ОҚИМЛИ ШИФРЛАШ АЛГОРИТМЛАРИНИ КРИПТОТАҲЛИЛ ҚИЛИШ УЧУН ЯНГИ ЁНДАШУВЛАР

Илҳом Рахматуллаевич РАҲМАТУЛЛАЕВ
Техника фанлари бўйича (PhD) фалсафа доктори
Доцент
докторант
Рақамли технологиялар ва сунъий интеллектни ривожлантириш илмий-тадқиқот институти
Самарқанд, Ўзбекистон
Ilhom9001@gmail.com
Бахтиёр Файзиевич АБДУРАХИМОВ
Профессор
Мирзо Улуғбек номидаги
Ўзбекистон миллий университети
Тошкент, Ўзбекистон
a_baxtiyor@mail.ru

Аннатация

Ушбу мақолада оқимли шифрлаш алгоритмларини криптотаҳлил қилиш учун чуқур ўрганишга (deep learning) асосланган янги методология таклиф этилади. Таклиф этилган ёндашувнинг асосий мақсади — оқимли шифрлар томонидан генерация қилинадиган *keystream*да юзага келиши мумкин бўлган статистик оғишларни (bias) аниқлашдир. Бунинг учун “Калит – Keystream” жуфтликларидан иборат катта ҳажмдаги маълумотлар тўплами яратилади, сўнгра ушбу тўплам асосида нейрон тармоқни ўқитиш ва синовдан ўтказиш амалга оширилади.

Таянч сўзлар: Deep Learning, Оқимли Шифрлаш, Криптотаҳлил, RC4, RC4A, Trivium, TRIAD-SC, Keystream, Bias, Нейрон Тармоқлар, Машинали Ўқитиш, Supervised Learning, Softmax, Adam Optimizer, Dataset, Black-Box Модели, Башоратлаш, Чуқур Ўрганиш Архитектураси.

НОВЫЕ ПОДХОДЫ К КРИПТОАНАЛИЗУ ПОТОКОВЫХ ШИФРОВ

Илхом Рахматуллаевич РАХМАТУЛЛАЕВ
Доктор философии по (PhD) техническим наукам
доцент
Докторант
Научно-исследовательский институт развития цифровых технологий и искусственного интеллекта
Самарканд, Узбекистан
Ilhom9001@gmail.com
Бахтиёр Файзиевич АБДУРАХИМОВ
Профессор
Национальный университет Узбекистана имени Мирзо Улугбека
Ташкент, Узбекистан
a_baxtiyor@mail.ru

Аннотация

В данной статье предлагается новая методология криптоанализа потоковых шифров, основанная на методах глубокого обучения (deep learning). Основная цель предлагаемого подхода заключается в обнаружении возможных статистических отклонений (bias), возникающих в keystream, генерируемом потоковыми шифрами. Для этого формируется крупномасштабный набор данных, состоящий из пар «Ключ – Keystream», после чего на основе этих данных проводится обучение и тестирование нейронной сети.

Ключевые слова: Deep Learning, Потоковое шифрование, Криптоанализ, RC4, RC4A, Trivium, TRIAD-SC, Keystream, Bias, Нейронные сети, Машинное обучение, Supervised Learning, Softmax, Adam Optimizer, Dataset, Модель «черного ящика», Прогнозирование, Архитектура глубокого обучения.

Замонавий рақамли технологиялар, глобал ахборот оқимлари ва юқори даражадаги тармоқлар кенгайиб бораётгани сабабли маълумотларни ишончли ҳимоялаш масаласи криптография фанининг энг муҳим йўналишларидан бирига айланди. Симметрик калитли шифрлаш алгоритмларининг, хусусан оқимли шифрлаш усулларининг самарадорлиги ва хавфсизлиги реал вақт режимида ишлайдиган тизимлар — IoT қурилмалари, мобил алоқа протоколлари, симсиз тармоқлар, киберхавфсизлик инфратузилмалари учун фундаментал аҳамият касб этади [1]. Шунинг учун оқимли шифрлаш алгоритмларининг математик барқарорлигини, статистик хусусиятларини ва криптотаҳлилга бардошлилигини чуқур ўрганиш бугунги кунда долзарб илмий масалалардан биридир.

Анъанавий криптотаҳлил усуллари одатда шифрнинг ички тузилмасига, мураккаб чизиқсиз функцияларига ёки калит–шифр–очиқ матн ўртасидаги боғланишларни математик таҳлил қилишга асосланади. Бундай ёндашувлар юқори ҳисоблаш мураккаблиги, аналитик моделларнинг ноаниқлиги ва кўп жиҳатдан қўлда танланадиган хусусиятларга таяниши сабабли ҳар доим ҳам етарли даражада самарали натижа бермайди. Айниқса оқимли шифрларда кузатилиши мумкин бўлган жуда кичик статистик

оғишлар (bias) анъанавий усулар орқали аниқланиши мураккаб бўлган масаладир [2].

Сўнгги йилларда машинали ўқитиш, хусусан чуқур ўрганиш (deep learning) технологияларининг ривожланиши криптография соҳасига янги ёндашувларни олиб кирди [3]. Deep learning хом маълумотлар билан бевосита ишлаши, фойдали хусусиятларни мустақил ўрганиши ва мураккаб боғланишларни автоматик равишда аниқлаши туфайли криптотаҳлилнинг турли йўналишларида — дифференциал криптоанализ, ён канал таҳлили, функционал заифликларни аниқлашда юқори самарадорлик кўрсатмоқда. Хусусан Gohr, Picek, Wang каби тадқиқотчилар нейрон тармоқларнинг криптотаҳлил жараёнини кучайтиришда сезиларли натижаларга эришган бўлиб, бу ёндашувнинг истиқболи юқори эканини тасдиқлайди [4].

Ушбу мақолада оқимли шифрлаш алгоритмларини — RC4, RC4A, Trivium ва TRIAD-SC — таҳлил қилиш учун чуқур ўрганишга асосланган янги методология таклиф этилади. Таклиф этилган ёндашувнинг асосий моҳияти шундан иборатки, оқимли шифрни “қора қути” (black-box) сифатида қараган ҳолда “Калит–Keystream” жуфтликларидан иборат катта ҳажмдаги маълумотлар тўплами яратилади ва нейрон тармоқлар ёрдамида keystream байтларининг башорат қилиниши ўрганилади. Агар модель маълум байтнинг ҳосил бўлиш эҳтимолини тасодифий қийматдан сезиларли даражада юқори башорат қила олса, бу байтда статистик оғиш мавжудлигини билдиради.

Ушбу мақолада оқимли шифрлаш алгоритмларини криптотаҳлил қилиш учун чуқур ўрганишга (deep learning) асосланган метод таклиф этилади. Ушбу методнинг асосий мақсади — оқимли шифрлар томонидан генерация қилинадиган калит оқимида (keystream) юзага келадиган эҳтимолий оғишларни (bias) аниқлашдир [5]. Таклиф этилаётган алгоритмнинг босқичма-босқич тафсилоти қуйидагича:

0. Дастлабки параметрлар қуйидагича белгилансин:

a) Бошланғич калит: **K**

b) Бошланғич ҳолат: **S**

c) Калит оқими: **KS**

(“Қора қути” — *black-box* — сценарийсида **S** нинг ўзгариш жараёни номаълум ҳисобланади).

1. Бошланғич калит *K* қуйидагича ифодаланади:

$\mathbf{K = (k_1, k_2, \ldots, k_n)}$

бу ерда **n** — калит узунлиги (байтларда).

2. ***K*** ни кириш сифатида бериб, оқимли шифрлаш алгоритми ишга туширилади ва натижада KS калит оқими ҳосил қилинади.

3. 1- ва 2-босқичлар ***N*** та турли қийматдаги ***K*** учун такрорланади ва натижада **N** та “Калит – Keystream” жуфтликлари (***K – KS***) тўплами ҳосил қилинади.

4. Ҳосил қилган барча **N** та жуфтликдан иборат маълумотлар тўплами (dataset) яратилади.

5. Dataset икки қисмга ажратилади:

Ўқув (Training) тўплами — 80%

Тест тўплами — 20%

6. Deep learнинг тармоғига кириш сифатида **K** берилади, чиқиш (*target*) сифатида эса ***KS*** нинг 1-байти олинади.

7. Ўқув тўплами ёрдамида нейрон тармоқ ўқитилиб, чуқур ўрганиш модели яратилади.

8. Тест тўпламида модель текширилади ва модельнинг KSнинг 1-байтини башорат қилиш эҳтимоли қайд этилади.

9. 6–8-босқичлар ***KS*** нинг 2-, 3-, 4-, … байтлари учун ҳам такрорланади.

10. Агар ***KS*** нинг ***i***-байти учун моделнинг башорат эҳтимоли 1/256 (≈ 0.0039) қийматидан сезиларли даражада юқори бўлса, у ҳолда ушбу ***i***-байт оғишга эга деб баҳоланади. Акс ҳолда, у тасодифий (*random*) ҳисобланади.

Таклиф этилган методология асосида оқимли шифрлагичларнинг таҳлили.

Юқорида баён этилган экспериментал созламалар RC4 алгоритмини криптотаҳлил қилишда қуйидаги параметрлар билан қўлланилди:

Бошланғич калит узунлиги (n) = 5 байт
Намуналар сони / жуфтликлар (N) = 400 000

Тажриба RC4 томонидан ҳосил қилинган калит оқимининг 1-, 2-, 3-, 4-, байтларини башорат қилишга қаратилган. Маълумки, RC4 ўзгарувчан узунликдаги бошланғич калитдан (одатда 5–256 байт оралиғида) фойдаланади ва натижада байтма-байт псевдо тасодифий *keystream* генерация қилади.

Бизнинг тажриба муҳитимизда кириш калити 5 байт қилиб белгиланди ва 400 000 та “initial key – keystream” жуфтлиги яратилди. RC4нинг ҳолат вектори (*State Vector*) ёки бошланғич ҳолати (*Initial State*) бошланғич калит асосида ташкиллаштирилиб, кейинчалик мунтазам янгиланишлар орқали keystream ҳосил қилади. Тажрибада ҳолат вектори узунлиги 256 қилиб олинган.

Таъкидлаш лозимки, ўзгарувчан бошланғич калит — кириш, *keystream*нинг j-байти эса — чиқиш (*target*) сифатида қабул қилинади.

RC4 учун чуқур ўрганишга асосланган башорат модели.

Биринчи босқичда нейрон тармоқ RC4 *keystream*нинг биринчи байтини башорат қилиш учун созланди. Ҳар бир намунада:

- **5 байтли калит** — кириш қатламидаги 5 та нейрон,
- **keystreamнинг биринчи байти** — чиқиш нейрони сифатида берилди.

Мисол: Агар бошланғич калит: 0xEB 9F 72 AE 1C ва генерация қилинган keystream: 0xBB F3 16 E8 D9 40 AF 0A D3… бўлса, унда:

- **Кириш нейронлари:** EB, 9F, 72, AE, 1C
- **Чиқиш (target):** BB

Жами 400 000 жуфтликдан иборат dataset ҳосил қилинди. У қуйидагича ажратилди:

- **80% (320 000)** — ўқув тўплами
- **20% (80 000)** — тест тўплами

Модель қуйидаги архитектурадан иборат бўлди:

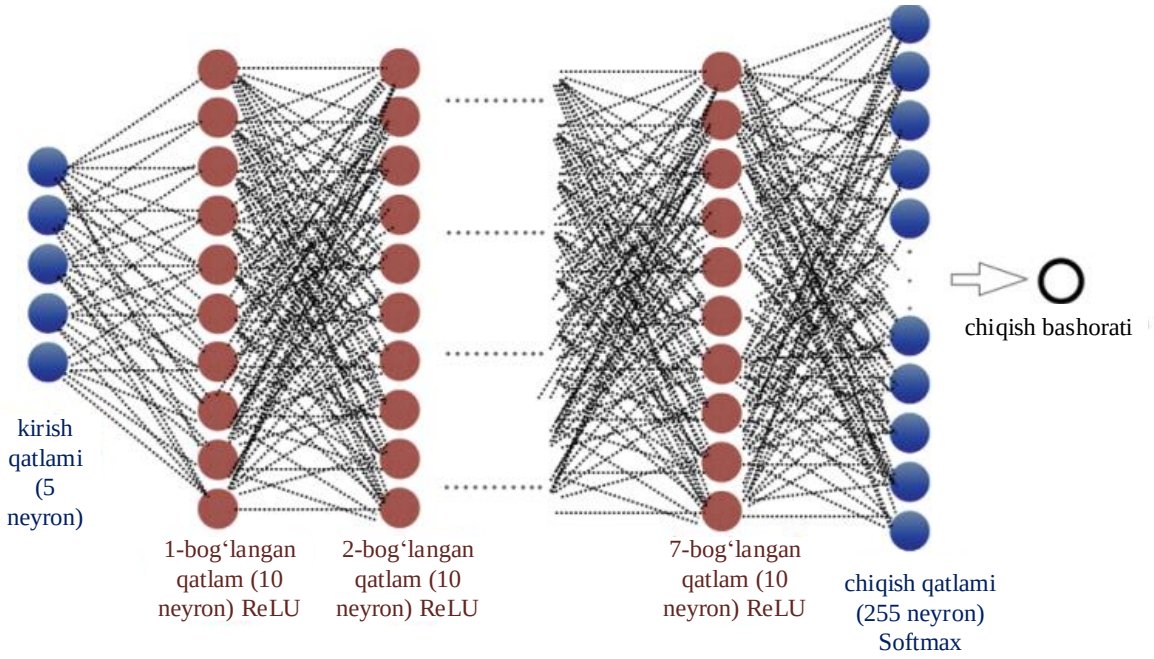
- **7 та зич (dense) қатлам**, ҳар бирида 10 та нейрон
- **Активация функцияси:** ReLU
- **Чиқариш қатлами:** 256 нейрон, активация — Softmax
- **Optimiser:** Adam
- **Йўқотиш функцияси:** categorical_crossentropy

Loss — ҳақиқий қиймат билан башорат орасидаги хатолик даражасини билдиради. Adam оптимизатори замонавий чуқур ўқитиш амалиётида жуда кенг қўлланилади. Categorical crossentropy эса кўп синфли классификация масалаларида қўлланиладиган стандарт йўқотиш функциясидир.

Ушбу масала назоратли (supervised) ўқитишга киради, бунда аввалдан тоифаси маълум бўлган намуналар асосида модель ўргатилади. Орқага тарқатиш (backpropagation) усули ёрдамида вазнлар оптималлаштирилади. Ўқитиш 5 epoch давомида бажарилди ва натижалар шу epochлар асосида баҳоланди.

Худди шу тарзда keystreamнинг дастлабки 10 байти учун алоҳида моделлар яратилди. Ушбу чуқур ўрганиш архитектураси 1-расмда келтирилган.

RC4A оқимли шифрлаш алгоритми бўйича таҳлил



1-расм. Экспериментда қўлланилган чуқурлашган ўрганиш (deep learning) тузилмаси

RC4A учун худди шундай экспериментал тартиб қўлланилди, бирок қуйидаги параметр фарқ қилади:

Бошланғич калит узунлиги (n) = 10 байт.

Бунинг сабаби — RC4A алгоритмида бошланғич калит иккита алоҳида калит компонентларидан ташкил топгани сабабли унинг узунлиги RC4никидан каттароқ бўлади.

Бундан ташқари, тажрибанинг барча бошқа жиҳатлари RC4 учун ишлатилган созламалар билан бир хил қолдирилди.

Trivium ва TRAID-SC оқимли шифрлагичлари бўйича таҳлил

Trivium ва TRAID-SC алгоритмларини баҳолашда:

- Ҳар бир шифр учун **2^{17} та keystream намунаси** генерация қилинди.
- Мақсад — keystream’нинг дастлабки **256 бити** учун башорат аниқлигини ҳисоблаш.
- Қўшимча равишда keystreamнинг дастлабки **32 байти** бўйича ҳам тажрибалар ўтказилди.

Ушбу мақолада оқимли шифрлаш алгоритмларини криптотаҳлил қилишда чуқур ўрганишга асосланган янги ёндашув ишлаб чиқилди ва унинг

самарадорлиги RC4, RC4A, Trivium ҳамда TRIAD-SC оқимли шифрлагичлари мисолида экспериментал тарзда баҳоланди. Таклиф этилган “Калит–Keystream” жуфтликларига асосланган модель оқимли шифрлагичларни “қора қути” (black-box) кўринишида таҳлил қилиш имконини беради ҳамда маълумотларнинг хусусиятларини олдиндан аниқлашга ҳожат қолдирмасдан, нейрон тармоқлар ёрдамида статистик оғишларни автоматик равишда аниқлашга хизмат қилади.

Тадқиқот натижалари RC4 оқимли шифрининг бир қатор байтларида сезиларли статистик оғишлар мавжудлигини кўрсатди. Ушбу оғишларнинг чуқур ўрганиш модели томонидан ишончли тарзда аниқланиши RC4 алгоритмининг ички механизмларида яширин заифликлар мавжудлигини яна бир бор тасдиқлайди. RC4A алгоритми бўйича ўтказилган тажрибалар эса такомиллаштирилган структура туфайли keystreamда сезиларли bias кузатилмаслигини намоён қилди. Trivium ва TRIAD-SC оқимли шифрлагичлари бўйича олинган натижалар ҳам ушбу алгоритмларда юқори даражадаги тасодифийлик сақланиб қолганини ва башорат қилиш эҳтимоли тасодифий қийматдан ошмаганини кўрсатди.

ФОЙДАЛАНИЛГАН АДАБИЁТЛАР:

1. Whitfield Diffie and Martin E Hellman. Privacy and authentication: An introduction to cryptography. Proceedings of the IEEE, 67(3):397–427, 1979.
2. S.W. Golomb. Shift Register Sequences. Holden-Day Series in Information Systems. Holden-Day, 1967.
3. El Groth. Generation of binary sequences with controllable complexity. IEEE Transactions on Information Theory, 17(3):288–296, 1971.
4. Edwin Key. An analysis of the structure and complexity of nonlinear binary sequence generators. IEEE Transactions on Information Theory, 22(6):732–736, 1976.
5. Vera S Pless. Encryption schemes for computer confidentiality. IEEE Transactions on Computers, 100(11):1133–1136, 1977.

6. Eli Biham and Adi Shamir. Differential fault analysis of secret key cryptosystems. In 17th International Cryptology Conference, Advances in Cryptology (CRYPTO 1997), volume 1294 of Lecture Notes in Computer Science, pages 513–525. Springer-Verlag, 1997.

7. Alex Biryukov and David Wagner. Slide attacks. In 6th International Workshop on Fast Software Encryption (FSE 1999), volume 1636 of Lecture Notes in Computer Science, pages 245–259. Springer-Verlag, 1999.

8. Andrey Bogdanov, Dmitry Khovratovich, and Christian Rechberger. Biclique cryptanalysis of the full AES. In 17th International Conference on the Theory and Application of Cryptology and Information Security, Advances in Cryptology (ASIACRYPT 2011), volume 7073 of Lecture Notes in Computer Science, pages 344–371. Springer-Verlag, 2011.

9. Dan Boneh, Richard A DeMillo, and Richard J Lipton. On the importance of checking cryptographic protocols for faults. In 16th International Conference on the Theory and Application of Cryptographic Techniques, Advances in Cryptology (EUROCRYPT 1997), volume 1233 of Lecture Notes in Computer Science, pages 37–51. Springer-Verlag, 1997.

10. Joan Daemen, Lars Knudsen, and Vincent Rijmen. The block cipher Square. In 4th International Workshop on Fast Software Encryption (FSE 1997), volume 1267 of Lecture Notes in Computer Science, pages 149–165. Springer-Verlag, 1997.