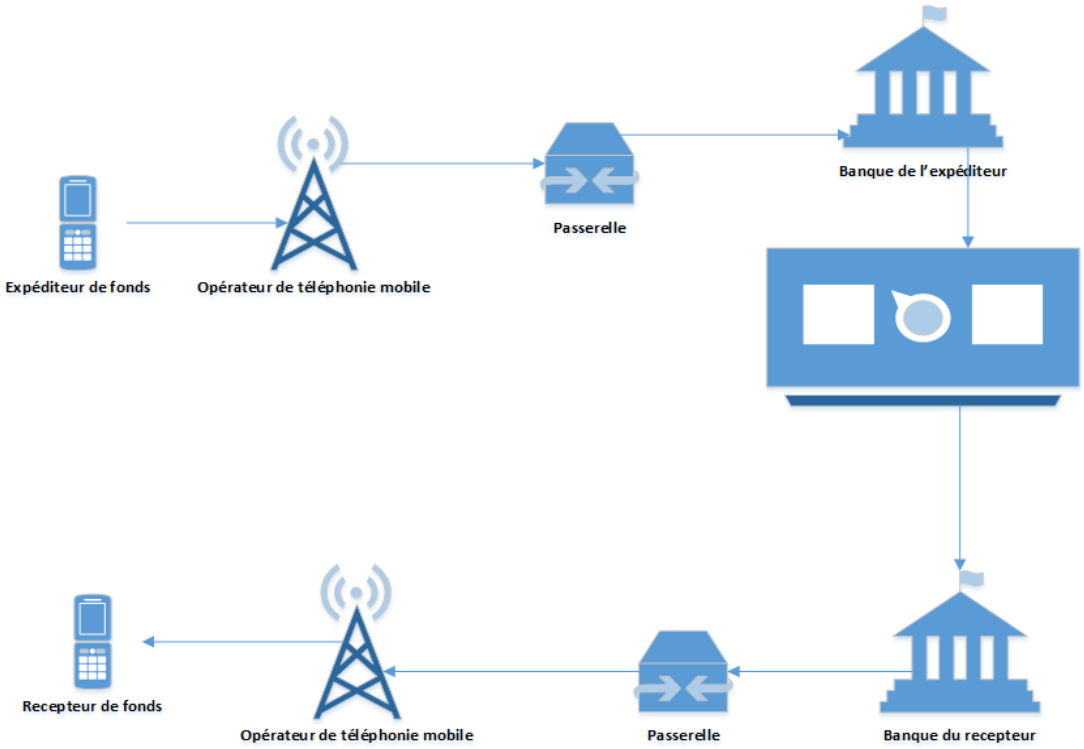


Figure 2.4 Architecture d'un service bancaire avec SMS

La deuxième sous-catégorie est l'ensemble de ces systèmes qui combinent le SMS et l'USSD : la session est ouverte jusqu'à la fin du service demandé et le SMS peut être utilisé pour notifier l'utilisateur (ou les utilisateurs) du résultat de la transaction. L'architecture générique d'un système de services bancaires avec USSD est montrée à la Figure 2.5 [13], [14]: Le fournisseur de service (généralement un MNO), gère les serveurs d'applications (Application Servers) qui doivent répondre aux requêtes des clients parvenues par la passerelle. Normalement, un compte assigné par l'opérateur de service à l'utilisateur est lié à son numéro de téléphone. La modalité d'affectation des comptes bancaires peut dépendre de protocoles d'accord entre banques et fournisseurs de services : en temps réel lors de la transaction ou en différé sous forme de traitement de lots de transactions.

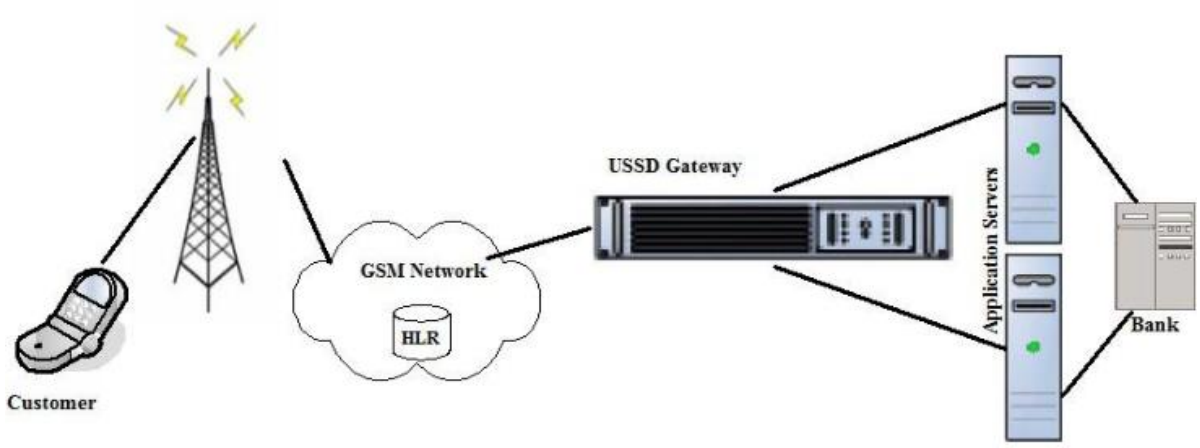


Figure 2.5 Architecture de services bancaires avec USSD [13]

Avec ces derniers systèmes, l'utilisabilité est améliorée par rapport aux plateformes se basant uniquement sur le service SMS. L'avantage est aussi le coût de mise en œuvre et d'exploitation pour un MNO déployant ce service et surtout le fait de toucher des régions dépourvues de grandes infrastructures technologiques. En effet, le service USSD n'utilise pas la bande de trafic et est librement offert par les opérateurs; de plus, le service est accessible depuis des téléphones bas de gamme. Côté sécurité, nous avons vu qu'une session est créée et pas de stockage d'information comme pour le SMS. Toutefois, le mécanisme manque de robustesse : souvent après enregistrement des pièces officielles du propriétaire du compte, juste un PIN (*Personal Identification Number*) de 4 caractères suffit pour accéder au compte; le service s'appuie sur les mécanismes de sécurité du réseau GSM. Or, la sécurité dans GSM se base sur une clé symétrique connue du HLR et installée sur la carte SIM du mobile; le BTS n'est pas authentifié, le clonage de la carte SIM est une menace importante ainsi que le fait que les algorithmes utilisés (A_3, A_5, A_8) par GSM ont déjà subi le processus de rétro-ingénierie [39]. L'utilisation des réseaux 3G apporte de grandes améliorations au niveau sécurité, certes; mais le service est souvent offert avec les réseaux 2G.

2.3.2 Les systèmes basés sur les protocoles de l'internet

Pour les systèmes précédents, la technologie de communication de proximité était la ligne de mire; il faut toutefois signaler que ces systèmes utilisent les protocoles de l'Internet pour la communication entre les participants distants traitant les transactions. Les systèmes dont il est

question maintenant s'appuient uniquement sur les protocoles de l'Internet : pas d'autres technologies de communication de proximité. Le client accède à ces services avec son équipement mobile sur lequel une application adéquate est installée ou à travers un navigateur dans tous les autres cas.

L'authentification des clients suit généralement le schéma classique multi-facteur : quelque chose qu'il possède (une carte par exemple), ce qu'il sait (des informations détenues par la banque lors de l'inscription ou un PIN). Souvent une infrastructure à clé publique est mise en œuvre pour renforcer l'authentification des participants et garantir d'autres propriétés de sécurité par la cryptographie. Une architecture dans le commerce mobile qui peut être associée à ces systèmes est proposée en [5]: cette étude utilise l'appariement bilinéaire symétrique pour des vérifications de signature, et assure l'authentification, la confidentialité et l'intégrité sans certificat émis par une autorité; de plus elle enregistre les transactions sur la Blockchain pour des cas de litige. L'auteur soutient que son schéma est assez robuste pour résister contre les attaques suivantes : l'homme du milieu, la personnification, le rejeu et le déni de service. Toutefois, un problème majeur avec ces systèmes est le manque de garantie de la vie privée et la fuite de données qui peut en résulter.

Nous trouvons en [40], un protocole d'échanges permettant de garantir le respect de la vie privée pour le paiement dans le cadre du transport en commun. Le système utilise les signatures de traçabilité d'identité et anonymes pour empêcher au système de paiement d'avoir accès aux données de voyage et aux compagnies de transit d'avoir accès aux informations de paiement du client. Cette logique peut être suivie dans le cas d'un schéma pour le commerce mobile de proximité; elle permettra de diminuer les informations privées partagées avec les systèmes des marchands.

2.3.3 Les Systèmes basés sur la Blockchain

En [26], plusieurs services financiers offerts par des systèmes basés sur la Blockchain sont notés : des crypto-monnaies, des services d'échange entre ces dernières et des monnaies fiduciaires et des services de transfert de fonds entre des banques. Si dans les cas précédents, le système de traitement de transactions entre expéditeur et récepteur de fonds était centralisé (système bancaire), dans la Blockchain, c'est un système décentralisé où chaque nœud du réseau est en charge de traiter la transaction. Dans la pratique, certains nœuds légers d'utilisation finale ne participent pas au processus de validation.

La première application financière de cette technologie et la plus notable des crypto-monnaies est le Bitcoin [6], [17]. Complètement ouvert et sans une autorité de confiance, ce réseau se base sur le protocole de preuve de travail ou *Proof-of-Work* (PoW), dans lequel un nœud doit fournir une preuve de computation pour valider un bloc de transactions. Or, ce protocole consomme beaucoup d'énergie et n'est pas applicable à certains services de commerce mobile, vu que seulement 7 transactions peuvent être validées par seconde [41]. De plus, les transactions exposées à tous les participants entravent la garantie du respect de la vie privée. Aussi, le Bitcoin est un réseau assez rigide, seulement le transfert des coins est possible avec éventuellement l'ajout d'informations supplémentaires dans un champ optionnel lors du transfert.

Plusieurs autres crypto-monnaies ont été proposés en vue d'améliorer des aspects négatifs du Bitcoin tels Zilliqa et Bitcoin Cash (BCH) pour le problème d'évolutivité, Monero et Zcach vise la sécurité et le respect de la vie privée par des techniques cryptographiques, le réseau Ethereum avec les contrats intelligents pour améliorer la flexibilité de la Blockchain en permettant de programmer d'autres services sur un tel réseau [17], [42]. Zilliqa cherche à améliorer la performance par la technique du partage de réseau; BCH le fait en augmentant la taille des blocs de transactions. Cependant, malgré des améliorations au niveau du respect de la vie privée ou de l'évolutivité, ces réseaux se basent tous sur le PoW, énergivore, non recommandée dans un contexte de luttes contre le changement climatique.

EOS est une autre crypto-monnaie construite sur le protocole de preuve d'enjeux déléguée ou *Delegated Proof of Stake* (DPoS) qui est une variation de la preuve d'enjeux ou *Proof of Stake* (PoS) dans lequel, les nœuds sont choisis pour valider les blocs de transactions en fonction de leurs biens dans le réseau [43]; le DPoS utilise un vote pour assigner la tâche de validation à des nœuds particuliers de l'ensemble des qualifiés [15], [16]. Ce dernier protocole permet de résoudre le problème de performance du Bitcoin. Cependant, on y note les faiblesses suivantes [20], [38]: une tendance vers la centralisation où les plus riches ont plus de chance de valider, l'attaque dit *Nothing-at-Stake* où le nœud malveillant n'a rien à dépenser pour miner sur plusieurs chaînes et augmenter la probabilité de succès de son attaque. De plus, certains nœuds peuvent s'amuser à augmenter les soldes de leurs comptes sans faire circuler leurs biens. L'une des variantes de ce protocole conçu pour augmenter sa robustesse est la preuve d'importance ou Proof of Importance (PoI). Ce dernier accorde une importance à chaque nœud dont la mesure inclut la participation au bon fonctionnement du réseau (biens acquis et échanges) [21]. Toutefois, tous ces protocoles sont

utilisés jusqu'à présent dans des Blockchains publiques avec la création des crypto-monnaies. Aussi, aucun mécanisme d'exclusion d'un fraudeur du système n'est prévu. Il est donc imprudent de déployer des services à grandes valeurs commerciales sur de telles plateformes sans aucune correction.

Hyperledger [24] et Ripple [44] sont des Blockchains non publiques (destinées à des consortiums d'entreprise) qui vont dans le sens de ce contrôle en environnement informatique distribué. Ils utilisent des protocoles qui s'appuient sur une connaissance des nœuds en communication. Ces protocoles sont la plupart du temps des variations du BFT dont un cas classique et répandu est le PBFT. Le client soumet sa transaction sur le réseau et chaque nœud connecté va la traiter avec l'un d'entre eux agissant comme le leader du processus; si un pourcentage de l'ensemble des nœuds (typiquement 2/3) s'accorde sur une réponse, ce sera la réponse acceptée par le client [45]. Ce dernier protocole peut souffrir aussi du problème d'évolutivité (une complexité de l'ordre $O(n^2)$ pour n validateurs), exige souvent l'homogénéité du matériel de chaque nœud pour une bonne performance [45], [46], et par-dessus tout, conduit à la centralisation où les nœuds doivent être connus d'avance pour mettre en place le système.

C. Berger et H. P. Reiser ont fait une synthèse les différentes directions d'améliorations constatées parmi les variations du BFT [27]. Certaines comme en [45], essaient d'améliorer la topologie des communications, mais ajoutent malheureusement trop de charge sur le leader du processus. Des variations de type FBA (*Federated Byzantine Agreement*) considèrent des sous-réseaux de confiance auxquels les différents nœuds adhèrent, mais l'évolutivité dépend dans ce cas de la structure concrète et des relations de confiance. Un troisième groupe utilise un matériel sécurisé (*TEE pour Trusted Execution Environnement par exemple)* pour améliorer l'évolutivité du PBFT. Enfin une dernière catégorie comme en [44] regroupe les requêtes des clients de façon à permettre des traitements parallèles. Toutefois, dans ce dernier cas, l'atteinte du consensus est plutôt probabiliste.

Enfin, nous devons considérer des schémas de paiement proposés pour le commerce mobile de proximité et construits sur la blockchain. L'une des dernières propositions est le schéma de Zhong et *al.* [47], dans lequel un règlement est préétabli entre l'expéditeur et le receveur de fonds : le premier a un engagement de paiement et le second de remboursement; le payeur prérègle aussi une information de paiement incluant les comptes concernés, les engagements de paiement et de

remboursement, des fonds bloqués à cette fin et une échéance. Il signe cette information, puis la diffuse sur la blockchain. Des nœuds de validation vont confirmer que cette transaction est correcte selon le protocole de consensus en vigueur (PoW, PoS, …). Une fois confirmée, expéditeur et récepteur de fonds peuvent finaliser le paiement en établissant le contact par une image qui avait été prise comme le mentionnent les auteurs; ils ont éventuellement travaillé pour le commerce de proximité. Ils ont démontré la robustesse en termes de sécurité qui s'appuie en grande partie sur la sécurité du protocole de consensus utilisé. Ils clament également que le système est flexible et peut s'exécuter sur des appareils légers comme les téléphones intelligents. Ce système malgré de bonnes caractéristiques, ne s'est pas attaqué au problème de gestion des clés, et sa performance dépend aussi du protocole de consensus du réseau. La figure 2.6 inspirée de leur description correspond à une vue globale de ce système.

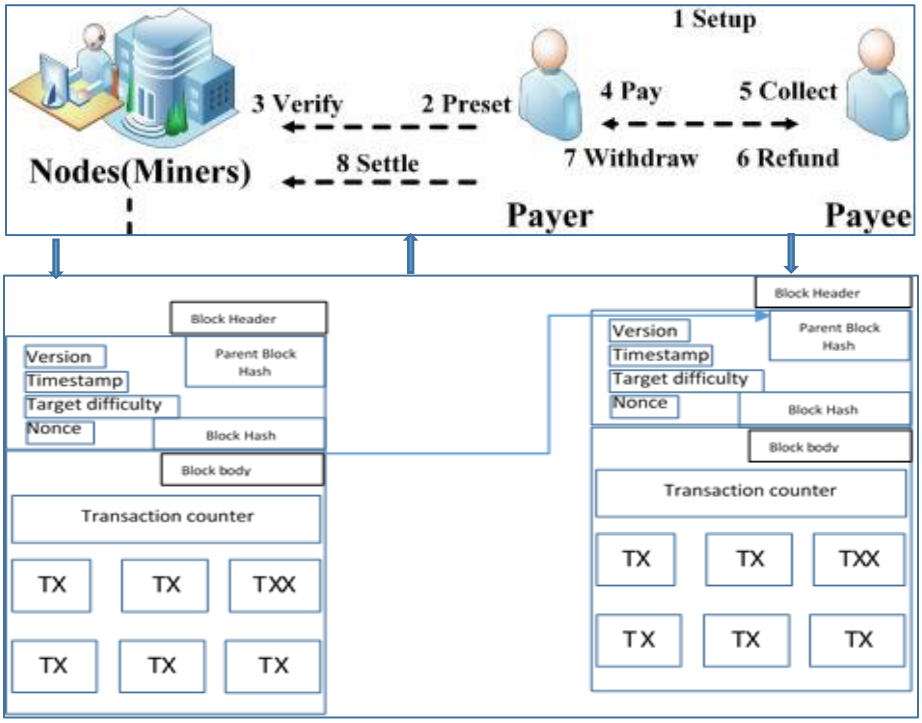


Figure 2.6 Vue globale d'un SPM basé sur la Blockchain [40]

2.3.4 Analyse des systèmes dans le paiement à distance et la Blockchain

Dans le paiement mobile à distance, nous retrouvons d'abord des systèmes bancaires offrant les services de paiement habituels via un équipement mobile, ils s'appuient principalement sur les

technologies de communication par internet. Puis, ceux qui utilisent le SMS ou une combinaison de ce service avec l'USSD. Les problèmes suivants relatifs à ces systèmes ont été identifiés :

- Avec les systèmes bancaires utilisant l'internet, le principal souci est le respect de la vie privée. En effet, l'institution financière dispose généralement de beaucoup d'informations qui certaines fois transitent par les systèmes des marchands. Si nous supposons une bonne implémentation des mécanismes traditionnels de sécurité, les propriétés telles la confidentialité, l'authentification, l'intégrité et la répudiation sont garanties. Cependant, le respect de la vie privée n'est pas garanti avec ces mécanismes; aussi, nous connaissons des cas de fuite de données d'institutions financières notables [25].
- Avec les systèmes s'appuyant sur le SMS, utilisant un module client installé avec un téléphone moyen de gamme, on a généralement le même problème de stockage de clé que pour les systèmes s'appuyant sur le BLE.
- Les PPM avec USSD/SMS – la plus accessible pour certaines régions à faible infrastructure technologique – souffrent des vulnérabilités liées aux réseaux GSM : le clonage de la SIM donnant accès aux clés de chiffrement, la non-authentification de la station de base rend le système vulnérable a une personnification de ce dernier composant.

La Blockchain, technologie relativement récente, apparait comme une alternative aux différentes technologies précitées pour la mise en œuvre de systèmes de services du commerce mobile. Elle est visée comme une plateforme garantissant une grande disponibilité, une reprise sur panne facile, la résistance à la falsification avec un grand potentiel pour faciliter l'interopérabilité. Toutefois, les limitations suivantes doivent être considérées :

- Dans les systèmes basés sur la Blockchain publique, la gestion des clés est laissée à l'utilisateur. En cas de perte de la clé privée, ses biens sur la Blockchain est perdue. De plus, avec un tel mécanisme, aucune liaison n'est garantie à priori entre la clé et une entité réelle.
- Les Blockchains publiques s'appuient sur des protocoles de consensus dont certains (PoW par exemple) sont énergivores, offrent une faible performance, non évolutifs et inadaptés pour le déploiement de services réels [41].
- Les protocoles de consensus pour les Blockchains à permission (PBFT, Ripple par exemple) peuvent offrir de bonnes performances, mais il faut savoir contrôler la tendance

vers la centralisation et adopter les meilleures approches si on veut garder les propriétés importantes d'une Blockchain.

- La plupart des schémas de paiements basés sur la Blockchain offrent seulement des fonctionnalités de portefeuille de crypto-monnaies. Les propositions qui vont au-delà – comme le schéma en [47] ne considèrent pas le problème de gestion de clés avec la Blockchain ou celui des protocoles de consensus.

2.3.5 Une cartographie des vulnérabilités des systèmes de paiement à distance

Dans cette section, nous mettons en évidence les différents types d'attaque possible sur une plateforme qui supporte le paiement à distance, les limitations des méthodes de sécurité existantes, et aussi l'avantage potentiel de la blockchain dans la sécurisation de ces plateformes.

2.3.5.1 Différents types d'attaques

- Déni de service ou Denial of Service (DoS): Les attaques DDoS visent à submerger les serveurs de paiement à distance avec un trafic excessif, entraînant une indisponibilité du service pour les utilisateurs légitimes. Ce qui engendre des pertes de revenus.
- Hameçonnage et ingénierie sociale: Les attaques d'hameçonnage et d'ingénierie sociale tentent de tromper les utilisateurs en leur faisant divulguer des informations sensibles, telles que les identifiants de connexion ou les informations de carte de crédit.
- Programmes malveillants: certaines applications de l'utilisateur sur son mobile peuvent être des vecteurs d'attaques. Par exemple, l'utilisation du SMS, du GPS peuvent donner accès à des programmes malveillants qui voudraient voler des données sensibles ou occasionner d'autres dommages. Nous pouvons également mentionner que dans le cadre de la collaboration entre plusieurs institutions, certaines d'entre elles peuvent ne pas avoir une infrastructure sécuritaire.
- Des attaques d'interceptions telles que l'homme du milieu ou *Man in the Middle* (MiTM), et le reniflage qui permettent de voler des données confidentielles aussi.

Les mesures pour protéger la plateforme de commerce mobile tournent autour de la cryptographie et l'authentification des sources de données [48], [74]. La cryptographie permet de chiffrer les échanges et d'assurer ainsi la confidentialité. Elle offre aussi des mécanismes permettant la non-répudiation des transactions avec les signatures électroniques. La blockchain dans sa conception s'appuie sur la cryptographie et offre déjà un bon début la sécurisation des

services. De plus, la décentralisation offre un rempart au système lorsque certains nœuds sont victimes d'attaque.

2.3.5.2 Méthodes d'authentification faibles

Sans une authentification adéquate, les systèmes de paiement à distance pourraient être vulnérables aux attaques d'usurpation d'identité. Ce qui pourrait engendrer des accès non autorisés aux comptes de l'utilisateur du système.

Plusieurs méthodes d'authentifications sont appliquées pour sécuriser les plateformes supportant les paiements à distance [72]-[73], [131]-[132]. Lors de l'accès à un système, l'utilisateur peut être invité à fournir soit un numéro d'identification personnel (NIP) soit un mot de passe. Ces informations correspondent à ce que l'utilisateur sait. Parfois, l'institution peut demander quelque chose que l'utilisateur possède, comme une carte ou un jeton. De plus, les utilisateurs peuvent s'authentifier en présentant des caractéristiques biométriques, telles que leurs empreintes digitales. Pour accroître la sécurité du système visé, les processeurs d'authentification combinent plusieurs facteurs d'authentification. Cette approche est connue sous le nom d'authentification multifactorielle (MFA).

Le PIN ou mot de passe peut avoir une faible entropie. Dans le cas contraire, l'utilisateur les oublie souvent. Dans le cas de l'authentification par une possession, la perte de l'objet engendre des troubles. Les limitations avec l'authentification biométrique sont nombreuses dont, la variation avec le temps, des problèmes de réversibilité et de révocabilité.

2.3.5.3 Mécanismes de chiffrement faibles

L'absence ou la faiblesse du chiffrement des données transmises entre les utilisateurs et les serveurs peut permettre à des attaquants d'intercepter et de lire les informations sensibles telles que les numéros de carte de crédit. Ce problème est exacerbé avec des systèmes basés (SMS/USSD), utilisant des anciennes générations de réseaux de télécommunications où il n'y a pas cryptage de bout en bout lors de l'envoie d'un message sur le réseau [39]. Cette faille peut engendrer le vol d'informations personnelles ou la compromission de données financières.

Un système basé sur la blockchain, bénéficierait déjà de l'intégration de la cryptographie dans cette dernière technologie.

2.3.5.4 Mauvaise gestion des clés de chiffrement

Une mauvaise gestion des clés de cryptage peut entraîner une vulnérabilité du processus de cryptage/décryptage, exposant ainsi les données à des risques d'accès non autorisé. Ce problème est particulièrement préoccupant dans le cadre de la Blockchain, puisqu'on n'est pas dans le paradigme de la confiance en une autorité centrale ; l'utilisateur a alors plus de responsabilité dans la gestion des clés [74].

2.3.5.5 Manque de respect de la vie privée et attaque d'initié

Les plateformes de paiement mobile font intervenir plusieurs participations [2] : l'envoyeur de fonds, le receveur de fonds, les banques, le MNO, les réseaux de cartes de paiement. Les données des utilisateurs transitent sur tous ces sous-systèmes et sont accessibles par plusieurs. Nous avons également cité plus haut un cas de vol de données venant d'un employé malintentionné. Ces problèmes de sécurité peuvent causer de grands dommages à l'utilisateur et aux institutions responsables des plateformes technologiques pour le commerce à distance.

À ce niveau, la blockchain pourrait réduire la dépendance de la plateforme par rapport à une institution particulière.

CHAPITRE 3 DÉMARCHE MÉTHODOLOGIQUE DE L'ENSEMBLE DU TRAVAIL DE RECHERCHE

Dans cette thèse, nous visons à concevoir une architecture basée sur la blockchain permettant d'offrir des services de commerce électronique : commerce de proximité, virements entre personnes incluant les transactions interbancaires, entre autres. Cette architecture contiendra plusieurs composants qui peuvent être déployés avec d'autres plateformes technologiques. D'abord, une fonction de dérivation de clé, permettant de générer une clé cryptographique à partir des données biométriques de l'utilisateur. Cette fonction est aussi liée à un mécanisme pour l'authentification biométrique ou un apparieur biométrique. Ensuite, un schéma de paiement décentralisé, permettant d'offrir des services de commerce électronique sur la blockchain en garantissant le respect de la vie privée. Finalement sont proposés : un modèle d'architecture intégrant les composants susmentionnés avec des protocoles de consensus existants, et une visualisation d'un protocole distribué plus adéquat pour le paiement interbancaire. L'évaluation de toute cette architecture passe par l'évaluation de chaque composant. Dans ce chapitre, nous préciserons d'abord les critères d'évaluation considérés. Ensuite, nous exposerons la méthodologie adoptée pour la mise en œuvre de chaque composant d'une part, puis pour le rassemblement de ces différents composants en un ensemble d'autre part. Nous aurons soin d'exposer également les liens entre les objectifs spécifiques et les articles scientifiques présentés dans les prochains chapitres.

3.1 Critère d'évaluation de l'architecture

Les critères peuvent être soit spécifiques à des composants, soit applicables à plusieurs d'entre eux en ayant une mesure ou une interprétation différente selon le composant (Tableau 3.1). Des propriétés voulues de la clé générée telles que : l'entropie de la chaine (longueur et caractère aléatoire), la révocabilité et l'irréversibilité sont spécifiques au KDF. Un critère de performance comme le nombre de transactions par seconde (TPS) est mesurable et a du sens au niveau du protocole de consensus entre les nœuds de la blockchain et de l'architecture globale. D'autres critères peuvent s'appliquer à toute l'architecture, mais ayant une mesure différente selon les composants considérés. C'est le cas de la privacité. L'on retrouve aussi dans cette catégorie, la flexibilité. Enfin, nous devons mentionner qu'un critère comme la durée d'une transaction est manifestement étendue à toute l'architecture avec un apport de chacun des composants. Ces différents critères sont décrits ci-dessous.

Flexibilité: s'appliquant à la fonction de dérivation de clé, ce critère considère la capacite de notre fonction à s'appliquer à plusieurs caractéristiques biométriques ou plusieurs techniques d'extraction de caractéristiques.

Aléatoire: cette caractéristique est nécessaire pour toute clé cryptographique. En effet, la clé ne doit pas être prévisible et ainsi prédictible.

Révocabilité : elle désigne la capacité à doter l'utilisateur d'une autre clé complètement différente avec les mêmes caractéristiques biométriques (empreinte digitale, face, etc.), en cas de corruption de la première clé.

Irréversibilité : s'appliquant encore à la fonction de dérivation de clé, elle désigne l'impossibilité à obtenir les données biométriques à partir des clés.

Exactitude / Taux d'acceptation faux / Taux de rejet faux : ces métriques permettent de savoir à quel point la mesure d'un système donné est proche de la réalité. Plus spécifiquement, dans le cas d'un système d'authentification qui doit répondre par oui l'utilisateur présent est authentique ou non, il ne l'est pas, ces grandeurs sont déterminées par les formules suivantes.

$$\text{Exactitude} = \frac{vrai\ positif + vrai\ negatif}{vrai\ positif + vrai\ negatif + faux\ positif + faux\ negatif},$$

$$\text{Taux d'acceptation faux} = \frac{faux\ positif}{vrai\ positif + vrai\ negatif + faux\ positif + faux\ negatif},$$

$$\text{Taux de rejet faux} = \frac{faux\ negatif}{vrai\ positif + vrai\ negatif + faux\ positif + faux\ negatif}$$

Avec :

- Vrai positif : le nombre de fois que le système détecte que l'utilisateur est authentique et il l'est effectivement.
- Vrai négatif : le nombre de fois que le système détecte que l'utilisateur n'est pas authentique et il ne l'est pas effectivement.
- Faux positif : le nombre de fois que le système détecte que l'utilisateur est authentique, alors qu'il ne l'est pas. Dans un système d'authentification, c'est une erreur à minimiser avant toute autre.
- Faux négatif : le nombre de fois que le système détecte que l'utilisateur n'est pas authentique alors qu'il l'est.

La privacité : de manière générale, elle évalue la quantité de données sensibles de l'utilisateur auxquelles ont accès certains participants à l'offre de services en question. Pour la fonction de dérivation de clé, la privacité peut faire référence à 2 caractéristiques distinctes de ces fonctions : la distinguabilité et la réversibilité. La distinguabilité réfère à la possibilité pour un agent externe de lier la donnée biométrique accessible aux autres informations sensibles de l'utilisateur; la réversibilité réfère à la capacité de l'agent externe à obtenir la donnée biométrique originale à partir de l'esquisse publique utilisée pour générer la clé. Dans le cas du service de paiement mobile, le numéro de compte de l'expéditeur de fonds, celui du récepteur de fonds et le montant de la transaction sont souvent accessibles à beaucoup de participants. La mesure de la privacité se réfère alors aux informations accessibles à tous.

Délai : c'est la durée entre l'instant d'une transaction sur la plateforme technologique par l'envoyeur de fonds et l'instant où cette transaction est disponible pour le receveur de ces fonds. Ce délai se mesure souvent en sous-multiple de secondes ou en secondes pour le commerce de proximité; mais en minutes, heures ou même jour pour des virements interbancaires et transfrontaliers.

Débit : cette valeur représente le nombre de transactions qui peuvent être traitées par unité de temps par la plateforme. Le débit s'exprime souvent en nombre de transactions par seconde (TPS).

Tableau 3.1. Critères d'évaluation de l'architecture

<table>
<tr><th>Composants</th><th>Propriétés spécifiques</th><th>Propriétés transversales à plusieurs composants</th><th>S'étendant sur toute l'architecture</th></tr>
<tr><td>Fonction de dérivation de clé (KDF)</td><td>Flexibilité
Aléatoire
Révocabilité
Irréversibilité</td><td rowspan="2">Exactitude
Taux d'acceptation faux
Taux de rejet faux
Privacité</td><td rowspan="3">Durée de validation d'une transaction
Nombre de transactions par seconde</td></tr>
<tr><td>Mécanisme d'authentification</td><td></td></tr>
<tr><td>Schéma de paiement sur la blockchain</td><td>Privacité</td><td rowspan="2">Nombre de transactions par seconde</td></tr>
<tr><td>Protocole de consensus</td><td></td><td></td></tr>
</table>

3.2 Volet 1 : une analyse des plateformes de paiement mobile avec intégration à la blockchain

Dans le but de proposer des composants de l'architecture qui résolvent effectivement les problèmes des plateformes de paiement existants et faciliter leur intégration avec la blockchain, nous avons commencé par analyser ces plateformes. Le travail est consigné dans la revue de littérature et a aussi donné naissance à l'article intitulé *« Mobile Payment Platforms : Taxonomy, Architecture, Security, and Integration with Blockchain »*. Deux grands points étaient considérés pour cette revue qui aura permis de mettre en lumière des directions de recherche futures.

3.3 Volet 2 : un appariteur biométrique et une fonction de dérivation de clé à partir des données biométriques

Dans l'optique de sécuriser l'architecture qui sera proposée dans le cadre de ce travail de recherche, notre premier objectif est de concevoir un mécanisme fiable permettant de lier une entité (personne physique dans ce cas) avec son identité sur la blockchain. L'identité sur la blockchain étant définie par une paire de clés (publique/privée), on peut chercher à générer la clé privée à partir des données biométriques de l'utilisateur. Les méthodes utilisées pour le faire sont traditionnellement : l'Extracteur Flou ou *Fuzzy Extractor* (FE), la KDF qui peut être basée sur des méthodes artisanales ou d'intelligence artificielle. Les FE ne peuvent pas engendrer des constructeurs de clé robustes; ils ne peuvent corriger un niveau acceptable de bruit. Les KDFs basées sur des méthodes artisanales échouent souvent à prendre en compte les spécificités de l'utilisateur. Celles qui sont basées sur l'intelligence artificielle requiert souvent l'utilisation d'une grande quantité de données, soit pour l'entrainement ou sauvegardées pour des comparaisons futures; ce qui engendre un manque de respect de la vie privée avec ces dernières. Ces raisons motivent la conception d'une méthode qui soit robuste en termes de prise en charge de la variabilité intra-user et qui ne nécessite pas une grande quantité de données pour générer la clé. Elle a été proposée dans l'article intitulé « *Privacy-Preserving Model for Biometric-based Authentication and Key Derivation Function* ». Cet article couvre le premier objectif spécifique et constitue le contenu du chapitre 5.

3.3.1 Rationnel de la méthode d'encodage

L'un des défis à relever pour créer une clé cryptographique à partir des données biométriques est la variabilité intra-utilisateur. En effet, pour une même personne, deux acquisitions du même

caractère biométrique sont rarement la même. Ayant acquis initialement un nombre *R* d'images de l'utilisateur, pour chaque caractéristique, les *R* valeurs constituent un ensemble flou. On remplace ces valeurs floues par une valeur moyenne. Ce qui est une méthode de défuzzification. Si les vecteurs biométriques acquis comprennent chacun *m* composants, on aura *m* ensembles flous et donc un autre vecteur final de *m* valeurs après défuzzification. Un algorithme de regroupement forme un nombre donné de groupes ou clusters selon la proximité de ces *m* valeurs. Chaque composant d'un vecteur biométrique dont la valeur tombe dans un cluster est représenté par le centroïde du groupe. Ce qui donne un code représentant la personne physique. La comparaison d'un code reproduit plus tard avec le code initial permet de conclure si l'utilisateur qui veut se connecter à l'appareil est authentique.

3.3.2 Application de la méthode d'encodage

À l'enrôlement d'un utilisateur, le système acquiert un nombre d'unités (*R* par exemple) du trait biométrique concerné. Puis, des méthodes de pré-traitement et/ou d'extraction de caractéristiques permettent d'avoir *R* vecteurs de nombres réels, représentant chacun l'utilisateur. Le processus de quantification permet d'avoir un vecteur *Q* des niveaux de quantification. De plus, en encodant plusieurs vecteurs biométriques, on trouve des composants qui tombent toujours dans le même cluster (donc admet le même niveau de quantification). Ce sont des caractéristiques biométriques consistantes; les positions du vecteur biométrique initial correspondant à ces positions sont gardées dans un vecteur *CP*. La reproduction d'un code dans le futur sera faite avec l'aide de *Q* et de *CP* : pour chaque composant du nouveau vecteur biométrique acquis de l'utilisateur, la valeur est remplacée par le niveau de quantification le plus proche et les éléments qui correspondent aux positions de *CP* forment le code final. La longueur de CP sera donc la longueur du code.

3.3.3 Génération de la clé cryptographique

Pour générer la clé cryptographique à partir du code créé. On adapte l'algorithme HMAC – SHA256 à notre cas. Notre code étant un vecteur avec les composants $(c_0, c_1, \ldots\ldots, c_m)$, nous calculons une première valeur K_0 de la clé avec la formule $K_0 = HMAC - SHA256(Extsalt \| c_0)$. Ensuite, nous itérons pour étendre cette clé selon la formule $K_i = HMAC - SHA256(K_{i-1} \| c_i)$, avec *i* entier variant de *1* jusqu'à une valeur finale qui peut être la longueur *m* du code. Pour des raisons de performance, notamment la durée de création de la clé, on peut choisir de grouper des

codes de même signe pour moins d'itérations. Si i_M est la plus grande valeur prise par *i, alors* K_{i_M} est la clé cryptographique recherchée.

3.3.4 Évaluation de performance

La KDF est évaluée sur trois ensembles de données sur les visages :

1. L'ensemble de données 1 (DS1) est un ensemble de données privé qui comprend 220 images provenant de 5 utilisateurs. Ces images ont été collectées à partir d'un téléphone portable.
2. L'ensemble de données 2 (DS2) est une version recadrée manuellement de la base de données de visages étendue de Yale. Il contient 2414 images.
3. L'ensemble de données 3 (DS3) est un sous-ensemble de la base de données de visages de Youtube. Il comprend 3425 vidéos provenant de 1595 utilisateurs.

Pour DS1 et DS3, nous utilisons FaceNet [48] pour détecter les visages, tandis que pour DS2, les visages sont déjà détectés. Des traitements supplémentaires tels que la création d'éléments structurants, des opérations d'érosion et d'effacement, ainsi que l'augmentation du contraste sont appliquées à DS2 et DS3.

Pour évaluer le processus d'enrôlement, nous avons divisé les images de chaque utilisateur en deux sous-ensembles d'échantillons. Le premier sous-ensemble, généralement représentant 75 % de tous les échantillons disponibles, contenant R1 échantillons, est utilisé pour l'apprentissage du vecteur *Q*. À partir du second sous-ensemble (R - R1), nous apprenons le vecteur *CP*. Nous construisons plusieurs vecteurs biométriques que nous encodons afin de déterminer le seuil minimal de similarité. Pour *DS1*, les codes ont une longueur au-dessus de 100 composants avec une similarité d'au moins 92 %; sur DS3, les codes dépassent une longueur de 82 pour une similarité d'au moins 93 %.

Le mécanisme d'authentification a été évalué selon les métriques standards telles que l'exactitude ou taux de reconnaissance, le rappel, le taux de vrais négatifs ou *True Negative Rate* (TNR), le taux de faux positifs ou *False Acceptation Rate* (FAR) et le taux de faux rejets ou *False Rejection Rate* (FRR). Sur DS1, le système authentifie 80 % des utilisateurs avec une exactitude supérieure à 0.85. Sur DS2, cette exactitude est obtenue pour plus de 87 % des utilisateurs. 80 % des utilisateurs testés de DS3 ont une exactitude qui est autour ou supérieure à 0.90.

Nous avons également comparé notre KDF avec des générateurs de clé à partir des données biométriques proposés respectivement par, Seo et *al.* [49], Anees et *al.* [50] et Sheng et *al.* [51]. Notre fonction de dérivation de clé montre la nouveauté sur des propriétés telles que : la flexibilité de la construction, la révocabilité de la clé et son caractère aléatoire.

3.4 Volet 3 : Un schéma de paiement sur la blockchain avec la garantie du respect de la vie privée

Avec les schémas de paiement courants, les participants d'une plateforme de paiement ont généralement accès aux informations suivantes de chaque transaction : compte de l'expéditeur de fonds, montant et compte du receveur de fonds. Ces participants étant souvent nombreux, ces plateformes font face à un sérieux problème de privacité. Dans sa conception, la blockchain a tenté de pallier aux problèmes en identifiant le compte par un pseudocode (dérivé d'une clé publique). Toutefois, on a pu trouver des méthodes pour remonter aux informations physiques liées à un pseudocode. De plus, le problème s'aggrave parce que les informations sont répliquées sur tous les nœuds du réseau de la blockchain qui valident les transactions. La plupart des solutions au problème du manque de privacité dans les plateformes de paiement font massivement utilisation de la cryptographie. Ce qui influence négativement leur performance et leur utilisabilité. Le travail du deuxième volet vise à fournir un protocole de paiement simple garantissant la privacité. Il a donné naissance à l'article intitulé : « *Blockchain and Mobile Payment : Assessment on Privacy and Usability and a Scheme for Enhancement* », qui couvre l'objectif de recherche 2 énoncé à la section 1.3. Le texte de cet article est présenté au chapitre 6.

3.4.1 Protocole de paiement décentralisé

Notre schéma de paiement s'appuie sur une collaboration entre plusieurs institutions financières interconnectées par une blockchain. Cette dernière doit permettre de programmer les règles pour l'enregistrement et la lecture des transactions. Le point clé est d'éviter de donner accès à toutes les données de la transaction aux différents participants. Le receveur de fonds (*R*) envoie une requête de paiement indiquant le montant (*M*) et un compte temporaire (*tA*); il envoie ces mêmes informations à son institution financière (IF_R). L'expéditeur de fonds (*E*) transmet cette requête de paiement à son institution financière (IF_E) après l'avoir reçu. IF_E déduit le montant du compte interne de l'expéditeur et envoie la transaction *{tA, M}* sur la blockchain. Lorsque la transaction est

disponible sur la blockchain, IF_R lit le montant, crédite le compte interne du receveur et notifie ce dernier. Une fois notifié, R met fin au cycle.

3.4.2 Évaluation de performance

Pour évaluer la performance d'un tel protocole de paiement, nous l'avons implémenté avec la blockchain Ethereum. L'analyse en profondeur a permis d'établir les bonnes propriétés de ce protocole : le respect de la vie privée, l'utilisabilité et la sécurité. Pour l'estimation de performance, nous avons considéré des expériences effectuées avec la blockchain Ethereum et des études de performance sur le réseau public principal; ces expériences et études sont reportées dans la littérature. La durée des traitements internes à l'institution financière (incluant signatures et vérifications de signatures) pour une transaction reste inférieure à 0.17 S. Toutefois, la validation de la transaction sur la blockchain augmente cette durée qui peut atteindre 15 S avec Ethereum. C'est pourquoi, une perspective de modification indiquée dans la thèse n'attend pas la validation de la transaction par le protocole de consensus pour envoyer la réponse au receveur de fonds. Cette astuce permettrait une durée de la transaction en dessous de 1 S.

3.5 Volet 4 : cadre et architecture basée sur la blockchain pour des services de commerce électronique

Dans le volet 3, le projecteur était fixé sur la partie frontale de l'architecture qui est le protocole de paiement. Il restait à spécifier tous les composants de cette architecture et leur agencement. De plus, il est motivant d'avoir un cadre facilitant la mise en œuvre d'architecture basée sur la Blockchain pour des services de commerce électronique. C'est le but fixé dans le volet 4, correspondant aux objectifs de recherche 3 et 4 et atteint dans l'article intitulé « *Framework and Architecture for Blockchain-based e-Commerce Services* ». Le texte de l'article est le contenu du chapitre 7.

3.5.1 Cadre et modèle d'architecture basée sur la Blockchain

Afin de proposer un cadre de conception pour une architecture de paiement électronique basée sur la blockchain qui tienne compte des limitations de cette technologie, nous avons entrepris les étapes suivantes. Tout d'abord, nous avons analysé la structure du système de paiement électronique conventionnel et l'architecture de la blockchain afin de mieux comprendre leurs fonctionnements respectifs. Ensuite, nous avons effectué une comparaison approfondie des principaux protocoles

de consensus existants, en tenant compte de leurs propriétés impactant la sécurité et la performance des applications basées sur ces protocoles.

Sur la base de cette analyse, nous avons conçu une architecture basée sur la blockchain, en intégrant quatre éléments essentiels. D'abord, la fonction de dérivation de clé (KDF) qui facilite la récupération des clés de manière sécurisée, pour les systèmes basés sur la blockchain. Ensuite, le mécanisme d'authentification basé sur la biométrie (ou l'apparieur biométrique), offrant ainsi une méthode de vérification de l'identité de l'utilisateur. En troisième lieu, le protocole de paiement décentralisé, qui assure la privacité des utilisateurs tout en garantissant la sécurité des transactions. Finalement, la couche de blockchain. Un connecteur léger a été également spécifié pour faciliter l'interconnexion entre les différents systèmes institutionnels participants.

Dans le but d'évaluer la faisabilité de l'architecture proposée, nous avons réalisé une étude approfondie de sa mise en œuvre sur la blockchain Ethereum. Pour sécuriser les échanges entre le client mobile et le serveur, nous avons utilisé un protocole existant [52]. De plus, nous avons spécifié des points d'amélioration pour une extension du protocole SCP, permettant ainsi d'optimiser la performance de l'architecture.

3.5.2 Évaluation du modèle d'architecture

Le modèle concret proposé est construit à partir des composants conçus dans le cadre de cette thèse : la KDF incluant le mécanisme d'authentification associé et le protocole de paiement décentralisé. Une évaluation a été effectuée avec la blockchain Ethereum comme réseau d'arrière-plan. Cette architecture s'appuie sur les propriétés de ses composants pour la sécurité et le respect de la vie privée. L'analyse des coûts donne : 1185 octets comme coût des mémoires et 388 octets comme coût de la communication. De plus, la comparaison avec les systèmes proposés respectivement par Wang et *al.* [53] et Kim et *al.* [54] a permis de mettre en lumière certaines bonnes propriétés distinctives du modèle : la flexibilité en ce qui a trait à la blockchain en arrière-plan, la privacité et la facilité à respecter la régulation.

CHAPITRE 4 ARTICLE 1: MOBILE PAYMENT PLATFORMS: TAXONOMY, ARCHITECTURE, SECURITY, AND INTEGRATION WITH BLOCKCHAIN.

Auteurs : Olson Italis, Samuel Pierre et Alejandro Quintero.

Statut : Publié dans la revue « *Academia Letters* », le 20 septembre 2022.

4.1 Introduction

Electronic commerce covers any relationship of exchange of goods, physical or virtual services, using technological infrastructures [7]. Mobile payment is often used when at least a mobile device transfers funds for goods or services [3]. It has become prominent today, and some reports predict that expansion will continue [55], [56]. In the meantime, the industry has observed an incredible growth of attacks against online transactions, while electronic payment systems remain vulnerable to many of them [56].

Moreover, there are many known challenges in existing mobile payment systems (MPS) [47], [57], [58], [5]. Among these challenges, we mention the limitations of existing payment methods, the costs of technological infrastructures for some architectures, the decrease in mobile payment adoption, the lack of standards, and the interoperability between the numerous existing MPS.

Blockchain technology is expected to solve some of the challenges that we cited above. This technology offers a new way of deploying, securing many applications, and connecting different systems [59], [60], [61], [62], [63]. However, Blockchain has challenges that prevent its complete adoption in the real world. Choo, Kim-Kwang Raymond *et al.* [59] highlighted the need for schemes for decentralization, privacy concerns, trust, and performance issues.

The rest of this paper is organized as follows. In Section 2, we present a panorama of the most prominent types of MPS. We give a short description of Blockchain with enlightenment on the limitations of this technology for mobile payment applications in Section 3. We conclude with existing challenges and suggestions for future research directions in Section 4.

4.2 Technological Landscape of Mobile Payment

Mobile payment services are designed and deployed on a technical infrastructure: hardware, and software, interconnected by communication networks, operating with pre-defined protocols. We designate this infrastructure by mobile payment platform (MPP). Figure 4.1 is an illustration of a MPP with many stakeholders: the sender of funds, the receiver of funds, the banks, the payment service provider (PSP), the national or the international switches, the credit cards networks, and the mobile network operators (MNO). This illustration presents components of the platforms that support two scenarios. The first scenario considers proximity commerce: a buyer who pays in-store with a mobile device. This buyer can communicate with the system of the merchant (front-end network) by one of the following technologies [57]: Near Field Communication (NFC), Bluetooth Low Energy (BLE), or Quick Response Code (QRC). With NFC-based architecture, sensitive data is stored on a tamper-proof subscriber identity module (SIM). This SIM can be integrated into the mobile device of the buyer or stored on a server in the cloud in the case of Host Card Emulation (HCE). When the SIM is integrated into the mobile device, the PSP depends on the mobile or the SIM provider and the cost of deployment and maintenance is relatively important. In the case of HCE [34], the system is vulnerable to a man-in-the-middle (MITM) attack. However, with BLE and QRC, sensitive data is stored on the phone of the buyer. In the second scenario, the user (labeled as the fund's sender) sends funds remotely to a receiver, where the Internet is often used to connect all stakeholders. In some developing countries, the communication between the user and the PSP can be realized by a Short Message Service (SMS) and an Unstructured Supplementary Service Data (USSD), through servers hosted by the Mobile Network Operator (MNO) [64], [13].

4.3 Blockchain Technology

Satoshi Nakamoto [6] described Blockchain for the first time when he presented Bitcoin. Users of Blockchain must have a pair of cryptographic keys: the public key where the account on the Blockchain is derived, and the private key that allows users to spend their assets on the system by signing transactions. The private key must be kept securely.

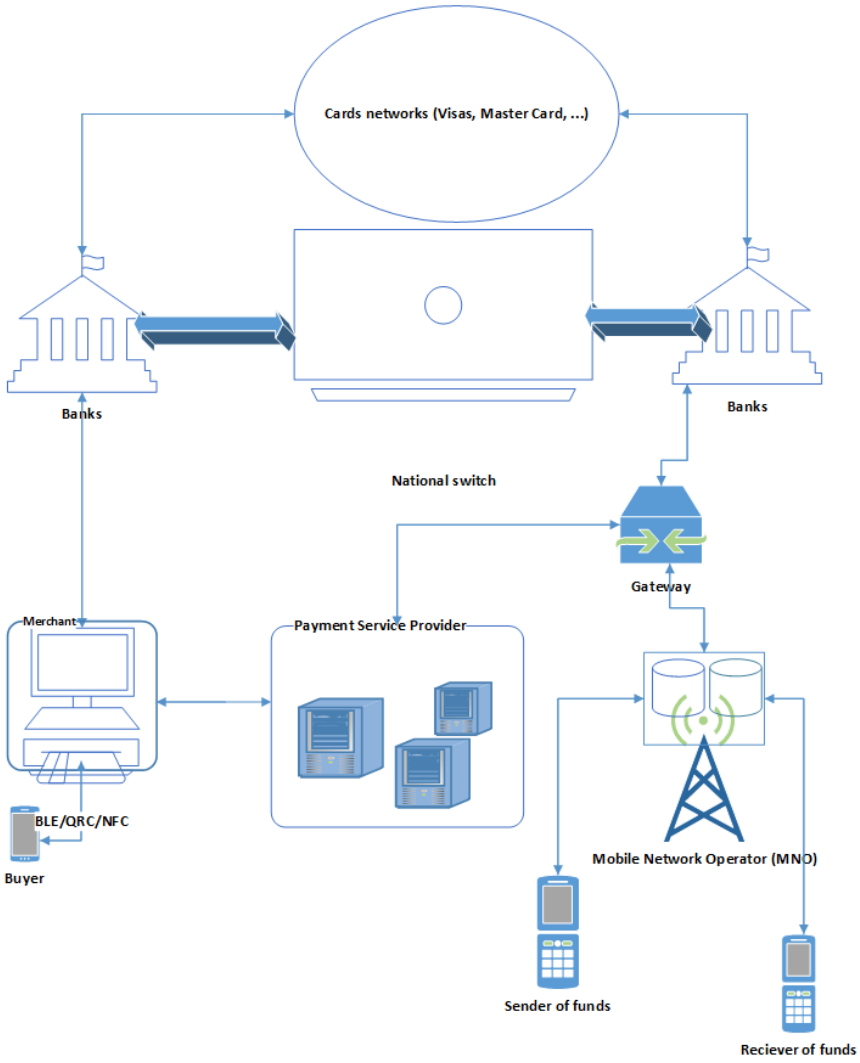


Figure 4.1 Integrated view of the current mobile payment platforms

Blockchain has many good properties. First, the technology is robust against falsification. Indeed, as the current block is attached to the preceding blocks, any previous data modification will imply a change in the data of the current block but not be accepted by the nodes which validate transactions. Decentralization is another attractive property of the technology, where a Blockchain system does not have a single point of failure. At last, we have easy traceability because transactions are available for all participating nodes.

Cryptocurrencies remain the most known and adopted applications of Blockchain, Bitcoin and Ethereum being the most notable ones [17], [65]. These popular and pioneering applications

of Blockchain are facing many issues that shape directions for improvement, such as performance and scalability, security, and privacy [15], [17].

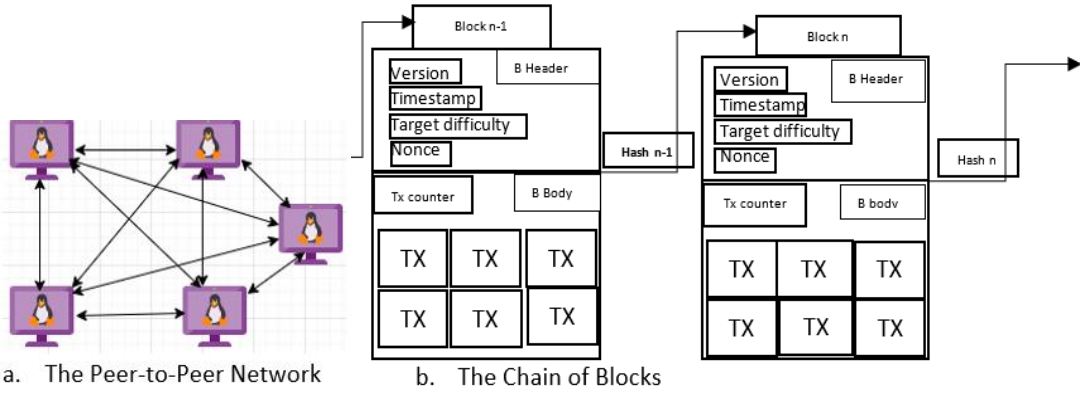


Figure 4.2 Blockchain: Illustration of the chain of blocks and the P2P model

4.3.1 Performance and Scalability

Bitcoin system can process less than ten transactions per second; Ethereum needs around 15 S to validate a transaction. These Blockchain networks are not suitable for mobile commerce, requiring a better throughput. Therefore, the poor performance of Bitcoin and Ethereum is due to the proof-of-work (PoW) protocol [6], [15]. Consequently, there have been many propositions to improve the performance of the Blockchain [15], [20]: many variants of PoW, Proof-of-stake (PoS), Delegated PoS (DPoS), and other similar protocols for public or permissionless Blockchains. These proposed protocols are of the type Proof-of-X (X is something like activity, importance, elapsed-time…). PoW is energy-consuming, but it is more costly for a malicious node to break the network. Others PoX-type protocols are less energy-consuming and also less secure. In PoS, the miner is chosen based on its assets in the system; DPoS (Delegated PoS) is a variant of PoS, where a selected group of nodes are pre-chosen to become miners [15]. With this kind of protocol, there is a risk of centralization, because the wealthiest nodes are more likely to be chosen as miners. Another flaw in these protocols is the lack of accountability for faulty nodes.

In the Practical Byzantine Fault Tolerance (PBFT) protocol, the validating nodes are supposedly known and have explicitly been granted the privilege to connect to the Blockchain network. Many protocols of the same class have been proposed [15], [20], [66], [67], [24]: Ripple, Tendermint, Libra, and other variants of PBFT. All are derived from the SMR-based distributed

computing concepts. PBFT protocol is not scalable. Indeed, experiences in [68] showed the difficulty of scaling above 16 participating nodes in the network. Moreover, there is a trend toward centralization.

4.3.2 Security and Privacy

Table 1 indicates that it is generally more expensive to break a public Blockchain than a permission Blockchain. It suggests that we have more security with public Blockchain. However, there is not any accountability mechanism for potential malicious nodes.

Table 4.1. Public and permissioned blockchains

	Permissionless Protocols	**Permissioned protocols**	**Comments for real applications in payment industry**
	PoW, and other variants	DPOS, PBFT, and variants	
Performance	≤ 200 tps	≥ 1000 tps	permissioned protocols are preferable
Tolerate power of the adversary	50% of computing power	≤ 33 % of fault nodes	Public protocols are more secure
Energy saving	No	Yes	Permissioned protocols preferable
Cost for users	Distributed, and there are incentives	No pre-defined incentive for the infrastructure	Public Blockchain is run by a great community. However, stakeholders must be maintaining the permissioned Blockchain network.
Decentralization	Yes	Partially	No accountability mechanism for faulty nodes in permissionless protocols

Although only the pseudocode identifies the user in the transactions, there is the possibility to link transactions of a user or even to find out his real-world identity [69], [70]. Therefore, Blockchain provides a limited privacy-preserving mechanism [71]. Moreover, it is not designed to fulfill transaction privacy requirements.

4.4 Conclusion

In this paper, we presented a taxonomy of the technological platforms for mobile payment services. These platforms can rely on BLE, QRC, or NFC to allow connecting a merchant with a buyer. With BLE and QRC-based MPP, sensitive data can be read by a malicious entity that has access to the mobile device. BLE-based systems are also vulnerable to Denial of Service (DoS) attacks. NFC

(with SIM integrated into mobile devices) implies a greater cost for exploitation, while protection of the link between the SIM in the cloud and the mobile is necessary with HCE. We also exposed some basic characteristics of Blockchain. Then, we highlighted challenges that would prevent using this new infrastructure for mobile payment services: key management problems, protocols with poor performance, security, and privacy.

These challenges in MPS and in Blockchain systems inspired by [55], [3], [57], [5], [2], [34], suggest many future research directions. We can aim at designing methods to protect sensitive user data in BLE and QRC-based MPP. New technology for easy exploitation of NFC-based systems is also welcome. For HCE architecture, it is essential to have a solid procedure to authenticate the mobile device of the buyer. It's worth mentioning the need for standards to ensure interoperability between many heterogeneous systems. Regarding Blockchain, the key management problem must be tackled; there must be a function to generate the key (without having it stored on an external server) in case of loss. There is room for improvement in Blockchain consensus protocols where we must find a good trade-off between security and performance. At last, ensuring transactions privacy in Blockchain may improve its applicability in the real-world payment industry.

CHAPITRE 5 ARTICLE 2: BLOCKCHAIN AND MOBILE PAYMENT: ASSESSMENT ON PRIVACY AND USABILITY AND A SCHEME FOR ENHANCEMENT

Auteurs: Olson Italis, Samuel Pierre, Alejandro Quintero.

Statut: Publié dans « *2022 Fourth International Conference on Blockchain Computing and Applications (BCCA)* », le 31 octobre 2022.

Abstract

Blockchain is a new paradigm to realize payment without a single Trusted-Third-Party. The technology exploits cryptography to secure all the transactions that are available to participating nodes for validation via distributed consensus algorithms. Blockchain-based architectures for mobile payment face challenges like transaction privacy and performance issues. We analyze these architectures and assess the privacy issues and the performance in the real world. Then, we propose a payment scheme to guarantee the privacy of the transactions with discussions on tricks to improve the performance. We study the feasibility of implementation of the proposed scheme, both on public and consortium Blockchains. The payment scheme ensures that a participating node has access to only a part of the meaningful data of a transaction. The expected performance with Hyperledger Blockchain (with less than 16 participating nodes) is more than 1000 TPS. We indicate updates to decrease the duration of a transaction from 15 S to less than 1 S with public Ethereum Blockchain.

5.1 Introduction

Mobile payment, a subset of electronic commerce refers to any payment transaction that implies buying goods or services realized with the help of a mobile device [3]. Operations in mobile payment are carried out by the mobile payment system (MPS): a technological platform composed of hardware; software interconnected by communication protocols. These systems rely on a Trusted-Third Party (TTP) to ensure the safety and liveness of transactions. We designate these systems as conventional mobile payment systems (CMPS). In CMPS, when user U_S makes a payment to user U_R, U_S sends to his financial institution (FI) his account, the amount and the

account of U_R. Much information is then available for the financial institution and other stakeholders of the payment process like hardware providers and network operators [57]. Therefore, CMPS has privacy issues.

With CMPS, payment operations are centralized into the servers of the TTP (therefore, a single point of failure). This is opposed to Blockchain-based systems that are decentralized. Nakamoto described Blockchain in [6] to enable peer-to-peer transactions without a TTP: cryptographic algorithms secure these transactions and all participants are involved in the validation process using consensus algorithms. However, the payment industry hesitates to use Blockchain for many reasons [107]: privacy issues, poor performance, lack of accountability and control mechanism. Regarding privacy concerns, this is the functional characteristic of Blockchain where transactions are available for participating nodes. The poor performance of Blockchain precludes its adoption in the real world [20], [108]: as a comparison, the Visa network can process above 65000 transactions per second (TPS). Several payment systems built on top of Blockchain have been proposed in the literature [61], [62], [109], [110]. While they avoid the performance issues by building a new payment channel with some trusted participating nodes (i.e., financial institutions), privacy concerns remained untouched.

In this paper, we analyze the limitations of the Blockchain for mobile payment applications and propose a scheme to overcome privacy concerns. Moreover, we indicate a direction of improvement in the performance of systems based on Ethereum Blockchain.

The main contributions of this paper are as follows:

1. Design a usable and privacy-preserving Blockchain-based payment scheme by proposing a trick to ensure that every participating node has access to the exact pieces of information it needs for its role, without revealing all the meaningful data of the transactions.
2. Conduct an in-depth analysis of the security and usability of the proposed system showing its robustness and privacy-preserving property.
3. Propose an algorithm to realize anti-money laundering (AML) functionality in a context where the single ordinary node has access to only a part of the transaction.
4. Assess the expected performance from such a payment scheme with the Ethereum Blockchain and indicate a direction for improvement.

The rest of the paper is organized as follows. Section II presents a literature review on existing blockchain-based MPS. The new proposed system is detailed in Section III. Analysis of the security and usability of the new system are discussed in section IV. Section V concludes the paper.

5.2 Related Work

In this section, we analyze privacy and usability in both CMPS and Blockchain-based systems. Then, we briefly describe some Blockchain-based mobile payment systems.

5.2.1 Privacy and Usability of CMPS

Researchers design privacy-preserving electronic cash (i.e., e-cash) systems with centralized authority using anonymous signatures (i.e., blind, ring) [111], [112]. There are many propositions of such systems in the academic world. However, they are not popular in the real world, where the great majority of CMPS are facing privacy-preserving issues as mentioned in the introduction section.

The centralized e-cash systems make extensive use of cryptography to ensure respect for privacy. They are built on protocols with encryption and anonymous signatures [111], [112]. These systems face the challenges of efficiency and scalability, as noted in [113]. The non-adoption in the real world may be a confirmation of this observation.

5.2.2 Privacy and Usability of Blockchain-based Systems

Blockchain is designed to fulfill transparency: the content of a transaction is available for each participating node. This property is opposed to transaction privacy property. First Blockchain-based systems are public, meaning that transactions are available for everyone on the internet wanting to connect to the network.

Although only a pseudocode identifies the user in the transactions, researchers trace and link transactions of a user or even find out his real-world identity by analyzing many transactions [69]. Traditional Blockchain-based systems provide identity privacy with cryptographic algorithms, but to a limited measure [71].

The poor performance of Blockchain-based systems (less than 300 TPS) precludes their usability [20], [108]: this is far below the capacity of the visa card network that can handle tens of

thousands of TPS. The lack of accountability for malicious nodes and control mechanisms makes it difficult for enterprises and regulators in mobile payment to adopt Blockchain.

5.2.3 Blockchain-based Systems

Reference [61] presents a payment scheme to transfer cryptocurrencies relying on a central bank or a set of financial institutions playing the role of the notary. This system has two merits. The first is to overcome the great delay for a transaction in public blockchains like Bitcoin (10 mn) by confirming the transaction with the central bank or the set of notary before its validation. The second is to provide Know Your Customer (KYC) and Anti-Money Laundering (AML) functionalities. However, using existing cryptocurrencies makes the system vulnerable to the price stability problem. Moreover, there is a need for a liquidity provider.

L. Zhong et al. [47] propose a secure and light payment system based on Blockchain. In the design, an agreement between payer and payee is established to be verified by participating nodes off-chain and before final validation. This is a manner to avoid the delay of the Blockchain validation process. In place of the digital signature algorithm, this system uses light one-way function to authenticate users. This system is then suitable for resources constraint devices.

The payment scheme described in [109] relies on Ethereum Blockchain and intents to tolerate networks with low QoS. It considers scenarios in remote rural villages where connexions are often intermittent. The village is connected to the online network by proxy via backhaul connections. Each village runs its proper Blockchain network where payment operators process the local transactions. The proxy saves account balance for the village and makes synchronization with banks when a connection is re-established.

The system in [110], uses a combination of Bitcoin and Hyperledger. A notary (for instance bank) has the privilege to mint tokens with the Hyperledger network and creates an unspent transaction (UTXO). The final user, who is the owner of the UTXO, can spend it later. He generally runs a light application on top of the Bitcoin network.

Table 5.1 Summary of the characteristics of Blockchain-based payment systems

Proposed systems in [61], [62], [[109], [110]	
Advantages	***Limitations***
▪ Overcome duration of transaction validation in popular Blockchain networks. ▪ They all promote the idea of relying on trusted financial institutions while using the Blockchain. ▪ Some offer functions to achieve compliance with financial laws like AML functions. ▪ Using light cryptographic functions.	▪ Use of popular and volatile cryptocurrencies. ▪ Need a liquidity provider. ▪ Some only allow cryptocurrency exchanges. ▪ They have all privacy preserving problems.

In all the preceding proposed Blockchain-based payment systems (BBPS), transactions are propagated and validated in the native Blockchain networks without any mechanism to hide the account of the sender or the receiver of funds. The threat that traces many transactions of a user is always present. We can summarize the advantages and limitations of these systems in Table 5.1.

5.3 The Proposed Payment Scheme

In this section, we describe our proposed system starting with the entities and their roles. Then, we present a normal payment scenario that explains the functional aspects of the system. This scenario highlights the components and the procedures of the system for which we give details at the end of this section.

5.3.1 Entities and their roles

5.3.1.1 User: physical person

The user is a sender or a receiver of funds. In local commerce scenario, the sender is a buyer and the receiver is a merchant. The user's module performs the following tasks:

- Generates a private key, then finds the corresponding public key according to a specific elliptic curve. The private key can be hosted on a Hardware Security Module (HSM) included in the user's mobile device. In this last scenario, the authentication step can give access to the generated key.
- Requests transactions when the user receives funds; and creates temporal accounts using a pseudorandom function to receive funds on the Blockchain.

- Signs and sends data to financial institutions: the sender of funds sends a request for payment to his financial institution, and the receiver of funds sends the temporal account to his own.

5.3.1.2 Merchant: a moral person

The merchant is a receiver of funds. He can have many public keys derived from many private tokens. He performs the following tasks:

- Creates temporal address and shares it with his financial institution.
- Requests payment from the buyer.
- Acknowledges received payment and delivers goods.

5.3.1.3 Financial Institution

A Financial Institution (FI) can be a bank or any Payment Service Provider (PSP). He runs one node or a group of nodes of a Blockchain network which can be a public or a permissioned Blockchain. A financial institution performs the following tasks.

- Creates and sends transactions on the Blockchain network on behalf of the sender.
- Maps temporal addresses of the receivers (e.g., its clients) to internal corresponding accounts.
- The financial institution of the sender of funds (FI_S) debits the internal account of its client according to the transaction sent on the Blockchain.
- The financial institution of the receiver of funds (FI_R) credits the internal account of its client according to the transaction received from the Blockchain.

5.3.2 A Normal Payment Scenario

Fig. 1 illustrates a scenario of mobile payment between a buyer and a merchant. The following steps describe the scenario:

1. Receiver of funds requests payment from the user. In local commerce, he can create a quick response code from the temporal address tA, the amount M, and optionally his signature for authentication $I_R=\{tA, M, P_{k,R}, SignR(tA, M, S_{k,R})\}$, where $P_{k,R}$ and $S_{k,R}$ are respectively the private and the public key of the receiver of funds. In remote payment, the signature is necessary; in this case, the receiver encrypts a set of information with the public key of the sender to transmit the data.

2. Receiver sends I_R to FI_R; he will generally encrypt this information with the public key of FI_R; FI_R maps tA to the internal account of the receiver IA_R.
3. The buyer decodes the quick response code (QRC) in local commerce or decrypts I_R and verifies the signature if applicable, then transmits the payment request to his financial institution FI_S. He sends I_S={tA, M, IA_S, Sign(tA, M, IA_S)}. It is recommended to encrypt this information with the public key of FI_S.
4. FI_S debits the internal account of the sender (IA_S). Then, it creates the Blockchain transaction (BTx) and broadcasts this transaction to the Blockchain.
5. At receiving the transaction, FI_R verifies the validity, updates IA_R, and confirms reception with the receiver.
6. Merchant delivers goods and/or receipts to the buyer.

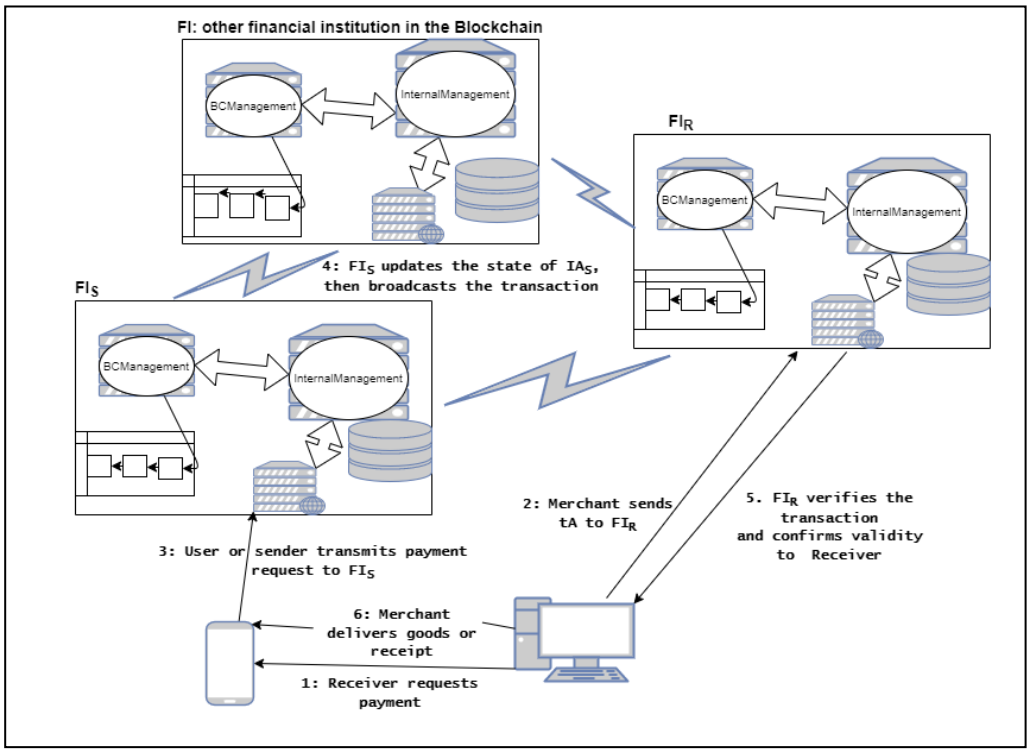


Figure 5.1 A normal payment scenario in local commerce

5.3.3 Management of the Payment Process

Financial institutions involved in the payment (FI_S and FI_R) run 2 modules to manage payment transactions: *internalManagement* and *BCManagement*. We describe these modules with illustration below (Fig. 5.2).

5.3.3.1 Module InternalManagement

When FI is the sender of funds, this module receives the request for payment from the buyer, verifies his authenticity, debits his internal account, stores information for traceability and sends the transaction for the Blockchain to *BCmanagement* (Fig. 2). When FI is a receiver of funds for a client, the module will credit the account of the client according to the amount of the transaction received on the Blockchain.

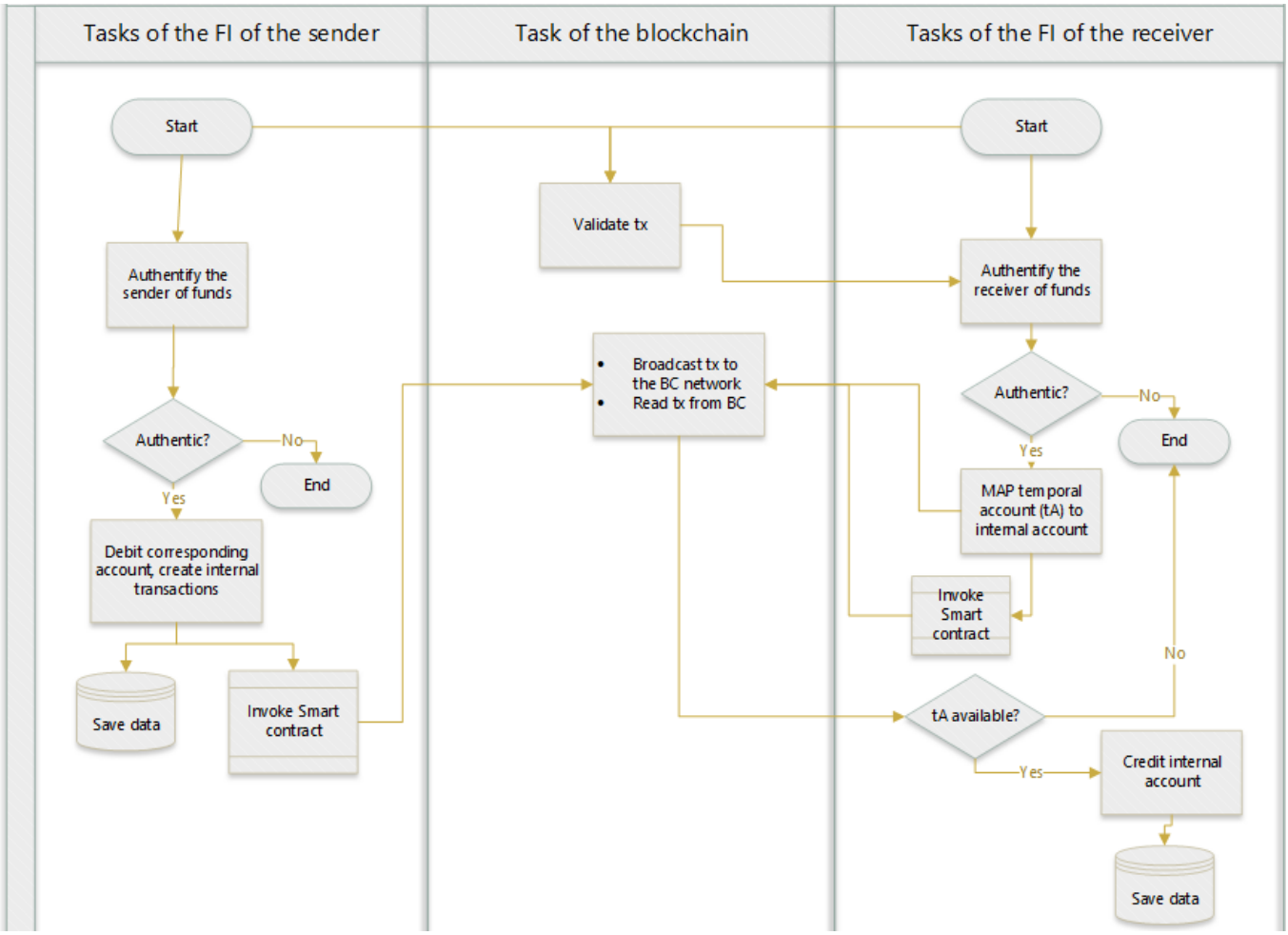


Figure 5.2 Cross flowchart for the management of the payment

5.3.3.2 Module BCManagement

This module acts as an oracle to feed and trigger the smart contract which manages the transactions on the Blockchain (Fig. 5.3). The data includes the information to send on the Blockchain, $I_{S,BC}$={tA, M}; tA is the temporal account to receive the value, M is the amount of the transaction. Optionally, we can add an address ($AF_{I,R}$) of the financial institution of the receiver. It could be used for future settlements when FI_R will require the total paid in fiat currency. We consider for this application, Blockchain systems empowered by programmable chain-codes, like Ethereum and Hyperledger Fabric [16], [17]. Chain-codes are typically called smart contracts in Ethereum and allow the registration of the assets on the underlined Blockchain (Figure 5.4).

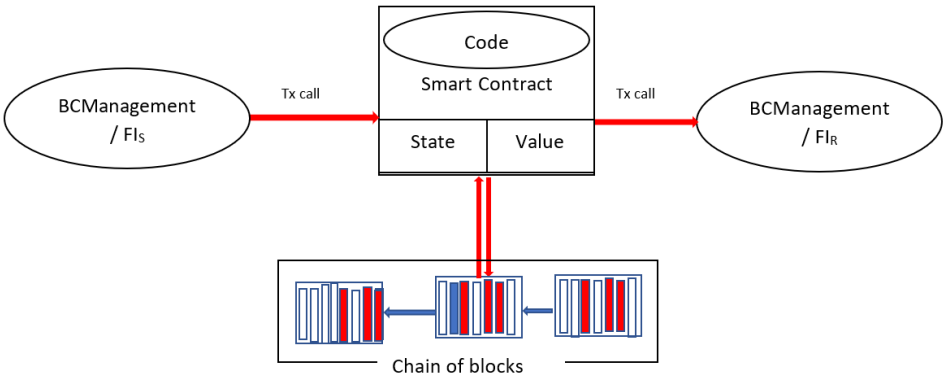


Figure 5.3 Interaction between the module *BCManagement* and the Blockchain to register or read the transactions

Algorithm 1: Contract TxManagement

```
1:  Transaction[] public lstTransactions
2:  Struct Transaction {
3:          string temporalAccount
4:          uint amount }
5:  EndStruct
6:  function CREATETRANSACTION(tAcnt, amnt)
7:      trans ← Transaction(tAccount, amnt)
8:      listeTransactions.push(trans)
9:  end function
10: function GETTRANSACTIONAMOUNT(tAcnt)
11:     p ← lstTransactions.length − 1
12:
13:     while p > 0 & lstTransactions[p].tAcnt! = tAcnt
    do
14:     p = p − 1
15:     return listeTransactions[p].amount;
16: end function
```

Figure 5.4 Basic operations of the Smart Contract

5.4 Analysis and Evaluation

We perform in this section an analysis of the security and privacy of the proposed scheme. Then, we give an estimation of the expected performance with an indication for improvement.

5.4.1 Security and Privacy

Attacks against a Blockchain-based system can be classified with respect to 4 layers [114], [115]: the P2P network is the first and most basic layer, the consensus mechanism is the second layer, the third layer is composed of the Blockchain data, and the fourth one is the application layer where we can find wallets, smart-contracts and off-chain programs for implementing logic. Our proposed scheme fits in the last layer. Then, the three first attack vectors are out of the scope of this study. We rely on existing security mechanisms for these layers. Specifically, our scheme profits of the security of proof-of-work when deployed with Ethereum.

We consider the misbehavior of the sender or the receiver of funds; they are at the layer of the client application. Our assumption is each honest user can trust his own institution (bank or other payment service provider), and there is a secure channel between an honest user and his financial institution. We also consider the known threats relative to smart contracts as mentioned in [42] and [115].

5.4.1.1 Hacking as a Sender of Funds

In the role of the sender of funds, a hacker can misbehave in two ways: pay less than what is required by the receiver or embezzle the funds to his own account. In the first scenario, FI_R can verify the amount of the transaction by comparing it with the amount sent by the receiver. Secondly, in the case of embezzlement, the transaction will not be completed, since the receiver waits for an acknowledgment from FI_R.

5.4.1.2 Hacking as a Receiver of Funds

In the role of the receiver of funds, the hacker can attempt to have a double payment for the same transaction. In this case, the system includes control measures at the sender side to avoid sending the same payment twice to FI_S: the sender module will always ask for confirmation from the user before transmitting any request for payment and waits for receipt before sending any other request.

5.4.1.3 Dealing with Contract Vulnerabilities

Smart contracts have several vulnerabilities, Chen et *al.* [115] presented 26; good development and deployment practices help prevent many of these vulnerabilities as well analyzed by the authors. Others are not applicable to our scheme because the scenarios which cause the vulnerability are not viable. We show here how the proposed scheme mitigates those of them that represent great risk:

- Vulnerabilities prevented by design

  DoS with unexpected Revert: a transaction is reverted if a caller contract fails to complete execution. A callee contract can induce this failure. In the proposed scheme, FI_S has the willingness to complete the transaction (assumption) and it cannot deceive FI_R who received the transaction from its client. *Confidentiality Failure* due to the public nature of the Blockchain is mitigated by sharing partial meaningful data on the Blockchain. *Insufficient Signature Information* is mitigated by the fact that FI_S sends a signature for each transaction.

- Vulnerabilities prevented by authentication

  Enforcing authentication of the financial institution before launching any function mitigate the following vulnerabilities: *DelegateCall Injection* where a malicious callee contract can manipulate the state variables of a caller contract, *Unprotected Suicide* where an attacker can use the self-destruct method to kill the contract, *Leaking Ether to Arbitrary Address*, and *Ether Loss to Orphan Address*. We keep authentic addresses in a list of structs {address, receivedAmount, sentAmount} to compare with the address triggering the smart contract.

- Other opened vulnerabilities

  Upgradable Contract: the issue is, that the contract developer responsible to upgrade a contract becomes malicious or is attacked. What protects the system against updates from a malicious source? We rely on the security of FI and good authentication to mitigate the vulnerability.

  Transaction Ordering Dependence: When several dependent transactions invoke the same contract, it can occur a kind of deadlock where a transaction being executed needs the output of a pending transaction. In our scheme, a request to read the amount of a transaction can be placed by FI_R while the struct {temporal account, amount} is not yet pushed in the

Blockchain. The consequence is the delay of the transaction. We will give an estimation of the maximum duration with directions for improvement later.

5.4.1.4 Robustness of the Infrastructure

Blockchain on its own has already the advantages of decentralization. However, as we trust a set of institutions, the tendency for centralization is present. Adopting a policy of more than one node by financial institution mitigates this threat. Also, this strategy increases the resilience of the system in case of an attack against a financial institution.

5.4.1.5 Privacy-Preserving Characteristics

When a user sends a transaction on the Blockchain, only the financial institution of the user (FI_S) knows that he pays someone. FI_S does not recognize the receiver. Likewise, FI_R receives the transaction on the Blockchain from the Blockchain account of FI_S and does not know the sender of funds.

5.4.2 Usability of the Proposed Scheme

5.4.2.1 Settlement between Financial Institutions

We can implement a distributed module for the clearing process. it holds a table with the Blockchain addresses of each FI and their corresponding identifiers hosted in the clearing database. The national sort code (NSC) is an eligible identifier for this task; other candidates are the bank identifier code (BIC) if the consortium wants to use the Swift system for a second channel payment. During a given period, a central bank or a committee of financial institutions can have the responsibility for the clearing process. All transactions emitted or received by each FI are collected. A report can be produced to give the total amounts of money to pay by or to each FI.

5.4.2.2 Facility of deployment

If the system is deployed on a public blockchain as Ethereum, the only task for the financial institution is to create a new account for the Blockchain and subscribe with the consortium by submitting its identifier. Then, it connects to the Ethereum network. It can have a full node or a group of nodes, but this is not necessary. The consortium can establish a private network using any permissioned Blockchain.

5.4.2.3 Accessibility to the Service

Final users are primarily clients of the participating financial institutions and identified here by U_S for the sender of funds or U_R for the receiver of the funds. We can prospect that any users can buy this service temporarily if financial institutions are willing to do so.

A mobile network operator offering a payment service can easily connect the network by setting a node or many nodes of the Blockchain to receive funds for its client. The module of the sender of funds can generate a pseudo-random number (the temporal account) and communicates it to the beneficiary using out of band channel, like short message service (SMS).

5.4.2.4 Compliance with the AML laws

The system is designed with the idea that each sender is known by the financial institution, which sends payment transactions on the Blockchain. Even if, no entity can access the identity of a sender or a receiver of funds by simply monitoring the Blockchain, each institution knows the source of the transaction it sends. Moreover, each institution knows the identity of the receiver when this last one is its client.

We design a distributed and autonomous module to be executed by legal authority to find the source and the destination of a given transaction. In a real scenario, the authority will have the date (dateT) of the transaction of interest and the identity of the sender (U_S) or the receiver (U_R) of funds. The following steps describe the algorithm of such a module, considering notations in Table 5.2.

- Let Tx_{FIs} = {dateT, Us, tA, M},
- let Tx_{FIr} = {dateT, U_R, tA, M},
- wT is found by a natural join on the 2 tables Tx_{FIs} and Tx_{Fir}: wT = {dateT, M, U_S, U_R}.

Table 5.2 Notations for the AML module

Symbols	**Description**
DT	Date (including the time) of the transaction.
M	Amount of the transaction
tA	Temporal Account
Tx_{FIs}	A transaction in the financial institution of the sender
Tx_{FIr}	A transaction in the financial institution of the receiver
U_R	Information identifying the receiver
U_S	Information identifying the sender

wT	The whole transaction, built by the AML module

5.4.3 Model and Estimation of the Performance

In this subsection, we give an evaluation of the number of transactions that can be processed by the system per unit of time, and we estimate the duration of a transaction. In all cases, we consider the state of the art in terms of the performance of the underlined Blockchain network.

5.4.3.1 Computational Cost

With d_t the total duration of a transaction, the notations and estimated value in Table III, we have the following expression for the duration of a transaction.

$$d_t \leq 2 * d_\sigma + 2 * d_{v,\sigma} + d_{IP,FIs} + d_{BC} \quad (1)$$

In expression 1, both signatures are performed by merchant and buyer, respectively. Their verifications are performed by involved financial institutions. They are not totally sequential processes, this is why we have inequality.

To estimate $d_{IP,FIs}$, which is the duration of the internal processing, we implement the methods in Python [21] and run them on a laptop with 12 GB of RAM and processor Intel (R) Core (TM) i7-4600U, CPU @ 2.10GHz.

With the maximum duration for a signature and the verification report from the experiments in [20], the value of d_t is more than 15 S for public Blockchain (because d_{BC} is around 15 S) and less than 1 S for some Ethereum consortium Blockchains (Table 5.3).

Table 5.3 Estimated values of the transaction delay

Symbols	**Description**	**Estimated values [S] Consortium Blockchain**
dσ	Maximum estimated time to sign a transaction.	$2.15*10^{-3}$ [116]
dv,σ	Maximum estimated time to verify a signature.	$4.25*10^{-3}$ [116]
$d_{IP,FIs}$	Duration of the internal processing at the financial institution of the sender	≤ 0.15
d_{BC}	Delay of the transaction propagation on the Blockchain	$4.5*10^{-3}$ [117]
d_t	Total duration of a transaction in seconds	0.1673

Table 5.4 Throughput of a consortium blockchain

Reference	Environment / Parameters		Ethereum	Hyperledger
Blockbench, a framework for analyzing Private Blockchain [117]	number of nodes	8 clients 8 servers	284 TPS	≥ 1000 TPS However, scalability issues # of nodes (servers)<16
	Processor/ Memory	3.5 GHz CPU, 32 GB RAM		
	OS	Ubuntu 14.04		
	Link between nodes	1 GB switch		
	Number of requests	8 TPS to 1024 TPS		

5.4.3.2 Prospection on Throughput

Let n, the number of servers (each server is running inside a FI) connected to Blockchain consortium, Bs the number of transactions that can be processed by a server per unit of times; we assume homogeneity. A transaction involves two servers (FI_S and FI_R).

In the most desirable scenario, we can have during a unit of times, each distinct group of 2 servers involved in processing a transaction, where the total transactions in processing would be Cs(n) = $n/2$ * Bs. As an example, if a server can process 1000 TPS and our system has 48 interconnected servers as in [117], it will process 24000 TPS. In the worst case (one server is involved in all transactions), our system will process 1000 TPS.

The preceding formula is intended to show the capacity available from the set of nodes; it's not complete without considering the effects of communication between nodes and above all, the limits imposed by the Blockchain consensus protocol. Then, at an instant, our system could have a throughput T, such as:

T = min{C_S, Throughput of the blockchain} (2)

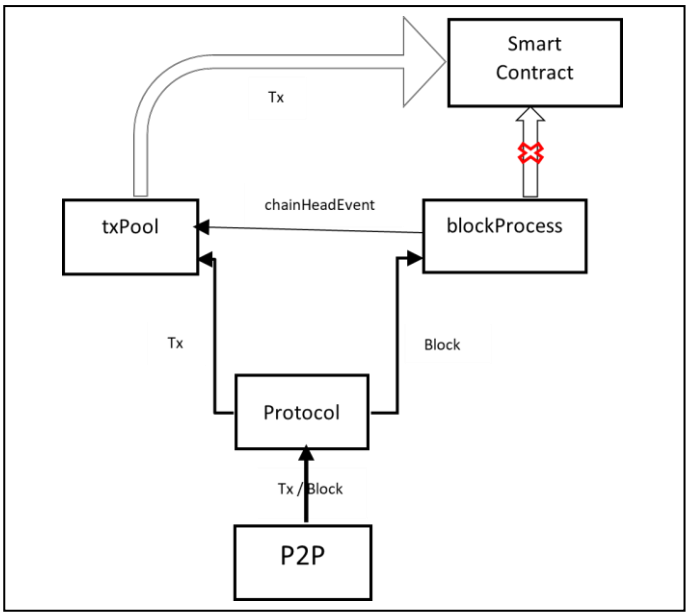


Figure 5.5 Reading data from the transaction pool with a smart contract

Many experiments estimated the performance of permissioned Blockchain, [20], [117], [108]. The best results come from [117] (Table 5.4): 284 TPS for Ethereum permissioned Blockchain and above 1000 TPS for Hyperledger. Parameters and results are reported in Table IV.

5.4.4 Indication for Performance Improvement

The preceding description of *BCManagement* is more suitable for consortium blockchain, where performance is acceptable with some applications. However, for public Blockchain, the broadcasted transaction is not instantaneously available in a block to be read by the smart contract. Then, to exploit the proposed scheme on a public Blockchain, we have to read the uncommitted transaction, just after the broadcast, as it is done in [118]. This process is illustrated in Figure 5.5, adapted from [119]. When the broadcasted transaction arrives, it is sent to *txPool* before reaching *blockProcess*, the module where validation and creation of blocks are processed. We can capture transactions in *TxPool*. In this scenario, the delay of a transaction depends only on the propagation delay, and not the consensus protocol. According to Wang et *al.* [119] a transaction is broadcasted to all nodes of the public Ethereum network at around 200 mS. With this broadcast delay and the estimation reported in Table III, the expected duration for a transaction in our scheme is less than 0.4 S. This is acceptable in mobile commerce.

5.5 Conclusion

In this paper, we designed a Blockchain-based payment scheme to overcome challenges that prevent the full adoption of Blockchain-based systems. Privacy and performance are the two main targeted challenges. We analyzed the privacy and performance issues in the existing systems: Blockchain-based or conventional mobile payment platforms in the real world. With the proposed scheme, a node of the Blockchain (FI) has access to the exact information it carried to process the transaction. We expect a performance of more than 1000 TPS with Hyperledger Blockchain and a transaction delay less than 1 S on consortium Blockchain (Ethereum and Hyperledger). We assess the usability of the proposed scheme by analyzing the security of the transactions, the accessibility for the final user (sender or receiver of funds), and presenting procedures to implement AML functionality and clearing between financial institutions. Finally, we indicate a modification at a high level for usability with a public Blockchain, to increase the throughput and decrease transaction delay from 15 S to 0.4 S.

One can extend this research work by formal and quantitative evaluation of the security and the privacy of the proposed scheme. Moreover, improvement of the performance when uncommitted transactions are read from both the public and consortium Blockchain is a direction for research.

CHAPITRE 6 ARTICLE 3: PRIVACY-PRESERVING MODEL FOR BIOMETRIC-BASED AUTHENTICATION AND KEY DERIVATION FUNCTION

Auteurs : Olson Italis, Samuel Pierre, Alejandro Quintero.

Statut : Soumis dans « *Journal of Information Security and Applications* », le 22 février 2023.

Abstract

Bio-cryptosystems often save the biometric template for authentication and generally employ randomly generated keys to encrypt and sign data. This method implies privacy issues. Furthermore, in Blockchain-based systems, the repercussions of losing a key are more severe for the owner compared to traditional systems. Indeed, in a Blockchain-based system, there is no reliance on a trusted third party for key recovery. To overcome the privacy issues, and allow the secure recovery of lost keys, we design a Key Derivation Function to extract a key from biometric data: a new method-based on clustering algorithms - detects consistent and discriminative features from biometric characteristics to create a code. Then, HMAC-SHA256 (as specified by the National Institute of Standards and Technology) generates a standard key from the code. To reproduce the code at future times (that also serves for authentication), the Key Derivation Function stores helper data with the guarantee of privacy. Indeed, with a private face dataset, the probability of generating the code with only the helper data is less than 2^{-300}, and less than 2^{-246} for a subset of the Youtube face database. Moreover, on the private database and the tested users from Youtube face database, our system has a false acceptance rate of 0%. It corrects up to 40.3% of noise levels on the private database and has good management of the inter-user variability.

6.1 Introduction

Existing mobile payment platforms and other systems are based on many methods for user authentication. Traditional authentication methods include the usage of a physical token or a logical password like a Personal Identification Number (PIN) [72]. Nowadays, biometric characteristics are very popular, because they are more convenient and imply a more perceived usefulness and

trust than the older methods [73]. Biometric-based authentication systems generally save a template to allow authentication. With this method, we have privacy concerns.

To improve security through encryption and signatures, such systems often use randomly generated cryptographic keys. However, in the scenario where the user saves his key on a mobile device, we cannot recreate this key if the device is lost. This problem relative to key management is more crucial for Blockchain-based systems [74] where the user has no means to recover his assets without the key. A way to overcome this challenge is to derive the key from some data given by the user: what he knows as passwords or biometric characteristics (what he is) [75], [76], [77]. Cryptographic key can be derived from biometric data in 2 ways [75], [76]: by *Fuzzy Extractor* (FE) constructions and *Key Derivation Function* (KDF). Many problems are known in the first category. Some of the systems of this category save too much data (helper data) to recreate the key, implying privacy-preserving issues [78], [79]. Moreover, they are unsuitable for mobile devices with limited memory resources [49]. Others are not robust enough in terms of security: practical implementation of *Robust Fuzzy Extractor* (RFE) can correct only 3% of the noise level [76]. Therefore, the RFE has a poor ability to handle intra-user variability. In the second category, (i.e., KDFs), some systems often fail to consider inter-user and intra-user variability [76], while others have privacy issues.

In this paper, we propose a model for biometric authentication with a privacy-preserving KDF to link a digital identity (for instance on the Blockchain) with a real physical person. Our model includes a new method for handling fuzziness in biometric data, by learning discriminative and consistent features for future authentication. The KDF is a light recovery mechanism in case of loss of the cryptographic key, intended to be used with many methods of feature extraction. It minimizes the size of the helper data saved for later authentication or reproduction of the key. The new KDF is designed to guarantee randomness, without assuming a specific biometric characteristic with enough entropy. The mechanism to generate the key also allows the revocability of the derived key. The main contributions of this paper are as follows:

- A method for coding biometric data; this method includes a quantization method for handling fuzziness in biometric data, using clustering algorithms.

- A privacy-preserving matcher for biometric authentication that incorporates two primary steps: learning a sketch to reproduce the code in the future and conducting a comparison between the first code and the new one for authentication.
- A KDF based on Hash Message Authentication Code (HMAC); this HMAC-based KDF (HKDF) integrates the previous biometric code and a user identifier as input data, allowing us to generate a random, irreversible, and revocable cryptographic key.

The rest of the paper is organized as follows. Section II gives a background on *Fuzzy Extractor* systems, *Key derivation Functions*, and other key concepts used in this paper. Section III presents a survey of existing systems allowing us to create a cryptographic key from biometric data. Section IV details the proposed system. Experimentations and results are reported in Section V. In Section VI, we perform general discussions. Section VII concludes the paper with indications for future research directions.

6.2 Background

One of the first and most well-known designs presented to convert noisy data into a cryptographic key is the so-called *Fuzzy Extractor* (FE) as defined by Doddis *et al.* [75]. More recently, *Key Derivation Function* (KDF) has been proposed to create a cryptographic key from some user data [80], [81]. In this section, we review definitions of these constructions and other concepts like k-means clustering and the Jenk Natural Break (JNB) algorithm.

6.2.1 Fuzzy Extractor

Considering a metric space M, integers m and l, real values t, and the *(M, m, l, t, ε) Fuzzy Extractor* is a pair of procedures *Gen.* and *Rec.* (Fig. 6.1) with the following properties [75]:

1. *Gen.* takes as input the noisy data $w \in M$ and generates an almost uniform random string $R \in \{0,1\}^l$. In other terms, Gen. is a function of w with a random string X that outputs a random string R, which is denoted by *Gen.(w, X)=R. FE* contains a helper and public data $P \in \{0,1\}^*$ which incorporates Secure Sketch *(SS)* allowing to recreate w from a closed w'; *(SS)* is often designed with codewords (C) and combined with the initial random string X to produce the public helper P.

2. *Rec.* takes as input $w' \in M$ and $P \in \{0,1\}^*$ to recreate w; the correctness of *FE* is defined by these implications: if $dis(w,w') < t$ and R, P generated by the procedure *Gen.*, so $Rec.(w', P) = w$; if $dis(w,w') > t$, no guarantee is given on the output.
3. Security property is guaranteed by the fact that R is nearly uniform even after observing P. In other words, with m, a guaranteed minimum level of entropy of w, R is uniformly distributed even given P. It can be expressed as: $Gen.(w) \rightarrow (R, P) \Rightarrow SD((R,P),(Ul,P)) < \epsilon$.

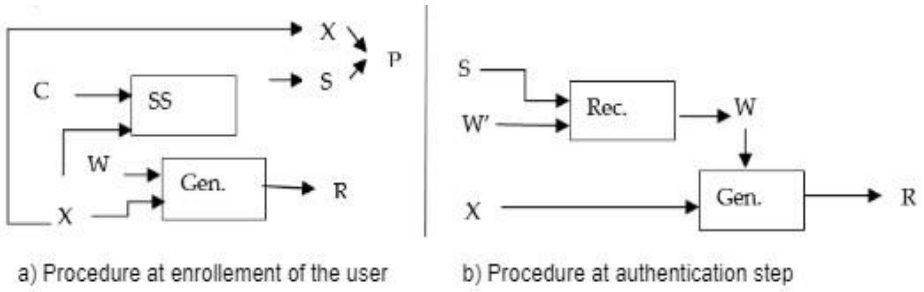


Figure 6.1 Fuzzy Extractor scheme in application

6.2.2 Key Derivation Function

A *KDF* is a procedure or a set of procedures to convert any string into a pseudo-random string. According to Ngel Paul Smart [80], it can be built using two basic methods:

1. Using a hash function: this method transforms the input string with a fixed hash function H which is assumed to be like a random oracle. H maps a string of arbitrary length to a fixed length string. It is approved if it has the two following properties [80]:

   (a) (One-way) Given x, it is computationally infeasible to find y such as $y = H(x)$.

   (b) (Collision resistant) It is computationally infeasible to find x and x' such as H(x) = H(x').

Given a string *mes.*, if H is of output length t bits and block size b, $KDF(mes.)$ is defined as:

$KDF(mes.) = trunc_n(H(mes.\|\langle cnt - 1\rangle_{64}) \| H(mes.\|\langle cnt - 2\rangle_{64}) \| \ldots\ldots \| H(mes.\|\langle 0\rangle_{64}))$;
Where,

- n is the number of output bits of the KDF,
- truncn, the truncation of a string to a substring of length n,
- cnt = $\lceil n/t \rceil$, and $\langle i\rangle_v$ the encoding of the integer i in v bits.

2. Using HMAC : HMAC is a keyed-hash message authentication code that acts like a pseudorandom function [82]. When building an HMAC function, *NIST* recommends a procedure of two steps: (1) key-derivation procedure consisting of randomness extraction, and (2) key expansion; with a message *mes.*, the i^{th} output block K_i is given by $K_i = HMAC(mes.||K_{i-1}||\langle i(mod256)\rangle_8)$, where K_0 is defined to be the zero string of length *t*.

The key generated by an HMAC-based KDF possesses desirable properties including strong randomness. This key is well suited for various cryptographic applications, such authentication and encryption [72], [73], [74].

6.2.3 K-means Clustering

K-means clustering also called Lloyd's algorithm is an algorithm to group *m* data points in *k* clusters, where each point is assigned to the closest centroïd [83]. The following steps define the procedure:

1. Choose *k* initial cluster centers at random.
2. Compute all distances from data points to centers (or centroids).
3. Assign each point to the closest center (or centroïd).
4. Compute new centroïds after assignments; a centroid for K points X_1, X_2, ,...., X_k is computed as $\frac{\sum_{i=1}^{k} X_i}{K}$, it will be the center for a new cluster.
5. Compute the sum of distances for all data points from the centroïds.
6. Repeat steps 2 through 4 until convergence or the number of maximum iterations is reached.

The given preceding steps form the standard k-means algorithm as often referenced in scientific literature. Arthur and Vassilvitkii [84] proposed a more efficient way to choose initial centers. They proposed a k-means++ algorithm, which will be used in our work. This algorithm processes as follows:

1. Choose uniformly at random a center c1 from the set X of data points.

2. Take a new center c_i, choosing $x \in X$ with probability $\frac{D(x)^2}{\sum_{x \in X} D(x)^2}$, where $D(x)$ is the shortest distance from the datapoint x to the closest center already chosen.
3. Repeat step 2 until having chosen all the centers.
4. Proceed with steps 2 – 5 of the standard k-means algorithm.

6.2.4 Jenk Natural Break Algorithm

JNB or Jenk optimal algorithm is an algorithm to classify many data points into clusters that best represent the breaks between these points [85]. The method separates two classes according to the best Goodness of Variance Fit (GVF). For array A of data points x, we calculate GVF to separate the data points at index i as follows. Let, SDAM the squared deviations from the mean of A and $SDMC_i$, squared deviations from the class mean:

1. $\text{SDMC}_i = \sum_{x \in C_i}(x - \mu_i)^2$, where μ_i is the mean of class C_i.
2. $SDMA = \sum_{x \in A}(x - \mu)^2$, where μ is the mean of A.
3. $GVF = \frac{SDAM - (SDMC_1 + SDMC_2)}{SDAM}$

We keep the separation that gives the maximal value of GVF.

6.2.5 Reliability

Our system produces a key from the biometric data of a user. The reliability is a measure of the capability of the system to correctly reproduce the key from biometric data captured at future times. If we attempt to generate the key on N samples (i.e., biometric data from the user), and NS is the number of samples for which the key has been successfully recreated, the reliability is computed as $\frac{N_S}{N}$.

6.3 Related Work

Among systems proposed to generate a key from biometric data, there are FE based ones and others known as KDF. KDF systems use numerical methods or artificial intelligence algorithms to derive a cryptographic key from biometric data. In this section, we describe one or two recent systems in each category with a statement on the limitations of each of them.

6.3.1 Key Generation Using Fuzzy Extractor

Constructions of the theoretical *Fuzzy Extractor* was proposed by Dodis *et al.* [75] for some distances (e.g., Hamming distance, set difference, edit distance, etc.) and codes (i.e., code-offset and syndrome). Practical implementation was reported later in [86]. This implementation was based on – Bose–Chaudhuri–Hocquenghem codes (BCH) – a family of cyclic codes in which one can set the level of correctable errors at designing – and Hamming distance [87], [88]. In systems generating a cryptographic key with FE, a secure sketch (SS) - also called helper data - allows recreating the key at future times.

By analyzing SS in *Fuzzy Extractor* based on code-offsets and bit-permutations, Koens Simoens [78] shows two important privacy weaknesses in such constructions:

1. (Distinguishability) Capacity to link SS to other sensitive and personal information of the user.
2. (Reversibility) Capacity to get the original biometric data from a few sketches.

Later, in 2012, a *Robust Fuzzy Extractor* based protocol is proposed to overcome manipulation attacks [89]. Authors aimed at allowing key agreement between two parties using close secrets (e.g., biometric data, etc.). Bridging the gap between known upper and lower bounds [87], [88], they showed for noisy data w of entropy m, and of n bits length, if $m > \frac{n}{2}$, it is possible to generate a key of length at least $m - \frac{n}{2}$ bits. In some cases, the length is $2m - n$. However, the theoretical constructions of *FE* often face entropy loss issues. In addition, Becker [76] recently showed the impossibility to build a provably secure RFE on the BCH construction proposed in [90] to correct noise levels above 3%. Moreover, the author mentioned that the systems making it possible to correct up to 15% of noise levels do not meet the conditions of the RFE. Another class of *Fuzzy Extractor* constructions are based on the resolution of a hard problem to secure the generated key [76], [91], [92]. Therefore, they are called computational FE and have practical usage in Physically Unclonable Functions (PUFs). However, under some conditions, the generating procedure *(Gen.)* is efficiently invertible by an injective function [92]. Moreover, Herder *et al.* [91] have expressed doubts relative to their applicability with biometric data to generate a cryptographic key.

6.3.2 Key Derivation Function

We consider in this section two classes of KDF. One class where mathematical handcraft methods build the key, and the another with methods from artificial intelligence.

6.3.2.1 KDF with Mathematical Handcraft Methods

The system proposed in [49] does not store helper data or any sample from the user to generate the key. In this design, the noise level to correct is chosen independently of the biometric data and the key is generated in two steps:

1. Set a public parameter as $PP = (t, Fv, H)$; where t is the threshold vector of components $t_i \in [0,1)$, F is a pseudo-random function, and v is a key from the key space of F selected uniformly at random.
2. Define the key K as $KDF(PP, b)$ that takes the public parameter and a biometric vector $b=(b_1,...,b_n) \in R_n$. First, the system computes $h_i = 1/2t_i$, $f_{hi}(b_i) = \lceil b_i * h_i \rfloor$ *for each* $i \in [\![1, n]\!]$. Then, it determines $X = Fv(fh1(b1)||.....||fhn(bn))$, and generates a key as $K=H(X)$.

Authors claim that the proposed function is privacy-preserving (i.e., no user data kept), light enough to be used for key generation in the IoT environment. However, they assume that the source has enough entropy for the key to be generated. Moreover, the level of noise percentage that can be corrected does not consider the biometric data. In other words, the threshold t is independent of the user data, which is not a guarantee for inter-user variability.

Anees and Chen [50] proposed a framework to generate cryptographic keys from facial features. This system has 3 different modules: learning the facial features, the quantization of the facial features, and key generation. A variant of Local Binary Pattern (LBP) called equalized LBP is designed to learn features from acquired data. The 3-bits quantization module helps to cater inter-user variations and to generate cryptographic keys. Key is generated by applying bit-wise rotation on the bitstring of the quantized data. Authors claim that the system can be used with many combined biometric traits. However, the computational complexity would increase.

6.3.2.2 KDF with Artificial Intelligence Algorithms

Another system that we analyzed in this paper exploits statistical features of biometric data to model inter-user and intra-user variability. The method in [51] is a semi-supervised clustering

process, and optimized with a niching algorithm to generate the key. However, the key is generated ultimately with a hash function that is assumed to be pseudo-random [80], [82]. Moreover, the number of effective bits is less than 50 in all results, which is relatively low.

Osadchy and Dunkelman [79] explain some of the problems we can face when using machine learning algorithms in transforming biometric raw data into binary strings. They detailed privacy issues in training-based feature extraction methods because of the set of data kept after the training process. In some Fisherfaces and Eighenfaces based systems [93], [94], [95], [96], it is possible to reconstruct the face of a user from the system parameters learned by the algorithms [79]. Moreover, auxiliary parameters stored for quantization of biometric features leak personal information in training data. Even though we have better accuracy, privacy leakages fade the board of results for these types of system.

Table 6.1 Generators of cryptographic key from biometric data

Systems	**Limitations**
Fuzzy Extractor Constructions by Dodis et al. [75]	Distinguishability and Reversibility.
Robust Fuzzy Extractor [89]	It is hard (impossible for some metrics) to realize practical implementation with the acceptable corrected noise level.
Computational Fuzzy Extractor [91], [92]	Not suitable for biometric data.
Biometric-based KDF by Seo et al. [49]	No guarantee for inter-user variability, assumption of a source with acceptable level of entropy.
Discriminative Binary Feature Learning and Quantization by Anees et al. [50]	Potential increase in computational complexity if we combine input biometric traits.
KDF base on Artificial Intelligence Algorithms [79], [51]	Need of multiple user data to learn parameters; this implies privacy leakages from sets of training data.

6.4 The Proposed System

In this section, we present the new proposed key derivation function. We describe first the elements of a procedure to encode the biometric data. Then, from this code we create the key. At last, we show how to authentify the user with this mechanism in the future.

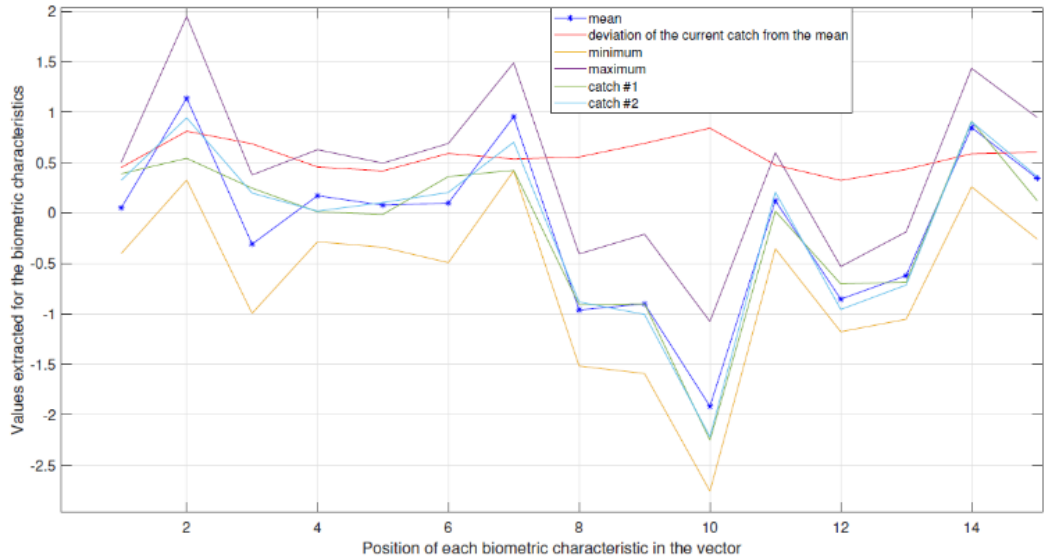
6.4.1 Requirements

Considering weaknesses of preceding systems (Table 6.1), we state some desirable properties of the proposed KDF:

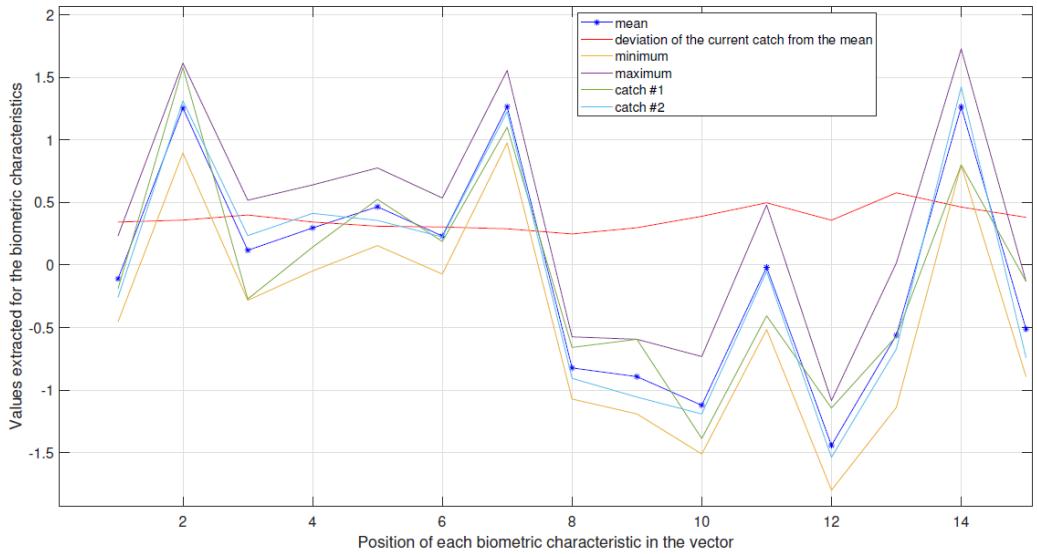
1. Robust against intra-user variability: the system can correct an acceptable level of noise, at least 15%.
2. Robust against inter-user variability: the system can always discriminate between two different users.
3. Reliability: as defined in section 2.4, we want the system to be able to reproduce the key with a high rate of success.
4. Not sensitive to entropy loss: entropy of the generated key is high, meaning it is almost uniform (randomness) and the string is long enough.
5. Privacy-preserving: a few data saved as a helper to recreate the key; does not allow recovery of the original biometric vector.

6.4.2 Assumptions

We assume for this construction that we get a vector of real numbers from the biometric characteristics of a user. This vector is the features extracted from fingerprint data, face, iris, other biometric traits, or a combination of them. Moreover, we consider two properties of these data: consistency and discrimination. These properties are basic requirements for using of biometric traits in an authentication system. Consistency means here, that two biometric vectors acquired from the same user have a great similarity. We mean by discrimination that biometric data from two different users can be separated by some criteria. The consistency is illustrated in Figure 6.2, where we have two vectors of features extracted from the same user: the violet line, and the yellow line; we observe that the values for the 10^{th} feature are between feature for user #1 are between -2.75 and -1, while these values are between -1.5 and -0.7 for user #2. The blue line is the mean vector, and the orange line represents the maximum variation of a specific feature from its mean value.



a) Biometric data for user #1



b) Biometric data for user #2

Figure 6.2 Illustration of the consistency in the biometric data with 2 different users

6.4.3 Rationale of the Coding Process

Let $B^r = (b_1^r, \dots, b_j^r, \dots, b_m^r)$, the biometric vector extracted at time t^r, and $B^s = (b_1^s, \dots, b_j^s, \dots, b_m^s)$, the biometric vector extracted for the same user at time t^s. If the feature b_j^r is consistent for the user with the applied image processing techniques, b_j^r and b_j^s is relatively closed. We consider all the values of this feature as a fuzzy set A_j, defined by $A_j = \{b_j^r / |b_j^r - \mu_j| \leq \sigma\}$, where μ_j is the mean of the elements in A_j and σ, a defined threshold [97], [98]. Replacing all elements in

A_j with μ_j can be considered as a defuzzification method. We can federate closed values according to the desired number of fuzzy sets.

6.4.4 The Coding Process

A functional view of the system (Figure 6.3) distinguishes two main stages: enrollment and authentication. During enrollment, we acquire the biometric data (with a mobile device typically), extract features from raw biometric data, apply quantization, and encode to create a representative code C for the user. Then from this code, we generate a key K using a pseudo-random function. Different techniques can be used for biometric data acquisition and feature extraction. Extraction of features (bloc 2 in Fig. 6.3) gives R biometric vectors. In the 3^{rd} bloc, we perform quantization and encoding. Quantizing is finding fuzzy sets of interest; each fuzzy set is a range of biometric values, an interval that we represent with its minimum value, its median, and its maximum value (we often consider the fuzzy set as a cluster represented by Cl={min, median, max}. The system's output is a vector where each component is a cluster of interest; this vector is the Quantizer Q.

Encoding is replacing each biometric value with the median of its cluster. Therefore, we form the code C for the user. We encode multiple vectors and keep the most consistent features. A consistent feature (a component of B^r) is a feature located in the same fuzzy set (the same interval) for all considered biometric vectors; positions of these consistent components are saved in a vector CP (consistent positions).

Later, at the authentication stage, the system uses (Q, CP) as helper data to reproduce the code C for a biometric vector, by replacing each consistent component with the median of the cluster in which it is located. All modules of the procedure are detailed in the coming sub-sections.

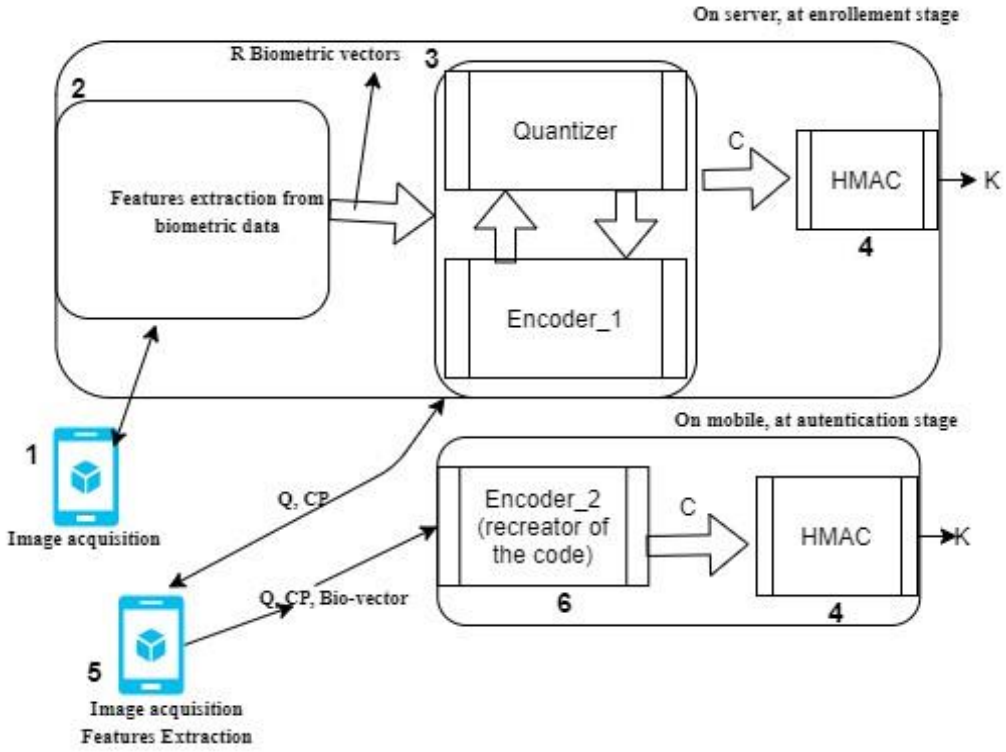


Figure 6.3 A view of the system with applicative characteristics

6.4.5 Quantization and Encoding during Enrollment

First, we generate Q from a vector of reference (we call it RV). Then we encode multiple biometric vectors; we also encode RV to get the code of reference C. The consistent positions (CP) is the vector of common components between all the code we formed. CP, and Q are saved on the server or mobile device for future authentication. An adaptation of the pseudo-random function defined in 6.2.2, generates the key K from C.

6.4.5.1 Code Generation Procedure

As it was inspired by Keykhaie *et al.* [99], to compute C, the system acquires enough biometric data (R) to find consistent and discriminative extracted values. This work is done at data acquisition (bloc 1 in Fig. 3). Then, we eventually apply preprocessing steps, and suitable feature extraction methods at bloc 2 of Fig. 3. The feature extraction process outputs a set of R biometric vectors $B^r = (b_1^r, \dots, b_j^r, \dots, b_m^r)$, with r ∈ {1, 2, ..., R}. The procedure to determine the code C - from the set of initial biometric vectors - consists of the following steps:

1. Split the biometric data acquired into *R1* samples for quantizing (like training) and *S* samples (used as test data) to find consistent positions.
2. Compute the reference vector *RV* (the mean is a good example) from the *R1* vectors. First, for each range of *R1* values of a component, we remove outliers. We consider as outliers any values whose distance from the median exceeds the Median Absolute Deviation (MAD) by a factor of three; then we take the mean of the remaining values,

$$\mu_j = \frac{\sum_{r=1}^{R1-Outliers} b_j^r}{R - Outliers} \qquad (1)$$

3. Form NC clusters from RV using quantizeWithJNB function (Algorithm 1) or quantizeWithKmeans (Algorithm 2), performed into the module named Quantizer in Figure 5.3; each cluster (Cl={min, median, max }) is a component of Q.
4. Determine the code for many biometric vectors from the test set; to encode a biometric vector Br, we replace each component b_j^r with the median (or centroïd) of its cluster (Algorithm 3), performed into the module Encoder_1 of Figure 5.3.
5. Keep the set of consistent positions (CP) between all S vectors; CP is a vector of positions (cp1,....,cpj,...,cpl); cpj is a component of CP means, ∃ Cl ∈ Q, with Cl={min, median, max} / $min \leqslant b_{cp_j}^r \leqslant max$, for all biometric vectors Br, with r ∈ {1, 2, ..., S}. The length of CP is the length of the generated code.
6. Recreate the code (considering only CP) for many biometric vectors formed from the test subset with algorithm 4.
7. Compare each new code with C (the reference) to find the percentage of components of this code that matched its corresponding in C; the minimum percentage is a threshold of similarity for future authentication.

When using the k-means algorithm, we often repeat the process of choosing randomly the centers to determine Q until we find the desired length for CP with a targeted value of reliability. We follow the algorithm in Figure 5.4.

6.4.5.2 Key Creation from Code

To convert the code *C* into a key *K*, we apply HMAC-SHA256 as specified by NIST [80], [82]. Our adaptation of this algorithm is based on the following reasoning. We calculate the first input of the

key as K_0. Then, we compute K_i considering K_{i-1} as input according to a process derived from the definition in 2.2.

1. $K_0 = HMAC - SHA256(Extsalt||c_0)$, where c_0 is the first component of code C and *Extsalt* is an external data, like a legal identification of the user.
2. $K_i = HMAC - SHA256(K_{i-1}||c_i)$, where c_i is the i^{th} component of C;
3. Iterate according to a chosen length $l_s \leq$ length of C.

For better performance in terms of duration of the key generation process, we often take sums of the same sign components of code C (for instance c_0+c_1 as input to compute K_0). When the number of iterations l_s increases, we have more randomness. An advantage of this mechanism is that the HMAC function acts as a pseudo-random function (implementation from NIST). In addition, it takes another external data (*Extsalt*) which can be a piece of information identifying the user, and allowing him to have different keys for different systems while using the same biometric traits for authentication.

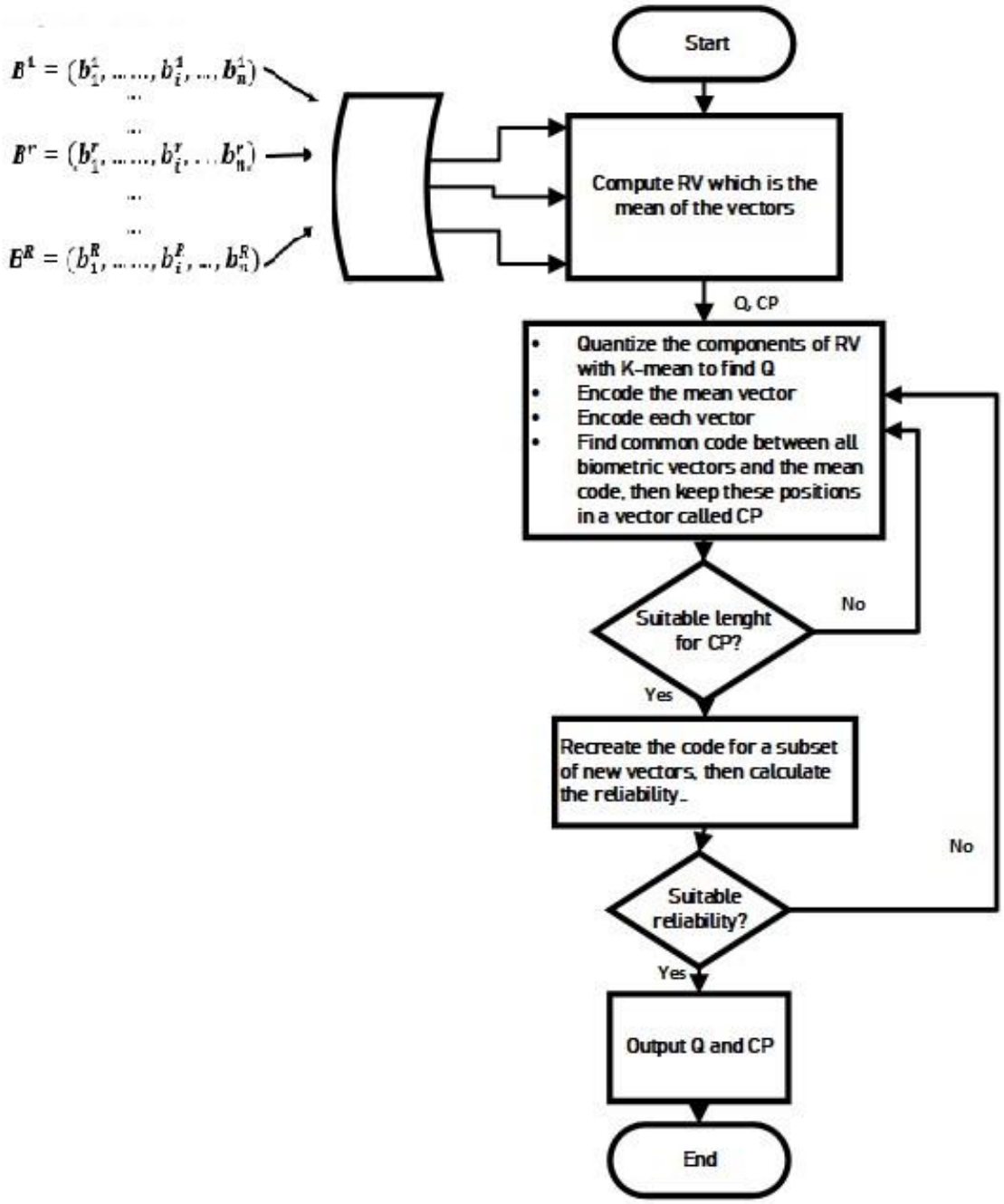


Figure 6.4 Flowchart of the coding process at enrollment with k-means

Algorithm 1: QuantizeWithJNB

Data: refVector (RV), number of clusters (NC)
Result: Quantizer Q is computed
/* First, finding interfaces that separate clusters using Jenks Natural Break algorithm */
$RV \leftarrow sort(RV)$;
$itf \leftarrow get\_Jenks\_Interfaces(RV)$; // itf for interface
$set\_of\_itf \leftarrow \{1, length(RV)\} \cup \{itf\}$;
while *length(set_of_itf) ; NC-1* **do**
  $sort(set\_of\_itf)$;
  $largest\_interval \leftarrow interval[1, itf]$; // first interval
  for $i \leftarrow 2$ **to** *length(set_of_itf)-1* **do**
    $current\_interval \leftarrow [set\_of\_itf(i), set\_of\_itf(i+1)]$;
    if card$(current\_interval) >$ **card**$(largest\_interval)$ **then**
      $largest\_interval \leftarrow current\_interval$;
    end
  end
  $itf \leftarrow get\_jenks\_interface(largest\_interval)$;
  $set\_of\_itf \leftarrow set\_of\_itf \cup \{itf\}$;
end
/* Second, creating the quantizer Q from the set of interfaces */
$interval \leftarrow RV(1 : set\_of\_interfaces(1))$;
$median \leftarrow median(interval)$;
$min \leftarrow min(interval)$;
$max \leftarrow max(interval)$;
$Q \leftarrow \{[median, min, max]\}$;
for $i \leftarrow 1$ **to** *length(set_of_interfaces)-1* **do**
  $begin\_id \leftarrow interfaces(i)$;
  $end\_id \leftarrow interfaces(i+1)$;
  $interval \leftarrow RV(begin\_id : end\_id)$;
  $median \leftarrow median(interval)$;
  $min \leftarrow min(interval)$;
  $max \leftarrow max(interval)$;
  $Q \leftarrow Q \cup \{[median, min, max]\}$;
end
$interval \leftarrow RV(set\_of\_interfaces(i) : length(RV))$;
$median \leftarrow median(interval)$;
$min \leftarrow min(interval)$;
$max \leftarrow max(interval)$;
$Q \leftarrow Q \cup \{[median, min, max]\}$;
Return Q ;

Algorithm 2: QuantizeWithKmeans

Data: refVector (RV), number of clusters (NC)
Result: Quantizer Q is computed
Find NC clusters with K-means++ algorithm, each cluster Cl_i primarily defined by its centroïd $centr_i$;
Sort the NC centroïds according to clusters density ;
foreach *2 adjacent clusters* Cl_i *and* Cl_j **do** ;
if Cl_i *and* Cl_j *are separated by a gap* **then**
  Compute radius r as $\frac{|centr_i - centr_j|}{2}$;
  Set r as the radius for Cl_i and Cl_j;
end
Return Q as a list of centroïds and corresponding radius ;

Algorithm 3: EncodeAfterQuantizing

Data: bio-Vector (B), Quantizer (Q)
Result: A vector C representing the code of the user
for $i \leftarrow 1$ **to** *length(B)* **do**
  Set $j \leftarrow 1$;
  while *j is smaller than length(Q) and B(i) not in the cluster Q(j)* **do**
    $j \leftarrow j + 1$;
  end
  if *j is smaller or equal to length(Q)* **then**
    C(i) = median of the cluster Q(j) ;
  else
    Find the closest median to represent the feature ;
  end
end
Return C ;

Algorithm 4: RecreateCode

Data: bio-Vector (B), Quantizer (Q), Consistent Positions (CP)
Result: code (C) for the user
for $i \leftarrow 1$ **to** *length(CP)* **do**
  Set $j = 1$;
  while *j is smaller than length(Q) and B(CP(i)) not in the cluster Q(j)* **do**
    $j = j + 1$;
  end
  if *j is smaller or equal to length(Q)* **then**
    C(i) = median of the cluster Q(j) ;
  else
    Find the closest median to represent the correspondent feature ;
  end
end
Return C as the list of closest median from each component of B in a position mentioned in CP ;

6.4.6 Authentication

At the authentication step, the system acquires a small subset of biometric data from the user (bloc 5 in Figure 5.3), extracts the features by the same techniques applied during enrollment, then reproduces the code C with the recreate function (Algorithm 4). This function takes as input: Q, and CP saved either on a server or on a mobile device. It produces the expected code by replacing each biometric feature corresponding to a position in CP with the median (or the centroïd) of its cluster in the vector Q. Then, the same pseudo-random function used at enrollment is applied to generate the key K.

We can authenticate the code without the key. In this scenario, we decide based on the threshold of similarity that we learned at enrollment. To learn this threshold of similarity, we encode multiple test data and keep the percentage of similar components between each code and the code of reference. We save the minimum value as this threshold of similarity.

6.5 Experiments and Results

We report experiments with three datasets of face data. (1) Private dataset 1 (DS1) composed of 220 images of 5 users, typically a smartphone with a 6.4-inch display and an octa-core CPU (2X1.6 GH-z + 6x1.35 GHz); (2) Dataset 2 (DS2) that is a manually cropped version of the extended Yale face database [100] composed of 2414 images of 38 users; and (3) the dataset 3 (DS3) that is a subset of Youtube face database [101], which is composed of 3425 videos of 1595 different users. Fig. 5 shows images of 2 individuals, one from DS2, and the other from DS3. Images have different resolutions: (1) face images that have been cropped to 160x160 pixels are good examples (in DS1) with the Facenet system [21], [28], [29], (2) 32x32 pixels for DS2, and (3) variable dimensions with 96x96 dpi for DS3. For DS1 and DS3, we detect faces from each picture with the method of FaceNet [48]; for DS2, faces are already detected in the cropped version. Additional processing on DS2 and DS3 include: creating rectangular shape structuring elements of the size of the image; then, applying erosion and deletion; at last, increasing the contrast of the image to overcome variation of illumination.

For all datasets, we consider a biometric vector of length $m = 128$ after feature extractions and 8 clusters. Facenet [48] detects the faces, extracts the features, and outputs the vector with 128 components. When processing methods give a biometric vector with more components (as for DS2

and DS3), we apply complementary processing techniques to keep the most consistent features. For instance, a vector from DS2 or DS3 has a number of components greater than 128. We compute the standard deviation on the j^{th} feature as,

$$\sigma_j = \sqrt{\frac{\sum_{r=1}^{R}(b_j^r - \mu_j)^2}{R}} \qquad (2)$$

we keep the 128 features with fewer standard deviations.

6.5.1 Performance at Enrollment

We evaluate the duration of the key generation procedure and the length of the generated code. We split each user into two subsets of samples: one (*R1* representing generally 75% of all available samples for the user) for training to learn *Q*; from the other subset (*R* - *R1*), we learn *CP*. We build many biometric vectors from the last subset; we encode all newly built biometric vectors and keep the sets of common components between all these codes and the reference *C*. This intersection is CP. We also keep the threshold of similarity.

To simulate variations in biometric captures for the same user, we generate test vectors by calculating the mean of a subset of five randomly selected samples from the test set. As a result, each vector in the test set is distinct from one another. This methodology allows us to replicate the natural fluctuations that occur in biometric captures of an individual. We always remove outliers before computing the mean. In Tables 6.2, 6.3, and 6.4, we present typical results for DS1 and DS2: the length of the code *C*, the duration to generate this code in seconds, and the minimum of threshold of similarity. We conduct these experiments with Matlab [102], on a laptop having 12 GB of RAM and a processor Intel (R) Core (TM) i7-4600U, CPU @ 2.10GHz.

a) Images of one user from DS2 (e.g., dimensions are 32x32 pixels), the main challenge is the variation of illuminations

b) Images of one user from DS3, in an uncontrolled environment

Figure 6.5 Images of two individuals, one from the extended Yale face database (DS2, images are typically of 32x32 pixels) [100], and the other from the Youtube face database [101]

Table 6.2 Results for code generation with DS1

DS1	Length of the codes	Duration	Similarity
U1	103	1.292 s	≥ 0.92
U2	103	1.115 s	≥ 0.97
U3	107	1.17 s	≥ 0.97
U4	109	1.175 s	≥ 0.90
U5	100	1.161 s	≥ 0.97

Table 6.3 Results for code generation with DS2

DS2	Length of the codes	Duration	Similarity
U1	56	20.018 s	≥ 0.80
U15	39	20.014 s	≥ 0.74
U18	42	20.833 s	≥ 0.69
U2	55	20.339 s	≥ 0.84
U21	39	20.316 s	≥ 0.79
U22	120	0.836 s	≥ 0.95
U23	29	20.544 s	≥ 0.41

U24	118	0.7 s	≥ 1.00
U3	46	20.066 s	≥ 0.46
U7	23	20.094 s	≥ 0.87

Table 6.4 Results for code generation with DS3

DS3	**Length of the codes**	**Duration**	**Similarity**
U412	82	1.83 s	≥ 1.00
U1389	107	0.9 s	≥ 0.97
U492	83	1.3 s	≥ 0.96
U1503	128	0.8 s	≥ 0.97
U20	101	0.98 s	≥ 0.98
U725	111	1.47 s	≥ 1.00
U497	106	0.98 s	≥ 0.93
U392	97	41.3 s	≥ 1.00
U959	112	0.92 s	≥ 0.99
U276	86	1.72 s	≥ 1.00

6.5.2 Performance at Authentication

To evaluate the performance at the authentication step, we consider a classification problem (two classes): we try to classify every biometric test vector as the authentic user (positive class) or an imposter (negative class). For DS1, we conduct experiments with 5 users (20 biometric vectors for the authentic user and 5 for each impostor).

For DS2 and DS3, we randomly choose 10 users to test. We generate 5 biometric vectors for each user. In this case, we generate the biometric vector for the test (as preceding) with randomly chosen samples, but from all available data of the user. So, we still simulate the variations in biometric captures. We recreate the code (with Algorithm 4) for each newly created biometric vector. We predict the true class if the similarity with the code of reference is above the minimum threshold found at the enrollment step. We consider many metrics for authentication performance: the accuracy also designated by recognition rate, the precision, the recall, the true negative rate

(*TNR*), the false acceptance rate (*FAR*), and the false rejection rate (*FRR*). We calculate these values with the following formulas.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$TNR = \frac{TN}{TN+FP} \quad (6)$$

$$FAR = \frac{FP}{TP+TN+FP+FN} \quad (7)$$

$$FRR = \frac{FN}{TP+TN+FP+FN} \quad (8)$$

Where,

TP (true positive): the number of samples of the positive class correctly classified by our system.

TN (true negative): the number of samples of the negative class correctly classified by our system.

FP (false positive): the number of samples of the negative class incorrectly classified by our system.

FN (false negative): the number of samples of the positive class incorrectly classified by our system.

Figure 6.6, Figure 6.7, and Figure 6.8 present accuracy, precision, recall, and TNR from experiments with DS1, DS2, and DS3 respectively. We find our best results with the quantization method that uses the JNB algorithm [85]. With DS1 and DS3, we often have an *FAR* of 0 %, even if the *FRR* is not null. In Table 6.5, we present results for DS1; each line represents the performance when the user in column 1 is fixed as the authenticated one and all others are impostors.

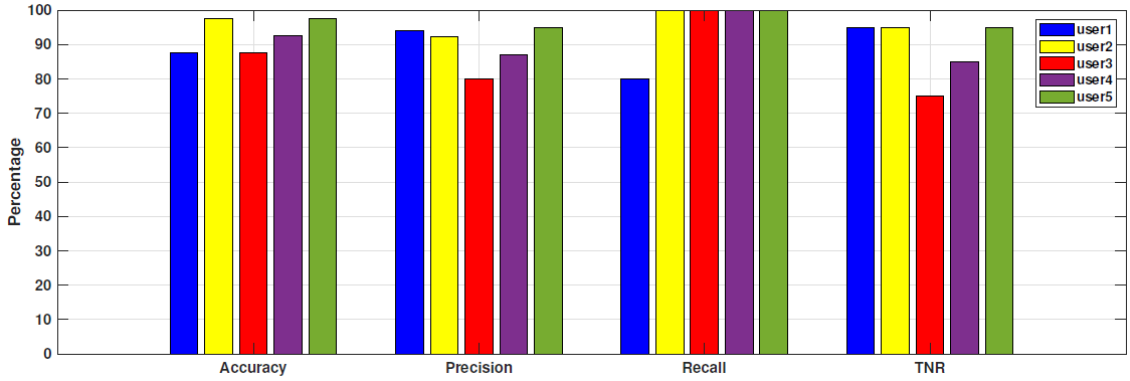


Figure 6.6 Performance of the authentication mechanism on DS1

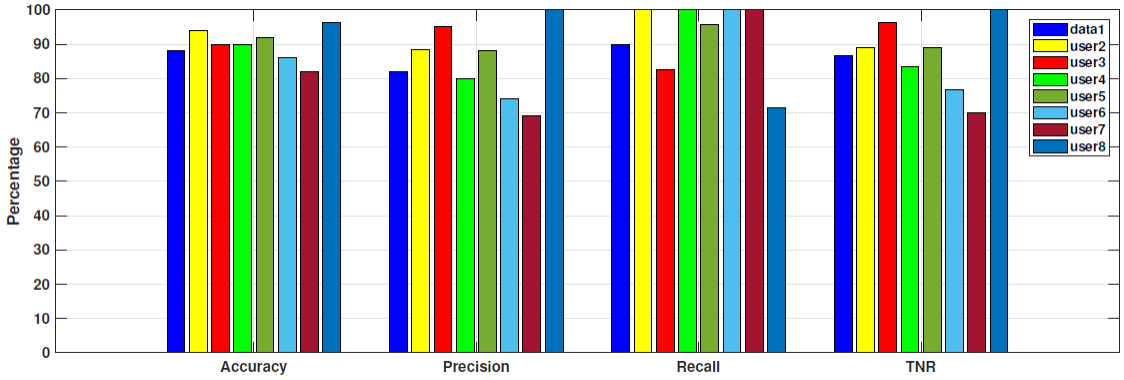


Figure 6.7 Performance of the authentication mechanism on DS2

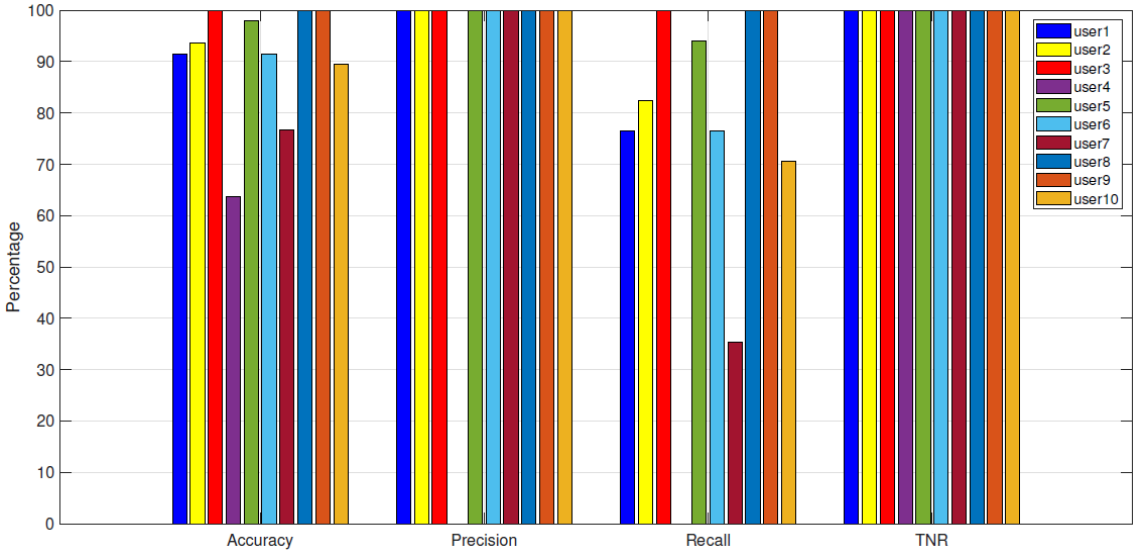


Figure 6.8 Performance of the authentication mechanism on DS3

To evaluate the duration of a session of authentication, we deploy the function to recreate the code on a mid-range smartphone: a pixel 2 having 1.5 GB of RAM with the android platform. We implement the algorithm to recreate the code using Java. We can recreate the code and generate the key with HMAC in around 400 ms (Table 6.6).

Table 6.5 Authentication performance with DS1

DS1	FAR	FRR	Accuracy
U1	5 %	10 %	85 %
U2	0 %	0 %	100 %
U3	0 %	0 %	100 %
U4	0 %	20 %	80 %
U5	0 %	5 %	95 %

As shown in Table 6.6, the procedure to generate the code at the authentication step is very efficient (1 ms). However, the duration to generate the key from the code with the adaptation of the HMAC algorithm depends on the number of iterations. We consider two iterations for reported results. We observe that the times to create the key K from the code C is a linear representation with respect to the number of iterations as shown in Figure 6.9.

Table 6.6 Duration of key generation on a mobile

Procedure	Duration[ms]
Recreate the code	1
Generate key from the code	399
Total	**400**

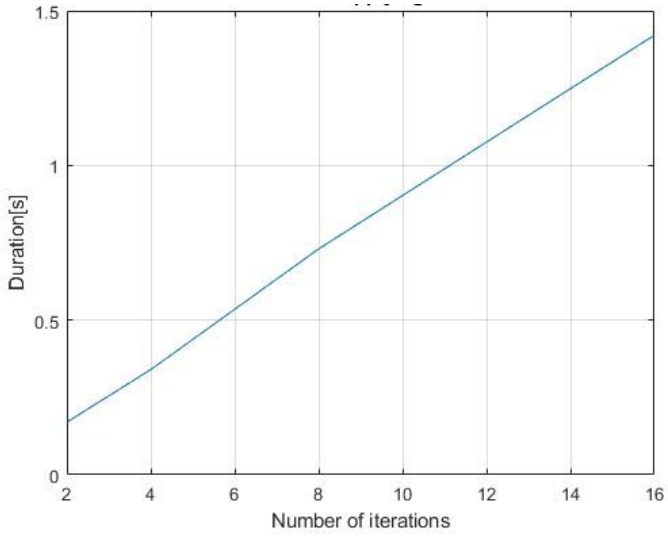


Figure 6.9 Duration of the key generation from code with HMAC

6.5.3 Robustness of the KDF

Using DS1, we evaluate the reliability of the code (this metric is relative to the capacity of the system to consider the inter-user variability), the corrected noise level CNL in biometric data (intra-user variability), and the entropy of the created key K. These values help to assess the robustness of this function.

6.5.3.1 Reliability of the Code

To generate the key at future times, the code must be the same at each reproduction. Therefore, the soft matching used for authentication becomes an exact matching (For two codes C and C', each component of C is equal to its correspondent in C'). To increase the probability of exact matching, we have to consider the most stable components of the biometric vector. Therefore, we encode a number of biometric vectors while minimizing aggregation, sometimes we consider every single sample of biometric data. We keep the common components among all the codes. We generally find a shorter final code than in the preceding procedure for authentication, but with higher reliability. We report in Table 6.7, the success rate of recreating a code (using Q and CP) at future times. We indicate in the first column the user and the length of his code. The intersection between line U_i and column U_j contains the rate of success at attempting to recreate the code of U_i with biometric data from U_j. In the experiments (with DS1), we take less than three biometric vectors to form a code for the test.

Table 6.7 Discriminative capacity of the KDF on DS1

	U1/88	**U2/81**	**U3/93**	**U4/95**	**U5/72**
U1/88	0.92	0	0	0	0
U2/81	0	1.00	0	0	0
U3/93	0	0	0.92	0	0
U4/95	0	0	0	0.96	0
U5/72	0	0	0	0	0.88

6.5.3.2 Intra-user Variability

We quantify the intra-user variability by the noise level in biometric data. It measures the deviation of a biometric vector of the user from the vector of reference. We calculate the corrected noise level (*CNL*) in biometric data for a user as,

$$CNL = mean_{r \in R_2, 1 \leq j \leq 128} \frac{\tau_j^r}{\mu_j}, \quad (9)$$

Where μ_j is the j^{th} component of the mean vector and is computed with relation (1). The value τ_j^r is the difference (absolute value) between the j^{th} component of biometric vector B^r and μ_j, $|b_j^r - \mu_j|$. The integer r varies from 1 to R_2, where R_2 is the number of samples in the test subset for which the procedure of recreating the code is successful. Our KDF corrects noise levels of up to a mean value of 40.3 % in DS1 (Table 6.8).

Table 6.8 Noise level corrected for each user in DS1

User	**CNL**
U1	14.3 %
U2	10.9 %
U3	5.28 %
U4	26.8 %
U5	40.3 %

6.5.3.3 Security and Privacy

In this subsection, we discuss the resilience of the KDF against adversarial activities, emphasizing its ability to withstand malicious attempts and maintain the security of the derived keys.

- Robustness against brute force attacks

  Let's consider the helper data, saved to recreate the code at a future step (Q, and CP). As a reminder, each element of *Q* is a cluster on the real axis, so an interval; we represent it as a triplet where the first component is the minimum value of the interval, the second component is the median value and the third is the maximum value: (min, median, max). Each element of *CP* is a position in the biometric vector of features for which all values are in the same cluster. We modelized *CP* as a tuple. We present here an example for Q and 20 first components of an instance of CP.

$$Q = \begin{pmatrix} \textit{Min.} & \textit{Median} & \textit{Max.} \\ -1.625 & -1.625 & -1.625 \\ -1.625 & -1.535 & -1.363 \\ -1.363 & -1.307 & -1.165 \\ -1.165 & -1.122 & -0.976 \\ -0.976 & -0.702 & -0.497 \\ -0.497 & -0.120 & 0.191 \\ 0.191 & 0.782 & 2.304 \\ 2.304 & 2.304 & 2.304 \end{pmatrix}$$

CP = (1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 19, 23, 25, 26, 28)

  If a hacker has access to *Q* and *CP*, he will be able to generate *C* by brute force attack after $length(Q)^{length(CP)}$ attempts. We considered 8 clusters, meaning *length(Q)* = 8. With DS1, we got length(CP)≥100 for authentication. In such a case, the probability to generate *C* is 8^{-100}, or 2^{-300}. For DS3, Table 5.4 shows that the length of the code is above 82; therefore, the probability to break the code by brute force attack is less than 8^{-82} or 2^{-246}.

- Privacy

  The previous point highlights the privacy-preserving property of the KDF. The utilization of a sketch to reproduce the key ensures that the original code cannot be easily deduced without authentic biometric data. Furthermore, it is important to note that the code itself does not represent the extracted biometric features directly. Instead, it is a compressed

representation derived from the extracted biometric values. This approach helps to maintain privacy by abstracting the underlying biometric information while still enabling the reproduction of the key and authentication.

- Robustness against cryptanalysis attacks

  The compliance of the code transformation with the standards of the Internet Engineering Task Force [39] for an HKDF ensures the resulting cryptographic key possesses the desired characteristics (i.e., randomness) to be robust against cryptanalysis attacks.

6.6 General Discussions

This section is a discussion on the potential application of this new system and a comparison with some other proposed systems.

6.6.1 Potential Applications

This function allows the creation of a cryptographic key K from biometric data. It is suitable for applications of HKDF listed in [103]. One of these applications is the derivation of cryptographic keys from a shared Diffie-Hellman value in a key-agreement protocol, like the one proposed by Loïc D. Tsobdjou *et al.* [52], to allow secure exchanges between a mobile client and a server. The KDF can be used as a key recovery mechanism for Blockchain-based systems. In such a case, we keep Q and CP in a safe environment like a Hardware Security Module *(HSM)*, and the user authenticates himself with K on a daily basis. If he loses his key, he will reproduce it with the help of Q and CP. This mechanism aims to prevent the loss of bitcoins or other assets on a blockchain. In addition, it links the cryptographic key (subsequently the address on the Blockchain) to the biometric data of the user, a desirable property for any Blockchain-based system.

At last, this coding process can be configured and deployed as a privacy-preserving matcher. For this last application, we only need the code without integrating the step of creating a key with HKDF. In this case, we save the Quantizer Q, the vector of consistent positions CP, the code of reference C, and the threshold of similarity at enrollment. These data are saved either on the mobile device of the user or on the server of a TTP. At the authentication, the mobile user recreates a code C′ to be compared with the code of reference C. Even in this scenario, less data will be saved than with the usual matcher. Indeed, the code is built with some central values representing a great range of biometric features, not the original biometric data.

6.6.2 Comparison with Other Systems

On the task of face recognition, we make a comparison with three other systems as shown in Table 9. The first one realizes unsupervised multimodal subspace clustering through approaches based on convolutional neural networks [104]; this is a Deep Multimodal Subspace Clustering Network (DMSC). The second one is a face recognition system based on an extended version of the Local Gradient Hexa Pattern (LGHP) [105]. The last on consists of two propositions using Doubly Stochastic Subspace Clustering methods in [106]: a Joint learning model (J-DSSC), and a model with sequential Approximation (A-DSSC). We consider experiments of these clustering methods with two users from the extended Yale face database; our method generally attempts to recognize one user among ten. Table 6.9 reports this comparison where our new system is PPM-BA/KDF (meaning privacy-preserving model for biometric authentication and key derivation function). DMSC has the best performance. However, it is based on deep learning methods that have privacy concerns. Our system shows a good trade-off between accuracy and privacy.

Table 6.9 Comparison of the KDF with existing models on face recognition

Model	**Accuracy**	
	extended-Yale	**Youtube**
DMSC [97]	0.992	-
LGHP [98]	0.750	-
J-DSSC [99]	0.920	-
A-DSSC [99]	0.917	-
PPM-BA/KDF	**0.898**	**0.90**

We also compare the key derivation function with three of the literature's most recent biometric-based *KDFs*. On many characteristics, our system brings novelty (Table 6.10):

- Flexibility: we show possible exploitations with different feature extraction techniques of face biometric data. Other systems do not demonstrate flexibility; they just claim it and sometimes do not speak about that property. In Table 10, when a system claims a property

or assumes it without demonstration, we report ”Yes u.a.”, meaning ”Yes, under an assumption”.

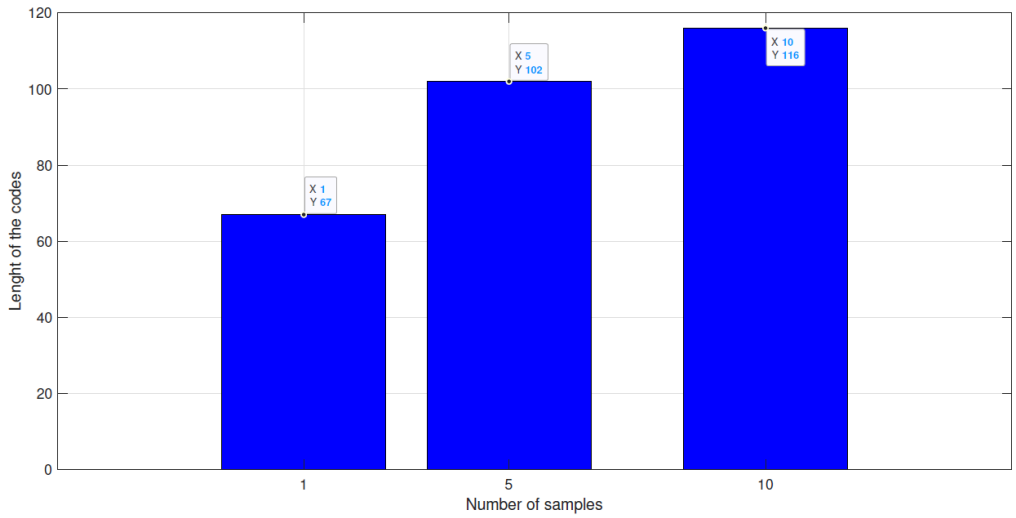
- Randomness: we build on the good properties of HKDF without assuming (as in [4] [49]) that the input is good enough in terms of entropy; other systems use a hash function that is assumed to transform a string to a random string.
- Revocability: Our system guarantees the revocability of the key; we ensure that by adding a salt (that we can control) in the process of creating the key. Moreover, using the k-means-based quantization method, we often get a different code at each enrollment for the same user. this point of view must be deepened for the revocability of the biometric-based code.
- Irreversibility: Our system uses the procedure of HKDF that ensures this property; other considered systems use a hash function that gives a guarantee for the property too.

Table 6.10 Comparison between the biometric-based KDF with other propositions

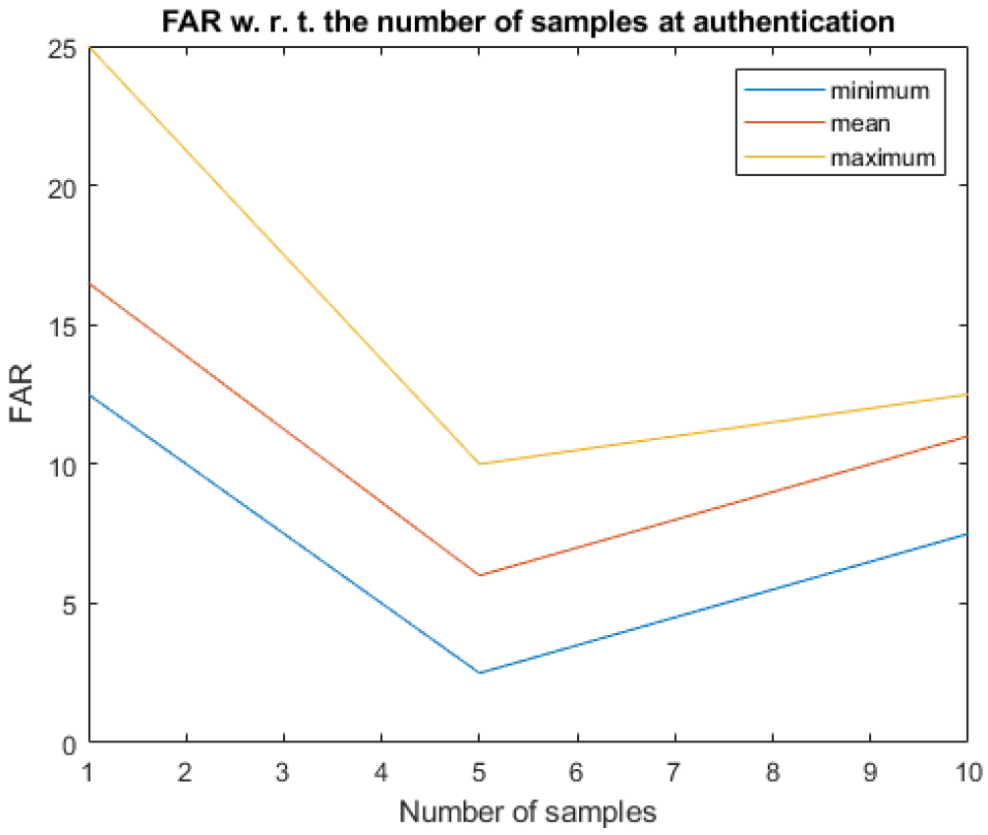
Property \ ***Study***	**[51]**	**[71]**	**[50]**	**New *KDF***
Flexibility	No	Unknown	Unknown	Yes
Revocalibility	No	No	No	Yes
Irreversibility	Yes	Yes	Yes	Yes
Randomness of the key	Yes u.a.	Yes u.a.	Yes u.a.	Yes

6.6.3 Role of the Data in the Learning Process

The number of samples (for aggregation) used at authentication impacts the performance of the system (Fig. 5.10). This parameter impacts the length of the code (Fig. 5.10.a) and the recognition rate. With DS1, we find the lowest false acceptance rate (FAR) when the number of samples at authentication is 5 (Fig. 10.b).



a) Influence on the length of the code.



b) Relations between the number of samples and the FAR.

Figure 6.10 Influence of the number of samples on code generation with DS1

6.7 Conclusion

In this paper, we designed and implemented a new KDF using clustering algorithms to deal with fuzziness and an adaptation of HMAC-SHA256. This function creates a code from the biometric data of a user and then transforms this code into a cryptographic key. We build a mechanism to authenticate the user with the created code. We evaluated this function using three data sets, where two of them (i.e., the extended Yale face database, and the YouTube face database) are publicly available and have great challenges. Evaluations and analysis show some good properties of our system: this is a privacy-preserving function. The latter performs good management of intra-user variability with up to 40.3 % of CNL. Moreover, it is reliable and has good discriminative capacity. Two great applications of such a KDF have been highlighted: a privacy-preserving matcher and a key generation function for key recovery with Blockchain-based systems.

However, we state that for some users, the code created is short or the similarity between different codes of the same user is low (Table 3, U3). These observations indicate room for improvement. Potential directions of our research work are: assessment of most suitable biometric traits and feature extraction methods for this function, improvement of the KDF to deal with changes in biometric characteristics, and finally, adapt and optimize the model to perform all processing on a sim card instead of a large server.

CHAPITRE 7 ARTICLE 4: FRAMEWORK AND ARCHITECTURE FOR BLOCKCHAIN-BASED E-COMMERCE SERVICES

Auteurs: Olson Italis, Samuel Pierre, Alejandro Quintero.

Statut: Soumis pour publication dans la revue « *IEEE Access* », le 23 juin 2023.

Abstract

Blockchain technology has great potential for solving the problems associated with centralized electronic payment systems. These include transaction reversibility, single point of failure, lack of inter- operability between heterogeneous systems, and delays, to name but a few. However, although blockchain supports many financial applications, its limitations are numerous (e.g., high energy consumption, privacy issues, lack of scalability, and poor performance). To establish a design framework for a blockchain-based e-payment architecture that considers these limitations, we first analyze the structure of the electronic payment system on the one hand and the architecture of the blockchain. We then compare the main consensus protocols and propose a taxonomy of them. We design a blockchain-based architecture composed of three notable components: (1) a Key Derivation Function (KDF) which is a light key recovery scheme for blockchain-based systems, (2) a biometric-based authentication model, and (3) a decentralized payment protocol to guarantee the privacy of the transactions. The evaluation results demonstrate that the proposed architecture is secure, preserves privacy, and allows interoperability between heterogeneous platforms for e-commerce services. The architecture is also flexible, allowing compliance with regulations by assigning distributed roles to different payment institutions. While existing blockchain-based architectures in the literature lack at least one of the aforementioned characteristics, the one we propose integrates all of them into a comprehensive solution.

7.1 Introduction

In the last few years, we have observed an increasing usage of e-commerce, with new trends in payment methods [55]. We have also testified of many failures, enlighting the limitations of the traditional payment systems [107], [120]: weaknesses of the trust model such as reversibility of the transactions, a single point of failure, and privacy-preserving issues. One particular aspect of the traditional payment system with specific limitations is cross-border payments which rely on

correspondent banking [121], [122]. However, correspondent banking requires money on hold [122]: in Figure 7.1.b, Bank1 must have an account with correspondent1 that holds the amount to pay when requested. Moreover, the number of intermediaries for a transaction rapidly increases, increasing the processing time and cost of a transaction; some intermediaries also have low-security infrastructure.

Blockchain technology has emerged as a potential solution to some of the previous challenges, including the old model based on trust in central authority and the high delay in transaction processing [6]. This technology can facilitate interoperability between many heterogeneous payment systems [63]. Blockchain's potential for e-commerce is great; well-established applications include cryptocurrencies [17] and inter-bank payments [121]. However, each proposed system inherits the technical challenges of blockchain technology without mentioning regulatory and business issues [20], [123].

Among the technical challenges are [20], [117]: the poor performance of the Proof-of-Work (PoW) protocol, security concerns in Proof-of-Stake (PoS) and other variants (e.g., nothing-at-stake), lack of scalability of protocols like Practical Byzantine Fault Tolerant (PBFT), to cite a few. The key management issues of blockchain-based systems are also of interest [124]. More specifically, in the case of a loss of the private key, we automatically lose our assets on the system without any key recovery mechanism. Meanwhile, there are many new propositions for consensus protocols trying to exploit the advantages of the first proposed ones or overcome their limitations [21], [20], [125], [126], [127], [128], [129] [125], [126]. The large number of available techniques and components makes it difficult to choose suitable components and protocols to form an optimized and usable blockchain-based architecture for e-commerce scenarios in the real world. Furthermore, it is necessary to define how to connect an existing centralized payment system with the blockchain in the context of e- commerce.

In our previous study [120], we designed a scheme to overcome privacy issues in blockchain-based systems. The focus was mainly on the front-end layer on the top of any blockchain network. This paper proposes a framework to build a blockchain-based architecture for e-commerce services, considering all layers of this architecture. We analyze how to integrate the existing components designed in the previous studies [120], [130] to build the architecture. More precisely, we make the following contributions.

1) We structure the components of the electronic payment ecosystems on the one hand and the blockchain on the other hand in a layered presentation.
2) We define a taxonomy of the leading consensus protocols, based on their functional architecture, the scope of their exploitation, and the limitations of each.
3) We specify essential constituent parts of a Blockchain- based architecture for e-commerce services.
4) We design a usable and privacy-preserving model of blockchain-based payment architecture and envision a consensus protocol for better throughput.
5) We evaluate an Ethereum-based scenario of the proposed architecture, demonstrating its usability, privacy guarantee, and acceptable performance in comparison with other existing models.

The originality of this study stems from the establishment of a connection between existing payment systems and emerging blockchain technology through the proposed framework and the resulting architecture. Moreover, the integrated architecture incorporates a streamlined biometric-based key recovery mechanism, and offers enhanced privacy compared to existing systems, without requiring supplementary cryptographic operations.

The remainder of this paper is organized as follows. Section II provides a structured view of centralized payment platforms. Section III describes blockchain as a new paradigm of e-commerce with more focus on the consensus protocol. Section IV analyzes some blockchain-based e-Commerce systems. Section V proposes a new blockchain-based architecture and Section VII presents a concrete model of this architecture. We evaluate a subset of this architecture in Section VII. Section VIII concludes the paper with indications of future research directions.

7.2 Centralized Payment Platforms

Figure 5.1.a presents a framework for Centralized Payment Platforms *(CPP)*. It consists of payment methods, Payment Gateways *(PGs)*, Institutions of Payment (*IoP*), and security layer. The payment methods are account-to-account (A2A) payment, Buy Now Pay Later (*BNPL*), credit/debit card payment, and digital wallet payments. The payment gateway is central to the *CPP*. It links all participants in the payment process: the sender of funds, the receiver of funds, and the IoPs to enable the payment process. Banks, Mobile Network Operators *(MNO)* that offer payment services to unbanked people, and independent Payment Service Providers (*PSPs*) are all processors of

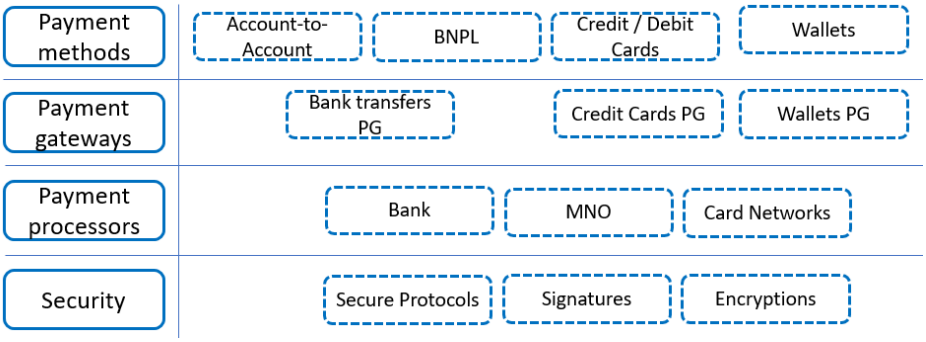
payment [8]; thus, we designate all these institutions with the umbrella expression: institution of payment. There are many types PGs according to the sup- ported payment methods. However, each PG includes a layer for authentication of the user, the Front-End Network (*FN*) that allows communication between final users (mainly in local commerce or relative services [40]), and the Back-End Network (*BN*) composed of payment processors involved in the payment.

7.2.1 Authentication Layer

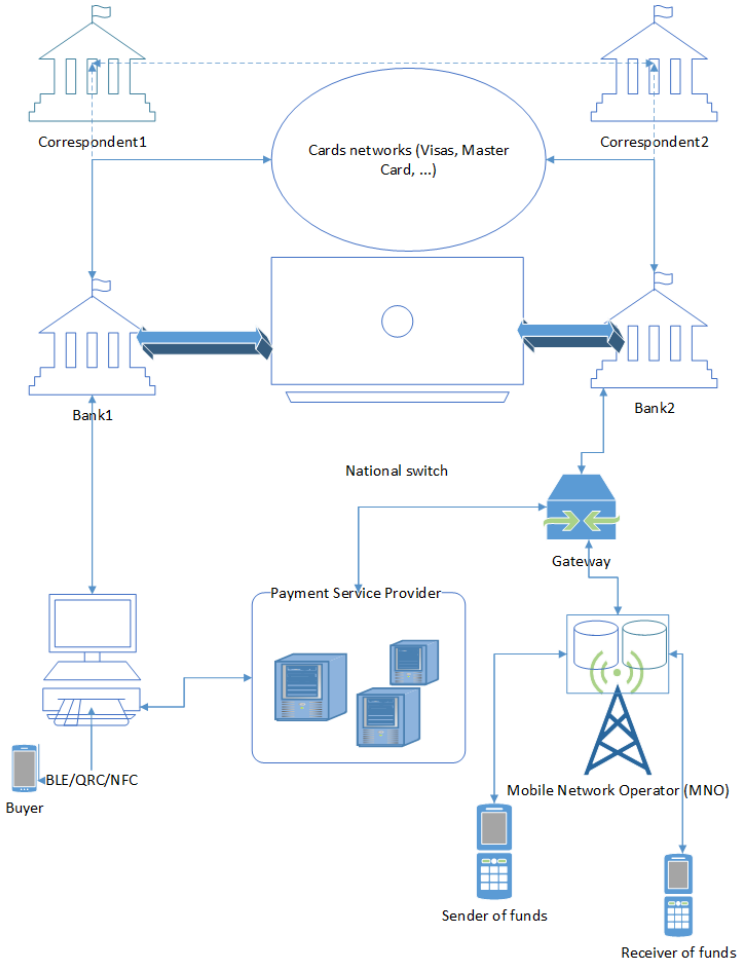
Each IoP has an authentication processor to ensure that only authorized users access its payment service. Several authentication methods (based on specific factors) are well-established in the industrial world [131], [132]. The user may be required to show a personal identification number (*PIN*) or any general password; this factor is what he knows. The IoP sometimes asks for something that the user possesses: a card or another token. Users may also present what they are (biometric characteristics) to authenticate themselves. To enhance the safety of the targeted system, authentication processors combine several authentication factors. This result is called multi-factor authentication (MFA). The use of biometric characteristics is becoming more prominent, and there are trends to authenticate users without a password for more convenience [132].

7.2.2 The Front-End Network

This layer allows the senders and receivers of funds to interact between them and with the processors of the transactions. The components of this layer depend on the type of e-commerce services or payment methods offered by the platform [133]. In remote payments such as Internet transactions, mobile banking, or person-to-person (*P2P*) transactions, the senders and receivers of funds communicate directly only to the central server. In this case, the Internet protocols may be the only communication technology between all participants of the payment process; sometimes, Short Message Service (*SMS*) and Unstructured Supplementary Service Data (*USSD*) are the communication technology to support interaction between the senders/receivers of funds and the central server [38].



a) The centralized payment framework.



b) Physical infrastructure of the centralized payment platform.

Figure 7.1 The centralized payment framework and infrastructure

In proximity payments such as local commerce, the senders of funds directly talk to the receivers of funds. In this case, there is a clear Front-End Network (*FN*) through, the communication is established between participants. Three main communication technologies are often used to establish this network [31], [11], [134]: Bluetooth Low Energy (*BLE*), Quick

Response Code (*QRC*), and Near Field Communication (*NFC*) to connect the sender to the receiver of funds (Figure 5.1.b). Both calls to the BN via the Transfer Control Protocol and Internet Protocol (*TCP/IP*) connection to process transactions.

7.2.3 The Backend Network

A BN is typically the interconnected payment system of all IoPs. It is composed of credit card networks, banks, several PSPs, and MNOs that offer payment services. The IoP authenticates the sender/receiver of funds, allows the debit or credit of funds, and acknowledges its actions to another IoP responsible for finalizing the transaction. Sometimes, verification is required before processing transactions.

7.3 Blockchain-Based Payment Framework

The layers constituting the blockchain ecosystem are illustrated in Figure 7.2. The first layer is the applications of the technology. The second allows the user to interact with the Blockchain: this is the user interface (*UI*). The third is a contract where we find algorithms and the mechanism that implements the business logic (mainly in the so-called smart contract). The incentive layer deals with the rewards of the participating nodes. The consensus layer has a set of rules that allow the validation of transactions or the acceptance of validation from other nodes of the blockchain network. The network layer describes the protocol used for the propagation of data between the nodes. Finally, the data layer defines the set of techniques adopted to organize and protect data stored at a node.

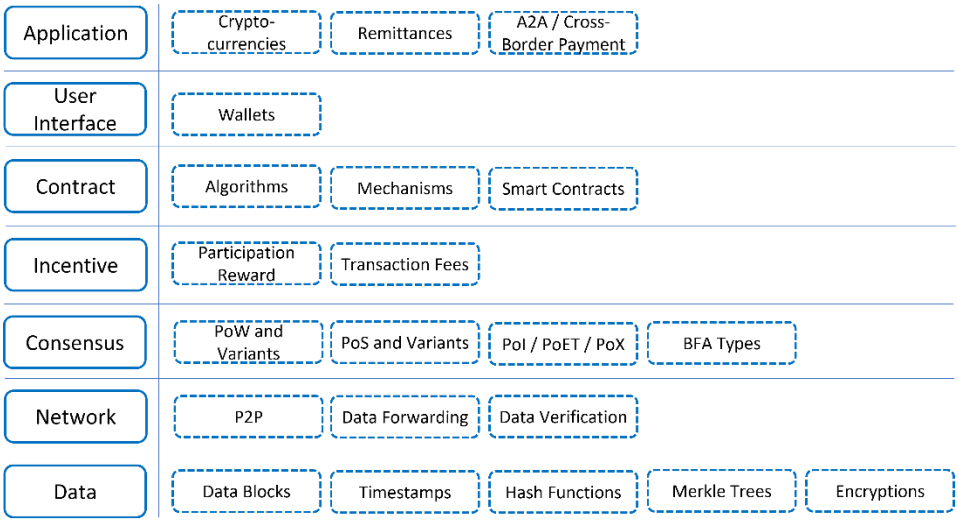


Figure 7.2 Blockchain-based payment framework

7.3.1 Known Applications

Cryptocurrencies are the most popular and established applications of blockchain [6], [17], [135]: Bitcoin, Ethereum, EOS, and Peercoin. However, blockchain has other applications in the financial world such as: cross-border payments [121], and inter-bank exchanges [53].

7.3.2 User Interface

The user has a crypto-wallet for interacting with the blockchain. This wallet is a container for cryptographic keys [135]. It can be software (e.g., web or mobile), hardware, or physical. Typically, a user of the blockchain generates a private key randomly (Figure 7.3). Even when we have a seed or bunch to determine a series of keys, the first piece of data comes from a random process. An issue with this method is the recovery of the key in the case of a loss.

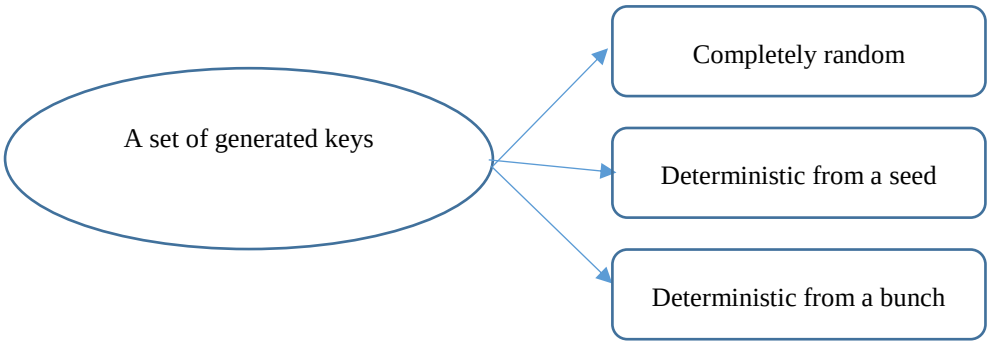


Figure 7.3 Key generation methods

7.3.3 Consensus Layer

The consensus protocol is one of the most important aspects of a blockchain-based system. This impacts the performance, security, and usability of the system. The characteristics of numerous existing protocols allow for a taxonomy when considering their scope of exploitation, performance, security, and limitations. This section presents the most representative and distinctive classes of consensus protocols.

7.3.3.1 Proof-of-Something Protocol

These are distributed consensus protocols that are opened to all unidentified computers connected to the Internet. These protocols are then completely decentralized. The Proof-of- Something (*PoX*) means that the participating node possesses a property that it provides to acquire the right to validate processing transactions in the system. This class's first known and well-established protocol is the

proof-of-work (*PoW*) proposed with the Bitcoin network [6]: each node provides work in resolving a mathematical problem. The system based on PoW (or variants of this protocol) performs poorly and is energy-consuming [20].

Proof-of-Stake (*PoS*) is an alternative to PoW in which the participating node acquires the right to validate transactions through its assets in the network. Many variants of this protocol have been proposed; Xiao et *al.* [20] put together these different variants and three sub-categories among them are: (1) the chain-based PoS where the system chooses the longest chain of blocks when many are proposed, (2) the committee-based PoS where only a committee can validate transactions, and (3) delegated PoS (*dPoS*) where validators are elected among the wealthiest nodes. The recorded performance of these systems is between hundreds and millions of TPS. Many other protocols of this class are proposed [21], [22], [136]: Proof-of-Importance (*PoI*), where the valid participation of the node to maintain the network gives the right to validate transactions, Proof-of-Reputation (*PoR*) where a measure of the reputation of the node is computed, Proof-of-Burn (*PoB*) where the node gains the privilege of validating transactions at the expense of its assets on the network, to name a few. While the latter proposals solve performance (low TPS) and energy waste problems, the network becomes more vulnerable. Indeed, the cost of a Sybil attack, for example, is much lower than that of a similar attack in the PoW case. In some scenarios with PoS, the attacker does not need to spend any assets, which is a nothing-at-stake attack. The designers of PoI compute a score and require a node's property level for participation, which tends to mitigate nothing-at-stake attacks.

7.3.3.2 BFT-Like Protocols

These consensus protocols are proposed solutions to the Byzantine general problem [23], searching for Byzantine agreements. One of the first implementations is the Practical Byzantine Fault Tolerance (*PBFT*) which supports the Hyperledger Fabric system [24]. All participants are identified, and their admissions to the network are generally centralized. Therefore, this protocol allows for partial decentralization. In PBFT, one of the nodes leads the validation of transactions. It starts each step of the vote by sending messages to all other servers. Consequently, when it receives sufficient confirmations (2/3 of all messages), it starts another step until the end of the last step (Figure 7.4). Libra Blockchain follows a similar consensus protocol [137]. These types of protocols are not scalable because of the large number of messages during the validation process.

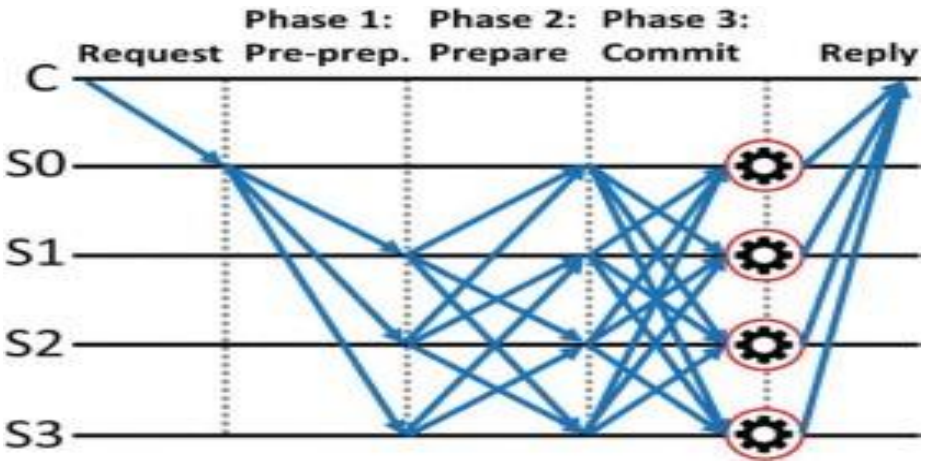


Figure 7.4 Normal operation of PBFT [20]

7.3.3.3 FBA-Like Protocols

A protocol is a Federated Byzantine Agreement (FBA) if it can reach an agreement with only a subset of participants. It allows a participant to choose a subset of other nodes to trust in order to validate transactions; this subset is a quorum. In the Stellar Consensus Protocol (SCP), each node can choose the pairs it trusts [127]. FBA-like protocols function as the Internet; therefore, there is no burdens on the leader of the process as it was the case with BFT protocols. They are also more decentralized than BFT-like protocols.

We present two consensus protocols that are FBA-like protocols [127], [129]: the Ripple Consensus Protocol (RCP) and the Stellar Consensus Protocol (SCP).

The following properties define RCP:

- The network consists of a set of interconnected servers or nodes $\{N_0, N_1, ..., N_i, ...\}$.
- Each node N_i trusts a subset of nodes UNL_i; N_i has a secure channel connection with all nodes of this subset.
- Each node stores a copy of a ledger L which represents accounts with balances in the systems, and a history of transactions $T_0, T_1, T_2, \ldots, T_n, \ldots$ from the beginning. At the point of agreement, each node had the same copy of the ledger.
- To update L, we have a deliberation step followed by the validation step.

Researches enlighten many limitations of RCP [138], [139]:

- The network does not always reach safety and liveness.

- Trends to centralization: analysis reveals the validation of transactions by a few servers in the network.
- It is easy to access the data of validated transactions in the system, which implies serious privacy issues.

SCP follows the same principle as RCP, based on FBA. However, with the RCP, the membership is nearly closed [128], [129]. This is one of the motivations for SCP. It is assumed that every node can join the network freely. Each well-behaved node relies on subsets of trusted nodes (quorum slices) in the network to validate transactions and believes that at any time, it can find the necessary information from one of its slices. A node forms quorum from its local and often static configurations. We do not know a clear dynamic mechanism to detect the well-behaved nodes among the groups to build their quorums.

In Table 7.1, we highlight the functional characteristics that allow a comparison of the most notable consensus protocols. These characteristics allow for taxonomic classification, as shown in Figure 7.5.

7.4 Related Work

In this section, we present two Blockchain-based Payment Systems (*BBPS*), aiming at offering e-commerce services. Next, we describe the main characteristics of these architectures.

7.4.1 Blockchain-based Payment Systems

Wang et *al.* [53] designed a Hyperledger-based system that implemented functionalities for Real-Time Gross Settlement (*RTGS*). It provides services of the traditional inter-bank payment system. Therefore, we designate it as a Blockchain- based Inter-bank Payment System (*IBPS-B*). Besides gross settlement, it supports gridlock resolution and reconciliation for inter-bank payment resolution. The system conducts transactions on bilateral channels (e.g., peer-to-peer links between banks) when it wants to guarantee privacy. A multilateral channel of Hyperledger Fabric [24] is used for distributed gridlock resolution algorithm. However, this method is not flexible enough (e.g., it is specific to Hyperledger Fabric). Moreover, as Hyperledger is based on PBFT, this system is not scalable.

Table 7.1 Characteristics of the most exploited consensus protocols in blockchain

	PoW and variants (Bitcoin, Ethereum, and Bitcoin-NG)	**PoS and variants**	**BFT-like protocols**	**FBA-like protocols**
Security level (*Price of a successful attack*)	High (*50 % of hash power*)	Variable (*rather high for 50 % of stakes, medium for 33 %, low for 0 %*)	Variable in terms of fault-tolerance (*medium for 33 %, rather low for 20 % of all nodes*)	Depends on the ability to form the quorums
Privacy	Partial	Partial	No	No
Performance (Throutput / Delay)	Low (*TPS* $\leq$ *200, delay around 600 S with Bitcoin*)	Variable (*TPS* $\leq$ *100 for chain-based PoS / 100* $\leq$ *TPS* $\leq 10^6$ *for BFT-based and delegated*)	Thousands of TPS / Instantaneous delay	Thousands of TPS / Instantaneous delay
Scalability	Good	Good for chain-based PoS / Bad for dPoS	Bad (*Generally, a complexity of* $O(n^2)$)	Good
Other good properties	Well-established (*many crypto-currencies, many applications on Ethereum network*) / Liveness and safety	Some variants have high throughput / Liveness and safety	Energy efficient / Suitable for inter-bank applications	Energy efficient / Flexibility given to each node to choose its trusted pairs
Other limitations	High energy consumption / No accountability mechanism	No accountability mechanism / Trends to centralization for some variants	Trends to centralization for some variants / Closed membership	Liveness not always guarantee / Closed membership with RCP

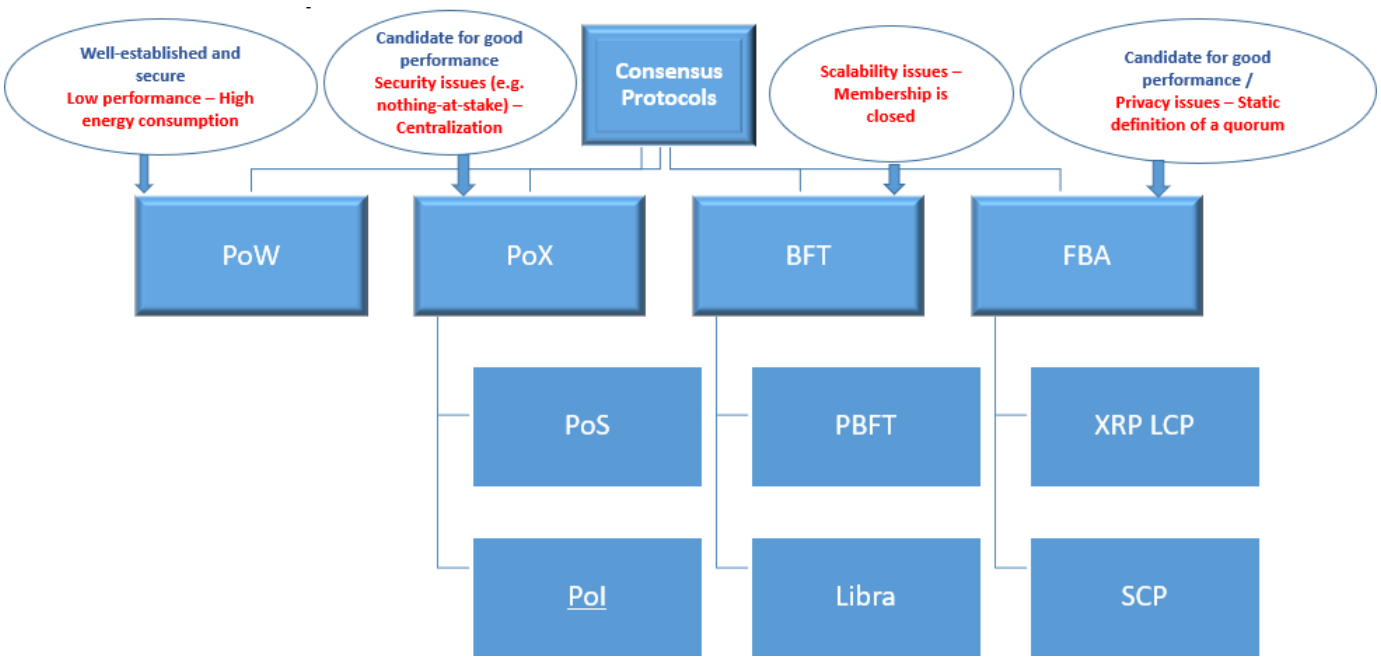


Figure 7.5 A taxonomy of consensus protocol in blockchain

The system proposed in [54] is an E-commerce Payment Model-based on Blockchain (*EPM-B*). It aims at realizing e-commerce payments without a Trusted Third Party (*TTP*). It relies on asymmetric cryptography inherent to the operation of the blockchain to guarantee authentication

and non-repudiation of transaction data. The described scenario makes use of direct communication between the sender and the receiver of funds through the Quick Response Code (*QRC*). First, the receiver of funds (e.g., the merchant) requests a payment that includes all information for the transaction to take place (e.g., transaction ID, address of the merchant, amount, timestamp, and signature with the private key of the merchant). Then, the sender of the funds (e.g., the buyer) asks the blockchain to pay the receiver of the funds. It sends all the necessary data with its signature.

7.4.2 Main characteristics of the BBPS

As main characteristics of IBPS-B, we notice the points hereunder:

- Lack of flexibility: The system is strongly connected to the design of Hyperledger;
- It intelligently ensures privacy at the level of a centralized system: only senders and receivers of funds have access to the entire data transaction;
- Gives a guarantee for Anti-Money Laundering (*AML*) functions: Here, this is a question mark because every- thing depends on the consortium.

As main characteristics of EPM-B, we notice the following points:

- Guarantee of Flexibility: the design is not specific to a particular Blockchain;
- No privacy: Every node has access to the whole data transaction;
- No guarantee for functionalities required by the regulation: the system is completely decentralized and the traceability of the transaction is Bitcoin-like.

7.5 The Proposed Architecture

The Blockchain-based architecture that we propose to offer electronic commerce services has four layers, as shown in Figure 7.6: (1) the authentication layer allows each IoP - which we also designate as a Financial Institution (*FI*) - to authenticate its own user, (2) the front-end (*FE*) layer gives the final user an interface to interact with the network of IoP to request services, (3) a connector that links the internal system at the IoP (e.g., CPP) to the blockchain, and (4) the blockchain layer.

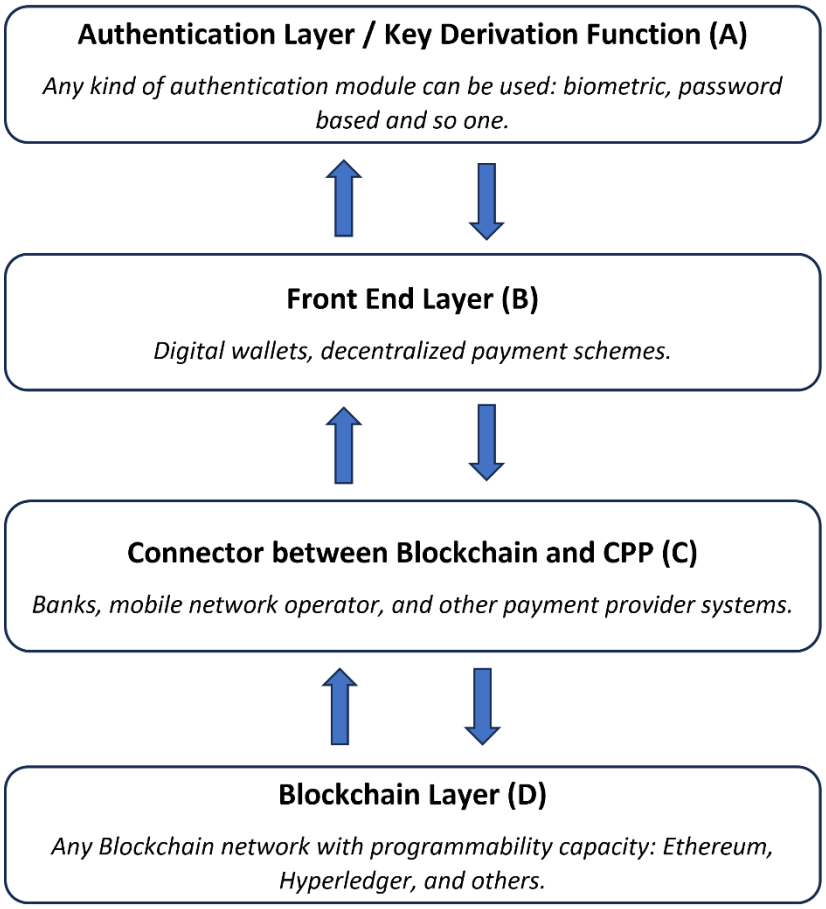


Figure 7.6 Blockchain-based e-commerce architecture

7.5.1 The Authentication Layer

An entity must be authenticated prior to accessing the system. The authentication module inside the IoP is of any type [131], [132] : password-based, biometric-based, or MFA. As part of this system, we provide a biometric-based authentication scheme for a physical person; with a non-physical person, a Physical Unclonable Function (*PUF*) may provide a personalized token [140]. We adopt the biometric-based key derivation function (*KDF*) described in chapter 5. The system allows user authentication before authorization to the mobile wallet. It also generates a cryptographic key based on the user's biometric characteristics. This aspect is essential for a blockchain-based system because it provides an efficient key recovery mechanism.

7.5.2 The Front-End Layer

The front-end (*FE*) layer is implemented with a digital wallet. A digital wallet manages the keys of the user in a blockchain-based system. However, with common existing wallets, transaction data are accessible to all the participating nodes in the blockchain network. This is a source of privacy leakage. We adopt for this architecture a decentralized scheme that protects privacy by preventing each node from having access to all meaningful transaction data [120]. This scheme allows for the transfer of funds between two users belonging to the institutions of payment (*IoP*) linked by a blockchain network. First, the receiver of funds sends a temporal account (*tA*) with the due amount (*M*) to the sender and his financial institution FI_R. The sender asks his financial institution FI_S to make the payment. The FI_S then debits the internal account of the sender IA_S and broadcasts the transaction {*tA*, *M*} to the blockchain network. When the transaction is validated on the blockchain, FI_R notifies the receiver to close the cycle. FI_S has no access to the account data of the receiver, and FI_R has no access to the account data of the sender.

7.5.3 Connector between the Front-End Layer with the Centralized Payment System

Inside the IoP, the functional payment infrastructure is represented by the simplified model shown in Figure 7.7. The main components are as follows. First, an authentication processor is used because the final user must be authenticated by the IoP. A transaction processor that receives the payment requests of a user, debits or credits the internal account of this user, and forms the transaction for the blockchain. The third component is the data saving feature (data store): IoP saves the movements on the internal account of the user. This ensemble often relies on a national or international switch to finalize a payment request in collaboration with other IoPs. In this paper, the group (processors/datastore) forms the Traditional Payment Infrastructure (*TPI*). In our architecture, TPI writes or reads transactions to the Blockchain network. Typically, {tA, M} is sent on the blockchain.

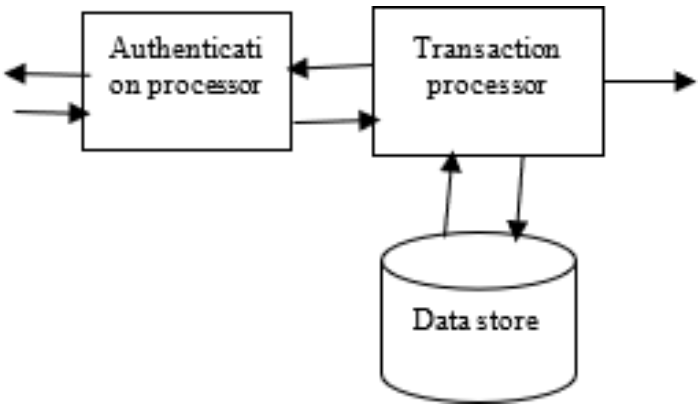


Figure 7.7 Simplified model of the centralized payment platform

7.5.4 The Blockchain Layer

The blockchain layer acts as a connector between different TPI. Blockchain characteristics secure inter-institution payments. The Blockchain for this architecture can be public (with open membership) or not; it must be programmable. At this point, we focus on the consensus protocol because it involves many other features such as performance and security. We present two models for this architecture: one with Ethereum [65] and the other with an extended version of the SCP [127].

7.6 Concrete Blockchain-based Models for the e-Commerce

This architecture can be implemented using Ethereum or the SCP. In both cases, the infrastructure includes the Blockchain network, the KDF, the decentralized payment scheme [120], and the connector with the internal system of the IoP.

7.6.1 The Concrete Model with the Ethereum Blockchain

With the Ethereum Blockchain, there are two possible scenarios: a consortium of IoPs running a private Ethereum network with the above-mentioned modules, or the involved IoPs connecting to the main public Ethereum network. In both cases, a protocol establishes a clearing period between the IoPs. As stated in [120], the process can be executed by an authority (e.g., a central bank), a committee, or using another network such as the Real Time Gross Settlement (*RTGS*), or the system of the Society for Worldwide Inter-bank Financial Telecommunication (*SWIFT*). The objective is to facilitate the secure transfer of funds between clients of IoPs that are loosely interconnected. Figure 8 illustrates the realization of two scenarios of the architecture with the Ethereum-based

model: (1) the scenario presented with red lines in dots is the protocol in [120] for local commerce with a decentralized payment scheme, and (2) the scenario with the yellow line shows P2P payment between users attached to 2 different types of IoP.

7.6.2 The Concrete Model with SCP-based Blockchain

With the extended version of SCP, we aim to facilitate A2A payment. It takes the following advantages from the SCP [127]:

- SCP is an internet-like protocol, meaning a federated agreement with open membership;
- SCP has the potential for good performance and instantaneous payment transactions;
- SCP has the potential for good decentralization.

However, some issues motivate the envision of an extension of the current protocol [127], [128]:

- The number of faulty nodes can increase, and a non-faulty node requires a mechanism to detect a faulty node and remove it from the quorum.
- Local configurations of nodes give birth to quorums, which often depend on a static mechanism to discover non-faulty nodes.
- Privacy remains of great concern for blockchain-based payment systems, thus we need to adapt the payment scheme in [120] with the SCP;

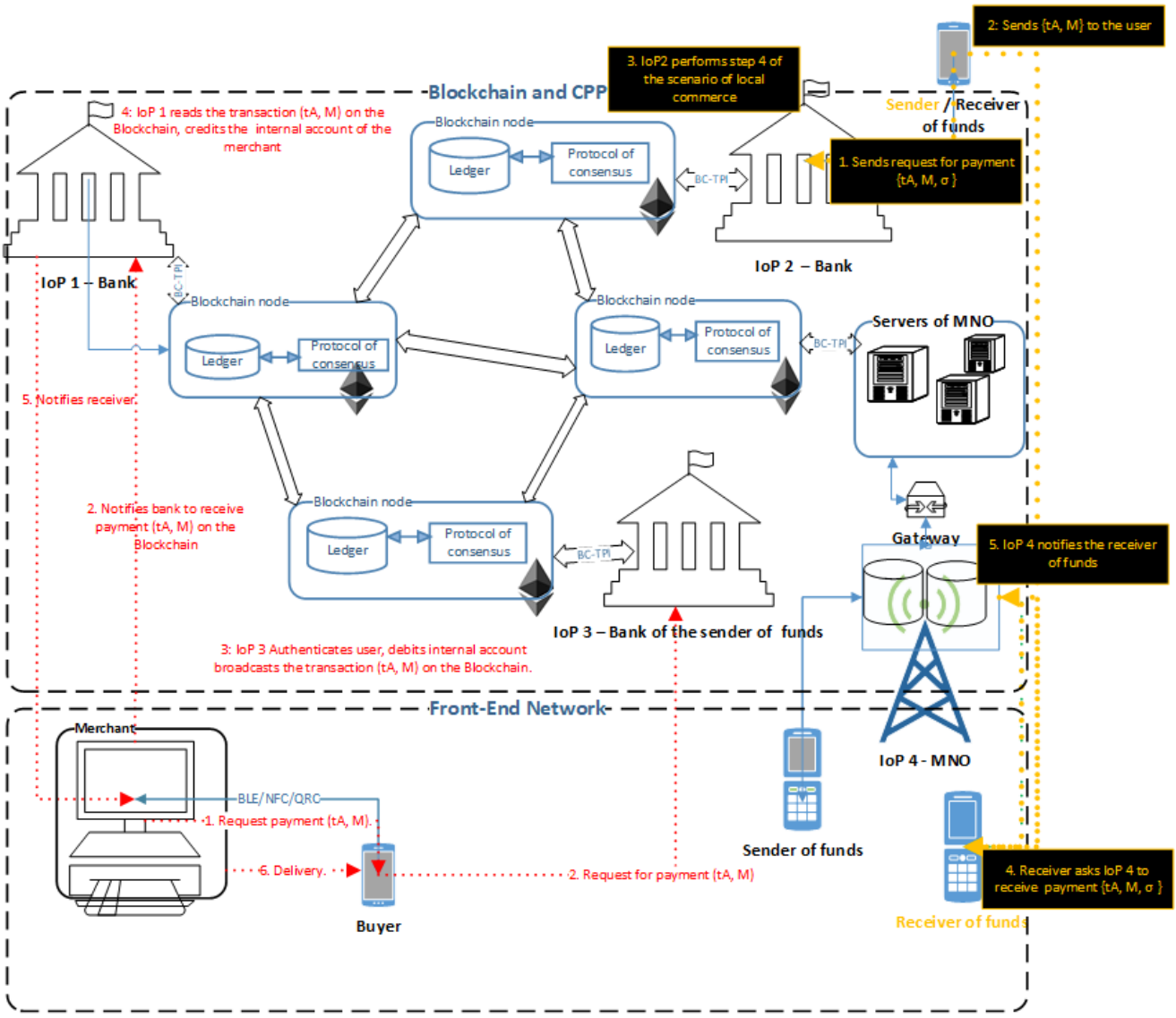


Figure 7.8 Ethereum-based payment model for e-commerce

7.6.3 On the Combination of SCP, PoI, and PoR

This section provides an envisionment of the potential extension of SCP for a candidate model with better throughput.

7.6.3.1 System Model

The consensus protocol operates according to the following rules:

- Each IoP has a dynamic set of confident institutions to which it is connected.

- Each node has a level of reliability that measures its importance in keeping the system alive; this reliability is also a measure of its reputation.
- Each IoP establishes a dynamic quorum (e.g., for a period) using the reliability level of each node.

7.6.3.2 Establishing Dynamic Quorum

To establish a quorum dynamically, the idea is to periodically compute a level of reliability for each discovered node. Some inspiration for this computation comes from the proof-of- reputation (*PoR*) and the PoI consensus [21], [22]. With PoR, each node computes a score for another node based on its interactions (Eq. 1) [28].

$$R_{ij} = \begin{cases} \sigma(i,j) * \dfrac{s(i,j)}{s(i,j) + f(i,j)}, & if\ s(i,j) + f(i,j) > 0 \\ p, & otherwise \end{cases}$$

Where:

- R_{ij} is the reliability of the node j perceived by i;
- s(i,j) measures the succeeded transactions between i and j;
- f(i,j) measures the failed transactions between i and j;
- σ(i, j) measures the similarity between i and j;
- p is an initial value of the trust of i in j;

With PoI, the assets of each node, the proximity, and its participation in flowing transactions, contribute to its reliability.

With respect to the computed score, a node classifies the most trusted participants to form its quorum. This score is updated regularly to have more reliable participants in a quorum at each point in time.

7.7 Analysis and Evaluation

In this section, we analyze the security and privacy of the Ethereum-based model. We then evaluate the costs of the payment protocol.

7.7.1 Security and Privacy

The characteristics (e.g., security and privacy) of the proposed architecture depend on those of the components. Therefore, we refer to the characteristics of the KDF and payment scheme as presented in previous articles [120], [130].

7.7.1.1 Authentication Layer

The final user (e.g., the sender/receiver of funds) has a key pair derived from the biometric characteristics using the system proposed in [130]. This system preserves privacy. In addition, the user is attached to an IoP (e.g., a bank, a MNO, or an independent PSP). With this last feature, we make available the Know Your Customer (*KYC*) functionalities, and any traceability required by the regulation.

7.7.1.2 Front-End Layer

In [120], we presented a detailed analysis of the security and privacy of the decentralized payment protocol that forms the front-end layer. An innovative aspect of this payment protocol is the provision of an elevated level of privacy beyond that of typical offerings. Specifically, the sender's financial institution remains unable to access the recipient's account, and vice versa. We have also studied how this system overcomes the vulnerabilities of smart contracts.

7.7.1.3 Internal System of the IoP and Blockchain

One of the weaknesses of correspondent banking is the lack of security guarantees for each intermediary. To mitigate this problem, the proposed system establishes connections be- tween the IoPs through Blockchain. Therefore, the entire system is immune to this vulnerability because the blockchain is designed to allow peer-to-peer exchanges without trust.

The security of the blockchain depends on the security of the underlined consensus protocol. The proposed concrete model is based on Ethereum. Therefore, it has the security of PoW which is one of the most secure distributed consensus protocols. We envision an SCP-based model that needs to be formally designed and evaluated. In this last scenario, the security depends on the security of the designed extension of the SCP.

7.7.2 Cost of the Payment Protocol

In this section, we analyze two types of costs: (1) the memory cost which is the memory required on the mobile device of the user to run the payment protocol, and (2) the communication cost which is the amount of data transmitted on the FN or from the FN to the BN.

As an initial setup, we learn consistent biometric characteristics from the user and keep helper data (e.g., *Q, CP*) and a code of reference (*C*) for the user; we generate the cryptographic key (*K*), and save all the necessary information on the mobile device of the user [130]. One can argue that the helper data, the code of reference, and the key are at risk on mobile devices. To mitigate this security problem, we suggest using BrainShield [141], a malware detection model. Before we begin to process a payment, we establish a secure channel by executing the protocol in [52].

7.7.2.1 Memory costs

The components of the memory cost are:

- Memory cost of authentication on the mobile device (auth_MC): auth_MC = size of Q (vector of 24 floats) + size of CP (vector of length less than 128 shorts) + size of C (vector of length less than 128 floats) + size of K $\leq$ 1120 bytes.
- Memory cost of the secure channel (sc_MC): sc_MC = 65 bytes [52].

The maximum value of the total memory cost is **tot_MC** = auth_MC + sc_MC = **1185 bytes**.

7.7.2.2 Communication costs

With the assumption of AES encryption, the components of the communication costs are:

- Secure channel cost (**sc_CC**): The cost of establishing the secure channel: **sc_CC = 162 bytes** [52];
- First request for payment (**req_p_1**): The cost of sending the request for payment ({tA, M}) from merchant to buyer: **req_p_1 = 16 bytes** (tA of size 12 bytes and M is a float).
- Second request for payment (**req_p_2**): The cost of sending encrypted ({tA, M}) from the merchant to the IoP of the receiver of funds: r**eq_p_2 = 16 bytes**;
- Third request for payment (**req_p_3**): the cost of sending encrypted ({tA, M}) from the buyer to the IoP of the sender of funds: **req_p_3 = 16 bytes**;

- The notification for payment (**note_p**): The cost of sending an encrypted notification from the IoP to the receiver of funds: **note_p = 16 bytes**.

The total communication cost is **tot_CC** =sc_CC*2 + req_p_1 + req_p_2 + req_p_3 + note_p = **388 bytes**.

7.7.3 Comparison with other Systems

We compare our architecture with 2 other systems that aim to offer e-commerce services on a blockchain. We consider three criteria: (1) flexibility which measures the ability to implement the architecture with multiple blockchain networks; (2) privacy-preserving which measures the capacity of the system to hide data transactions from a non-authorized user; and (3) regulations which measure the capacity of the plat- form to provide functions that are required by the regulation.

Table 7.2 Comparison of three Blockchain-based architecture for e-commerce

	EPM-B	**IBPS-B**	**Our work**
Flexibility	Yes	No	Yes
Privacy	No	Yes *(like existing systems)*	Yes
Regulation	No	Yes	Yes

7.8 Conclusion

In this research, we introduce a framework for facilitating blockchain-based payments. We identified crucial components from established electronic payment systems and blockchain networks. We have designed an architecture that leverages blockchain technology to connect diverse payment platforms. A concrete model of our architecture includes two new designed components: 1) a Key Derivation Function (*KDF*) that allows authentication and key recovery, 2) a distributed payment scheme to tackle privacy concerns commonly associated with blockchain systems. We conducted performance and security analyses using the Ethereum network and envisioned extending the Stellar Consensus Protocol (SCP) for improved performance. This system is well-suited for establishing interoperability among various heterogeneous e- commerce

systems and represents a significant step towards integrating blockchain technology into the financial industry, with legal acceptance.

Moving forward, future research will concentrate on formally defining, modeling, and evaluating the extended version of SCP. Subsequently, a more detailed examination of the e-commerce system based on this protocol will be conducted.

CHAPITRE 8 DISCUSSION GÉNÉRALE

Dans ce chapitre, nous présentons les résultats de la thèse eu égard aux objectifs de recherche énoncés à la section 1.3. Nous analysons également la méthodologie de l'ensemble de la recherche. Finalement, nous évaluons les résultats obtenus et l'applicabilité de cette nouvelle architecture dans le monde du commerce électronique.

8.1 Analyse de la méthodologie

L'objectif de la thèse étant de concevoir une architecture basée sur la Blockchain pour le commerce électronique, nous avons visé certains composants avec des enjeux connus. Ces enjeux ont été détectés de l'analyse des travaux existants. D'abord, une fonction pour lier la personne physique à son identité sur le système et un apparieur biométrique. Ensuite, un schéma de paiement pour préserver la vie privée. Finalement, une intégration intelligente des composants fournis avec d'autres composants existants pour réaliser l'architecture, avec la proposition d'un cadre adéquat. Nous analysons ces différentes étapes de la conception en vue d'évaluer l'ensemble de la méthodologie de la recherche.

8.1.1 Un apparieur biométrique et une fonction de dérivation de clé

Pour lier l'utilisateur qui est une personne physique à son identité sur le système conçu, nous avons créé la clé privée de l'utilisateur à partir de ses caractéristiques biométriques. Il existe d'autres approches pour créer une clé cryptographique à partir des données floues. Nous les avons classifiées dans 3 grandes catégories : les extracteurs flous ou *Fuzzy Extractor* (FE), des méthodes mathématiques artisanales et celles qui se basent sur l'intelligence artificielle.

L'une des premières approches, l'extracteur flou (FE) ne permet pas de reproduire la clé avec un niveau acceptable de variabilité intra-utilisateur. Certains ne prennent pas en compte les spécificités biométriques liées à un utilisateur donné. D'autres n'ont pas les caractéristiques (ex. bon niveau d'entropie) qui font qu'une clé cryptographique générée soit robuste.

La KDF fournie dans ce travail utilise dans un premier temps des méthodes de regroupement (qui sont des méthodes de l'intelligence artificielle). Avec un algorithme léger (comparativement à des méthodes qui utilisent l'apprentissage profond), elle crée d'abord un code à partir des données biométriques. Ce code pouvant être reproduit ultérieurement à l'aide d'esquisse apprise initialement et des données biométriques de l'individu, une comparaison avec

un code de référence permet d'authentifier ce dernier. À partir de ce code, la clé cryptographique est créée, utilisant les normes de l'IETF pour un HKDF. Pour appliquer l'adaptation de l'algorithme, un identifiant externe de l'individu forme l'entrée avec le code biométrique.

Nous avons détaillé le rationnel derrière notre méthode d'encodage. Elle s'appuie sur le fait que toute caractéristique biométrique pour pouvoir être utilisée pour l'authentification a 2 propriétés de base : consistance et discrimination. La consistance mesure la similarité entre des caractéristiques extraites d'un même individu à des instants différents. Quant à la discrimination, elle mesure le fait que 2 caractéristiques extraites de deux individus différents peuvent être séparées par certains critères.

L'essentiel du processus d'encodage est de trouver des groupes qui peuvent être représentés par des valeurs typiques, telles que les éléments d'un groupe soient toujours proches de la valeur qui représente son groupe pour toute lecture des caractéristiques biométriques. L'implémentation avec un algorithme comme k-moyen peut ne pas donner toujours de bonnes solutions car, il est probabiliste. Ce qui peut impliquer qu'il tombe sur des représentants tellement proches qu'il y ait chevauchement des éléments de groupes différents. Cependant, avec un algorithme qui considère l'interface de rupture naturelle comme la proposition de *Jenk* [85], on a la garantie que les éléments d'un groupe sont plus proches les uns des autres qu'ils ne le sont de tout élément d'un autre groupe.

Enfin, le fait que la transformation du code en clé cryptographique obéit aux normes de l'IETF pour un HKDF, elle permet d'avoir les caractéristiques voulues : irréversibilité et aléatoire. Cette clé peut être utilisée avec n'importe quelle variation du protocole de Diffie-Hellmann.

8.1.2 Un protocole de paiement décentralisé

Une autre phase de cette conception a pris en compte le défi lié à la privacité dans le schéma de paiement sur la blockchain. Le fait que plusieurs nœuds du réseau valident les transactions aggravent le problème du respect de la vie privée déjà présent dans les systèmes conventionnels. Alors que la plupart des propositions pour préserver la privacité se basent surtout sur les algorithmes cryptographiques, notre méthode n'ajoute pas des opérations cryptographiques supplémentaires pour donner la garantie du respect de la vie privée. Nous avons plutôt adopté un mécanisme intelligent qui utilise un nonce temporaire pour permettre l'envoie de fonds entre un expéditeur et un récepteur. Ce nonce associé au montant est envoyé sur la blockchain. Ce qui empêche aux différents nœuds d'avoir accès à l'ensemble des données d'une transaction. Ce

mécanisme permet une meilleure préservation de la vie privée que dans le cas conventionnel où plusieurs institutions de paiement (par exemple des banques, des opérateurs de téléphonie mobile ou d'autres fournisseurs de service de paiement) peuvent avoir accès à toutes les données de la transaction.

8.1.3 Un cadre pour des architectures basées sur la blockchain pour le commerce électronique

Dans cette dernière phase, nous avons mis en lumière les composants qui formeront l'architecture visée ainsi que leurs interactions. Notre approche pour arriver à une bonne intégration des composants consiste à analyser d'abord l'architecture des systèmes conventionnels, puis celle des systèmes basés sur la blockchain. Cette analyse a permis de tirer une architecture hybride qui connecte les institutions de paiement conventionnelles à travers une blockchain. Ensuite, la proposition de modèles concrets de cette architecture avec Ethereum ou base sur SCP permet d'expliciter un cas d'application de ce cadre d'architecture.

8.2 Analyse des résultats

Les expérimentations sont menées d'abord sur chaque composant séparément, faisant abstraction des autres parties de l'architecture : le mécanisme d'authentification, la fonction de dérivation de clé et le protocole de paiement. Par la suite, nous avons implémenté un prototype pour évaluer certains critères de performance de l'ensemble de l'architecture dans un contexte plus réel.

8.2.1 Évaluation du mécanisme d'authentification et du KDF

Le banc d'essai pour évaluer le mécanisme d'authentification se compose des éléments suivants:

- Un ordinateur portable équipé d'un processeur Core i7-4600U @ 2,10 GHz et de 12 Go de mémoire fonctionnant sous Windows 10.
- Un smartphone du modèle pixel 2 having 1.5 GB of RAM with the android platform.
- Une application mobile développée avec java pour Android utilisant le modèle Facenet [48] pour détecter le visage d'un individu et extraire des caractéristiques biométriques.
- Les algorithmes sont implémentés avec Matlab et s'exécutent sur l'ordinateur portable.

Notre approche a été évaluée sur trois jeux de données de visages avec des défis différents (variation de la luminosité, environnement ambiant). La durée de l'apprentissage des

caractéristiques biométriques à l'initialisation, la longueur du code produit et le seuil minimal de similarité entre différentes lectures biométriques d'un individu ont permis d'évaluer l'étape d'enrôlement d'un utilisateur. Des métriques tels que la précision, le taux de faux positifs et le taux de faux négatifs entre autres, permettent d'évaluer les performances du mécanisme d'authentification. Finalement, nous avons évalué la clé générée sur des propriétés telles que : l'entropie, la révocabilité, l'irréversibilité et la flexibilité du processus.

8.2.1.1 Les résultats relatifs à l'enrôlement d'un utilisateur

La durée de l'apprentissage des caractéristiques biométriques à l'enrôlement est souvent inférieure à 5 S. C'est une durée acceptable dans un contexte d'exploitation réel. De plus, sur deux des jeux de données, incluant le sous-ensemble de la base de données de visages de Youtube, les longueurs de code obtenus sont toutes supérieures à 82. Ce qui implique une faible probabilité (plus petit que 2^{-246}) de pouvoir générer le code en utilisant uniquement l'esquisse enregistrée pour permettre la reproduction. On dit alors que le processus d'encodage possède la propriété d'irréversibilité. Aussi, le code de référence sauvegardé sur le mobile est constitué de valeurs moyennes qui représentent chacune un ensemble flou de valeurs biométriques initiales. En d'autres mots, ce n'est pas un patron complet ou Full Template en anglais qui est enregistré comme c'est souvent le cas avec les systèmes d'authentification biométrique.

8.2.1.2 Les résultats relatifs à l'authentification d'un utilisateur

Le système authentifie avec une exactitude égale ou supérieure à 0,9 pour plus de 80 % des individus des jeux de données 1 et 3 (DS1 et DS3). De plus, sur DS1, on obtient souvent un taux de faux positif nul. Ce qui est très bon pour un système d'authentification. La comparaison sur 2 classes (aléatoirement tirées de la base de données des visages de Yale) avec quatre autres méthodes récentes, nous place proche de deux des systèmes, mais dépassant de loin l'un d'entre eux : 0.898 contre 0.99, 0.917 et 0.92 pour les 3 premiers et 0,75 pour le dernier classé. Un point remarquable de ces résultats, c'est que le premier du groupe avec une précision qui dépasse de loin tous les autres systèmes. Cette grande précision vient du fait que c'est un système basé sur l'apprentissage profond. Ces systèmes ont souvent de sérieux problèmes de respect de la vie privée. Tandis que notre système offre un bon équilibre entre le respect de la vie privée et la précision.

8.2.1.3 Les résultats relatifs à la clé cryptographique générée

La clé cryptographique générée combine de bonnes propriétés : l'irréversibilité, le caractère aléatoire et la révocabilité. Notons ce qui suit en rapport à l'irréversibilité. Les données d'aide pour pouvoir recréer le code sans l'avoir sauvegardé avant (ce qui est souhaitable) ne permettent de générer le code original qu'avec une probabilité très faible. Et même le code est une version compressée des données originales. De plus, utilisant un processus similaire au HKDF, la clé bénéficie des caractéristiques liées à un tel processus, notamment le caractère aléatoire. Quant à la révocabilité, elle est obtenue en intégrant une donnée externe comme entrée avec le code pour créer la clé cryptographique. En cas de corruption de cette dernière, on peut garder les mêmes caractéristiques biométriques avec un autre identifiant de l'utilisateur pour générer une clé complètement différente.

8.2.1.4 Mesures de la robustesse de la KDF

Plusieurs résultats témoignent de la robustesse de cette KDF relativement à d'autres systèmes de sa catégorie : un processus d'apprentissage qui ne requièrent un grand ensemble de données, la fiabilité du processus de reproduction de la clé, la capacite à corriger un niveau de bruit acceptable. En effet, les expériences avec DS1 montrent que 30 images d'un individu suffisent pour avoir les caractéristiques nécessaires et l'esquisse permettant d'authentifier l'individu ou de recréer la clé ultérieurement. La clé se reproduit avec un taux de fiabilité d'au moins 0.88 sur DS1. Aucun des KDF analysées dans la littérature ne parlent de cette caractéristique. Le niveau de bruit (ou variation des données biométriques) mesurées avec DS1 est de 19.5 %. Un FE qui n'est pas robuste peut corriger un niveau maximal de bruit (variation des données) de 15 %, et 3 % si c'est un RFE [76].

Analyse cryptographique de la fonction proposée (HKDF)

La fonction que nous proposons est robuste contre les principales attaques qui visent les KDFs:

- Attaques par force brute/ Entrée trop courte et prédictible : Notre système utilise 2 données d'entrées et l'une d'entre elles est le code biométrique produit précédemment. La probabilité de générer ce code sans les caractéristiques authentiques est très faible. De plus la combinaison de ces 2 entrées devient encore moins prédictible.
- Réutilisation du sel (Salt Reuse): Si le même sel est utilisé entre différentes applications ou sessions, un attaquant qui en a connaissance peut l'intercepter et dériver la clé. Notre

système est robuste contre cette attaque puisque le sel n'est pas la seule donnée d'entrée pour notre HKDF et le code ne peut être facilement généré.

- Appauvrissement de l'entropie (Entropy Depletion): nous avons conçu un HKDF en conformité avec les normes de l'IETF [82], [104]. L'entropie de la clé ne dépend pas de celle de l'entrée.
- Fonction de hachage faible (Weak Hash Function): nous utilisons HMAC-SHA256 qui est appropriée pour construire des KDFs robustes.
- Nombre insuffisant d'itérations: une HKDF utilise des itérations pour ralentir délibérément le processus de dérivation de la clé, ce qui rend les attaques par force brute plus difficiles. Néanmoins, si le nombre d'itérations est trop faible, cela peut compromettre la force de la clé obtenue. Le choix peut être fait pour avoir un nombre d'itérations adéquates. Dans notre cas, nous avons la longueur du code biométrique qui détermine le nombre d'itérations. Les expérimentations nous donnent des codes d'une longueur dépassant 82 composants, voir 100.

8.2.2 Évaluation du protocole de paiement

Le banc d'essai pour évaluer le protocole de paiement décentralisé comprend les modules suivants :

- Un ordinateur portable équipé d'un processeur Core i7-4600U @ 2,10 GHz et de 12 Go de mémoire fonctionnant sous Windows 10.
- Dans l'environnement de cet ordinateur s'exécutent les modules de gestions du paiement, implémenté en langage Python et simulant le système interne de l'institution de paiement.
- Un serveur implémenté avec le langage Python pour la réception des requêtes du client de l'institution; ce client peut être un expéditeur de fonds ou un receveur de fonds.
- La Blockchain est simulée à l'aide du cadre de développement Truffle.
- Un contrat intelligent ou *smart contract* pour la gestion des paiement est développé avec le langage *Solidity* et déployé sur la blockchain fournie par Truffle.
- L'interconnexion entre la blockchain et le système interne de l'institution de paiement est réalisé à l'aide de la librairie *Web3.py*.

Ce banc d'essai permet d'évaluer la durée d'une transaction avec ce protocole de paiement : les traitements internes à l'institution de paiement durent en général moins de 0.15 S; avec les signatures pour former la transaction à soumettre sur la blockchain, la durée reste en dessous de

0.17 S. La durée totale dépend de la blockchain. Avec le réseau public d'Ethereum, ce sera autour de 15 S. Toutefois, certaines expérimentations avec des réseaux de blockchain privée (8 nœuds) [117] montrent un délai très inférieur à 1 S, plus précisément 0.1673 S. Sur la blockchain publique, l'exploitation peut se faire pour des virements de fonds qui n'exigent pas une validation instantanée.

8.2.3 Évaluation du modèle proposé

Le modèle concret avec Ethereum intègre la KDF pour l'authentification et une base pour la cryptographie avec la génération de la clé privée. De plus, le protocole en [52] permet de sécuriser la communication entre l'expéditeur (respectivement le receveur) de fonds et son institution. Le coût de mémoire du protocole de paiement incluant le protocole sécurisé est de 1185 bytes, et le coût de communication est de 388 bytes. Ce sont des coûts acceptables pour une exploitation du système dans le commerce électronique mobile.

De plus, ce modèle permet de connecter les systèmes de paiement internes des institutions de paiement à travers la blockchain, facilitant ainsi une interopérabilité entre des systèmes hétérogènes.

8.2.4 Difficultés d'intégration de la solution avec les standards et normes actuels

En ce qui a trait à l'intégration aux normes et standards, il faut considérer le HKDF d'un côté et l'ensemble de l'architecture qui est un système basé sur la blockchain de l'autre. Dans d'autres cas, les normes techniques pour la communication et la cryptographie sont bien intégrées dans la solution, puisque la blockchain fait usage des standards de communications et des normes de la cryptographie.

8.2.4.1 HKDF

La HKDF est construite selon les normes admises par l'IETF [82], [104]. Son exploitation réelle aurait une limitation lorsque le client a un équipement de faibles ressources. En effet, les normes indiquent que l'entrée doit être assez longue pour donner lieu à une clé robuste. Dans notre cas, le code produit doit être assez long. Or, avec un équipement mobile a ressource limitée, la durée de génération de clé augmente considérablement avec le nombre d'itérations.

8.2.4.2 Le modèle d’architecture basé sur la Blockchain

Les systèmes basés sur la blockchain font face à la régulation. Dans notre cas, les autorités publiques sont réticentes à adopter les systèmes basés sur la blockchain pour des raisons d’audit dans la lutte contre le blanchiment d’argent. Assurer le respect de la vie privée et en même temps permettre les règles d’audit est un défi. Notre système considère cet aspect en permettant à un régulateur de retracer un ensemble de transactions ciblées.

CHAPITRE 9 CONCLUSION

Dans ce dernier chapitre de la thèse, nous présentons un récapitulatif des différents travaux réalisés. Nous y inclurons également les limitations de ces travaux, avant de terminer en précisant des directions de recherche qui pourraient aider à dépasser les limites.

9.1 Synthèse des travaux

Cette thèse visait la conception d'une architecture basée sur la blockchain pour des services de commerce électronique. Le premier défi considéré ici est lié à la gestion des clés pour de tels systèmes. En effet, le paradigme de la blockchain envisage une gestion des clés par l'utilisateur; ce qui est désavantageux pour le recouvrement en cas de perte. L'architecture proposée inclut une KDF qui permet non seulement d'authentifier un utilisateur de la plateforme, mais aussi de le lier à son compte sur la blockchain, et de recréer la clé ultérieurement à partir des caractéristiques biométriques. D'un autre côté, la blockchain ayant été initialement conçu avec l'idée de la transparence des données pour les différents nœuds validateurs de transactions, l'exploitation pour le commerce électronique pose un problème de privacité. Ainsi, un composant de cette architecture est un protocole de paiement qui s'attaque à cet enjeu. Finalement, la multiplicité et l'hétérogénéité des plateformes dans le commerce électronique engendrent un besoin d'interopérabilité. C'est pourquoi cette thèse propose aussi un cadre (pour l'architecture) qui envisage de coupler les systèmes internes aux institutions de paiement avec la blockchain. Ce qui facilite la connexion entre différents systèmes à travers la blockchain.

Les différents travaux réalisés durant cette thèse couvrent l'authentification biométrique, le commerce électronique et la blockchain. Elles peuvent être résumées à travers les points suivants :

1. **Un apparieur biométrique** : le mécanisme inclut une **méthode de codification** qui permet de créer un code pour l'utilisateur à partir de ses données biométriques; le code initialement créé reste comme une référence auquel d'autres codes créés ultérieurement sont comparés pour authentification de l'individu concerné. La création de ce code et sa reproduction fait face à la variabilité intra-utilisateur, puisque les données biométriques ne sont pas forcément les mêmes à chaque acquisition. C'est là le rôle joué par la méthode de quantification. Pour pouvoir reproduire le code, des données auxiliaires sont sauvegardées. Le coût en mémoire de cet apparieur reste inférieur à 1120 octets. De plus, les données

sauvegardées sont compressées et ne permettent pas de reconstituer les données biométriques captées initialement.

2. **Une fonction de dérivation de clé** : cette fonction inclut un algorithme qui suit un processus similaire au HKDF et transforme le code créé avec l'appareur biométrique précédent en une clé cryptographique. L'entrée de cette fonction est constituée non seulement du code créé à partir des données biométriques, mais aussi d'un identifiant externe de l'individu. Cette clé est donc révocable, puisqu'on peut toujours changer l'identifiant. Elle est une chaine aléatoire et irréversible. Cette KDF est aussi un mécanisme simple et efficace pour recouvrir la clé privée de l'utilisateur en cas de perte.
3. **Un protocole de paiement basé sur la blockchain garantissant le respect de la vie privée :** c'est une interface fournie à l'utilisateur pour permettre l'interaction avec la plateforme formée du système de son institution de paiement et de la blockchain. Elle offre un mécanisme intelligent de préservation de la vie privée en ne permettant à aucun nœud ordinaire de la blockchain d'avoir accès aux données complètes d'une transaction. En effet, chaque institution financière a accès à un ensemble partiel des données de la transaction : l'institution financière de l'expéditeur de fonds n'a pas accès au compte du receveur et l'institution de ce dernier n'a pas accès au compte de l'expéditeur.
4. Une taxonomie des protocoles de consensus de la Blockchain eu égard à leur applicabilité dans le commerce électronique :

   Après avoir mis en lumière les principaux protocoles de consensus établis dans l'écosystème de la blockchain et leurs principales propriétés, une analyse de leurs caractéristiques permet de faire une comparaison selon certains critères clés. Ces critères sont la sécurité, la performance en termes de nombre de transactions par seconde et de durée d'une transaction, la privacité, d'autres avantages particuliers ou limitations. Cette taxonomie a permis de définir des classes de protocoles qui facilitent le choix selon les services à offrir.

5. **Un cadre d'architecture basée sur la blockchain pour des services de commerce électronique** : ce cadre spécifie les composants à intégrer pour former l'architecture et la manière de les intégrer. Le cadre tire avantages des composants des plateformes de paiement conventionnelles et ceux des écosystèmes de blockchain pour proposer une

architecture permettant un couplage faible et sécuritaire entre les différentes plateformes dans le commerce électronique.

6. **Un modèle concret de l'architecture basée sur Ethereum avec implémentation et évaluation d'un sous-ensemble de composants :** nous avons intégré l'apparieur biométrique, la KDF, le protocole de paiement, avec la blockchain Ethereum pour concevoir un modèle concret de l'architecture en considérant le cadre proposé**.** Avec l'implémentation du protocole de paiement (environnement d'expérimentation précisé au chapitre 8), la durée d'une transaction a été évaluée. Les couts en mémoire et en communication ont été établis en associant l'ensemble avec un protocole d'établissement d'un canal sécurisé.
7. **Visualisation d'une extension du SCP** :

   L'analyse de la performance du modèle concret montre un débit limité par rapport aux systèmes existants qui offrent les mêmes services de commerce électronique. Cela est dû au protocole de consensus sous-jacents. Comme le SCP est candidat pour de meilleures performances et fonctionnent de façon similaire au réseau financier mondial, la thèse indique de s'appuyer sur ce protocole pour de meilleures performances. Toutefois, SCP définit ses quorums de manière statique; ce qui motive la visualisation d'une dynamisation de ce protocole (en utilisant les notions d'importance et de réputation d'un nœud) pour le rendre plus robuste dans un réseau ouvert à tous.

9.2 Limitations des travaux réalisés

Les travaux réalisés dans le cadre de cette thèse présentent les limitations suivantes :

- Bien que le modèle conceptuel de l'apparieur biométrique et de la KDF ne suppose aucun trait biométrique particulier (physiologique ou comportemental), il requiert une longueur du code basé sur les données biométriques pour avoir de la robustesse. Les expérimentations ont été effectuées avec les données du visage. Elles ne permettent pas de conclure quant à l'utilisation d'autres traits biométriques avec ce modèle. De plus, si le modèle offre une certaine flexibilité eu égard à l'utilisation de plusieurs techniques d'extraction de caractéristiques biométriques avec la méthode de codage, toutes les techniques d'extraction ne permettent pas d'avoir de bons résultats. C'est encore une autre limitation du modèle.

- L'appareur biométrique et la KDF sauvegardent les données auxiliaires et de référence pour authentification ultérieure sur le mobile de l'utilisateur. C'est généralement un environnement avec beaucoup d'autres applications qui peuvent être sources de vulnérabilités. Ce qui expose ces données qui doivent être confidentielles. Une mesure de mitigation a été proposée avec la réalisation concrète de l'architecture : il s'agit de l'utilisation d'un modèle de détection de logiciels malveillants. Toutefois, cette mesure ne protège pas contre une lecture à distance de ces données (par Bluetooth par exemple).
- Limitation due à la performance du contrat intelligent : le protocole de paiement conçu se base sur un contrat intelligent (avec Ethereum) ou toute autre structure correspondante sur une autre blockchain. Une bonne analyse de la sécurité du système a été effectuée pour cette couche. Toutefois, la performance du contrat intelligent n'a pas été formellement étudiée en ce qui a trait à l'évolutivité du système et à la consommation de ressources (gas par exemple) par ce contrat.
- Limitation de la performance par le réseau de la blockchain : une autre limitation du modèle concret proposé est une conséquence directe de la blockchain sous-jacente. Le débit avec des expériences proches du réel restaient en dessous de 1500 TPS. De plus, des transactions instantanées ne sont pas garanties. L'exploitation dans le commerce local ou d'autres services qui nécessitent une durée très faible de la transaction ne peut être offerte par le modèle concret dans son état.
- Pour un système avec de meilleures performances en termes de débit et de durée d'une transaction, la thèse indique des modifications de fonctionnement relatives au protocole de consensus sous-jacent : lire des données non encore validées avec Ethereum ou étendre le SCP pour construire un nouveau modèle. Toutefois, ces indications n'ont pas été étudiées à fonds pour évaluer l'amélioration apportée et la faisabilité technique de telles conceptions.

9.3 Orientations de recherches futures

Se basant sur les travaux effectués et leurs limitations présentées précédemment, nous indiquons certaines directions de recherche qui peuvent faire l'objet de travaux futurs dans le but de dépasser ces limites :

- Une caractérisation du modèle conceptuel d'appariement biométrique et de la KDF qui permettrait de trouver les traits biométriques et des techniques d'extraction de

caractéristiques associées pouvant être exploitées avec ce modèle. Pour être plus profonde, cette recherche pourrait aller jusqu'à la caractérisation éventuelle des couples (traits biométriques, techniques d'extractions de caractéristiques) souhaitables pour l'authentification.

- L'utilisation d'une carte SIM/eSIM comme élément de sécurité pour garder les données confidentielles biométriques peut être envisagée. Dans ce cas, le travail de recherche consisterait à concevoir des modèles optimisés pour les cartes SIM/eSIM qui feraient l'appariement biométrique sur la carte en évitant l'exposition des données au niveau du système d'exploitation du mobile.
- L'analyse formelle de la performance de l'architecture eu égard au contrat intelligent est à effectuer; elle peut donner lieu à une conception d'un schéma de paiement basé sur la blockchain offrant une meilleure performance.
- Une autre direction de recherche intéressante est la conception d'un modèle de cette architecture garantissant une meilleure performance en termes de débit et de durée de la transaction. Cette recherche peut inclure l'analyse de la perspective indiquée dans la thèse de lire des données non encore validées par le protocole de consensus pour diminuer la durée d'une transaction.
- Il faudra aussi définir formellement et modéliser la version étendue de SCP selon la perspective de modification envisagée dans la thèse; cette perspective indique d'accorder une fiabilité à chaque nœud, calculée à partir d'une combinaison des notions d'importance et de réputation dans le sens des protocoles PoI et PoR.
- La conception complète d'un modèle concret (respectant le cadre fourni) basé sur la version étendue de SCP est une bonne piste de recherche. Ce nouveau modèle est candidat pour de meilleures performances compte tenu des performances affichées par le SCP.

RÉFÉRENCES

[1] A. Sanchez and B. Carro, "Emerging Markets: Mobile Money for the Unbanked," in *Digital Services in the 21st Century: A Strategic and Business Perspective*, Wiley Online Library, 2017, pp. 117-142.

[2] Y. Wang, H. Christen and S. Kruttika, "Mobile Payment security threats and challenges," in *2016 second international conference on mobile and secure services (MobiSecServ)*, Gainesville, FL, 2016.

[3] J. Téllez et S. Zeadally, Mobile Payment Systems, Springer, 2017.

[4] O. Italis, *Étude comparative des plateformes de paiement mobile,* Port-au-Prince: Mémoire de maitrise à l'Institut des Sciences, des Technologies et des Études Avancées d'Haïti, 2018.

[5] K.-H. Yeh, C. Su, J.-L. Hou, W. Chiu and C.-M. Chen, "A Robust Mobile Payment Scheme With Smart Contract-Based Transaction Repository," *IEEE Access*, vol. 6, pp. 59394 - 59404, 2018.

[6] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized business review,* pp. 21260-21268, 2008.

[7] M. H. Sherif, Protocols for secure electronic commerce, CRC press, 2017.

[8] T. Lerner, "Mobile Technology and Security," in *Mobile Payment*, Mainz, Germany, Springer Vieweg, 2013, pp. 39-59.

[9] N. Gupta, Inside Bluetooth low energy, Artech house, 2016.

[10] J. Padgette, K. Scarfone and L. Chen, "Guide to bluetooth security," *NIST Special Publication,* vol. 800, no. 121, pp. 25-93, 2017.

[11] V. Coskun, B. Ozdenizci and K. Ok, "A survey on near field communication (NFC) technology," *Wireless personal communications,* vol. 71, no. 3, pp. 2259-2294, 2013.

[12] J. Jumić et M. Vuković, «Analysis of credit card attacks using the NFC technology,» chez *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, 2017.

[13] W. B. Nyamtiga, A. Sam et L. S. Laizer, «Security Perspectives for USSD versus SMS in conducting mobile transactions: A case study of Tanzania,» *International journal of technology enhancements and emerging engineering research,* vol. 1, n° %13, pp. 38-43, 2013.

[14] K. K. Lakshmi, H. Gupta et J. Ranjan, «USSD—Architecture analysis, security threats, issues and enhancements,» chez *2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS)*, Dubai, 2017.

[15] Z. Zheng, S. Xie, H. Dai, X. Chen and H. Wang, "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," in *2017 IEEE International Congress on Big Data (BigData Congress)*, Honolulu, 2017.

[16] X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. P. C. Bass and P. Rimba, "A Taxonomy of Blockchain-Based Systems for Architecture Design," in *2017 IEEE International Conference on Software Architecture (ICSA)*, Gothenburg, 2017.

[17] Y. Yuan and F.-Y. Wang, "Blockchain and Cryptocurrencies: Model, Techniques, and Applications," *IEEE Transactions on Systems, Man, and Cybernetics: Systems* , vol. 48, no. 9, pp. 1421-1428, 2018.

[18] M. Nonaka, J. Konko and C. Gaffney, "FinCEN issues guidance to synthesize regulatory framework for virtual currency," *Journal of Investment Compliance,* vol. 20, no. 3, pp. 54-56, 2019.

[19] "Today's Cryptocurrency Prices by Market Cap," CoinMarketCap, [Online]. Available: https://coinmarketcap.com/. [Accessed 10 October 2019].

[20] Y. Xiao, N. Zhang, W. Lou and Y. T. Hou, "A Survey of Distributed Consensus Protocols for Blockchain Networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1432-1465, 2020.

[21] New Economy Movement, "Proof-of-Importance," in *NEM technical reference*, 2018, pp. 26-43.

[22] New Economy Movement, "A reputation system for nodes," in *NEM Technical Reference*, 2018, p. 21 – 25.

[23] S. Robert, M. Pease and L. Lamport, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 382-401, 1982.

[24] E. Androulaki, A. Barger, V. Bortnikov and C. Cachin, "Hyperledger fabric: a distributed operating system for permissioned blockchains.," in *Proceedings of the Thirteenth EuroSys Conference (EuroSys'18)*, New York, USA, 2018.

[25] F. Halin, «Le vol des données de Desjardins touche tous ses membres,» Le journal de Montréal, 2 November 2019. [En ligne]. Available: https://www.journaldemontreal.com/2019/11/02/le-vol-des-donnees-de-desjardins-touche-tous-ses-membres. [Accès le 2019].

[26] F. Casino, T. K. Dasaklis and C. Patsakis, "A systematic literature review of blockchain-based applications: Current status, classification and open issues," *Telematics and informatics*, vol. 36, pp. 55-81, 2019.

[27] C. Berger et H. P. Reiser, «Scaling Byzantine Consensus: A Broad Analysis,» chez *Proceedings of the 2nd Workshop on Scalable and Resilient Infrastructures for Distributed Ledgers*, Rennes, 2018.

[28] S. Pierre, «Interopérabilité, architectures et plates-formes,» chez *Réseaux et systèmes informatiques mobiles*, Presses internationales Polytechnique, 2003, pp. 269-316.

[29] A. Yohan, N.-W. Lo and D. Winata, "An indoor positioning-based mobile payment system using Bluetooth low energy technology," *Sensors,* vol. 18, no. 4, pp. 974-999, 2018.

[30] G. Lackner, "A Comparison of Security in Wireless Network Standards with a Focus on Bluetooth, WiFi and WiMAX," *International Journal of Network Security,* vol. 15, no. 6, pp. 420-436, 2013.

[31] A. T. Purnomo, Y. S. Gondokaryono et C.-S. Kim, «Mutual authentication in securing mobile payment system using encrypted QR code based on public key infrastructure,» chez *2016 6th International Conference on System Engineering and Technology (ICSET)*, Bandung, Indonesia, 2016.

[32] R. Focardi, F. L. Luccio and H. A. Wahsheh, "Security Threats and Solutions for Two-Dimensional Barcodes: A Comparative Study," in *Computer and network security essentials*, Springer, 2018, pp. 207-219.

[33] S. Bojjagani and V. N. Sastry, "A secure end-to-end proximity NFC-based mobile payment protocol," *Computer Standards & Interfaces,* vol. 66, pp. 103348-103368, 2019.

[34] M. Pasquet and S. Gerbaix, "Fraud on host card emulation architecture: Is it possible to fraud a payment transaction realized by a mobile phone using an" Host Card Emulation" system of security?," in *2016 Second International Conference on Mobile and Secure Services (MobiSecServ)*, Gainesville, FL, 2016.

[35] P. Urien, "Innovative mobile payments in the cloud for connected citizen: The MobiSIM project," in *2016 18th Mediterranean Electrotechnical Conference (MELECON)*, Lemesos, Cyprus, 2016.

[36] I. Turk, P. Angin and A. Cosar, "RONFC: A novel enabler-independent NFC protocol for mobile transactions," *IEEE Access* , vol. 7, pp. 95327-95340., 2019.

[37] I. Mas and K. Kumar, "Banking on mobiles: why, how, for whom?," *CGAP Focus note,* vol. 48, 2008.

[38] S. Bojjagani and V. N. Sastry, "A secure end-to-end SMS-based mobile banking protocol," *International journal of communication systems,* vol. 30, no. 15, p. e3320, 2017.

[39] K. Krishna Prakasha and B. Muniyal, "Security issues and challenges in mobile Computing and m-commerce," *International Journal of Computer Science & Engineering Survey,* vol. 6, no. 2, pp. 29-45, 2015.

[40] J. Kang and D. Nyang, "A privacy-preserving mobile payment system for mass transit," *IEEE Transactions on Intelligent Transportation Systems* , vol. 18, no. 8, pp. 2192-2205, 2017.

[41] L. Cocco, A. Pinna and M. e Marchesi, "Banking on blockchain: Costs savings thanks to the blockchain technology," *Future internet,* vol. 9, no. 3, pp. 25-44, 2017.

[42] S. Wang, L. Ouyang, Y. Yong, X. Ni, X. `Han and F.-Y. Wang, "Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 49, no. 11, pp. 2266 - 2277, 2019.

[43] J. Brown-Cohen, A. Narayanan, A. Psomas and S. M. Weinberg, "Formal Barriers to Longest-Chain Proof-of-Stake Protocols," in *Proceedings of the 2019 ACM Conference on Economics and Computation*, New York, USA, 2019.

[44] F. Armknecht, G. O. Karame, A. Mandal, F. Youssef et E. Zenner, «Ripple: Overview and Outlook,» chez *Trust and Trustworthy Computing*, Heraklion, Greece, 2015.

[45] J. Liu, W. Li, G. O. Karame and N. Asokan, "Scalable Byzantine Consensus via Hardware-Assisted Secret Sharing," *IEEE Transactions on Computers,* vol. 68, no. 1, pp. 139-150, 2019.

[46] M. Eischer and T. Distler, "Scalable Byzantine fault-tolerant state-machine replication on heterogeneous servers," *Computing,* vol. 101, no. 2, pp. 97-118, 2019.

[47] L. Zhong, Q. Wu, J. Xie, J. Li and B. Qin, "A secure versatile light payment system based on blockchain," *Future Generation Computer Systems,* vol. 93, pp. 327-337, 2019.

[48] F. Schroff, D. Kalenichenko and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in , 2015, pp. 815-823," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.

[49] M. Seo, J. H. Park, Y. Kim, S. Cho, D. H. Lee, Hwang and J. Yeon, "Construction of a new biometric-based key derivation function and its application," *Security and Communication Networks,* vol. 2018, pp. 1-14, 2018.

[50] A. Anees and Y. P. P. Chen, "Discriminative Binary Feature Learning and Quantization in Biometric Key Generation," *Pattern Recognition,* vol. 77, pp. 2890-305, 2018.

[51] W. Sheng, S. Chen, G. Xiao, J. Mao and Y. Zheng, "A Biometric Key Generation Method Based on Semi-supervised Data Clustering," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 45, no. 9, pp. 1205-1217, 2015.

[52] L. D. Tsobdjou, S. Pierre and Q. Alejandro, "A New Mutual Authentication and Key Agreement Protocol for Mobile Client—Server Environment," *IEEE Transactions on Network and Service Management,* vol. 18, no. 2, p. 1275–1286, 2021.

[53] X. Wang, X. Xu, L. Feagan, S. Huang, L. Jiao and W. Zhao, "Inter-Bank Payment System on Enterprise Blockchain Platform," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*, San Francisco, CA, 2018.

[54] S. I. Kim and S. H. Kim, "E-commerce payment model using blockchain," *Journal of Ambient Intelligence and Humanized Computing,* vol. 13, no. 3, pp. 1673-1685, 2022.

[55] P. Lown, "The Ultimate Guide to Mobile Payments," Pay Vision, 2020.

[56] K. J. Sprake, G. Oosthuizen et N. Jones, «The Current Landscape and Emerging Frauds Trends in Commerce,» chez *Fraud Prevention in Ecommerce Report 2020 / 2021*, The PAYPERS, 2020, pp. 8-14.

[57] S. Pierre et O. Italis, «Les systèmes de paiement mobile à l'ère de la Covid-19: sécurité, vie privée et confiance numérique,» *ISTE OpenScience: Technologie et Innovation,* vol. 6, Janvier 2021.

[58] B. Zlatko, «The future of the mobile payment as electronic payment system,» *European Journal of Business and Management,* vol. 8, n° %18, pp. 127-132, 2016.

[59] K.-K. R. Choo, Z. Yan and W. Meng, "Editorial: Blockchain in Industrial IoT Applications: Security and Privacy Advances, Challenges, and Opportunities," *IEEE Transactions on Industrial Informatics,* vol. 16, no. 6, pp. 4119-4121, 2020.

[60] L. Zhang, Y. Xie, Y. Zheng, W. Xue, X. Zheng and X. Xu, "The challenges and countermeasures of blockchain infinance and economics," *Systems Research and Behavioral Science,* vol. 37, no. 4, pp. 691-698, 2020.

[61] L. Xu, ChenLin, Z. Gao, L. Carranco, X. Fan, N. Shah, DialloNour and W. Shi, "Supporting Blockchain-Based Cryptocurrency Mobile Payment With Smart Devices," *IEEE Consumer Electronics Magazine,* vol. 9, no. 2, pp. 26-33, 2020.

[62] L. Zhong, Q. Wu, J. Xie, J. Li et B. Qin, «A secure versatile light payment system based on blockchain,» *Future Generation Computer Systems,* vol. 93, pp. 327-337, 2019.

[63] X. Xu, C. Pautasso, L. Zhu, V. Gramoli, A. Ponomarev, A. B. Tran and S. Chen, "The Blockchain as a Software Connector," in *2016 13th Working IEEE/IFIP Conference on Software Architecture (WICSA)*, Venice, Italy, 2016.

[64] Vodafone Group, "Consumer products and services," [Online]. Available: https://www.vodafone.com/about-vodafone/what-we-do/consumer-products-and-services. [Accessed April 2022].

[65] G. Wood, "Ethereum: A secure decentralized generalized transaction ledger," in *Ethereum project yellow paper, vol. 151,no 2014*, 2021, pp. 1-32.

[66] N. Lagaillardie, M. A. Djari and Ö. Gürcan, "A computational study on fairness of the tendermint blockchain protocol," *Information,* vol. 10, no. 12, pp. 378-391, 2019.

[67] D. Schwartz, N. Youngs and A. Britto, "The ripple protocol consensus algorithm," 2014.

[68] T. T. A. Tien Tuan Dinh, J. Wang, G. Chen, R. Liu, B. C. Ooi et K.-L. Tan, «Blockbench: a framework for analyzing private blockchains,» chez *Proceedings of the 2017 ACM International Conference on Management and Data*, Chicago, IL, 2017.

[69] D. Ron et A. Shamir, «Quantitative analysis of the full bitcoin transaction graph,» chez *Financial Cryptography and Data Security: 17th International Conference, FC 2013*, Okinawa, Japan, 2013.

[70] F. Reid and M. Harrigan, "An Analysis of Anonymity in the Bitcoin System," in *Security and Privacy in Social Networks*, New York, NY., Springer, 2013, pp. 197-223.

[71] N. Andola, V. K. Yadav, S. Venkatesan and S. Verma, "Anonymity on blockchain based e-cash protocols—A survey," *Computer Science Review,* vol. 40, pp. 100394-100411, 2021.

[72] S. Saxena, S. Vyas, B. S. Kumar and S. Gupta, "Survey on Online Electronic Paymentss Security," in *2019 Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, 2019.

[73] O. Ogbanufe and D. J. Kim, "Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment," *Decision Support Systems,* vol. 106, pp. 1-14, 2018.

[74] F. Dai, Y. Shi, N. Meng, L. Wei and Z. Ye, "From Bitcoin to cybersecurity: A comparative study of blockchain application and security issues," in *2017 4th International Conference on Systems and Informatics (ICSAI)*, Hangzhou, China, 2017.

[75] Y. Dodis, R. Ostrovsky, L. Reyzin and A. Smith, "Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data," *SIAM journal on computing,* vol. 38, no. 1, pp. 97-139, 2008.

[76] G. T. Becker, "Robust fuzzy extractors and helper data manipulation attacks revisited: Theory versus practice," *IEEE Transactions on Dependable and Secure Computing,* vol. 16, no. 5, pp. 783-795, 2017.

[77] K. Moriarty, B. Kaliski and A. Rusch, *Pkcs 5: Password-Based Cryptography Specification Version 2.1,* RFC Editor, 2017.

[78] K. Simoens, P. Tuyls and B. Preneel, "Privacy weaknesses in biometric sketches," in *2009 30th IEEE Symposium on Security and Privacy*, Oakland, CA, 2009.

[79] M. Osadchy and O. Dunkelman, "It is All in the System's Parameters: Privacy and Security Issues in Transforming Biometric Raw Data into Binary Strings," *IEEE Transactions on Dependable and Secure Computing,* vol. 16, no. 5, pp. 796-804, 2018.

[80] P. N. Smart, "Hash Functions, Message Authentication Codes, and Key Derivation Functions," in *Cryptography made simple*, Berlin, Springer, 2016, pp. 271-294.

[81] E. Barker, L. Chen and R. Davis, *Recommendation for Key-Derivation Methods in Key-Establishment Schemes*, NIST Special Publication, 2018.

[82] J. M. Turner, "Recommendation for Key-Derivation Methods in Key-Establishment Schemes, SP 800-56C Rev. 1," *Federal Information Processing Standards Publication,* vol. 198, no. 1, pp. 1-13, 2008.

[83] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory,* vol. 28, no. 2, pp. 129-137, 1982.

[84] D. Arthur and S. Vassilvitskii, "K-means++ the advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. 2007*, 2007.

[85] G. F. Jenks, "Optimal data classification for choropleth maps," Department of Geographiy, University of Kansas Occasional Paper, Lawrence (Kansas), 1977.

[86] H. Kang, Y. Hori, T. Katashita and M. Hagiwara, "The implementation of fuzzy extractor is not hard to do: An approach using puf data," in *Proceedings of the 30th Symposium on Cryptography and Information Security*, Kyoto, Japan, 2013.

[87] V. Guruswami, A. Rudra and M. Sudan, "The Fundamental Question," in *Essential Coding Theory*, CRC Press, 2012, pp. 19-41.

[88] W. C. Huffman and V. Pless, "BCH and Reed-Solomon codes," in *Fundamentals of Error-Correcting Codes*, Cambridge, United Kingdom, Cambridge university press, 2010, p. 168–207.

[89] Y. Dodis, B. Kanukurthi, J. Katz, L. Reyzin and A. Smith, "Robust Fuzzy Extractors and Authenticated Key Agreement From Close Secrets," *IEEE Transactions on Information Theory,* vol. 58, no. 9, pp. 6207-6222, 2012.

[90] X. Boyen, Y. Dodis, J. Katz, R. Ostrovsky and A. Smith, "Secure Remote Authentication Using Biometric Data," in *Annual international conference on the theory and applications of cryptographic techniques*, Berlin, 2005.

[91] C. Herder, L. Ren, M. van Dijk, M.-D. Yu and S. Devadas, "Trapdoor computational fuzzy extractors and stateless cryptographically-secure physical unclonable functions," *IEEE Transactions on Dependable and Secure Computing,* vol. 14, no. 1, pp. 65-82, 2017.

[92] K. YASUNAGA and K. YUZAWA, "On the Limitations of Computational Fuzzy Extractors," *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences,* Vols. E106-A, no. 3, pp. 350-354, 2023.

[93] Y. C. Feng, P. C. Yuen and A. K. Jain, "A Hybrid Approach for Generating Secure and Discriminating Face Template," *IEEE Transactions on Information Forensics and Security,* vol. 5, no. 1, pp. 103-117, 2010.

[94] A. B. J. Teoh, A. Goh and D. C. L. Ngo, "Random Multispace Quantization as an Analytic Mechanism for BioHashing of Biometric and Random Identity Inputs," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 28, no. 12, pp. 1892-1901, 2006.

[95] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk and T. Toft, "Privacy-preserving face recognition," in *International symposium on privacy enhancing technologies symposium*, Berlin, 2009.

[96] A. R. Sadeghi and T. W. I. Schneider, "Efficient privacypreserving face recognition," in *International Conference on Information Security and Cryptology*, Berlin, 2009.

[97] S. N. Sivanandam, S. Sumathi and S. N. Deepa, "Defuzzification," in *Introduction to fuzzy logic using MATLAB*, Berlin, Springer, 2007, pp. 95-112.

[98] J. T. Starczewski, "Uncertainty in Fuzzy Sets," in *Advanced Concepts in Fuzzy Logic and Systems with Membership Uncertainty*, Berlin, Springer, 2013, pp. 1-31.

[99] S. Keykhaie and S. Pierre, "Mobile Match on Card Active Authentication Using Touchscreen Biometric," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 4, pp. 376-385, 2020.

[100] K. Lee, J. Ho and D. Kriegman, "Acquiring Linear Subspaces for Face Recognition under Variable Lighting," *IEEE Trans. Pattern Anal. Mach. Intelligence,* vol. 27, no. 5, pp. 684-698, 2005.

[101] L. Wolf, T. Hassner and I. Maoz, "Face recognition in unconstrained videos with matched background similarity," CVPR 2011, Colorado Springs, CO, USA, 2011, pp. 529-534, doi: 10.1109/CVPR.2011.5995566.," in *CVPR 2011*, Colorado Springs (Colorado), 2011.

[102] The MathWorks, Inc., "Matlab for Artificial Intelligence," The MathWorks, Inc., 2023. [Online]. [Accessed 2023].

[103] H. Krawczyk and P. Eronen, *HMAC-based extract-and-expand key derivation function (HKDF) RFC 5869,* Internet Engineering Task Force (IETF), 2010.

[104] M. Abavisani and V. M. Patel, "Deep Multimodal Subspace Clustering Networks," *IEEE Journal of Selected Topics in Signal Processing,* vol. 12, no. 6, pp. 1601-1614, 2018.

[105] S. Chakraborty, S. K. Singh and K. Kumar, "Facial Biometric System for Recognition Using Extended LGHP Algorithm on Raspberry Pi," *IEEE Sensors Journal,* vol. 20, no. 14, pp. 8117-8127, 2020.

[106] L. Derek, V. René and H. Benjamin D., "Doubly stochastic subspace clustering," *arXiv preprint arXiv:2011.14859,* 2020.

[107] O. Italis and S. Q. A. Pierre, "Mobile Payment Platforms: Taxonomy, Architecture, Security, and Integration with Blockchain," 2022.

[108] S. Wan, M. Li, G. Liu and C. Wang, "Recent advances in consensus protocols for blockchain: a survey," *Wireless Networks,* vol. 26, no. 8, pp. 5579-5593, 2020.

[109] Y. Hu, A. Manzoor, P. Ekparinya, M. Liyanage, K. Thilakarathna, G. Jourjon and A. Seneviratne, "A delay-tolerant payment scheme based on the Ethereum blockchain," *IEEE Access,* vol. 7, pp. 33159-33172, 2019.

[110] Q. Meng, S. Hou, Z. Li et S. Lu, «A uniform payment system for Hyperledger Fabric blockchain,» chez *CCF China Blockchain Conference*, Singapore, 2019.

[111] D. M. Konidala, M. H. Dwijaksara, K. Kim, D. Lee, B. Lee, D. Kim and S. Kim, "Resuscitating privacy-preserving mobile payment with customer in complete control," *Pers Ubiquit Comput* , vol. 16, pp. 643-654, 2012.

[112] J. Ni, M. H. Au, W. Wu, X. Luo, X. Lin and X. S. Shen, "Dual-anonymous off-line electronic cash for mobile payment," *IEEE Transactions on Mobile Computing,* vol. 22, no. 6, pp. 3303-3317, 2021.

[113] C. Lin, D. He, X. Huang, M. K. Khan and K.-K. R. Choo, "DCAP: A Secure and Efficient Decentralized Conditional Anonymous Payment System Based on Blockchain," *IEEE Transactions on Information Forensics and Security,* vol. 15, pp. 2440-2452, 2020.

[114] T. Guggenberger, V. Schlatt, J. Schmid and N. Urbach, "A structured overview of attacks on blockchain systems," in *PACIS 2021 Proceedings*, Dubai, 2021.

[115] H. Chen, M. Pendleton, L. Njilla and S. Xu, "A Survey on Ethereum Systems Security: Vulnerabilities, Attacks, and Defenses," *Association for Computing Machinery,* vol. 53, no. 3, 2020.

[116] B. Warner, "Pure-Python ECDSA and ECDH," Python Sofware Foundation. [Online]. [Accessed 28 July 2022].

[117] T. T. A. Dinh, J. Wang, G. Chen, R. Liu, B. C. Ooi and K.-L. Tan, "Blockbench: a framework for analyzing private blockchains," in *Proceedings of the 2017 ACM International Conference on Management of Data*, Chicagi, IL, 2017.

[118] V. Cook, Z. Painter, C. Peterson and D. Damian, "Read-uncommitted transactions for smart contract performance, pp. 1960-1970, 2019.," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, Dallas, TX, 2019.

[119] T. Wang, C. Zhao, Q. Yang, S. Zhang and S. C. Liew, "Ethna: analyzing the underlying peer-to-peer network of Ethereum blockchain," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 3, pp. 2131-2146, 2021.

[120] O. Italis, S. Pierre and A. Quintero, "Blockchain and Mobile Payment: Assessment on Privacy and Usability and a Scheme for Enhancement," in *2022 Fourth International Conference on Blockchain Computing and Applications (BCCA)*, San Antonio, TX, 2022.

[121] Q. Deng, "Application analysis on blockchain technology in cross-border payment," in *5th International Conference on Financial Innovation and Economic Development (ICFIED 2020)*, 2020.

[122] C. o. P. a. M. Infrastructures, "Correspondent Banking," Bank for International Settlements, 2016.

[123] R. Lewis, J. W. McPartland and R. Ranjan, "Blockchain and financial market innovation," *Economic Perspectives*, vol. 41, no. 7, p. 1 – 17, 2017.

[124] O. Pal, B. Alam, V. Thakur and S. Singh, "Key management for blockchain technology," *ICT Express*, vol. 7, no. 1, pp. 76-80, 2021.

[125] V. Arasev, "POA Network Whitepaper," 28 September 2018. [Online]. Available: https://github. com/poanetwork/wiki/wiki/POA-Network-Whitepaper. [Accessed 2023].

[126] K. Olson, M. Bowman, J. Mitchell, S. Amundson, D. Middleton and C. Montgomery, "Sawtooth: An Introduction," The Linux Foundation, New York, NY, 2018.

[127] D. Mazieres, "The stellar consensus protocol: A federated model for internet-level consensus," *Stellar Development Foundation*, vol. 32, pp. 1-45, 2015.

[128] M. Lokhava, G. Losa, D. Mazières, G. Hoare, N. Barry, E. Gafni, J. Jove, R. Malinowsky and J. McCaleb, "Fast and Secure Global Payments with Stellar," in *Association for Computing Machinery*, Huntsville, Ontario, 2019.

[129] D. Schawrtz, N. Youngs and A. Britto, "The Ripple Protocol Consensus Algorithm," Ripple Labs Inc. , 2014.

[130] O. Italis, S. Pierre and A. Quintero, "Privacy-Preserving Model for Biometric-based Authentication and Key Derivation Function," *Journal of Information Security and Applications.*

[131] M. V. Dalsen, S. Cohen, A. Maritz, G. Oosthuizen, R. Capps and S. Armstrong, "New Directions in Authentication," Entersekt Ideas Lab, 2022.

[132] D. Tirfe and V. K. Anand, "A Survey on Trends of Two-Factor Authentication," Singapore, 2022.

[133] A. Kania, "Payment Methods Report 2022 : Latest Trends in Payment Preferences," THE PAYPERS, 2022.

[134] S. N. S. King, "Peercoin Whitepaper," 2016.

[135] A. M. Antonopoulos, "Keys, Addresses, Wallets," in *Mastering Bitcoin: unlocking digital cryptocurrencies*, O'Reilly Media, Inc., 2014, p. 61–106.

[136] A. Singh, G. Kumar, R. aha, a. Alazab, M. Conti and R. Thomas, " A survey and taxonomy of consensus protocols for blockchains," *Journal of Systems Architecture,* vol. 127, 2022.

[137] M. Baudet, A. Ching, A. Chursin, G. Danezis, F. Garillot and Z. Li, "State Machine Replication in the Libra Blockchain," 2019.

[138] A. Di Luzio, A. Mei and J. Stefa, "Consensus Robustness and Transaction De-Anonymization in the Ripple Currency Exchange System," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, Atlanta, GA, 2017.

[139] B. Chase and E. MacBrough, "Analysis of the XRP ledger consensus protocol," arXiv preprint arXiv:1802.07242, 2018.

[140] H. Ning, F. Farha, A. Ullah and L. Mao, "Physical unclonable function: architectures, applications and challenges for dependable security," *IET Circuits, Devices & Systems*, vol. 14, no. 4, pp. 407-424, 2020.

[141] C. Rodrigo, S. Pierre, R. Beaubrun and F. El Khoury, "Brain- Shield: A Hybrid Machine Learning-Based Malware Detection Model for Android Devices," *Electronic*, vol. 10, no. 20, 2021.

[142] K. Sayood, "Scalar Quantization," in *Introduction to Data Compression*, Burlington (Massachusetts), Morgan Kaufmann, 2006, pp. 227-271.

[143] T. Lerner, "Strategies," in *Mobile Payment*, Springer, 2013, pp. 23-37.

[144] F. Gai, B. Wang, W. Deng and W. Peng, "Proof of reputation: A reputation-based consensus protocol for peer-to-peer network," in *Database Systems for Advanced Applications: 23rd International Conference, DASFAA 2018*, Gold Coast, QLD, 2018.