A Concrete Treatment of Efficient Continuous Group Key Agreement via Multi-Recipient PKEs

<u>Keitaro Hashimoto</u> Tokyo Tech/AIST, JP

Shuichi Katsumata AIST, JP

Eamonn W. Postlethwaite CWI, NL

Thomas Prest PQShield SAS, FR Bas Westerbaan Cloudflare, NL

ACM CCS 2021



Efficient post-quantum CGKA protocol

- 1. Background
- 2. Our solution: Chained CmPKE
- **3.** More efficient PQ multi-recipient PKEs
- 4. Comparison and implementation

Background: Secure (Group) Messaging

Secure (Group) Messaging

Recently, a lot of people use secure (group) messaging apps.

Applications	Num. of monthly active users
WhatsApp	2.0 billion
Facebook Messenger	1.3 billion
Telegram	550 million
Snapchat	514 million

Ref: https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/

Secure (Group) Messaging

Recently, a lot of people use secure (group) messaging apps.

Applications	Num. of monthly active users
WhatsApp	2.0 billion
Facebook Messenger	1.3 billion
Telegram	550 million
Snapchat	514 million

Ref: https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/

Because governments and hackers try to gather personal information.

- "NSA Prism program taps into user data of Apple, Google and others", The Guardian, 2013 https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data
- "Al Jazeera journalists 'hacked via NSO Group spyware'", BBC, 2020 https://www.bbc.com/news/technology-55396843
- "Grand jury subpoena for Signal user data, Central District of California", Signal, 2020 https://signal.org/bigbrother/central-california-grand-jury/

Secure (Group) Messaging

Recently, a lot of people use secure (group) messaging apps.

Applications	Num. of monthly active users
WhatsApp	2.0 billion
Facebook Messenger	1.3 billion
Telegram	550 million
Snapchat	514 million

Ref: https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/

Existing secure (group) messaging:

- 2-party messaging: Signal protocol
 - Analyzed by a lot of works [CGC+17, ACD19, BFG+20, HKKP21]
- Group messaging : Continuous Group Key Agreement (this talk)

Group key agreement protocols that concentrate the cryptographic mechanisms of secure group messaging protocols:

- Add a party to the group
- Remove a party from the group
- Update key materials (Ratcheting)



Continuous Group Key Agreement (CGKA) [ACDT20]

CGKA achieves strong security properties by updating key materials

Forward secrecy (FS)



Post-compromise security (PCS)



Existing CGKA protocols

- TreeKEM [BBR18, BBN19, ACDT20, ACJM20, AJM20...]
 - Used in IETF Messaging Layer Security (MLS) [OBR+21, BBM+20]
- Chained mKEM [BBN19]
 - Based on multi-recipient PKE (mPKE)
 - Starting point of our study

Bandwidth cost for key update (*N*: group size)

Scheme	Upload cost	Download cost	Total cost (upload + N-1 download)
TreeKEM	$\Omega(\log N)$	$\Omega(\log N)$	$\Omega(N\log N)$
Chained mKEM	O(N)	O(N)	$O(N^2)$

As the group size N increases,

- the <u>size</u> of key update messages also increases
- the <u>frequency</u> of key update also increases
 - Likelihood of key compromise is higher for large group

As the group size N increases,

- the <u>size</u> of key update messages also increases
- the <u>frequency</u> of key update also increases
 - Likelihood of key compromise is higher for large group

This tension is amplified by two factors:

- Messaging apps target <u>mobile devices</u>
 - Data cap per month is limited (e.g., 1GB)
- Post-quantum cryptography
 - Consume x10 or more bandwidth than classical counterpart
 - Example: TreeKEM with Classic McEliece [ABC+20] used in 256 users' group.
 If each user updates its key material twice, it costs 1 GB for each user.

(((¢)))

As the group size N increases,

- the <u>size</u> of key update messages also increases
- the <u>frequency</u> of key update also increases
 - Likelihood of key compromise is higher for large group

This tension is amplified by two factors:

- Messaging apps target <u>mobile devices</u>
 - Data cap per month is limited (e.g., 1GB)
- Post-quantum cryptography
 - Consume x10 or more bandwidth than classical counterpart
 - Example: TreeKEM with Classic McEliece [ABC+20] used in 256 users' group.
 If each user updates its key material twice, it costs 1 GB for each user.

Smaller key update costs are desirable in the real-world!



As the group size N increases,

- the <u>size</u> of key update messages also increases
- the <u>frequency</u> of key update also increases
 - Likelihood of key compromise is higher for large group

This tension is amplified by two factors:

- Messaging apps target <u>mobile devices</u>
 - Data cap per month is limited (e.g., 1GB)
- Post-quantum cryptography
 - Consume v10 or more bandwidth than classical counternart

Purpose

Design PQ CGKA protocol with small key update costs



Chained CmPKE: CGKA with asymmetric bandwidth cost

Scheme	Upload cost	Download cost	Total cost (upload + N-1 download)
TreeKEM	$\Omega(\log N)$	$\Omega(\log N)$	$\Omega(N\log N)$
Chained mKEM	O(N)	O(N)	$O(N^2)$
Chained CmPKE	$O(N)^{\star}$	0 (1)	O (N)

*: When N is about hundreds, the concrete upload cost is smaller than TreeKEM.

Chained CmPKE is based on Chained mKEM with two new ideas:

- **1.** Committing mPKE \Rightarrow achieve O(1) download cost
- **2.** More efficient PQ mPKE \Rightarrow reduce the concrete size of key update messages



New CGKA: Chained CmPKE

Racap: Multi-recipient PKE (mPKE)

 ek_1



- The same message M can be efficiently encrypted to N parties
- Recently, [KKPP20] has revisited mPKE in the post-quantum setting
 - $|\hat{ct}_i| \ll |ct_0|$ in this setting

 $\operatorname{mEnc}(M, (ek_i)_{i \in [N]}) \to (ct_0, (\widehat{ct}_i)_{i \in [N]})$

Starting point: Chained mKEM [BBN19]

CGKA protocol based on **mPKE**

Key update on *N* parties' group









Chained mKEM [BBN19]



Chained mKEM [BBN19]



Drawback of Chained mKEM [BBN19]



Drawback of Chained mKEM [BBN19]



Sender signs individually messages for each user

 $(ek_1, ct_0, (\hat{c}t_i)_{i \in [N]}, sig_1, (sig_{2,i})_{i \in [N]})$

1. Gen new public key ek_1

2. Gen new group key *K*

- 3. Gen $(ct_0, (\widehat{ct}_i)_{i \in [N]}) \leftarrow \operatorname{mEnc}(K, (ek_i)_{i \in [N]})$
- 4. Gen $sig_1 \leftarrow Sign(sk_1, ek_1)$

5. For $i \in [N]$, $sig_{2,i} \leftarrow \text{Sign}(sk_1, (ct_0, \hat{ct}_i))$



 ek_2

 ek_3









Our solution: Committing mPKE (CmPKE)



 $\operatorname{CmEnc}(M, (ek_i)_{i \in [N]}) \to (T, (ct_i)_{i \in [N]})$



Our solution: Committing mPKE (CmPKE)



 $\operatorname{CmEnc}(M, (ek_i)_{i \in [N]}) \to (T, (ct_i)_{i \in [N]})$



Commitment-binding: T is linked to a unique message M \Rightarrow If parties receive the same T, they decrypt the same M or \bot

Our solution: Committing mPKE (CmPKE)



 $\operatorname{CmEnc}(M, (ek_i)_{i \in [N]}) \to (T, (ct_i)_{i \in [N]})$



 $\operatorname{CmDec}(dk_N, (T, ct_N)) \to M \text{ or } \bot$

Propose IND-CPA mPKE ⇒ **IND-CCA CmPKE** transformation

- CmEnc runs mEnc $(k, (ek_i)_{i \in [N]}) \rightarrow (ct_0, (\widehat{ct_i})_{i \in [N]})$ and SKE.Enc $(k, M) \rightarrow c$
- Outputs $T = (ct_0, c)$ and $ct_i = \hat{c}t_i$, |c| = 32 bytes

Use key-committing AEADs [FOR17, GLR17, ADG+20] as SKE











Chained CmPKE is as secure as TreeKEM version 10 in MLS

- Adopt the UC security model in [AJM20] used to analyze TreeKEM
 - It considers active adversaries and malicious insiders
- Extend this model to capture selective downloading of messages
 - Our model is the strict generalization of the model in [AJM20]



More efficient post-quantum mPKEs

Existing post-quantum mPKE

[KKPP20] proposed efficient PQ mPKEs based on LWE, LWR, and SIDH. Example scheme based on [LPR10, LP11]:

 $Enc(ek = \mathbf{B}, M)$:

- 1. Sample short matrixes **R**, **E**', **E**''
- **2.** $\mathbf{U} \leftarrow \mathbf{RA} + \mathbf{E}'$
- **3.** $\mathbf{V} \leftarrow \mathbf{RB} + \mathbf{E}'' + \text{Encode}(M)$

 $4. \quad ct \coloneqq (\mathbf{U}, \mathbf{V})$

mEnc($\{ek_1, \dots, ek_N\}, M$):

1. Sample short matrixes **R**, **E**'

2.
$$\mathbf{U} \leftarrow \mathbf{R}\mathbf{A} + \mathbf{E}'$$

3. For
$$i = 1, ..., N$$

1. Sample short matrix \mathbf{E}_i''

2.
$$\mathbf{V}_i \leftarrow \mathbf{RB}_i + \mathbf{E}_i'' + \text{Encode}(M)$$

4.
$$(ct_0, (\widehat{ct_i})_{i \in [N]}) \coloneqq (\mathbf{U}, (\mathbf{V}_i)_{i \in [N]})$$

[KKPP20] proposed efficient PQ mPKEs based on LWE, LWR, and SIDH. Example scheme based on [LPR10, LP11]:

 $Enc(ek = \mathbf{B}, M)$:

- 1. Sample short matrixes **R**, **E**', **E**''
- **2.** $\mathbf{U} \leftarrow \mathbf{RA} + \mathbf{E}'$
- **3.** $\mathbf{V} \leftarrow \mathbf{RB} + \mathbf{E}'' + \text{Encode}(M)$
- $4. \quad ct \coloneqq (\mathbf{U}, \mathbf{V})$

mEnc($\{ek_1, \dots, ek_N\}, M$):

1. Sample short matrixes **R**, **E**'

$$2. \quad \mathbf{U} \leftarrow \mathbf{R}\mathbf{A} + \mathbf{E}'$$

3. For
$$i = 1, ..., N$$

Sample short matrix \mathbf{E}_i''

2.
$$\mathbf{V}_i \leftarrow \mathbf{RB}_i + \mathbf{E}_i'' + \text{Encode}(M)$$

4.
$$(ct_0, (\widehat{ct_i})_{i \in [N]}) \coloneqq (\mathbf{U}, (\mathbf{V}_i)_{i \in [N]})$$

Two shortcomings of [KKPP20]:

- 1. Not optimize parameters to make $\widehat{ct_i}$ smaller
 - In CGKA setting, small $\widehat{ct_i}$ is desirable to reduce upload cost (~ $|\widehat{ct_i}| \cdot N$)
- 2. Not analyze the hardness of underlying problems in mPKE setting

[KKPP20] proposed efficient PQ mPKEs based on LWE, LWR, and SIDH. Example scheme based on [LPR10, LP11]:

 $Enc(ek = \mathbf{B}, M)$:

- 1. Sample short matrixes **R**, **E**', **E**''
- **2.** $\mathbf{U} \leftarrow \mathbf{R}\mathbf{A} + \mathbf{E}'$
- **3.** $\mathbf{V} \leftarrow \mathbf{RB} + \mathbf{E}'' + \text{Encode}(M)$
- $4. \quad ct \coloneqq (\mathbf{U}, \mathbf{V})$

mEnc($\{ek_1, \dots, ek_N\}, M$):

1. Sample short matrixes **R**, **E**'

2.
$$\mathbf{U} \leftarrow \mathbf{R}\mathbf{A} + \mathbf{E}'$$

B. For
$$i = 1, ..., N$$

1. Sample short matrix \mathbf{E}_i''

2.
$$\mathbf{V}_i \leftarrow \mathbf{RB}_i + \mathbf{E}_i'' + \text{Encode}(M)$$

4.
$$(ct_0, (\widehat{ct_i})_{i \in [N]}) \coloneqq (\mathbf{U}, (\mathbf{V}_i)_{i \in [N]})$$

Two shortcomings of [KKPP20]· 1. No • We fix this two shortcomings ③ 2. No

Designing Lattice-Based mPKEs: Attacks and Toolkit

Attacks with O(1) samples

- Lattice (primal)
- Lattice (dual)
- Decoding

Attacks with many samples

- Arora-Ge: requires n^{O(d)} samples
 (d = cardinality of error support)
- BKW

Toolkit

- Bit dropping
 - + Decrease $|V_i|$
 - + Increase the LWE noise
 - Increase decryption failure
- Coefficient dropping
 - + Decrease $|\mathbf{V}_i|$
- Increase the modulus q
 - + Pack more bits / coefficient
 - Increase $|\mathbf{U}_i|$
 - Decrease the LWE noise

Designing Lattice-Based mPKEs: Attacks and Toolkit

- Attacks with O(1) samples
 - Lattice (primal)
 - Lattice (dual)
 - Decoding

Good for security!

Attacks with many samples

- Arora-Ge: requires n^{0(d)} samples
 (d = cardinality of error support)
- BKW

Toolkit

- Bit dropping Good for efficiency! + Decrease $|V_i|$
 - + Increase the LWE noise
 - Increase decryption failure
- Coefficient dropping
 + Decrease |V_i|
- Increase the modulus q
 - + Pack more bits / coefficient
 - Increase $|\mathbf{U}_i|$
 - Decrease the LWE noise

Comparison: new parameters vs. existing parameters

Bandwidth of mPKE based on existing parameters (blue) and new parameters (blank) Size in byte. Security level is NIST I (\geq AES-128).

mPKE scheme	<i>ek</i>	<i>ct</i> ₀	$ \widehat{ct_i} $
Kyber512 [SAB+ 20]	768(+32)	640	128
llum512	768	704	48
LPRime653 [BBC+ 20]	865(+32)	865(+32)	128
LPRime757	1076	1076	32
Frodo640 [NAB+ 20]	9600(+16)	9600	120
Bilbo640	10240	10240	24
SIKEp434 [JAC+ 20]	330	330	16

 $|\hat{ct_i}|$ is reduced by 60-80% at the cost of slightly increase in |ek| and $|ct_0|$ \Rightarrow Minimize the concrete size of key update messages ($\sim |\hat{ct_i}| \cdot N$)

Comparison and Implementation

Chained CmPKE vs. TreeKEM: upload and download cost

Size of key update messages in Kilobyte (y-axis) depending on the group size (x-axis)



Chained CmPKE vs. TreeKEM: total cost (normalized by N)

Total cost of key update in Kilobyte (y-axis) depending on the group size N (x-axis)



Chained CmPKE: computation cost



Execution time in nanoseconds of some procedures as a function of group size for Ilum512 (___), LPRime757 (__), Bilbo640 (__), SIKEp434 (_+). <u>Log-scale</u>. Times are obtained on Apple M1@3.2 GHz.

Chained CmPKE: CGKA with asymmetric bandwidth cost

Scheme	Upload cost	Download cost	Total cost (upload + N-1 download)
TreeKEM	$\Omega(\log N)$	$\Omega(\log N)$	$\Omega(N\log N)$
Chained mKEM	O(N)	O(N)	$O(N^2)$
Chained CmPKE	$O(N)^{\star}$	0 (1)	O (N)

*: When *N* is about hundreds, the concrete upload cost is smaller than TreeKEM.

Chained CmPKE is based on Chained mKEM with two new ideas:

- **1. Committing mPKE** \Rightarrow achieve O(1) download cost
- **2.** More efficient PQ mPKE \Rightarrow reduce the size of key update messages

References

- [CGCD+17] K. Cohn-Gordon, C. Cremers, B. Dowling, L. Garratt, and D. Stebila. A Formal Security Analysis of the Signal Messaging Protocol. EuroS&P 2017.
- [ACD19] J. Alwen, S. Coretti, and Y. Dodis. The double ratchet: Security notions, proofs, and modularization for the signal protocol. EUROCRYPT 2019.
- [BFG+20] J. Brendel, M. Fischlin, F. Günther, C. Janson, and D. Stebila, "Towards Post-Quantum Security for Signal's X3DH Handshake. SAC 2020.
- [HKKP21] K. Hashimoto, S. Katsumata, K. Kwiatkowski, and T. Prest. An Efficient and Generic Construction for Signal's Handshake (X3DH): Post-Quantum, State Leakage Secure, and Deniable. PKC 2021.
- [ACDT20] J. Alwen, S. Coretti, Y. Dodis, and Y. Tselekounis. Security analysis and improvements for the ietf mls standard for group messaging. CRYPTO 2020.
- [BBN18] K. Bhargavan, R. Barnes, and E. Rescorla. TreeKEM : Asynchronous Decentralized Key Management for Large Dynamic Groups. A protocol proposal for Messaging Layer Security (MLS). 2018.
- [BBN19] K. Bhargavan, B. Beurdouche, and P. Naldurg. Formal Models and Verified Protocols for Group Messaging: Attacks and Proofs for IETF MLS. Research Report. 2019.
- [ACJM20] J. Alwen, S. Coretti, D. Jost, and M. Mularczyk. Continuous Group Key Agreement with Active Security. TCC 2020.
- [AJM20] J. Alwen, D. Jost, and M. Mularczyk. On The Insider Security of MLS. IACR ePrint. 2020.

References

- [OBR+21] E. Omara, B. Beurdouche, E. Rescorla, S. Inguva, A. Kwon, and A. Duric. The Messaging Layer Security MLS Architecture - draft-ietf-mls-architecture-06. *Internet Engineering Task Force Draft*. 2021.
- [BBM+20] R. Barnes, B. Beurdouche, J. Millican, E. Omara, K. Cohn-Gordon, and R. Robert. The Messaging Layer Security - draft-ietf-mls-protocol-11. *Internet Engineering Task Force Draft*. 2020.
- [FOR17] P. Farshim, C. Orlandi, and R. Rosie. Security of symmetric primitives under incorrect usage of keys. IACR Transactions on Symmetric Cryptology, pages 449–473, 2017.
- [GLR17] P. Grubbs, J. Lu, and T. Ristenpart. Message franking via committing authenticated encryption. CRYPTO 2017.
- [ADG+20] A. Albertini, T. Duong, S. Gueron, S. Kölbl, A. Luykx, and S. Schmieg. How to abuse and fix authenticated encryption without key commitment. To appear in USENIX Security 2022.
- [KKPP20] S. Katsumata, K. Kwiatkowski, F. Pintore, and T. Prest. Scalable Ciphertext Compression Techniques for Post-quantum KEMs and Their Applications. ASIACRYPT 2020.
- [LPR10] V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. EUROCRYPT 2010.
- [LP11] R. Lindner and C. Peikert. Better key sizes (and attacks) for LWE-based encryption. CT-RSA 2011.

References

- [ABC+20] M. R. Albrecht, D. J. Bernstein, T. Chou, C. Cid, J. Gilcher, T. Lange, V. Maram, I. von Maurich, R. Misoczki, R. Niederhagen, K. G. Paterson, E. Persichetti, C. Peters, P. Schwabe, N. Sendrier, J. Szefer, C. Jung Tjhai, M. Tomlinson, and W. Wang. Classic McEliece. Technical report, National Institute of Standards and Technology, 2020.
- [SAB+20] P. Schwabe, R. Avanzi, J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, G. Seiler, and D. Stehlé. CRYSTALS-KYBER. Technical report, National Institute of Standards and Technology, 2020.
- [BBC+20] D. J. Bernstein, B. Bob Brumley, M. Chen, C. Chuengsatiansup, T. Lange, A. Marotzke, B. Peng, N. Tuveri, C. van Vredendaal, and B. Yang. NTRU Prime. Technical report, National Institute of Standards and Technology, 2020.
- [NAB+20] M. Naehrig, E. Alkim, J. Bos, L. Ducas, K. Easterbrook, B. LaMacchia, P. Longa,
 I. Mironov, V. Nikolaenko, C. Peikert, A. Raghunathan, and . Stebila. FrodoKEM. Technical report, National Institute of Standards and Technology, 2020.
- [JAC20+] D. Jao, R. Azarderakhsh, M. Campagna, C. Costello, L. De Feo, B. Hess, A. Jalali,
 B. Koziel, B. LaMacchia, P. Longa, M. Naehrig, J. Renes, V. Soukharev, D. Urbanik, G. Pereira, K. Karabina, and A. Hutchinson. SIKE. Technical report, National Institute of Standards and Technology, 2020.