



HAUTE ÉCOLE
D'INGÉNIERIE ET DE GESTION
DU CANTON DE VAUD
www.heig-vd.ch

Département TIC
Filière Télécommunications
Orientation Sécurité de l'information

Travail de Bachelor

Chiffrement / Signature d'Emails

Étudiant

Enseignant responsable

Année académique

Mickael Bonjour

Prof. Alexandre Duc

2019-2020

Yverdon-les-Bains, le 28 juillet 2020

Département TIC
Filière Télécommunications
Orientation Sécurité de l'information
Étudiant Mickael Bonjour
Enseignant responsable Prof. Alexandre Duc

Travail de Bachelor 2019-2020
Chiffrement/Signature d'Emails

Résumé publiable

Dans ce travail de bachelor je vais analyser les besoins classiques d'un système de messagerie électronique. De plus, une analyse des solutions actuelles de messagerie sécurisée est faite afin d'identifier les propriétés cryptographiques mise en avant dans de tels systèmes. De ces propriétés je détermine une primitive cryptographique adéquate au problème et qui permettrait d'avoir les propriétés cryptographiques mentionnées auparavant. Ensuite j'établirais un *Proof Of Concept* amenant la technologie choisie dans un cadre de messagerie électronique sécurisée. Ce Proof of Concept se base sur la Certificateless Public Key Cryptography qui a été la primitive choisie grâce à ses propriétés intéressantes dans un contexte de mail sécurisés. J'évalue ma solution par rapport à d'autres solutions présentes sur le marché au niveau du temps de chiffrement/déchiffrement, de l'overhead introduit et des propriétés cryptographiques présentes.

Étudiant :	Date et lieu :	Signature :
Mickael Bonjour
Enseignant responsable :	Date et lieu :	Signature :
Prof. Alexandre Duc

Préambule

Ce travail de Bachelor (ci-après TB) est réalisé en fin de cursus d'études, en vue de l'obtention du titre de Bachelor of Science HES-SO en Ingénierie.

En tant que travail académique, son contenu, sans préjuger de sa valeur, n'engage ni la responsabilité de l'auteur, ni celles du jury du travail de Bachelor et de l'Ecole.

Toute utilisation, même partielle, de ce TB doit être faite dans le respect du droit d'auteur.

HEIG-VD

Vincent Peiris
Chef de département TIC

Yverdon-les-Bains, le 28 juillet 2020

PRÉAMBULE _____

vi _____

Authentification

Le soussigné, Mickael Bonjour, atteste par la présente avoir réalisé ce travail et n'avoir utilisé aucune autre source que celles expressément mentionnées.

Yverdon-les-bains, le 28 juillet 2020

Mickael Bonjour

AUTHENTICATION _____

Cahier des charges

Résumé du problème

Les outils de chiffrement et de signature d'email actuels se résument principalement à S/MIME et à PGP.

Ces deux solutions sont anciennes, souffrent assez régulièrement de nouvelles vulnérabilités et ne proposent pas certaines propriétés cryptographiques qui pourraient être utiles (par exemple, la "forward secrecy". Le but de ce travail de bachelor est d'étudier quelles propriétés seraient utiles pour la sécurité des emails, de proposer un nouveau protocole les implémentant et de développer un proof of concept.

Problématique

Les systèmes de mails sécurisés souffrent de manque de praticité quant à leur implémentations, de plus elles ont été prouvées vulnérables à plusieurs reprises.

Solutions existantes

Les solutions existantes sont représentées majoritairement par S/MIME et PGP. Cependant des nouveaux protocoles émergent tel que PEP et des fournisseurs proposent des implémentations transparentes de PGP p. ex. comme le fait Protonmail. de plus, l'on pourrait s'orienter aussi sur la messagerie instantanée qui bénéficie de protocoles sécurisés comme Signal.

Solutions possibles

Un début de solution est proposée dans ce papier à l'aide d'un nouveau système qui pourrait être mis facilement en place et qui bénéficierait de meilleures propriétés que les protocoles actuellement utilisés. L'autre solution serait de rester avec PGP et S/MIME malgré le manque d'intégration dont ils font preuves.

Cahier des charges

Voici un résumé du cahier des charges sous formes d'une liste d'objectifs à atteindre :

- Analyser les besoins d'un système d'E-mails actuel.
- Analyser et étudier les solutions de sécurité existantes.
- Comprendre et évaluer les propriétés cryptographiques défendues.
- Établir une liste des propriétés cryptographiques voulues pour un système de mails sécurisés.
- Trouver une primitive cryptographique satisfaisant les besoins énoncés et l'étudier pour en comprendre les bases et les besoins nécessaires en termes de sécurité.
- Établir la spécification pour un nouveau protocole en utilisant la primitive choisie.
- Faire un Proof Of Concept du protocole proposé.

Si le temps le permet :

- Comprendre plus en détails les mathématiques derrière la primitive utilisée.
- Faire un prototype de client mail utilisant une architecture mise en place pour le POC.

Déroulement

Tout d'abord je vais m'intéresser à faire une évaluation des concepts existants en messagerie sécurisée, tel que PGP et S/MIME pour les emails ou encore Signal pour la messagerie instantanée. Ayant vu ce qu'il se fait j'essaie de trouver une solution alternative pour le chiffrement et la signature d'emails. De là je vais conceptualiser un protocole et l'implémenter au sein d'un *Proof Of Concept*.

Livrables

Les livrables seront les suivants :

1. Une documentation contenant :
 - Une analyse de l'état de l'art
 - La décision qui découle de l'analyse
 - Spécifications
 - L'implémentation faites et les choix faits
 - Proof Of Concept
 - Les problèmes connus
2. Le code du *Proof Of Concept* fait, expliqué à l'aide de commentaires.

Table des matières

Préambule	v
Authentification	vii
Cahier des charges	ix
1 Introduction	1
2 Analyse - État de l'art	3
2.1 Besoins d'un système de messagerie	3
2.2 Protocoles existants	4
2.3 Implémentations existantes	10
2.4 Attaques existantes	13
2.5 Signal	13
2.6 Compromis	15
2.7 Primitives	16
2.8 Recherches sur la primitive	17
2.9 État de l'art	21
3 Architecture / Design du protocole	25
3.1 Architecture globale	25
3.2 Acteurs	26
3.3 Fonctionnement Certificateless PKC	26

3.4	Design du protocole	27
4	Implémentation	31
4.1	Choix d'implémentations	31
4.2	Implémentation clés de chiffrement	33
4.3	Fonctionnement global POC (KGC)	33
4.4	Fonctionnement global POC (Client)	36
4.5	Comparaisons avec état de l'art	37
5	Conclusion	39
	Bibliographie	41
A	Outils utilisés pour la compilation	49
A.1	RELIC Toolkit	49
A.2	Libsodium	49
A.3	binn	50
A.4	libetpan	50
A.5	libcurl	50
B	Fichiers	51
B.1	Code du POC	51
B.2	Tableaux comparatifs	51

Chapitre 1

Introduction

Ce travail de Bachelor a pour but de sensibiliser à la vulnérabilité dans les systèmes actuels de messagerie électronique. Il propose aussi un nouveau protocole permettant de sécuriser ce type de messagerie à l'aide d'une primitive cryptographique peu répandue, le *Certificateless Public Key Cryptography*. Ma démarche dans ce travail de bachelor est de voir si des solutions s'offrent à nous en considérons ce qui se fait sur le marché actuellement. En essayant d'améliorer les solutions actuelles proposées qui peuvent souffrir d'un manque de sécurité assez souvent ou un manque de simplicité d'utilisation.

Ce travail est découpé en plusieurs parties. En effet, on commence par une analyse de l'état de l'art, ce qui existe et analyser pourquoi il faudrait trouver de nouvelles solutions. Puis une présentation de la primitive cryptographique utilisée pour ma proposition dans ce travail ainsi de la raison qui amène à ce choix. Enfin, la présentation de l'architecture de mon protocole et une implémentation proposée en *Proof Of Concept* ainsi que les choix importants qui ont été faits en rapport à cette implémentation. En commentant évidemment les problèmes connus et les améliorations qui seraient possible et envisageable si l'on voulait faire de ce *Proof Of Concept* une réalité.

Chapitre 2

Analyse - État de l'art

Dans ce chapitre je vais m'intéresser aux différentes propositions d'implémentations de système de messagerie sécurisée afin de voir où en est l'état de l'art. Pour cela j'ai recherché les systèmes les plus connus tel que PGP et S/MIME mais aussi les implémentations de ces protocoles dans des clients mails tel que Protonmail ou Tutanota. J'élargis l'analyse à des protocoles plus orientés vers la messagerie instantanée comme Signal.

2.1 Besoins d'un système de messagerie

Cette section permettra de lister les besoins principaux d'un système de messagerie et l'utilisation qui en est faite habituellement.

2.1.1 Besoins principaux

Dans un système de mail les besoins principaux sont surtout de pouvoir consulter sa boîte mail à tout moment avec les anciens et nouveaux mails reçus. De plus il est préférable de pouvoir envoyer des mails aussi. Ces envois peuvent avoir plusieurs propriétés et fonctionnalités. L'on pourrait envoyer un mail à de multiples destinataires et même sans que les uns et les autres sachent exactement à qui est envoyé le mail exactement (Copie cachées). Un client mail permet aussi d'envoyer des pièces jointes, celles-ci seront encodées au sein du message et envoyée avec. Un mail a aussi un état pour savoir s'il a déjà été lu ou non. Parmi l'utilisation simple d'un système de messagerie il y aussi le fait que l'on veut pouvoir voir ses mails de n'importe quel appareil à n'importe quel moment.

2.1.2 Détails techniques

Afin d'établir les futures notations utilisées ci-après et montrer le fonctionnement global d'un système de messagerie électronique je vais utiliser la figure 2.1. Dans cette figure l'on peut voir que 3 protocoles différents sont utilisés pour la gestion des emails ; SMTP, IMAP et POP3. Ces 3 protocoles sont utilisés par différents acteurs, le MUA (Mail User Agent), le MTA (Mail Transfer Agent) et le MDA (Mail Delivery Agent). Le MUA est en fait un client mail qui va s'occuper d'envoyer des mails ou de les recevoir (rechercher sur le MDA). Les MTAs sont les serveurs mails responsables du bon acheminement des mails. Ainsi les 3 protocoles énoncés plus hauts sont soit pour l'envoi et la transmission (SMTP) soit pour la récupération des messages (IMAP et POP3). Lors de l'envoi un MUA va simplement renseigner les destinataires du message ainsi que sa source, son sujet et son message puis le serveur va transmettre ces informations au MTA du domaine de destination qui s'occupera de le transmettre au MDA (souvent les 2 à la fois), celui-ci stocke les mails en attendant qu'un MUA fasse une demande via POP3 ou IMAP. IMAP est souvent préféré car les mails restent ainsi sur le serveur mail et est donc consultable depuis un autre appareil utilisant aussi IMAP. POP3 va plutôt télécharger les mails et les enlever du serveur et ils ne seront donc plus disponibles par le biais d'un autre appareil. Les MTAs sont des serveurs de transmission de données, transmises en clair jusqu'à l'introduction d'ESMTP et de la directive STARTTLS qui permet un niveau basique de sécurité entre 2 MTAs pour le transfert de mails. Cependant, si un MTA est mal configuré et ne permet pas cette directive le mail transitera en clair. C'est dans ce contexte là et celui du stockage des mails en clair par le MDA que des solutions de chiffrement de mail en E2E (End to End - Chiffrement de bout-en-bout) ont vu le jour.

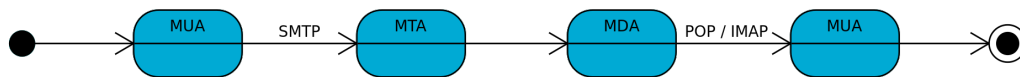


FIGURE 2.1 – Le fonctionnement d'un système de mail [6]

2.2 Protocoles existants

Dans cette section je vais analyser les différents protocoles existants afin de sécuriser la messagerie électronique, ainsi que leur implémentation au sein de certains clients mails. De plus je m'intéresserai à la messagerie instantanée afin de voir s'il est possible d'implémenter cela dans un système de mails.

2.2.1 PGP

Fonctionnement. PGP (Pretty Good Privacy ou Assez bon niveau de confidentialité) est un moyen de chiffrer des données (mails, fichiers, ...). C'est une méthode de chiffrement hybride (utilise le chiffrement symétrique et asymétrique) qui fonctionne comme montré sur la Figure 2.2. Comme on peut le voir, on tire une clé symétrique aléatoirement qui permettra de chiffrer notre mail avec un chiffrement symétrique comme AES. Ensuite, l'on va chiffrer cette clé symétrique à l'aide d'un chiffrement asymétrique, en utilisant la clé publique du destinataire.

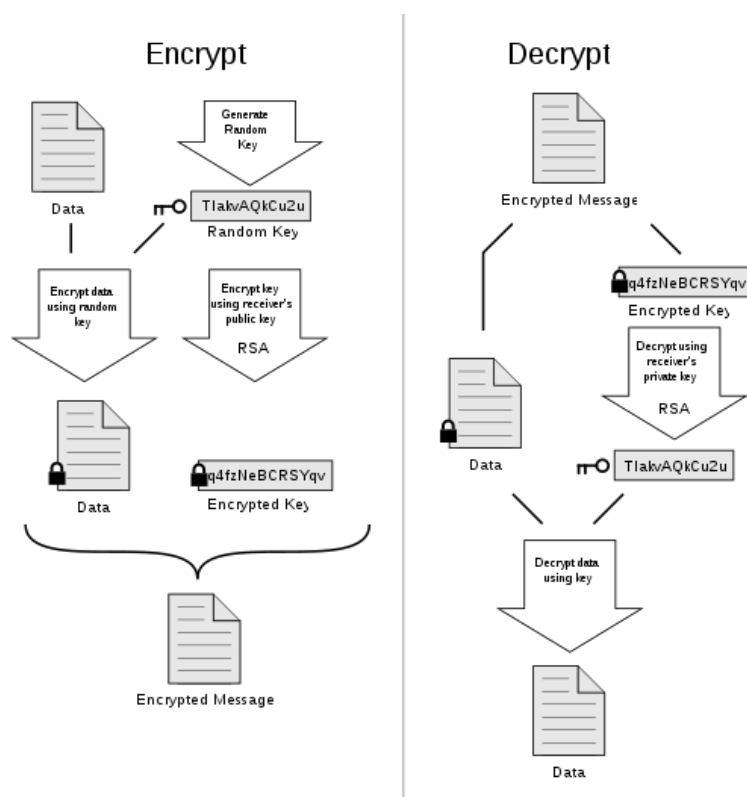


FIGURE 2.2 – Le fonctionnement global de PGP

Ce fonctionnement hybride est expliqué par la lenteur d'un chiffrement asymétrique sur un certain nombre de données. Ainsi en chiffrant uniquement la clé symétrique qui a servi à chiffrer le message le déchiffrement est bien plus rapide et simple et est effectué typiquement avec un chiffrement symétrique tel qu'AES qui a le droit à des instructions dédiées dans certains processeurs. Contrairement à des chiffrements asymétriques qui sont plus contraignants. Mais il est nécessaire de passer par cette phase asymétrique, on a en effet besoin d'un secret partagé dès le début de la communication si cette méthode n'est pas utilisée.

Pour ce qui est des primitives cryptographiques proposées dans la RFC4880 [5], elle sont listées ci-après dans la table 2.1. IDEA, TripleDES, CAST5, Blowfish, AES-128, AES-192,

Symetric	Asymetric	Hash	Compression
IDEA	RSA	MD5	ZIP

TABLE 2.1 – Table des algorithmes utilisés par PGP

AES-256, Twofish-256. Pour les primitives asymétrique : RSA, ElGamal, DSA, ECDSA, Diffie-Hellman. La RFC spécifie aussi les formats de compression : ZIP, ZLIB, BZip2. Puis enfin, les algorithmes de hachage : MD5, SHA-1, RIPE-MD160, SHA256, SHA384, SHA512, SHA224 mais MD5 a été annoncé déprécié. Puis il faut savoir que pour chacune de ces catégories il y a 10 éléments réservés pour des primitives privées/expérimentales.

Lorsqu'un mail est chiffré et signé avec PGP, il est d'abord hashé puis ce hash est signé avec la clé privée de l'utilisateur afin de faire une signature digitale. Le message et la signature sera alors chiffrée à l'aide la clé symétrique.

L'organisation d'un message PGP se fait via des "paquets" d'informations encodés en base64. La RFC définit bien ces types de paquets, leur fonctionnement et les différents codes associés. Sur le site <https://cirw.in/gpg-decoder/> l'on peut entrer un message et ainsi voir l'organisation d'un message, de clés publiques et de clés privées. Ainsi, je montre un exemple de mail envoyé à plusieurs destinataires dans la figure 2.3. Cela démontre comment fonctionnent PGP, en effet le message étant chiffré avec une clé symétrique, le chiffré sera le même pour tout le monde. Mais afin que tout les destinataires puissent avoir la clé symétrique, la clé est chiffrée à l'aide des clés publiques des différents destinataires (et de la source, pour pouvoir la déchiffrer à l'avenir et ne pas conserver le mail en clair dans la boîte d'envoi). À noter qu'avec l'option de *blind copy* (option permettant d'envoyer à un utilisateur un mail sans qu'il sache qu'il a aussi été envoyé à un autre utilisateur), PGP *leak* les receveurs des mails avec leur KeyID qui sera présent dans le message PGP¹.

PGP utilise donc un système de clés publiques afin d'envoyer des clés symétriques. Comment obtient-on une clé publique pour envoyer notre message ?

Selon certaines utilisations les clés peuvent être obtenues via un serveur de clés, mais cela implique un parti tiers auquel il faut faire confiance, ainsi ce qui peut être fait c'est aussi d'avoir son propre serveur de clés. Mais pour être sûr que tel clé appartienne bien à tel utilisateur, PGP a introduit dès ces débuts un système décentralisé de confiance appelé le "Web of Trust". Ainsi, un utilisateur pourra faire confiance à certaines clés d'autres utilisateurs et les signer, puis chacun des utilisateurs aura des clés de confiance et ainsi de suite. Cela permet d'avoir une toile de confiance entre les utilisateurs. Cependant, ce système est difficile à utiliser, il est nécessaire de faire attention à quelles clés les utilisateurs signent et approuvent. Surtout pour les nouveaux utilisateurs qui ne peuvent faire vérifier leurs clés publiques facilement et ne seront donc pas vérifiées, c'est pour cela que des "fêtes" de signature de clés sont organisées afin de se rencontrer personnellement et vérifier que tel utilisateur

1. <https://crypto.stanford.edu/portia/papers/bb-bcc.pdf>

```

04 44 03 f4 8a 01 b0 9d 52 95 5b 12 01 07 40 4c Public-Key Encrypted Session Key Packet (0x1)
af 00 72 9e 7e ec 1a 1f 07 06 50 c5 68 5b f8 20
77 00 10 25 a4 8c 1f 28 c3 2a 4a 00 00 3b f2 20
70 f2 f4 0b a9 92 5c 2b ea 21 4b b3 22 63 52 0e
10 b0 a3 cf d4 04 b7 3c 65 c0 00 00 f9 f2 50 ea
cipherTypeByte: "132"
length: "78"
version: "3"
keyId: "F48A81890529558"
publicKeyAlgorithm: "Reserved for Elliptic Curve (0x12)"

04 40 03 ed 0d 96 a7 50 7b a9 00 12 01 07 40 c5 Public-Key Encrypted Session Key Packet (0x1)
0b 00 90 f1 f7 88 73 f6 43 11 6c 64 07 a7 cb a8
4b 01 11 23 3e a4 00 38 55 20 3f bc d7 30 7c 20
42 01 00 a2 c0 6c bf 99 46 bf 39 a7 26 35 82 a8
9e 07 7f d0 d7 31 91 7e bd 23 a6 79 a2 00 08 1b
cipherTypeByte: "132"
length: "78"
version: "3"
keyId: "E0D096E75F8E86A"
publicKeyAlgorithm: "Reserved for Elliptic Curve (0x12)"

04 40 03 20 c4 34 9a 6c bd a2 dc 12 01 07 40 ed Public-Key Encrypted Session Key Packet (0x1)
ab 7a ab 09 f6 c6 69 8c f0 a0 07 d0 47 13 6b
4b 43 71 85 9a 90 c9 b2 c0 c0 c0 e7 05 48 20
04 b2 00 9e 61 3d ff 44 04 1a 0e 00 b2 3c f9
04 a5 01 36 3f 43 4a f7 cf 4d a3 db cd 2f f6 2a
cipherTypeByte: "132"
length: "78"
version: "3"
keyId: "20C4548A6C3DA20C"
publicKeyAlgorithm: "Reserved for Elliptic Curve (0x12)"

02 e9 03 af b2 02 ec bd 20 c0 61 b4 e9 9a 41 ea Sym. Encrypted and Integrity Protected Data Packet (0x12)
2e a0 2a f1 10 43 01 b5 c0 8e 93 b0 05 70 2b a8
30 b0 25 c5 a3 a0 05 25 70 a3 af 60 93 2c ca 12
02 90 1a 04 4e 7f 05 04 fc 71 1a aa 0b 20 3c d7
cf 10 6a c0 5f 9a c7 10 26 c2 10 e7 73 b0 0b
01 52 00 02 a1 76 f6 d0 c0 51 a6 0b 54 1a dc c7
70 50 50 49 07 f1 f6 a2 05 7a 3e 10 00 7a
ed 0a 04 7e 19 10 f2 bc c7 f9 73 b0 05 c0 f6 d7
50 0a f6 47 6a 05 13 a0 07 7a 3e 10 00 7a
43 1a dc cd c0 05 13 10 b0 06 02 03 03 7f
2e 73 20 90 00 27 0c 6a 79 aa b2 a3 a2 00 40
2a b0 05 3d c7 0b 2b 3a 9a 00 3a 59 4f 05 f3
01 0b 20 4c 85 14 2c 07 a5 02 20 72 04 a4 0d bc
dc 00 6a 70 82 04 17 0b 3a aa 00 00 20 2d a0
52 c0 7a 3d f4 11 2a 20 a8 af a4 f1 15 0b 03 ff
f0 01 60 0f 70 5c 27 01 a9 75 07 2c 00 55 20
-- -- -- -- -- -- -- -- -- -- -- -- -- -- -- --

```

FIGURE 2.3 – Exemple décodage d'un message PGP

est bien telle personne. Ces événements sont d'ailleurs possible grâce aux Fingerprints des clés publiques, ce sont des chaînes hexadécimales de 20 bytes permettant d'identifier une clé publique plus facilement.

SHA-1 est encore beaucoup utilisé pour les certificats d'identité et il a été prouvé que des attaques sur cet algorithme au niveau des collisions peuvent être faites [13] donc à éviter d'utiliser également. Maintenant modifié dans les nouvelles version de GnuPG, ce n'est plus le standard pour signer les clés publiques d'autres utilisateurs dans le Web Of Trust.

Dans la dernière spécification de PGP des moyens de créer des certificate authorities ont été ajoutés afin d'avoir un système décentralisé de *trust signatures* à différents niveaux et ainsi pouvoir se baser sur un système qui ressemble à un PKI mais avec une flexibilité sur les CAs (utilisateurs) auxquels on fait confiance ou non.

Propriétés cryptographiques. PGP a surtout été créé pour fournir du **chiffrement de bout-en-bout** afin de résoudre les problèmes de transmissions en clair entre les MTAs et le stockage des mails en clair dans les MDAs. Ainsi même lors d'une récupération de mails sur un serveur les mails seraient chiffrés. Ensuite PGP propose de signer ou non ses mails ce qui amène donc de la **répudiation** (si non-signé, le mail ne pourra pas être utilisé pour prouver qu'il a été envoyé par telle personne) et **non-répudiation** (mail signé, ainsi l'on peut prouver que l'expéditeur a bien envoyé le mail).

Le problème qui est souvent reproché à PGP c'est qu'il n'implémentes pas de **Forward Secrecy**. La *Forward Secrecy* permet d'affirmer que si l'on a une brèche à un instant t , et qu'un attaquant récupère notre clé privée, il ne pourra pas déchiffrer les anciens messages chiffrés avant l'instant t .

Utilisation. Lors des tests l'utilisation la plus simple possible a été utilisée pour voir si un utilisateur lambda pouvait arriver à mettre en place ce genre de sécurité. Il s'est avéré que

cela était assez simple au départ, mais dès lors que l'on veut envoyer un mail chiffré à un correspondant cela se complique. Uniquement l'installation d'un Add-On sur le logiciel de messagerie (Thunderbird dans ce cas) s'appelant Enigmail a été nécessaire. Ensuite Enigmail a généré les clés PGP (de manière totalement automatisée). Puis l'envoi d'un mail se fait simplement et rapidement via des icônes et des options dans le client mail. Cependant c'est très opaque et on ne sait pas ce qu'Enigmail fait réellement derrière les décors. L'utilisateur doit encore choisir s'il veut chiffrer ses mails ou non. De plus, Enigmail utilise Autocrypt, un système permettant d'envoyer la clé publique directement dans le mail. Pour les clés, Enigmail les envoie sur des serveurs de clés par défaut. Il va aussi interroger ces serveurs si une clé publique pour un destinataire est disponible dessus lors du chiffrement d'un message. Ces serveurs sont les suivants : keys.opengpg.org (vks), hkps.pool.sks-keyservers.net (hkps), pgp.mit.edu (hkps). vks (Verifying Keyserver) et hkps (HTTP keyserver protocol over TLS) sont des interfaces avec des serveurs de clés afin d'enregistrer des nouvelles clés ou trouver une clé sur le serveur par différents moyens (email, key-id, fingerprint). Les clés générées ont été générées avec les algorithmes de courbe elliptique EdDsa 4096bits par défaut, et dans les paramètres avancés il est possible de choisir entre cet algorithme et RSA.

2.2.2 S/MIME

Fonctionnement. S/MIME (Secure/Multipurpose Internet Mail Extensions) se base sur un système de PKI (Public Key Infrastructure) pour chiffrer et signer les mails. Dans une telle infrastructure les CAs (Certificate Authorities) garantissent que les certificats décrivent bien l'entité.

MIME est un standard qui étend le format de mail standard afin de pouvoir transmettre des données plus complexes que le format ASCII. En effet, cela permet de transmettre d'autres sets de caractères et la possibilité de transmettre des fichiers joints audio, vidéo, images, programmes aux emails. MIME permet de séparer le message en partie, c'est d'ailleurs là dessus que l'attaque EFMAIL (c.f. section 2.4.2) s'appuie. Ces parties ont différent format de données et le message MIME est défini selon un type parmi : mixed, digest, alternative, related, report, signed, encrypted, form-data, mixed-replace, byteranges. Ces types de contenu sont là pour définir de quoi est composé le message et comment le décoder, fonctionnant via une notion de frontières entre les différentes parties.

Dans ce contexte S/MIME vient ajouter un type MIME *application/pkcs7-mime*, un format de donnée qui enveloppe une entité MIME afin de la chiffrer, puis cette enveloppe chiffrée est ensuite le contenu de ce nouveau type. Cette enveloppe Pour les signatures S/MIME utilisera plus le signatures détachées et le type *multipart/signed* pour cela ou encore *application/x-pkcs7-signature*. Anciennement c'était effectivement le format de message PKCS#7 utilisé pour le format des messages chiffrés. Cependant, à l'heure actuelle c'est la spécification de CMS (Cryptographic Message Syntax) qui est utilisé dans ces types MIME. Afin de commencer à signer des messages avec S/MIME et pouvoir recevoir des messages chiffrés il faut un certificat. Ce certificat peut être obtenu soit par une autorité de

certification interne ou une autorité externe. Ces certificats peuvent être de classe 1 (vérification que le propriétaire du certificat peut recevoir des messages envoyés au "From : " de ses messages) ou de classe 2 avec plus de précisions sur le propriétaire du certificat.

Pour signer un message, S/MIME va utiliser la clé privée liée au certificat de la source, ainsi le message va être signé avec la primitive adéquate par rapport aux informations du certificat. Puis la signature sera envoyée avec le certificat, afin que le destinataire puisse vérifier la signature à l'aide du certificat. Et vérifier le certificat avec un CA.

Comme le présente la RFC8551 [17], le chiffrement effectué par S/MIME s'approche de celui fait par PGP, en effet S/MIME va chiffrer la clé de chiffrement symétrique (Content Encryption Key dans CMS) une fois par destinataire en utilisant la clé publique authentifiée par leur certificat respectif et aussi une fois pour la source du message (afin de pouvoir relire le message envoyé à l'avenir). Les algorithmes qui doivent être pris en charge dans CMS sont les suivants selon la RFC8551 :

Hash : SHA-256, SHA-512

Signature : ECDSA (courbe P-256 et SHA-256), EdDSA (courbe 25519 avec PureEdDSA mode), RSA PKCS #1 v1.5 avec SHA-256, RSASSA-PSS avec SHA-256

Asymétrique : ECDH sur P-256, ECDH avec HKDF-256, RSA, RSAES-OAEP

Symétrique : AES-128 CBC pour le non-authentifié, ChaCha20-Poly1305, AES128-GCM, AES256-GCM pour l'authentifié

Propriétés cryptographiques. S/MIME est aussi crée pour établir un **chiffrement bout-en-bout** et ainsi éviter de révéler trop d'informations lors d'une brèche tel que dans un MDA. **Authentification** et **intégrité** du message en utilisant la signature digitale, en plus de **non-répudiation** grâce à celle-ci. Le fait d'avoir le message chiffré de bout en bout permet un certain respect de la vie privée, on ne peut voir les données envoyées.

Utilisation. Pour utiliser S/MIME et avoir des exemples et références de certificats j'ai essayé 2 fournisseurs **gratuits** pour les certificats S/MIME. J'ai testé un plugin firefox qui permet d'avoir des certificats pour GMail et envoyer des mails signés et chiffrés à l'aide de S/MIME. La clé privée est générée par l'extension localement et n'est pas sauvegardée dans un cloud, il est possible de la sécuriser à l'aide d'une passphrase. Cependant les certificats ne sont pas vérifiés correctement comme on peut le voir à la réception dans MeSince (le prochain programme testé) à la figure 2.5. Le certificat utilise RSA avec SHA256 pour la signature des messages. Pour tester S/MIME, j'ai aussi lié un compte <https://www.mesince.com>. En effet, ce service permet d'utiliser S/MIME afin de chiffrer et signer ses mails, et ils fournissent les certificats, seulement il ne fonctionnait pas avec Gmail. De plus le service fourni n'a pas fonctionné pour se loguer et récupérer son certificat, ainsi le certificat a été généré et peut être utilisé par leur application mobile pour envoyer des mails signés et chiffrés mais je ne peux pas le voir. Sur l'application mobile il n'y a pas moyen de vérifier la clé générée. Attention cependant le même problème qu'avec Fossa arrive comme je le montre dans la

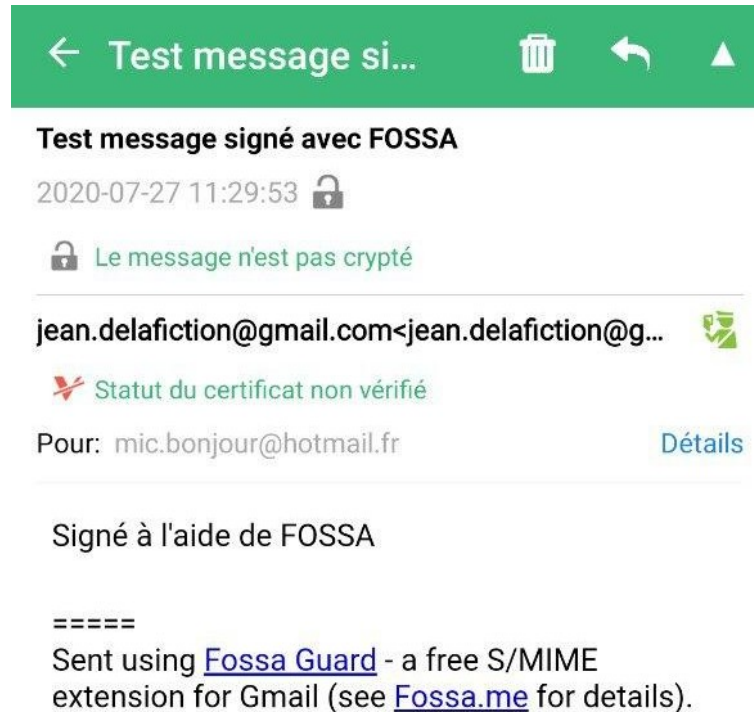


FIGURE 2.4 – Erreur de vérification pour Fossa

figure 2.5, par contre l'avertissement n'est pas très voyant au sein de Gmail. Les certificats utilisent RSA avec SHA256 pour la signature des messages. MeSince informe que la clé privée est automatiquement sauvegardée dans leur cloud sécurisé, ce qui n'est pas forcément une bonne nouvelle, de plus la clé est générée automatiquement donc aucune vérification de la part de l'utilisateur. Cependant ces 2 tests effectués ne représentent pas vraiment une utilisation réelle de S/MIME, en effet, le meilleur moyen de tester S/MIME aurait été d'avoir un nom de domaine à soi et de générer des certificats S/MIME pour une adresse privée. Ensuite d'importer ses certificats et ses clés dans le client mail utilisés.

2.3 Implémentations existantes

Dans cette section je présentes certaines implémentations des protocoles discutés dans la section précédente, plus particulièrement PGP.



FIGURE 2.5 – Erreur de vérification pour MeSince

2.3.1 Protonmail

Revendications. Protonmail revendique beaucoup de propriétés cryptographiques, tel que le zero-access encryption (lors de la réception d'un message externe chiffré ou non Protonmail le chiffrera avec la clé publique de l'utilisateur pour ne plus y avoir accès dans le futur). Et l'end-to-end chiffrement + zero-knowledge pour les messages sécurisés, même avec leur fonctionnalité de chiffrement vers l'extérieur utilisant AES256-GCM. Pour l'authentification Protonmail utilise une manière fortement sécurisée (SRP) pour ne pas avoir d'informations direct sur le mot de passe de l'utilisateur.

Fonctionnement. Protonmail a plusieurs modes de fonctionnement dépendant du destinataire final. En effet de Protonmail à Protonmail les mails sont chiffrés à l'aide de PGP automatiquement. L'on peut utiliser Protonmail pour utiliser PGP si l'on a la clé de notre destinataire par exemple. Et l'on peut écrire un mail chiffré à quelqu'un qui n'utilise pas PGP grâce à une fonctionnalité de chiffrement vers l'extérieur. Cette fonctionnalité enverra une URL au destinataire qui, en la consultant, pourra déchiffrer le mail en utilisant un mot de passe communiqué auparavant de manière sécurisées entres les deux partis.

Open Source. Leur code est open-source afin d'avoir une validation externe, de plus ils ont un programme de Bug Bounty pour les chercheurs. Ils ont largement contribué au projet

d'intégration openpgp en javascript et en Go.

Utilisation. Lors de l'utilisation de protonmail, l'analyse des clés PGP générées à montrer qu'elles utilisent RSA. Pour envoyer ou recevoir des messages via PGP, Protonmail le fait de manière très transparente. Apparemment utilisant les mêmes serveurs de clés cités plus haut pour Enigmail, en effet, les 2 emails de tests ont pu s'envoyer des messages sans s'envoyer les clés au préalable.

Sensation de sécurité en utilisant Protonmail, en effet un mot de passe est utilisé pour chiffrer notre boîte mail, pratique si une brèche survenait au niveau du stockage protonmail. De plus, la technologie utilisée pour l'authentification et les mots de passe est SRP. Permettant de ne pas avoir de hashs de mots de passe stockés chez Protonmail.

Par contre la génération de la clé privée à la création du compte est assez obscure mais de l'information est disponible sur le site de protonmail².

2.3.1.0.1 Critiques. Protonmail est très critiqué de manière générale sur les réseaux sociaux. De plus une réponse au chercheur Nadim Kobeissi [12] avait été formulée par Protonmail suite à son papier sur l'insécurité de leur webmail. Cependant, ils sont surtout critiqué sur les réseaux sociaux par des chercheurs en sécurité³ pour leur publicité mensongère, en effet, leur page d'accueil indiquait une sécurité pour **tous** les email sortants. Ce qui n'est pas le cas, uniquement les mails intentionnellement chiffré à l'aide de PGP ou chiffré vers l'extérieur sont chiffré. L'annonce ne représentait donc pas la réalité, ce qui peut être mal interprété par les utilisateurs. Après vérification la page d'accueil a bien été modifiée pour ne plus inclure la mention de chiffrement automatique pour tout les emails. Le commentaire fait référence aussi à d'autres points :

- Ingère les emails en plaintext - par rapport à leur politique de zero-access, ils ont au moins une fois accès au plaintext d'un email non chiffré.
- Pas de réels solutions pour du chiffrement vers l'extérieur - La solution amenant l'utilisateur distant à se connecter à leur service pour déchiffrer un message n'est pas une solution viable.
- Utilisation de la cryptographie dans le navigateur non validée - Modules cryptographiques servis par les serveurs de Protonmail, on revient ici aux revendications de ce papier [12]. Malgré cela une solution existe désormais, utiliser un bridge disponible à <https://github.com/ProtonMail/proton-bridge>.
- Agit comme un serveur de clés, sans transparence de clés - Suite à la mise en place d'un serveur de clés pour chercher les clés PGP d'utilisateur Protonmail pas de transparence sur ce serveur n'a été communiqué.

2. <https://protonmail.com/support/knowledge-base/how-is-the-private-key-stored/>

3. <https://twitter.com/FiloSottile/status/1277068367728435202>

2.4 Attaques existantes

Présentation des attaques faites sur les différents systèmes présentés jusqu'ici ainsi que leur fonctionnement global.

2.4.1 Défauts webmails

Selon un chercheur [12] l'infrastructure de Protonmail aurait des failles via son webmail. Mais son papier est en fait plus général et parle des webmails en règle générale. Il part du principe que les serveurs de Protonmail ne sont pas des serveurs à faire confiance, pour ainsi prouver le zero-knowledge de Protonmail. Par contre, le fait qu'il ne peuvent pas être mis en confiance est un problème selon lui, car c'est ces serveurs qui vont délivrer le code d'OpenPGP afin de faire le chiffrement. Cela indique que si Protonmail était corrompu le fait d'avoir le code délivré par Protonmail pourrait avoir des effets néfastes. Comme p.ex l'extraction de la clé privée PGP. La conclusion est que dès le moment où vous avez utilisé une fois le webmail de protonmail la clé PGP pourrait être corrompue ou connue de Protonmail.

2.4.2 EFAIL

Malgré ces sécurités qui pourraient être mises en place à l'heure actuelle, une attaque nommée EFAIL [15] a été faite en 2018. En effet cette attaque a seulement été mitigée en évitant d'afficher les contenus HTML et les images dans les boîtes mails de base. Car le problème vient de là principalement, des problèmes sont liés aussi aux modes de chiffrement utilisé (typiquement CBC et CFB pour S/MIME et PGP respectivement) grâce à des "gadgets". Cette attaque permet en fait d'injecter une image dans l'HTML du message, puis faire en sorte de récupérer le contenu du message déchiffrer dans l'URL. Ceci est possible grâce au multi parties de MIME et en les abusant afin d'entourer le message chiffré d'une balise *img* et de mettre en source le message qui sera déchiffré selon les règles de MIME. Ainsi l'attaquant aura le message déchiffré dans l'url qu'il peut contrôler. Cette attaque exploite en fait une erreur par rapport à la gestion des messages utilisant HTML et multipart/mixed, il faudrait en effet vérifier dans ces cas que le document HTML est un document entier. Ainsi que de ne pas traiter le contenu chiffré de la même origine que du contenu non protégé. La spécification [17] S/MIME a aussi mise à jour pour implémenter des chiffrements authentifiés et remplacer CBC qui permettait des attaques par gadgets.

2.5 Signal

L'analyse s'est faite aussi pour la messagerie instantanée à cause de sa ressemblance avec la messagerie électronique.

2.5.1 Fonctionnement

Le fonctionnement de Signal est complexe à expliquer, ainsi la figure 2.6 sera plus parlante pour expliquer le Diffie-Hellman Ratchet. Ce premier Ratchet permet d'utiliser Diffie-Hellman afin d'avoir une première sortie synchronisée entre l'envoi et la réception d'un utilisateur. Au début de la discussion Bob va envoyer sa clé publique et Alice va pouvoir commencer son premier Ratchet en générant elle aussi une paire de clés Diffie-Hellman. On verra dans la figure suivante ce qu'il se passe ensuite pour le chiffrement des messages mais une fois cette opération finie, Alice envoie sa clé publique précédemment générée à Bob pour qu'il puisse générer la même clé partagée à l'aide de Diffie-Hellman et ainsi déchiffrer le contenu en s'aidant du Deuxième Ratchet. Par contre l'image est une simplification, la clé DH ne va pas

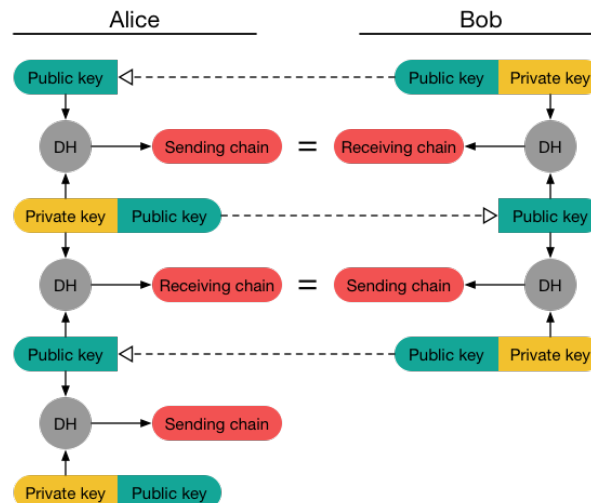


FIGURE 2.6 – Schéma fonctionnement du DH Ratchet[14]

être utilisée tel quel, elle sera insérée dans une KDF avec une Root Key (partagée à l'aide de X3DH) qui sortira la prochaine Root Key et la Receiving/Sending Chain Key qui sera donnée au deuxième Ratchet.

Le deuxième ratchet va générer une clé pour chaque message/batch de messages afin d'envoyer un message chiffré avec le moins d'utilisation de la même clé possible. Ainsi la forward secrecy est plus forte. Dans cet exemple on voit qu'Alice a envoyé un message chiffré à l'aide d'A1 puis qu'elle a reçu un nouveau DH Ratchet de Bob qui lui a permis de générer sa prochaine Receiving Chain Key et de déchiffrer les messages que Bob a envoyé (B1).

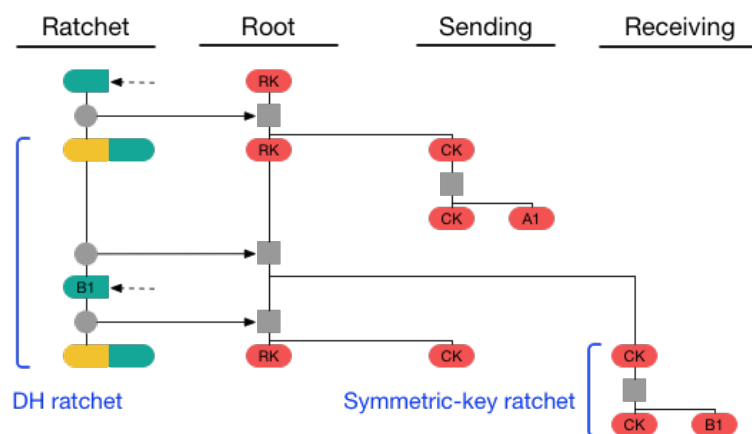


FIGURE 2.7 – Schéma fonctionnement du deuxième Ratchet[14]

2.5.2 Propriétés cryptographiques

Signal a de nombreux avantages concernant les propriétés cryptographiques qu'il promet. En effet le *Double Ratchet* permet de la **Break-in recovery**, a une **Perfect Forward Secrecy**, **End-to-End encryption**, **Non-Répudiation**.

2.5.3 Problèmes d'intégrations

Le problème avec le protocole Signal quant à mes besoins niveaux mails est la *perfect forward secrecy* qui est très forte. En effet comme vu dans le chapitre précédent il utilise une clé par message grâce au *Double Ratchet*. Cependant ce fonctionnement comporte un gros problème en rapport aux mails, en effet si l'on veut pouvoir récupérer les anciens mails reçus/envoyés cela devient vraiment compliqué. En effet, la *forward secrecy* est une propriété utile dans un système de mail, mais il faut pouvoir aussi récupérer les messages facilement si l'on connaît la clé privée. Alors qu'avec ce système de Double Ratchet il nous faudrait enregistrer chaque Diffie-Hellman ratchet pour reconstruire toute les clés utilisées.

2.6 Compromis

Pour passer à l'implémentation concrète d'un nouveau protocole il faut faire des compromis et aller chercher dans des primitives moins connues.

Je suis tout de même rester sur un système de clés publiques comme PGP le fait. Cependant

cette primitive a une identité propre à chaque clé publique ce qui évite un système de certificat trop complexe comme S/MIME.

De plus pour avoir une *forward secrecy* l'on peut ajouter une notion de temps ou de token à l'ID pour chaque batch de messages.

2.6.1 Résultats des recherches

Comme mentionné avant les recherches ont beaucoup été orientées sur le protocole Signal qui a une très bonne forward secrecy, résilience et break-in recovery. Cependant le problème avec l'utilisation des mails c'est d'avoir envie de consulter tout ces mails depuis n'importe quel appareil. Ce n'est malheureusement pas le cas avec Signal à moins de conserver une *root key* quelque part qui ferait s'effondrer les caractéristiques principales du protocole.

S/MIME est la solution prédominante pour s'envoyer des mails chiffrés cependant il est compliqué de l'utiliser. Il est en effet difficile d'obtenir un certificat pour envoyer des mails et échanger avec une autre personne ayant S/MIME. De plus la complexité d'un système de PKI et l'overhead induit est assez conséquent.

En faisant quelques essais PGP de mon côté je me suis heurté à beaucoup de difficultés et de problèmes avec les clés PGP, notamment pour se les échanger mais pour envoyer ensuite ce n'est pas si complexe, le problème étant de bien voir les primitives utilisées pour chiffrer/-signer notre email, en effet les solutions *plug-and-play like* ne permettent pas une gestion précise des primitives cryptographiques utilisées, ce qui pourrait induire à des primitives par défaut non sécurisées, comme l'a démontré EFAIL(c.f. 2.4.2).

2.7 Primitives

Présentation des primitives considérées pour implémenter un système de messagerie sécurisée. Cette section présente un aperçu des primitives analysées, pourquoi je l'ai ou non retenue et leur fonctionnement en quelques mots.

2.7.1 Primitives analysées

Les primitives analysées sont présentées ci après. Elles ont en commun de s'appuyer sur un principe basé sur l'identité. C'est très pratique dans un système de mail car une identité peut-être très vite définie par le biais d'une adresse email. Voici donc les technologies auxquelles les recherches se sont portées :

- Identity based encryption [18], primitive instaurant en premier le binding entre un ID et le chiffrement, ceci afin d'éviter les infrastructures complexes de PKI. Cependant, cette primitive introduit le *key escrow problem*, un problème inhérent de la construction

qui nécessite un KGC générant les clés privées pour les utilisateurs et qui en a par la conséquence connaissance.

- HIBE - Hierarchical Identity Based encryption [10]
- Certificateless PKC [1]

2.7.2 Primitive choisie

- Certificateless PKC [1]

Cette primitive a été choisie car elle est similaire à de l'identity based encryption avec un ID pour désigner une clé publique. Le problème avec l'identity based encryption c'est le fait que le serveur central génère la clé publique et la clé secrète de l'utilisateur, cela amène ce qu'on appelle le *key escrow* problème. C'est le fait qu'une entité connaisse à elle seule toutes les clés de tout les utilisateurs. Ce problème est résolu dans le certificateless en introduisant des *Partial Private Keys* permettant d'avoir une clé secrète partiellement générée par le serveur (KGC - Key Generation Center) et par l'utilisateur puis assemblée pour former la clé privée seulement connue de l'utilisateur. De plus, cette primitive a l'avantage de ne pas introduire de certificats et ainsi évite la complexité d'une infrastructure de PKI (Public Key Infrastructure).

2.8 Recherches sur la primitive

Dans cette section je vais introduire les détails techniques et les principes mathématiques utilisés. De plus, le choix de schéma parmi tous ceux analysés est détaillé ici. Ainsi que des précisions sur certains principes introduit par ce schéma.

2.8.1 Principes mathématiques

Les variantes de *Certificateless Cryptography* choisies utilisent un concept appelé les *pairings* ou *bilinear map groups*. Des groupes tels que $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T$ d'un ordre premier p pour lesquels il existe un mapping $e : \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ avec les propriétés suivantes :

1. Bilinéarité : $e(g^a, h^b) = e(g, h)^{ab}$ pour tout $(g, h) \in \mathbb{G}_1 \times \mathbb{G}_2$ et $a, b \in \mathbb{Z}$;
2. Pas de dégénérescence : $e(g, h) \neq 1_{\mathbb{G}_T}$ tant que $g, h \neq 1_{\mathbb{G}_{1,2}}$;

2.8.2 À savoir

Avant d'analyser les différents schémas, différents facteurs présents dans les tableaux doivent être expliqués.

Types. Les types présentés peuvent être soit **concret** soit **générique**. Les types concrets sont des schémas qui présentent leurs algorithmes en utilisant des calculs bien établis et présentent l'entiereté du fonctionnement de leur schéma, tandis que les schémas présentés génériques peuvent s'appuyer sur d'autres problèmes et se baser sur des algorithmes déjà existants.

Modèle de sécurité. Ces modèles définissent sur quoi le schéma va se reposer pour établir sa sécurité et comment il va l'évaluer face à un adversaire. À nouveau il existe deux modèles présents dans les schémas analysés, le *Random Oracle Model* et le *Standard Model*. Le *Random Oracle Model* se base sur des oracles aléatoires mais est un peu controversé, en effet l'aléatoire cryptographiquement sûr est difficile à atteindre, ainsi habituellement le *Random Oracle Model* implémente ces oracles via des fonctions de hachage. Le modèle standard se base lui sur des problèmes mathématiquement difficiles tel que DDH (Decisional Diffie Hellman).

Modèle d'adversaires. Pour évaluer les schémas de certificateless public key cryptography il y a différents niveaux de sécurité établis pour 2 types d'adversaires différents. Ces adversaires ont été décrits dans le papier d'Al-Riyami-Paterson [1] pour la première fois afin de prouver que leur schéma était IND-CCA sûr dans le modèle Standard. Ces deux adversaires ont été définis comme suit :

- Type I (*outsider adversaries*) est appelé *outsiders* et est permis de remplacer des clés publiques, obtenir des clés partielles privées, et des clés privées puis faire des requêtes de déchiffrements.
- Type II (*honest but curious KGC*). L'adversaire de Type II est en fait un KGC connaissant la Master Secret Key et qui peut donc générer des PPK, obtenir des clés privées et faire des requêtes de déchiffrement tout en faisant confiance à ce KGC pour pas qu'il ne remplace de clés publiques.

2.8.3 Schémas Certificateless de Chiffrement

Pour choisir parmi les nombreux schémas existants en certificateless pour le chiffrement j'ai établi un tableau comparatif des différentes manières de faire, inspiré de [19]. En suivant ce tableau je me suis rendu compte que la construction de Dent-Libert-Paterson [7] était probablement la plus adaptée en vue des propriétés qu'elle présentait. Le tableau se trouve en annexe B.

2.8.4 Détails techniques

Les détails techniques sur le chiffrement avec la *Certificateless Cryptography*. Le chiffrement se base sur le problème difficile *The Decision 3-Party Diffie-Hellman Problem* (3-DDH).

C'est de décider si $T = g^{abc} \text{ayant}(g^a, g^b, g^c, T) \in \mathbb{G}_4$.

Pour expliquer les détails techniques je vais ici montrer les calculs faits dans le schéma choisi [7] et les expliquer, cependant dans le schéma il est noté les calculs avec $\mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ mais il est mentionné que c'est facilement adaptable pour $\mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ ce que j'ai fait. De plus, la conversion vers un groupe additif (travaillant sur les courbes elliptiques) est faite afin que les calculs ici puissent être lus avec mon implémentation :

Setup($1^k, n$) : Avec $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T$ avec un ordre $p > 2^k$. g est un générateur de \mathbb{G}_1 . Ensuite $g_1 = g * \gamma$ pour un $\gamma \leftarrow \mathbb{Z}_p^*$ aléatoire. Puis $g_2 \leftarrow \mathbb{G}_2$. Deux vecteurs (U, V) seront tirés aléatoirement dans \mathbb{G}_2^{n+1} en tant que fonctions de hash notés :

$$F_u(ID) = u' \sum_{i=1}^n u_j^{i_j} \quad \text{and} \quad F_v(w) = v' \sum_{i=1}^n v_j^{w_j}$$

L'on va aussi prendre une fonction de hash résistante aux collisions : $H : \{0, 1\}^* \rightarrow \{0, 1\}^n$. Au final notre *mpk* (master public key) est :

$$mpk \leftarrow (g, g_1, g_2, U, V)$$

Et le *msk* (master secret key) est $msk \leftarrow g_2 * \gamma$.

Extract(mpk, γ, ID) : On prend $r \leftarrow \mathbb{Z}_p^*$ puis on retourne $d_{ID} \leftarrow (d_1, d_2) = (g_2 * \gamma + F_u(ID) * r, g * r)$

SetSec(mpk) : Retourne un secret aléatoirement choisi $x_{ID} \leftarrow \mathbb{Z}_p^*$.

SetPub(x_{ID}, mpk) : Retourne $pk_{ID} \leftarrow (X, Y) = (g * x_{ID}, g_1 * x_{ID})$.

SetPriv(x_{ID}, d_{ID}, mpk) : On choisit $r' \leftarrow \mathbb{Z}_p^*$ puis on reprends $(d_1, d_2) \leftarrow d_{ID}$ et l'on va prendre en secret key :

$$sk_{ID} \leftarrow (s_1, s_2) = (d_1 * x_{ID} + F_u(ID) * r', d_2 * x_{ID} + g * r')$$

Avec sk_{ID} étant la clé secrète de l'utilisateur, donnée par l'Extract (notre Partial Private Key) et la valeur secrète de SetSec.

Encrypt(m, pk_{ID}, ID, mpk) : Pour chiffrer $m \in \mathbb{G}_T$, l'on va reprendre $(X, Y) \leftarrow pk_{ID}$. Pour chiffrer ce message on va tiré aléatoirement $s \leftarrow \mathbb{Z}_p^*$ puis calculer :

$$C = (C_0, C_1, C_2, C_3) \leftarrow (m + e(Y, g_2) * s, g * s, F_u(ID) * s, F_v(w) * s)$$

Où $w \leftarrow H(C_0, C_1, C_2, ID, pk_{ID})$.

Decrypt(C, sk_{ID}, mpk) : L'on peut reprendre $(C_0, C_1, C_2, C_3) \leftarrow C$ et la clé privée $(s_1, s_2) \leftarrow sk_{ID}$. Afin d'accélérer le déchiffrement le calcul suivant peut être fait en tirant une valeur aléatoire $\alpha \leftarrow \mathbb{Z}_p^*$:

$$m = C_0 + \frac{e(s_2 + \alpha * g, C_2) * e(\alpha * g, C_3)}{e(C_1, s_1 + F_u(ID) * \alpha + F_v(w) * \alpha)}$$

Qui donnera m le texte en clair si le chiffré était bien formaté ou un élément aléatoire dans G_T .

2.8.5 Schémas Certificateless de Signature

Pour choisir parmi les nombreux schémas certificateless pour la signature j'ai établi un tableau comparatif des différentes manières de faire inspiré de [19]. En analysant les différentes possibilités dans ce tableau il y a peu de solutions se dégage, en effet l'on peut voir que beaucoup de schémas de signature sont cassés, mon choix s'est porté au final sur la construction de Zhang et Zhang [22] pour des signatures robustes en Certificateless. J'ai pris cette construction car elle est résistante au Malicious KGC (si le KGC a été setup avec des paramètres vulnérables) datant de 2006 et n'a pas été cassée depuis. Le tableau se trouve en annexe B.

2.8.6 Détails techniques

Les détails techniques sur la signature avec la *Certificateless Cryptography*. La signature se base sur le problème difficile *The Computational Diffie-Hellman Problem* (CDH).

Ayant P, aP, bP où a, b aléatoires $\in \mathbb{Z}_q^*$ il n'est pas possible de trouver abP .

Pour expliquer les détails techniques je vais ici montrer les calculs faits dans le schéma choisi [22] et les expliquer. Cependant dans le papier original le groupe bilinéaire de couplage choisi est de forme $\mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ avec un *pairing* $e(g, h)$ avec $g \in \mathbb{G}_1, h \in \mathbb{G}_2$ alors que le papier annonce une construction tel que $\mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$.

Setup(1^k) : Tout d'abord l'on va prendre les groupes d'ordre q énoncés auparavant. Puis on choisit un générateur $P \in \mathbb{G}_1$. La *master secret key* va être choisie aléatoirement $s \in \mathbb{Z}_q^*$. Puis la clé publique calculée : $P_{pub} = sP$. Finalement, trois fonctions de hash distinctes H_1, H_2, H_3 vont être choisies, chacune d'elle *mappant* de $\{0, 1\}^*$ à \mathbb{G}_2 . Pour cela j'ai choisi de faire du *Hash Domain Separation* comme expliqué dans le Chapitre 4. L'on définit les **params** = $(\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, e, q, P, P_{pub}, H_1, H_2, H_3)$

Partial-Private-Key-Extract(*params*, s, ID_A) : Pour avoir la *Partial Private Key* (D_A) de l'utilisateur A avec l'identité ID_A . Calculer $Q_A = H_1(ID_A)$. Alors $D_A = sQ_A$.

Set-Secret-Value : La valeur secrète $x \in \mathbb{Z}_q^*$ est tirée aléatoirement.

Set-Public-Key(*params*, x) : La clé publique PK_A de l'utilisateur A est $PK_A = xP$.

Set-Private-Key(*params*, D_A, x) : La clé privée SK_A de l'utilisateur A est calculée comme ceci $SK_A = (D_A, x)$.

CL-Sign(*params*, m, ID_A, SK_A) : Tout d'abord $r \in \mathbb{Z}_q^*$ est tiré aléatoirement puis on calcule les 2 composantes de la signature :

$$U = rP$$

$$V = D_A + rH_2(m, ID_A, PK_A, U) + xH_3(m, ID_A, PK_A)$$

Ainsi ces composantes forment la signature $\sigma = (U, V)$.

CL-Verify($params, PK_A, m, ID_A, \sigma$) : Tout d'abord l'on va calculer $Q_A = H_1(ID_A)$ puis effectuer ce calcul afin de vérifier la signature :

$$e(V, P) == e(P_{pub}, Q_A) * e(U, H_2(m, ID_A, PK_A, U)) * e(PK_A, H_3(m, ID_A, PK_A))$$

2.9 État de l'art

Dans cette section je vais analyser et comparer les différentes solutions trouvées utilisant la *Certificateless Cryptography* dans une implémentation concrète pour sécuriser des envois d'emails. J'ai trié les articles suivants par date de publication afin de voir comment les suivants ont repris les technologies etc.

2.9.1 Email Encryption System Using Certificateless Public Key Encryption Scheme

Analyse de l'article [8].

Description. Cet article présente une façon de faire pour chiffrer les mails à l'aide de *Certificateless Cryptography*. Il va d'abord comparer 6 schémas pour choisir celui à utiliser par rapport à ses propriétés. Ensuite il va comparer les différents algorithmes au niveau du temps avec une implémentation simple en J2SE.

Détails techniques. Les détails techniques ne sont pas très fourni dans cet article, en effet, il est mentionné uniquement le choix du schéma (Whang-Huang-Yang). Puis une comparaison des temps entre les différent algorithmes de la primitive. Finalement ils présentent la différence de temps entre le chiffrement du message via le certificateless et via une clé AES qui est chiffrée avec le certificateless.

Conclusion. Ce papier nous conforte dans l'idée de l'utilisation d'AES pour la rapidité du chiffrement qui va avec cette primitive. Cependant, ils n'expliquent pas comment la clé AES est prise et chiffrée réellement. Une implémentation existe en J2SE mais je ne l'ai pas trouvée. Le schéma choisi l'a été pour son avantage de ne pas utiliser les *pairings* et est donc plus rapide. Puis parmi les autres schémas qui n'utilisent pas les *pairings* à ce moment là, un est de type générique (c.f. sous-section 2.8.2) et l'autre est vulnérable aux *outsider attacks* (c.f. sous-section 2.8.2).

2.9.2 An End-To-End Secure Mail System Based on Certificateless Cryptography in the Standard Model

Analyse de l'article ⁴.

Description. Cet article présente une façon de chiffrer et signer dans un système de mail avec le schéma original de *Certificateless Cryptography* à savoir le schéma d'Al-Riyami et Paterson [1]. Un article complet définissant bien le contexte de mails et formalisant pour la première fois un moyen de chiffrer et signer des mails avec de la *Certificateless Cryptography*. Cela en expliquant dans les détails comment ils feraient, sans implémentations citées de ce schéma.

Détails techniques. Les détails techniques intéressant dans ce papier est la manière d'encapsuler la clé de chiffrement du message. Sinon le reste s'appuie sur le schéma d'Al-Riyami et Paterson. Pour établir une clé de chiffrement symétrique afin de chiffrer le mail l'n va tout d'abord tirer une valeur aléatoire $t \in \mathbb{Z}_p$ puis la chiffrer avec CL-PKC en utilisant la clé publique du destinataire $t^* = Enc_{P_B}$. Ce t^* sera envoyé avec l' email. Pour en tirer une clé symétrique on va établir : $K_{AB} = tx_AP_B$ à l'aide de la clé privée de la source et la clé publique du destinataire et enfin la valeur aléatoire tirée auparavant. Puis l'on va calculer la clé symétrique $K = H_2(Q_A || Q_B || K_{AB})$.

Conclusion. Ce papier est assez complet concernant la partie fonctionnement des mails en globalité et offres une bonne idée pour la construction d'une clé symétrique par mail envoyé. Cependant la mise en place de la clé symétrique et la preuve de son fonctionnement n'est pas très explicitée. D'ailleurs il y a selon moi une erreur dans le papier original pour la logique de déchiffrement de t^* et de la récupération de la clé symétrique. De plus, le système de signature d'Al-Riyami et Paterson a été cassé par [11].

2.9.3 Practical Implementation of a Secure Email System Using Certificateless Cryptography and Domain Name System

Analyse de l'article [4].

Description. Ce papier traite le problème de la même façon que le précédent mais essaies d'aller plus loin dans les détails d'un implémentation à plus grande échelle (utilisation DNS). Il reprend le même schéma et les mêmes principes pour la création de la clé symétrique de chiffrement. Le même schéma de signature est présent aussi, qui est cassé rappelons-le. Le but

4. <https://www.ijcsi.org/papers/IJCSI-10-2-3-264-271.pdf>

serait d'avoir une entrée DNS similaire au DKIM déjà utilisé pour les emails afin d'informer les utilisateurs quelle adresse donne es clés publiques du domaine en question.

Détails techniques. Beaucoup de détails concernant les domain policies qui pourraient être appliqués aux domaines pour la distribution des clés publiques. Proposition d'utiliser les headers d'emails pour transmettre la signature de l'email et informer le destinataire si l'email est chiffré ou non et de transmettre les IDS utilisés et le TImestamp utilisé. En effet, l'introduction d'un timestamp est proposé ici pour avoir un temps d'expiration au mail. Les Domain Policies sont là pour informer les utilisateurs si les emails de ce domaines doivent être signés/chiffrés ou non.

Conclusion. Une implémentation est citée utilisant la librairie MIRACL et en utilisant le C++ comme langage de programmation. L'implémentation est citée comme extension *Thunderbird* en C++ / Javascript. Mais il n'y a pas de réel guide pour implémenter cela au monde réel avec des exemples de configuration DNS et autres. Pas vraiment d'explications sur l'utilisation d'une multitude de KGC ou un seul central, comment les synchroniser et autres... Par contre beaucoup d'explications sur comment pourrait fonctionner une entrée DNS afin d'informer aux utilisateurs où aller pour récupérer les clés publiques des utilisateurs du domaine en question et des policies qui pourraient s'appliquer à ce domaine.

2.9.4 PriviPK : Certificate-less and secure email communication

Analyse de l'article [2].

Description. Cet article propose une implémentation très concrète utilisant CL-PKC pour communiquer de manière sécurisée dans la messagerie électronique. Il attaque beaucoup d'aspects que les autres papiers n'ont pas mentionnés comme la *key transparency*. Le papier insiste sur la transparence du protocole pour l'utilisateur afin qu'il n'ait pas d'opérations fastidieuses à faire (comme c'est le cas dans PGP et S/MIME par exemple). Ce papier s'appuie sur un système de *key agreement* proposé dans la littérature de la cryptographie basée sur l'identité.

Détails techniques. CONIKS serveur, authentification via les clients mails déjà existants (gmail et yahoo), mise en place d'un système de key agreement id-based repensée pour le certificateless.

Conclusion. Ce papier est assez intéressant et c'est la seule véritable implémentation que j'ai trouvée, il y a un repo sur github. Cependant il s'appuie sur du *key agreement*. Comme

le prochain système d'ailleurs. Par ailleurs, il insiste sur la transparence et sur l'utilisation des authentifications déjà présentes sur les clients emails.

2.9.5 A certificateless one-way group key agreement protocol for end-to-end email encryption

Analyse de l'article [20].

Description. Dans cet article les auteurs présentent un moyen d'avoir une clé partagée entre n -partis et avec un seul message, ce qui permet dans un système de mail d'avoir qu'à envoyer un mail avec les informations nécessaires pour recomposer la clé partagée. Cette clé partagée est utilisée afin de chiffrer le mail et de l'envoyer ensuite avec les informations nécessaires à la création de la clé partagée. De plus, le système est n -parti, cela veut dire que l'on peut envoyer le mail à n personnes et le chiffrer avec la même clé. On enverra juste pas les mêmes informations de créations de la clé partagée à tous.

Détails techniques. Pour ce qui est des détails techniques on peut voir que le principe est de créer une clé partagée à l'aide des différents ID et clés publiques des destinataires. On aura une sous-clé x_i pour chaque utilisateur i . L'on va construire un $y_i = x_0 \dots x_{i-1} + x_{i+1} \dots x_n$ pour un utilisateur où l'on additionnera tout les x des utilisateurs sauf de l'utilisateur i . Ainsi à la réception du message l'utilisateur pourra recréer la clé partagée en faisant $y_i + x_i = x_0 + \dots + x_n = K$. Ce K sera ensuite utilisé pour chiffrer le mail.

Conclusion. Ce système est simple et efficace mais ne permet pas la signature des éléments nécessaires à la création de la clé partagée, l'on peut donc envisager des DOS afin qu'un utilisateur ne puisse plus lire ces messages. Cependant c'est une construction intéressante se basant sur un *key-agreement* via le *Certificateless Cryptography* et non pas sur ses possibilités de chiffrement/signature.

Chapitre 3

Architecture / Design du protocole

Dans ce chapitre, je vais m'intéresser à expliquer le fonctionnement de la *certificateless cryptography* et démontrer comment je l'ai utilisée afin de l'intégrer à un protocole de chiffrement de mail.

3.1 Architecture globale

Dans cette Figure 3.1, je présente uniquement l'architecture globale pour bien représenter les différents acteurs présents dans le protocole et ainsi avoir une vue d'ensemble pour faciliter la compréhension.

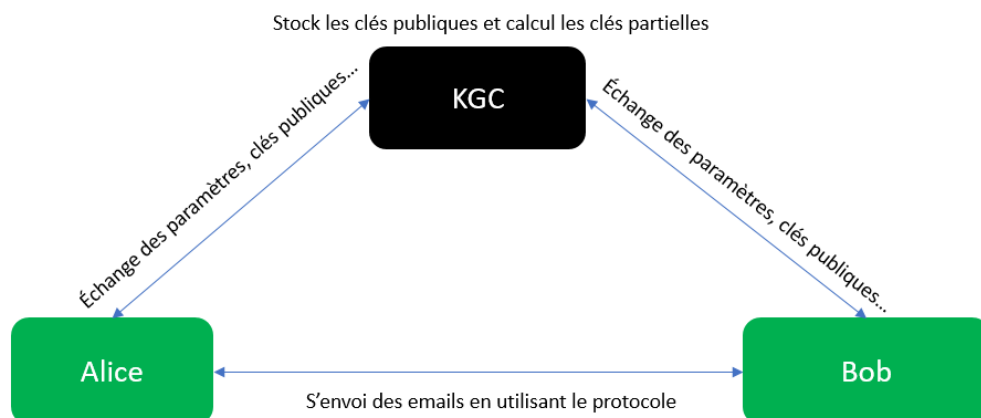


FIGURE 3.1 – Schéma global du protocole

3.2 Acteurs

Les parties impliquées sont les suivantes comme vu à la figure 3.1.

- Alice : L'envoyeur du mail en direction de Bob. Alice doit discuter avec le KGC pour construire sa clé privée (afin de signer) et récupérer la clé publique de Bob.
- Bob : Le destinataire du message, communique uniquement avec le KGC en ayant reçu le message d'Alice afin de récupérer sa clé publique pour vérifier la signature et de construire sa clé privée pour déchiffrer le message.
- KGC : Permet aux différents acteurs de pouvoir récupérer les clés publiques des clients, mais aussi de recevoir les *Partial Private Keys* qui permettent aux acteurs de construire leur clé privée.

Ces parties sont les principaux présents dans un exemple de *Certificateless Cryptography* dans mon système de mail.

3.3 Fonctionnement Certificateless PKC

Je vais ici découper les différents algorithmes présent dans le certificateless public key cryptography. En passant par le chiffrement et la signature. Ces algorithmes seront accompagnés d'explications sur leur utilité. Les noms donnés aux algorithmes seront réutilisés ensuite pour les schémas afin de démontrer l'architecture du protocole mis en place. L'on peut voir des définitions spécifiques dans l'article sur lequel je me suis appuyé pour ce travail [7].

3.3.1 Chiffrement

Liste des différents algorithmes de *Certificateless Cryptography* et leur description, les détails techniques de leurs implémentations sont disponibles à la Section 2.8.

- *Setup*. (seulement une fois par le KGC).
- *Partial-Private-Key-Extract*. Calcul d'une clé privée partielles lorsque qu'un client le demande pour identité donnée.
- *Set-Secret-Value*. Le client ne le fait qu'une fois pour tirer sa valeur secrète.
- *Set-Private-Key*. Le client combine ses clés partielles et sa clé secrète pour obtenir une clé privée afin de déchiffrer les message reçus, chiffrés avec une certaine identité.
- *Set-Public-Key*. Le client ne le fait qu'une fois, il calcule sa clé publique en fonction de sa valeur secrète.
- *Encrypt*. Chiffre un message avec la clé publique du destinataire et son identité.
- *Decrypt*. Déchiffre un message utilisant sa clé privée et l'identité utilisée pendant le chiffrement.

3.3.2 Signature

Pour la signature les algorithmes sont les mêmes avec une différence dans leur conception et évidemment le *Encrypt* et *Decrypt* sont remplacé par *Sign* et *Verify*. Dans la littérature certificateless les schémas de signatures sont beaucoup plus cassés que ceux de chiffrement apparemment (voir tableau). Il faut donc faire attention à suivre les schémas afin de vérifier que le schéma choisi ne soit pas mis à mal.

3.4 Design du protocole

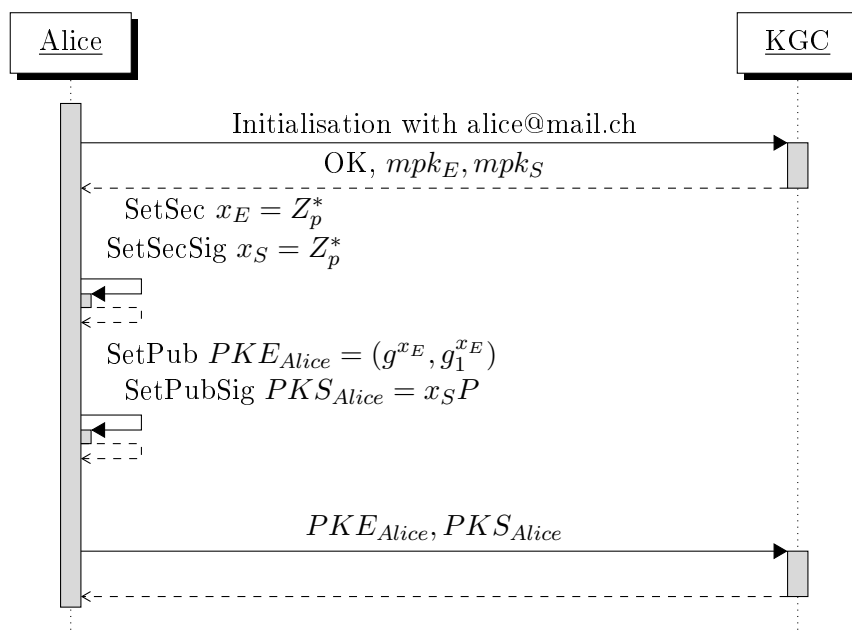


FIGURE 3.2 – Schéma de la première connexion

Dans la Figure 3.2 l'on voit la première connexion d'un utilisateur. Alice veut s'enregistrer auprès du KGC, ainsi le KGC lui renvoi les paramètres publics (mpk_S et mpk_E) si aucun utilisateur n'a déjà cet email. Ces paramètres publics sont assez lourd, en effet, ils font environ 52 kB au total, c'est donc un gros paquet qui est envoyé là. Dans mes expérimentations je n'ai jamais eu de problèmes, mais il est préférable de le mentionner tout de même. L'utilisateur va alors crée sa valeur secrète tirée aléatoire modulo p puis générer sa clé publique. Pour finir Alice envoie sa clé publique au KGC afin qu'il l'associe à son ID et puisse le donner aux personnes qui veulent envoyer un mail à Alice.

Dans la Figure 3.3 l'on voit comment se déroulerait l'envoi d'un message à Bob :

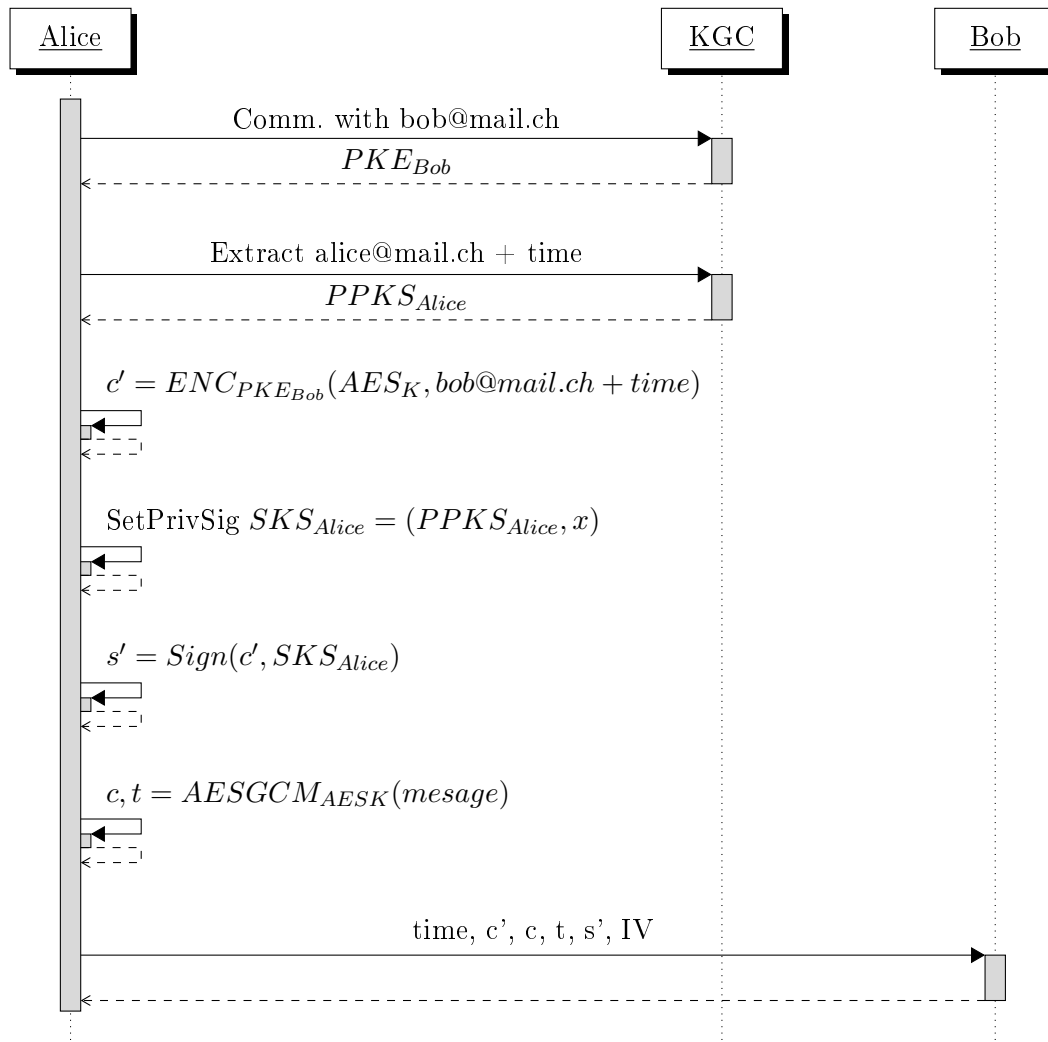


FIGURE 3.3 – Alice envoie un message à Bob

- Tout d'abord, Alice va récupérer le clé publique de Bob via son ID (aka email).
- Elle devra aussi récupérer sa clé privée partielle de signature pour créer ses clés privées afin de signer le message. Elle va le faire à l'aide de son ID et du même timestamp qu'utilisé pour la suite.
- Elle va ensuite tirer une valeur aléatoire dans G_t qui représentera sa clé AES pour la suite, elle va chiffrer cet élément à l'aide de la clé publique de Bob et de son ID complété par un timestamp. Ce timestamp sert à garder une certaine Forward Secrecy. Le cipher sera c' .
- Elle va calculer la signature du cipher donné (s' sur la figure)

- Alice utilisera un chiffrement authentifié comme AES_GCM pour chiffrer et authentifier son mail à Bob, t pour le tag et c pour le cipher.
- Finalement elle va envoyer tout ces éléments à bob (à savoir, l'ID utilisé, c , c' , t , s' et l'IV utilisé pour AES_GCM).

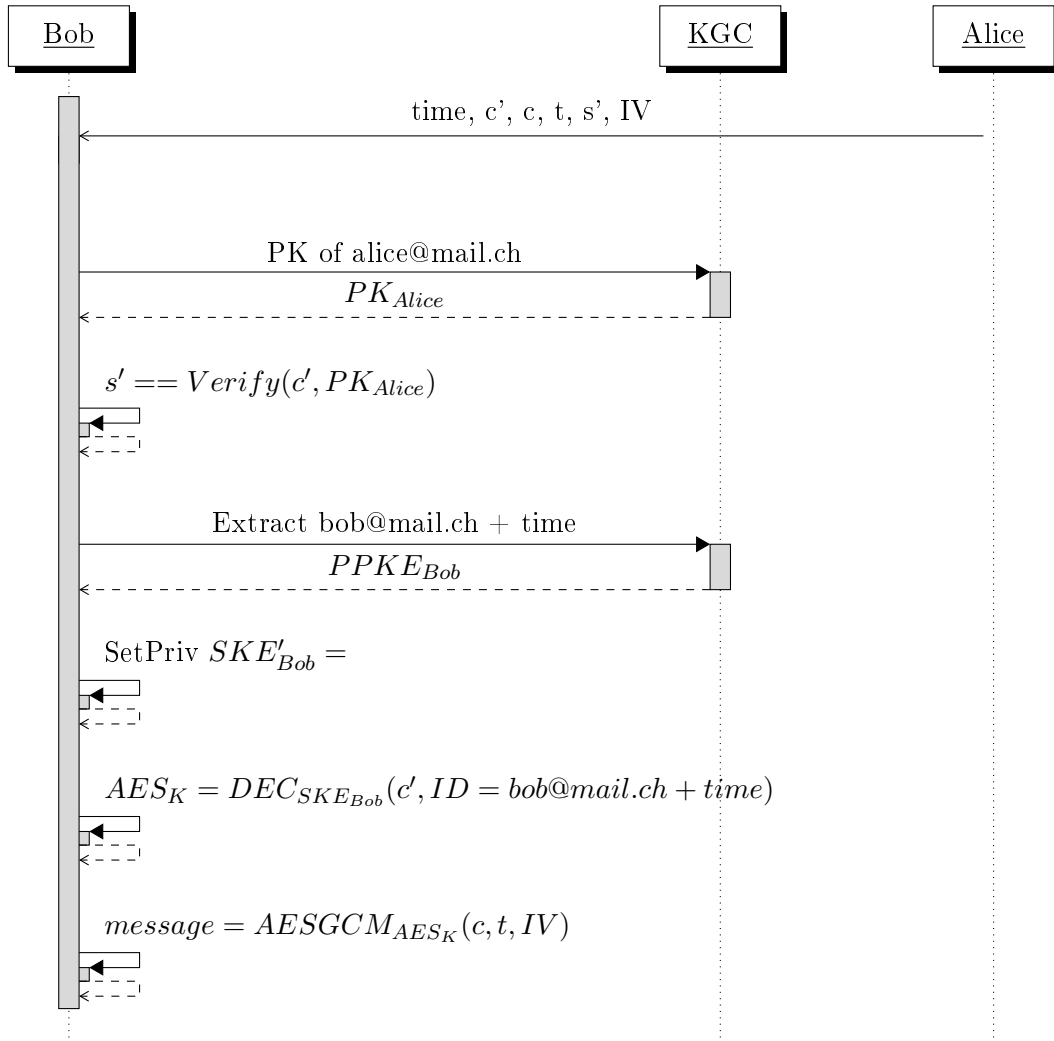


FIGURE 3.4 – Bob reçoit le message

Mais dans la Figure 3.4 l'on voit comment la réception du côté de Bob se déroulerait :

- A la réception la première chose à faire est de vérifier le cipher de la clé AES. Pour cela l'on va demander la clé publique d'Alice au KGC. Puis on va vérifier ce cipher c' à l'aide de sa signature s' .
- Ensuite Bob va récupérer sa clé privée partielle via le KGC en fournissant son ID avec

le timestamp envoyé par Alice. Il va ainsi pouvoir former sa clé privée.

- Avec sa clé privée il va pouvoir déchiffrer c' et obtenir la clé AES pour la suite.
- Une fois que l'on a la clé AES l'on peut simplement déchiffrer à l'aide d'AES_GCM c pour obtenir le message initial.

Chapitre 4

Implémentation

4.1 Choix d'implémentations

4.1.1 Langage

Au départ le choix du langage s'est porté sur sagemath (framework python) afin de mieux comprendre les différents calculs et faire un premier POC du chiffrement/déchiffrement. Cependant l'implémentation du POC était lente et le changement d'algorithme pour les pairings était difficile. Je me suis donc orienté sur le C pour avoir de meilleures performances et pouvoir mieux gérer la mémoire de mon implémentation. Pour pouvoir faire facilement des calculs sur les courbes elliptiques et les pairings en C il me fallait une librairie ce que je décris dans la section suivante. Comme on peut le voir sur la table 4.1, une différence des temps d'exécution entre les deux langages. Il faut cependant mettre en lumière que les temps sont calculés avec des courbes différentes et des couplages différents (Sage avec des Weil pairings, C avec ate pairings).

Temps des algorithmes entre langages [s]		
Algorithms	C	Sage
Setup	0.2856898	6.5858234
Encrypt	0.0061584	7.6450206
Decrypt	0.00951	3.3274426

TABLE 4.1 – Table de comparaison des temps d'exécution pour les différents algorithmes de Certificateless Cryptography

4.1.2 Librairie cryptographique

La librairie utilisée est RELIC Toolkit [3], c'est une librairie en cours de développement qui se veut efficiente. Sa concurrence avec MIRACL m'a fait hésiter dans mon choix, mais MIRACL est plus codée en C++ avec des équivalences en C j'ai donc choisi RELIC. De plus j'ai trouvé par exemple que RELIC était généralement plus adapté dans le domaine universitaire pour des POC puis il est plus efficient que d'autres librairies [16].

4.1.3 Courbe utilisée

La courbe utilisée pour le POC est la BLS12-P381, en effet cette courbe est assez efficiente et compatible avec les pairings. De plus RELIC l'a dans ses options et fonctionne bien, elle a un niveau de sécurité de 128bits. Je voulais prendre une courbe avec une plus grande sécurité cependant RELIC ne l'a pas encore totalement implémenté (certains tests concernant \mathbb{G}_2 ne passent pas), mais la librairie étant toujours en cours de développement il faudrait suivre ça de près, le code ne changerait en effet pas.

4.1.4 Dérivation de la clé AES

Le but de mon schéma certificateless est de chiffrer puis signer une clé AES qui permettra à mon message d'avoir un chiffrement authentifié. Pour cela il me faut dériver un élément de \mathbb{G}_t en clé AES, en effet le chiffrement dans le schéma certificateless se fait sur un élément de \mathbb{G}_t .

Pour cela j'ai utilisé une fonction permettant d'écrire sous forme compressée cet élément en bytes (fourni par la librairie RELIC utilisé et la fonction `gt_write_bin()`). Puis j'ai effectué un hachage avec SHA256 dessus, ainsi le résultat du hachage est une clé de 256 bits utilisable par AES-256-GCM. La fonction de hachage doit être par conséquent cryptographiquement sûre.

4.1.5 Fonctions de hachage - signature

Pour le schéma de signature il nous faut plusieurs fonctions de hachage différentes, en effet ce schéma est basé sur le *Random Oracle Model* comme définit dans le chapitre 2. Pour appliquer cela j'ai utilisé la même méthode de mapping disponible dans RELIC pour mapper une char array (tableau de byte) à un point sur G_2 à savoir `g2_map`. Pour H1, la première fonction de hachage j'ai simplement utilisé cette fonction directement, mais pour H2 et H3 j'ai ajouté un byte devant les données à mapper respectivement les bytes '01' et '02'. Ceci afin de séparer les domaines des résultats des hashes, cela s'appelle du *Hash Domain Separation*. En effet l'on peut voir dans ce draft [9] définit comme une simulation pour prendre en compte plusieurs *Random Oracle*.

4.1.6 Sérialisation des données

Pour la sérialisation des données, typiquement les clés publiques et les clés privées partielles envoyées en réseau ou les clés publiques enregistrées dans les fichiers par exemple, j'ai utilisé la librairie binn¹. Cela permet de packer facilement des données binaires, pour cela RELIC met à disposition des méthodes `gl_write_bin` `gl_read_bin` qui a permis de faire ces enregistrements binaires. Ainsi les transferts de données sont simplifiés. Cependant il faut faire attention à certaines choses, on ne peut lire et écrire simultanément à l'aide de binn, si l'on crée un objet via un buffer on ne pourra modifier cet objet. Cela m'a posé des problèmes pour l'enregistrement des données secrètes, j'ai donc du copier l'objet lu pour pouvoir le modifier et sauver les nouveau paramètres.

4.1.7 Enregistrement des clés publiques (serveur)

Pour l'enregistrement j'ai utilisé une petite base de données NoSQL stockant les clés publiques des utilisateurs sur le KGC. Cela permet de facilement récupérer une clé publique pour un utilisateur si besoin. Pour implémenter cela j'ai utilisé la librairie UnQlite². J'ai stocké les clés publiques pour le schéma de signature et de chiffrement séparément, en effet, l'entrée pour la signature porte le nom "signature/ID" et le chiffrement "encryption/ID".

4.1.8 Récupération via IMAP

4.2 Implémentation clés de chiffrement

Pour pouvoir implémenter ce schéma de chiffrement et signature certificateless dans un système hybride il a fallu penser à une manière d'encapsuler la clé et les données. Pour cela j'ai essayé de faire un système comparable à la figure 4.1.

4.3 Fonctionnement global POC (KGC)

Ici je présente le fonctionnement global de mon implémentation du KGC pour mon POC. De plus je présente les problèmes connus et des propositions d'améliorations.

1. <https://github.com/liteserver/binn>

2. <https://unqlite.org/>

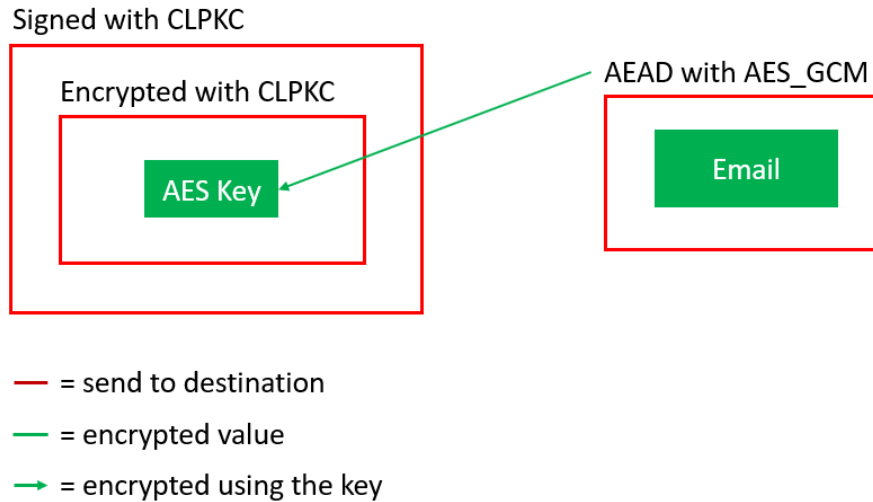


FIGURE 4.1 – Schéma encapsulation des données

4.3.1 Fonctionnement

Le KGC est un élément important du protocole, en effet c'est lui qui va fournir une partie de la clé privée de l'utilisateur. De plus dans mon cas il permet de distribuer les clés publiques de ses utilisateurs, mais habituellement on fera plutôt appel à un serveur dédié pour la gestion des clés.

Par mesure de généricité j'ai établi des codes d'opérations (arbitraires) afin de définir les opérations demandées au serveur par le client. Cela à l'aide de la librairie binn et les constructions d'objets proposés.

Structure des paquets reçus. Pour la structure des paquets que le KGC va traiter ils se présentent sous la forme d'un objet binn qui a comme propriétés au moins un code d'opération **opCode** et un **ID** associé. Cela permet de trier le paquet et de l'associer à une opération afin de traiter la donnée amenée avec le paquet. L'**ID** sert aussi différemment en fonction des codes employés.

Ainsi un paquet typique sera :

```
{opCode : PK, ID : alice@wonderland.com, payload : xxx}
```

Codes d'opérations. Les différents codes d'opérations sont :

- **HELO** : Permet de s'annoncer au KGC pour la première fois, le KGC répondra systématiquement avec les paramètres globaux du système. Cela implique la Master Public

Key de chiffrement du KGC ainsi que celle du schéma de signature.

- PK : Permet d'annoncer les clés publiques de l'utilisateur avec un certain ID. Le paquet est composé de {ID : alice@wonderland.com, opCode : PK, PKE : base64 de la PKE, PKS : base64 de la PKS}. La PKE est encodée en base64 par le client et envoyée au serveur. Elle est aussi représentée à l'aide d'un objet binn mais le serveur n'a pas besoin d'en prendre connaissance, il l'a stocke donc tel quel dans la base de donnée NoSQL. La même chose est faite pour la PKS, la clé publique pour la signature.
- GPE : Permet de récupérer la clé publique de chiffrement (utile pour chiffrer un message) d'un utilisateur ayant l'ID mentionné. Ainsi le serveur va simplement regarder dans la base de donnée pour "encryption/ID" et récupérer la clé encodée en base64 et la renvoyer à l'utilisateur. Si le serveur ne trouve pas cette clé il va renvoyer une erreur dans l'objet et ainsi à la réception on va d'abord regarder cette erreur.
- GPS : Fonctionne de la même manière que "GPE" mais pour les clés publiques de Signature (utile pour vérifier une signature).
- SE : Permet de faire la "Signature Extraction" et donc de demander la Partial Private Key pour l'utilisateur ID. Utilisé lors de la signature d'un message afin de construire sa clé privée et de signer le message avec.
- EE : Permet de faire la "Encryption Extraction" et donc de demander la Partial Private Key pour l'utilisateur ID. Utilisé lors du déchiffrement d'un message pour reconstruire sa clé privée.

4.3.2 Problèmes connus

4.3.3 Améliorations

Quelques améliorations qui seraient possibles mais dont j'ai pas eu le temps de m'occuper :

Vérification email. Implémentation d'une vérification par email afin d'être sûr que l'adresse email annoncée appartient bien à l'utilisateur qui s'authentifie pour la première fois. Typiquement lors du "HELO", l'on pourrait envoyer un mail de vérification avec un code sur l'email annoncé (s'il n'est pas dans la base de données) et demander le code envoyé avant de pouvoir uploader sa clé publique. Ainsi le client devrait pouvoir implémenter cette fonctionnalité aussi. Cela permettrait d'être sûr que tel utilisateur a effectivement tel email.

Zero-Knowledge proof pour PPK. On a vu dans les modèles d'adversaires définis par Al-Riyami et Paterson [1] qu'un adversaire pourrait obtenir les PPK des utilisateurs et ne pas réussir à déchiffrer des messages. Mais avec du Zero Knowledge l'on pourrait ajouter une sécurité supplémentaire pour ne pas délivrer à n'importe qui des PPK.

4.4 Fonctionnement global POC (Client)

Ici je présente le fonctionnement global de l'implémentation du client mail sécurisé pour mon POC, des améliorations possibles et des problèmes connus.

4.4.1 Fonctionnement

Je vais dépeindre le fonctionnement de mon POC ici et les différentes fonctionnalités que j'ai implémentée au client de mon POC.

Sécurité connexion mail. Pour la connexion au serveur SMTP / IMAP j'ai fait attention à la connexion sécurisée pour éviter de *leak* des mots de passe des utilisateurs du POC. En effet, cela permet de chiffrer les communications avec les serveurs de Gmail.

No.	Time	Source	Destination	Protocol	Length	Info
379	26.730664952	10.192.109.108	74.125.143.108	TCP	74	53902 → 465 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
381	26.748371123	74.125.143.108	10.192.109.108	TCP	74	465 → 53902 [SYN, ACK] Seq=0 Ack=1 Win=65535 Len=0 MSS=1430 S
382	26.748391401	10.192.109.108	74.125.143.108	TCP	66	53902 → 465 [ACK] Seq=1 Ack=1 Win=64236 Len=0 TSval=339349673
383	26.801969464	10.192.109.108	74.125.143.108	TLV1.3	419	Client Hello
384	26.820859839	74.125.143.108	10.192.109.108	TCP	66	465 → 53902 [ACK] Seq=1 Ack=354 Win=66816 Len=0 TSval=1380497...
385	26.822464881	74.125.143.108	10.192.109.108	TLV1.3	2675	Server Hello, Change Cipher Spec, Application Data
386	26.822480994	10.192.109.108	74.125.143.108	TCP	66	53902 → 465 [ACK] Seq=354 Ack=2610 Win=62336 Len=0 TSval=3393...
387	26.822700306	10.192.109.108	74.125.143.108	TLV1.3	72	Change Cipher Spec
388	26.823397880	10.192.109.108	74.125.143.108	TLV1.3	140	Application Data
389	26.841426422	74.125.143.108	10.192.109.108	TCP	66	465 → 53902 [ACK] Seq=2610 Ack=434 Win=66816 Len=0 TSval=1380...
390	26.843413707	74.125.143.108	10.192.109.108	TLV1.3	140	Application Data
391	26.843424392	10.192.109.108	74.125.143.108	TCP	66	53902 → 465 [ACK] Seq=434 Ack=2684 Win=64128 Len=0 TSval=3393...
392	26.843572915	10.192.109.108	74.125.143.108	TLV1.3	106	Application Data
395	26.865372848	74.125.143.108	10.192.109.108	TLV1.3	311	Application Data
396	26.865409214	10.192.109.108	74.125.143.108	TCP	66	53902 → 465 [ACK] Seq=474 Ack=2929 Win=64000 Len=0 TSval=3393...
397	26.865808987	10.192.109.108	74.125.143.108	TLV1.3	100	Application Data
398	26.884621928	74.125.143.108	10.192.109.108	TLV1.3	94	Application Data
399	26.885204201	10.192.109.108	74.125.143.108	TLV1.3	142	Application Data
400	26.908503711	74.125.143.108	10.192.109.108	TCP	66	465 → 53902 [ACK] Seq=2957 Ack=584 Win=66816 Len=0 TSval=1380...
405	27.142411781	74.125.143.108	10.192.109.108	TLV1.3	108	Application Data
406	27.142545835	10.192.109.108	74.125.143.108	TLV1.3	128	Application Data
408	27.160620151	74.125.143.108	10.192.109.108	TCP	66	465 → 53902 [ACK] Seq=2999 Ack=646 Win=66816 Len=0 TSval=1380...
409	27.161942631	74.125.143.108	10.192.109.108	TLV1.3	128	Application Data
410	27.162095361	10.192.109.108	74.125.143.108	TLV1.3	126	Application Data
413	27.181224853	74.125.143.108	10.192.109.108	TLV1.3	128	Application Data
414	27.181385704	10.192.109.108	74.125.143.108	TLV1.3	94	Application Data
415	27.204666916	74.125.143.108	10.192.109.108	TCP	66	465 → 53902 [ACK] Seq=3123 Ack=734 Win=66816 Len=0 TSval=1380...
426	27.578453148	74.125.143.108	10.192.109.108	TLV1.3	129	Application Data
427	27.578976106	10.192.109.108	74.125.143.108	TLV1.3	121	Application Data
428	27.579137777	10.192.109.108	74.125.143.108	TLV1.3	123	Application Data
429	27.579327130	10.192.109.108	74.125.143.108	TLV1.3	121	Application Data
430	27.579517979	10.192.109.108	74.125.143.108	TLV1.3	173	Application Data, Application Data
431	27.579790851	10.192.109.108	74.125.143.108	TLV1.3	1239	Application Data, Application Data
432	27.580017048	10.192.109.108	74.125.143.108	TLV1.3	156	Application Data, Application Data
433	27.580581205	10.192.109.108	74.125.143.108	TLV1.3	93	Application Data
434	27.580801400	74.125.143.108	10.192.109.108	TCP	66	465 → 53902 [ACK] Seq=3400 Ack=700 Win=66816 Len=0 TSval=1400...
▶ Frame 382: 66 bytes on wire (528 bits), 66 bytes captured (528 bits) on interface wlp58s0, id 0						
▶ Ethernet II, Src: IntelCor 4a:04:a7 (f8:59:71:4a:04:a7), Dst: Cisco ff:fc:3c (00:08:e3:ff:fc:3c)						
▶ Internet Protocol Version 4, Src: 10.192.109.108, Dst: 74.125.143.108						
▶ Transmission Control Protocol, Src Port: 53902, Dst Port: 465, Seq: 1, Ack: 1, Len: 0						
0000	00 08 e3 ff fc 3c f8 59 71 4a 04 a7 08 00 45 00<Y qj....E				
0010	00 34 b4 c7 40 00 40 06 33 e7 0a c0 6d 6c 4a 7d	4...@.@.3...mLJ}				
0020	8f 6c d2 8e 01 d1 68 86 6c 9e 50 ba 1b f5 80 10	1.....h. 1P.....				
0030	01 f6 8f ef 00 00 01 01 08 0a ca 44 a6 9a 52 48D...RH				
0040	b9 6e	.n				

FIGURE 4.2 – Connexion SSL/TLS avec le serveur email

4.4.2 Problèmes connus

Les problèmes qu'il reste à résoudre à ce jour :

Développement sécurisé. Lors du développement j'ai fait attention au maximum d'avoir le moins possible de fuite mémoires à l'aide de sanitizers et de valgrind. Cependant cela ne suffirait pas pour une application correctement sécurisée, il faudrait mettre à 0 les structures utilisées et qui stockent des informations confidentielles.

4.4.3 Améliorations

Les améliorations à amener dans le client :

Multiples destinataires. Pour le moment l'implémentation ne prends pas en compte une situation où un mail doit être envoyé à des destinataires multiples, c'est une fonctionnalité importante à mettre en œuvre dans une implémentation de client mail. Pour ce faire il faudra spécifier dans les headers l'utilisateur ciblé par tel cipher et signature. Ainsi lors de la réception le client prendra les options X-ID-CIPHER-B64 p.ex. Ou alors trouver un moyen d'envoyer un mail différent à chaque utilisateur, sans perdre la possibilité qu'un destinataire puisse choisir de répondre à tous.

Possibilité d'ajouter des pièces jointes. Pour le moment la possibilité d'ajout de pièces jointes n'a pas été pris en considération. Cependant une des bibliothèques choisies pour la réception des emails pourrait composer des messages contenant des pièces jointes. Il faudrait ainsi les chiffrer avec la clé symétrique avant de l'ajouter dans le mail.

GUI. Mettre en place une interface utilisateur pour le client mail, cela aiderait à rendre le chiffrement plus transparent et plus simple pour l'utilisateur. En effet demander à l'utilisateur d'écrire son mail au terminal n'est pas spécialement agréable.

4.5 Comparaisons avec état de l'art

Dans cette section je vais présenter les différents protocoles et implémentations existantes présentées au chapitre 2 et les comparer à mon implémentation. Tout d'abord en présentant les différentes propriétés cryptographiques puis les temps d'exécution.

4.5.1 Propriétés cryptographiques

Ici je fais un comparatif sur les différentes propriétés cryptographiques que les systèmes de mails sécurisés proposent avec mon implémentation Certificateless. On peut le voir dans la table 4.2.

Comparaisons des propriétés cryptographiques proposées		
Implémentations	E2EE	Forward Secrecy
CLPKC-POC	Oui	Oui
PGP	Oui	Non
S/MIME	Oui	Non

TABLE 4.2 – Table de comparaison des différentes propriétés cryptographiques

4.5.2 Temps des différentes implémentations

Comparasion du temps mis pour chiffrer et signer / déchiffrer et vérifier un mail entres les différentes implémentations existantes. Dans la table 4.3 l'on voit les calculs faits.

Comparaisons des temps d'exécution entres différentes implémentations proposées		
Implémentations	Chiffrement	Déchiffrement
CLPKC-POC	0.0061584s	0.00951s
PGP	0	0
S/MIME	0	0

TABLE 4.3 – Table de comparaison des temps d'exécution entres les implémentations de mails chiffrés

4.5.3 Overhead induit

Ici je présente les différents *overhead* que j'ai remarqué en utilisant les différents systèmes de mails sécurisés analysés au chapitre 2. Dans le tableau 4.4 on voit la taille d'overhead induit par les différents systèmes testés.

Comparaisons de l'overhead induit dans un mail		
Implémentations	Taille overhead	Contenu
CLPKC-POC	Environ 1200 bytes	Signature, timestamp, nonce, Encrypted Session Key
PGP	Environ 300 bytes	Encrypted Session key
S/MIME	0	TODO

TABLE 4.4 – Table de comparaison des différents overhead en rapport avec les solutions existantes

Chapitre 5

Conclusion

Bibliographie

- [1] Sattam S. Al-Riyami and Kenneth G. Paterson. Certificateless public key cryptography. In Chi-Sung Lai, editor, *Advances in Cryptology - ASIACRYPT 2003, 9th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, November 30 - December 4, 2003, Proceedings*, volume 2894 of *Lecture Notes in Computer Science*, pages 452–473. Springer, 2003.
- [2] Mashael AlSabah, Alin Tomescu, Ilia A. Lebedev, Dimitrios N. Serpanos, and Srinivas Devadas. Privipk : Certificate-less and secure email communication. *Comput. Secur.*, 70 :1–15, 2017.
- [3] D. F. Aranha, C. P. L. Gouvêa, T. Markmann, R. S. Wahby, and K. Liao. RELIC is an Efficient Library for Cryptography. <https://github.com/relic-toolkit/relic>.
- [4] Suresh Kumar Balakrishnan and V. P. Jagathy Raj. Practical implementation of a secure email system using certificateless cryptography and domain name system. *I. J. Network Security*, 18(1) :99–107, 2016.
- [5] J. Callas, L. Donnerhacke, H. Finney, D. Shaw, and R. Thayer. Openpgp message format. RFC 4880, RFC Editor, November 2007. <http://www.rfc-editor.org/rfc/rfc4880.txt>.
- [6] Wikimedia Commons. File :e-mail.svg — wikimedia commons, the free media repository, 2014. [Online ; accessed 26-July-2020].
- [7] Alexander W. Dent, Benoît Libert, and Kenneth G. Paterson. Certificateless encryption schemes strongly secure in the standard model. In Ronald Cramer, editor, *Public Key Cryptography - PKC 2008, 11th International Workshop on Practice and Theory in Public-Key Cryptography, Barcelona, Spain, March 9-12, 2008. Proceedings*, volume 4939 of *Lecture Notes in Computer Science*, pages 344–359. Springer, 2008.
- [8] Yee-Lee Er, Wei-Chuen Yau, Syh-Yuan Tan, and Bok-Min Goi. Email encryption system using certificateless public key encryption scheme. In James Jong Hyuk Park, Jongsung Kim, Deqing Zou, and Yang Sun Lee, editors, *Information Technology Convergence, Secure and Trust Computing, and Data Management - ITCS 2012 & STA 2012, Gwangju, Korea, September 6-8, 2012*, volume 180 of *Lecture Notes in Electrical Engineering*, pages 179–186. Springer, 2012.

- [9] Armando Faz-Hernandez, Sam Scott, Nick Sullivan, Riad Wahby, and Christopher Wood. Hashing to elliptic curves. Internet-Draft draft-irtf-cfrg-hash-to-curve-09, IETF Secretariat, June 2020. <http://www.ietf.org/internet-drafts/draft-irtf-cfrg-hash-to-curve-09.txt>.
- [10] Jeremy Horwitz and Ben Lynn. Toward hierarchical identity-based encryption. In Lars R. Knudsen, editor, *Advances in Cryptology - EUROCRYPT 2002, International Conference on the Theory and Applications of Cryptographic Techniques, Amsterdam, The Netherlands, April 28 - May 2, 2002, Proceedings*, volume 2332 of *Lecture Notes in Computer Science*, pages 466–481. Springer, 2002.
- [11] Xinyi Huang, Willy Susilo, Yi Mu, and Futai Zhang. On the security of certificateless signature schemes from asiacrypt 2003. In Yvo Desmedt, Huaxiong Wang, Yi Mu, and Yongqing Li, editors, *Cryptology and Network Security, 4th International Conference, CANS 2005, Xiamen, China, December 14-16, 2005, Proceedings*, volume 3810 of *Lecture Notes in Computer Science*, pages 13–25. Springer, 2005.
- [12] Nadim Kobeissi. An analysis of the protonmail cryptographic architecture. *IACR Cryptol. ePrint Arch.*, 2018 :1121, 2018.
- [13] Gaëtan Leurent and Thomas Peyrin. SHA-1 is a shambles - first chosen-prefix collision on SHA-1 and application to the PGP web of trust. *IACR Cryptol. ePrint Arch.*, 2020 :14, 2020.
- [14] Trevor Perrin and Moxie Marlinspike. The Double Ratchet Algorithm. Technical report, 2016.
- [15] Damian Poddebniak, Christian Dresen, Jens Müller, Fabian Ising, Sebastian Schinzel, Simon Friedberger, Juraj Somorovsky, and Jörg Schwenk. Efail : Breaking S/MIME and openpgp email encryption using exfiltration channels. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 549–566. USENIX Association, 2018.
- [16] Lucian Popa, Bogdan Groza, and Pal-Stefan Murvay. Performance evaluation of elliptic curve libraries on automotive-grade microcontrollers. In *Proceedings of the 14th International Conference on Availability, Reliability and Security, ARES '19, New York, NY, USA, 2019*. Association for Computing Machinery.
- [17] J. Schaad, B. Ramsdell, and S. Turner. Secure/multipurpose internet mail extensions (s/mime) version 4.0 message specification. RFC 8551, RFC Editor, April 2019.
- [18] Adi Shamir. Identity-based cryptosystems and signature schemes. In G. R. Blakley and David Chaum, editors, *Advances in Cryptology, Proceedings of CRYPTO '84, Santa Barbara, California, USA, August 19-22, 1984, Proceedings*, volume 196 of *Lecture Notes in Computer Science*, pages 47–53. Springer, 1984.
- [19] Hu Xiong, Zhen Qin, and Athanasios V. Vasilakos. *Introduction to Certificateless Cryptography*. CRC Press, Inc., USA, 2016.

- [20] Jyh-haw Yeh, Srisarguru Sridhar, Gaby G. Dagher, Hung-Min Sun, Ning Shen, and Kathleen Dakota White. A certificateless one-way group key agreement protocol for end-to-end email encryption. In *23rd IEEE Pacific Rim International Symposium on Dependable Computing, PRDC 2018, Taipei, Taiwan, December 4-7, 2018*, pages 34–43. IEEE, 2018.
- [21] Lei Zhang and Futai Zhang. A new provably secure certificateless signature scheme. In *Proceedings of IEEE International Conference on Communications, ICC 2008, Beijing, China, 19-23 May 2008*, pages 1685–1689. IEEE, 2008.
- [22] Zhenfeng Zhang, Duncan S. Wong, Jing Xu, and Dengguo Feng. Certificateless public-key signature : Security model and efficient construction. In Jianying Zhou, Moti Yung, and Feng Bao, editors, *Applied Cryptography and Network Security, 4th International Conference, ACNS 2006, Singapore, June 6-9, 2006, Proceedings*, volume 3989 of *Lecture Notes in Computer Science*, pages 293–308, 2006.

. BIBLIOGRAPHIE _____

Table des figures

2.1	Le fonctionnement d'un système de mail [6]	4
2.2	Le fonctionnement global de PGP	5
2.3	Exemple décodage d'un message PGP	7
2.4	Erreur de vérification pour Fossa	10
2.5	Erreur de vérification pour MeSince	11
2.6	Schéma fonctionnement du DH Ratchet[14]	14
2.7	Schéma fonctionnement du deuxième Ratchet[14]	15
3.1	Schéma global du protocole	25
3.2	Schéma de la première connexion	27
3.3	Alice envoi un message à Bob	28
3.4	Bob reçoit le message	29
4.1	Schéma encapsulation des données	34
4.2	Connexion SSL/TLS avec le serveur email	36

Liste des tableaux

2.1	Table des algorithmes utilisés par PGP	6
4.1	Table de comparaison des temps d'exécution pour les différents algorithmes de Certificateless Cryptography	31
4.2	Table de comparaison des différentes propriétés cryptographiques	38
4.3	Table de comparaison des temps d'exécution entre les implémentations de mails chiffrés	38
4.4	Table de comparaison des différents overhead en rapport avec les solutions existantes	38

Annexe A

Outils utilisés pour la compilation

A.1 RELIC Toolkit

Pour pouvoir faire des calculs de *Pairings* et sur des courbes elliptiques je me suis fier à RELIC Toolkit [3] qui est une librairie C permettant ce genre de calculs assez simplement. Cette librairie demande à être compilée avec une certaine courbe et certaines options (typiquement fonction de hachage et autres...). Des presets existent et c'est donc ce que j'ai utilisé pour ce POC. Cela demande donc de fournir la librairie précompilée avec les bonnes options pour l'utilisateur. L'inconvénient c'est donc que pour mettre à jour une courbe il va falloir recompiler toute la librairie et la fournir à l'utilisateur, néanmoins on n'aura pas à changer de code.

A.2 Libsodium

Pour faire du chiffrement authentifié j'ai utilisé libsodium¹, en effet, m'étant un peu familiarisé avec la librairie il m'a semblé être le choix le plus évident en plus de fournir des méthodes de chiffrement simples à mettre en place. Nécessite d'avoir libsodium en librairie liée.

1. <https://libsodium.gitbook.io/doc/>

A.3 binn

A.4 libetpan

A.5 libcurl

Annexe B

Fichiers

Je liste ici les fichiers annexes à mon rapport, ce qu'ils contiennent et comment les utiliser si besoin.

B.1 Code du POC

Le code est en annexe du rapport avec le nom *POCCertificatelessCryptography*. Le *README.md* présent à la source devrait être suffisant pour compiler soit même le code. Le code est très commenté au niveau du *main.c* pour bien montrer les différentes étapes telles qu'elle pourraient arriver dans une implémentation finale.

B.2 Tableaux comparatifs

Les tableaux comparatifs cité dans le chapitre 2 apparaissent sous forme de feuille dans un fichier excel se nommant *ComparatifsCLPKCSchemes.xlsx*. La feuille nommée CLEs contient un comparatif des schémas de chiffrement tandis que la feuille CLSs contient un comparatif des schémas de signatures.