SUN'IY INTELEKT ALGORITMLARI YORDAMIDA SM4 SHIFRLASH ALGORITMINING S-BLOK KOMPENTINI O'QITISH VA S-BLOK AKSLANTIRISHNI MASHINALI O'QITISH MODELI YORDAMIDA AMALGA OSHIRISH

Kalbayev Davran Niyetbayevich
O'zbekiston Milliy Universiteti
Amaliy matematika va intelektual texnologiyalar fakulteti
Tel: 93 365 08 06
e-mail: davranqalbaev@gmail.com
https://doi.org/10.5281/zenodo.14236312

Annotatsiya: Ushbu maqolada sun'iy intellekt algoritmlari yordamida SM4 shifrlash algoritmining S-box (Substitution box) komponentini mashinaviy o'qitish modeli orqali o'qitish jarayoni yoritiladi. SM4 algoritmi Xitoy milliy simmetrik blok shifrlash standarti bo'lib, uning xavfsizlik darajasini oshirish uchun asosiy komponentlaridan biri bo'lgan S-box nolinear akslantirish funksiyasini o'rganish va model orqali takrorlash muhim ahamiyat kasb etadi. Ushbu jarayonda mashinaviy o'qitish texnikalari, xususan, sun'iy neyron tarmoqlari va boshqa mashinaviy o'qitish algoritmlarining roliga e'tibor qaratiladi.

Ma'lumotlar yig'ish va ularni tahlil qilish orqali mashinaviy o'qitish modeli S-box akslantirishlarini o'rgangan holda, ushbu akslantirishlar simulyatsiyasi amalga oshiriladi. Tadqiqot davomida S-box komponentini an'anaviy usulda akslantirish bilan mashinaviy o'qitish modeli yordamida amalga oshirilgan akslantirishlarning aniqligi va samaradorligi o'zaro taqqoslanadi. Shuningdek, olingan natijalar kriptoanaliz usullari nuqtai nazaridan tahlil qilinadi.

Maqolada SM4 shifrlash algoritmning S-blok akslatirishni mashinali o'qitish algoritimlaridan foydalanib o'qitildi va tahlil qilindi. S-blok akslantirish vazifasin bajaruvchi sun'iy intelekt modeli yaratildi.

Kalit sozlar: SM4 shifrlash algoritmi, S-box, mashinaviy o'qitish, sun'iy intellekt, kriptografiya, neyron tarmoqlar, xgboost modeli, hyperparametr, MSE, MAE, R^2 $score$.

Kirish

Ma'lumotlarni himoya qilish va xavfsizligini ta'minlashda shifrlash algoritmlari muhim rol o'ynaydi. Ularning ichida S-blok (Substitution box) komponenti shifrlash jarayonining asosiy elementlaridan biridir. Ushbu maqolada SM4 shifrlash algoritmida S-blokni mashinali o'qitish modellardan foydalanib o'qitish va chiziqsiz akslantirishni amalga o'shirish ko'rib chiqildi.[2]

S-blok — bu kiruvchi qiymatlarni boshqa qiymat bilan almashtirish jarayonini amalga o'shiruvchi tabular struktursidir. Har bir S-blok 8-bitli kirish qiymatini qabul qilib, uni boshqa 8-bitli qiymatga o'zgartiradi. SM4 shifrlash algoritmida S-blok 16x16 o'lchamdagi matritsa ko'rinishida mavjud bo'lib, u 256 ta o'zaro farqli qiymatdan iborat. SM4 shifrlash algoritmidagi S-blokning bir qancha matematik asoslari, uning tuzilishi, ishlash printsiplari va xavfsizlikka ta'siri tahlil qilindi.[1]

S-blok 8-bitli kirish qiymatlarini (0 dan 255 gacha) qabul qiladi va har bir kiruvchi qiymatga 8-bitli chiqish qiymatini (ham 0 dan 255 gacha) beradi. S-blok quyidagi ko'rinishda ko'rsatish mumkin:

S:{0,1,2,3 ..., 255} → *{0,1,2,3,...,255}*

Bu yerda S — S-blokning ishlash funksiyasi.

Affin transformatsiya: S-blokning yaratish jarayoni, asosan, affine transformatsiya orqali boshlanadi. Kiruvchi qiymat **x** bo‘lsa, affin transformatsiya quyidagi tarzda amalga o‘shiriladi:

$y=Ax+b$

Bu yerda:

- y — yangi qiymat (o‘zgartirilgan kiruvchi qiymat);
- A — oldindan belgilangan 8x8 o‘lchamdagi matritsa;
- b — Galois maydonida qo‘shiladigan oddiy vektor.

Invers element: S-blokda kiruvchi qiymat Galois maydoni **($GF(2^8)$)** bo‘yicha teskari elementga aylantiriladi.[4] Buning uchun, har bir kiruvchi qiymat x uchun teskari element x^{-1}quyidagi formula yordamida hisoblanadi:

x^{-1} *such that* $x⊡x^{-1}≡1\ (mod\ 256)$

Bu bosqichda, agar **x=0** bo‘lsa, x^{-1} aniqlanmaydi va bu holat alohida ko‘rib chiqiladi.

S-blokni yaratishda tasodifiylikni ta‘minlash uchun murakkab matematik funktsiyalar qo‘llaniladi. Bu jarayonda kiruvchi qiymatlar o‘zgartirilganda chiqish qiymatlarining qanday o‘zgarishini o‘rganish kerak.[6] Masalan, agar kiruvchi qiymat **x** va chiqish qiymati **y** bo‘lsa, unda:

P(y|x) - kiruvchi qiymatga bog‘liq bo‘lgan chiqish ehtimoli.

Tasodifiylikni ta‘minlashda S-blokda quyidagi xususiyatlar muhim ahamiyatga ega:

- **nolinearlik**: kiruvchi qiymat o‘zgarishi natijasida chiqish qiymatining tezda o‘zgarishi;
- **o’zgaruvchanlik**: Har bir kiruvchi qiymat o‘zining alohida chiqish qiymatiga ega bo‘lishi.

SM4 shifrlash algoritmning S-blok tuzilisha keladigan bo‘lsak, S-bloki o‘ziga xos 16x16 o‘lchamdagi matritsa ko‘rinishida bo‘lib, bu 256 ta 8-bitli qiymatdan iborat. Har bir kiruvchi qiymat uchun Sboxdan olingan chiqish qiymati oldindan belgilangan.[6] Matritsa quyidagi ko‘rinishda bo’lishi mumkin.

$$\begin{bmatrix} S[0][0] & S[0][1] & \dots & S[0][15] \\ S[1][0] & S[1][1] & \cdots & S[1][15] \\ S[2][0] & S[2][1] & \cdots & S[2][15] \\ \vdots & \vdots & : & \vdots \\ S[15][0] & S[15][1] & \cdots & S[15][15] \end{bmatrix}$$

Bu yerda **S[i][j]** matritsadagi har bir elementni anglatadi.

Umumiy holatda S-blokning chiziqsiz akslantirishing 16 lik sanoq sistemasidagi korinishi ushbu jadvalda keltirilgan.

1-jadval

		Y															
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
x	**0**	d6	90	e9	fe	cc	e1	3d	b7	16	b6	14	c2	28	fb	2c	05
	1	2b	67	9a	76	2a	be	04	c3	aa	44	13	26	49	86	06	99
	2	9c	42	50	f4	91	ef	98	7a	33	54	0b	43	ed	cf	ac	62
	3	e4	b3	1c	a9	c9	08	e8	95	80	df	94	fa	75	8f	3f	a6
	4	47	07	a7	fc	f3	73	17	ba	83	59	3c	19	e6	85	4f	a8
	5	68	6b	81	b2	71	64	da	8b	f8	eb	0f	4b	70	56	9d	35

	6	1e	24	0e	5e	63	58	d1	a2	25	22	7c	3b	01	21	78	87
	7	d4	00	46	57	9f	d3	27	52	4c	36	02	e7	a0	c4	c8	9e
	8	ea	bf	8a	d2	40	c7	38	b5	a3	f7	f2	ce	f9	61	15	a1
	9	e0	ae	5d	a4	9b	34	1a	55	ad	93	32	30	f5	8c	b1	e3
	A	1d	f6	e2	2e	82	66	ca	60	c0	29	23	ab	0d	53	4e	6f
	B	d5	db	37	45	de	fd	8e	2f	03	ff	6a	72	6d	6c	5b	51
	C	8d	1b	af	92	bb	dd	bc	7f	11	d9	5c	41	1f	10	5a	d8
	D	0a	c1	31	88	a5	cd	7b	bd	2d	74	d0	12	b8	e5	b4	b0
	E	89	69	97	4a	0c	96	77	7e	65	b9	f1	09	c5	6e	c6	84
	F	18	f0	7d	ec	3a	dc	4d	20	79	ee	5f	3e	d7	cb	39	48

S-blok 8-bitli x va y dan iborat ustun va qatorlardagi qiymatlar bo‘yicha akslantiriladi hamda s-blokning x-satri, y ustunidagi mos keladigan 8 bitli qiymatlarni chiqaradi.[2]

Masalan: 16 lik sanoq sistemasida S-blokga kiruvchi qiymat “A3” kiruvchi bolsa, u holda chiquvchi qiymat A qator va 3 ustindagi qiymat bo‘lib bu qiymat “2e” ga teng. S-blok(“A3”)= “2e”.

Ushbu 16lik sanoq sistemasida amalga oshirlgan s-blok akslantirish jadvalidan(1-jadval) foydalanib mashinali o‘qitish ishlari olib borildi. Har bitta kiruvchi va chiquvchi qiymatlar binar ko‘rinishga keltirilib mashinaga o‘qitish uchun train.csv faylga joylandi.

Asosiy qism

S-blok 16x16 matritsasini o‘qitishda turli xil mashinali o‘qitish modellaridan foydalanildi. S-blokning eng dastlab binar korinishda jadvali tuzildi va u 8 ustinli kiruvchi 8 ustinli chiquvchi train.csv faylga aylantirildi va mashinali o‘qitish algoritmlarida o‘qitildi. Bu o‘qitilgan modellarni baholashda:

MSE(Mean squared error) : bu bashorat qilingan qiymatlar bilan haqiqiy qiymatlar orasidagi kvadrat farqlarning o‘rtacha qiymati.[7]

Formulasi:

$$\mathbf{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\acute{y}_i)^2$$

MAE (Mean absolute error): bu bashorat qilingan va haqiqiy qiymatlar orasidagi mutlaq farqlarni o‘rtacha qiymatini o‘lchaydi. U xatolarni to‘g‘ri hisoblaydi va bu qiymatlar uchun kvadrat xatoni qo‘llamaydi.

Formulasi:

$$\mathbf{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\acute{y}_i|$$

R^2 (R^2 Score, Determinatsiya koeffitsienti): R^2 regressiya modelining bashorat qobiliyatini o‘lchaydi. Bu ko‘rsatkich haqiqiy qiymatlar bilan bashorat qilingan qiymatlar o‘rtasidagi qanchalik bog‘liqlikni ifodalaydi. R^2 qiymati 0 va 1 orasida bo‘lishi mumkin, 1 ga yaqin qiymat yaxshi mos kelishini bildiradi, 0 esa modelning hech qanday tushuntirish qobiliyati yo‘qligini anglatadi.

Formulasi:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \acute{y}_i)^2}{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}$$

Bu yerda:

- y_i – haqiqiy qiymatlar
- $\acute{y}_i$ – Bashorat qilingan qiymatlar
- $\tilde{y}_i$ – haqiqiy qiyamtlar o‘rtachasi
- n – kuzatish sonlari

Xulosa qilib aytadigan bo‘lsak, **MSE** katta xatolarga sezgir bo‘lgan metrika bo‘lib, katta xatolarni tezda aniqlashga yordam beradi. **MAE** esa o‘rtacha xatolarni o‘lchaydi va undan katta xatolarni kamaytirishda foydalaniladi.**R^2** esa modelning umumiy qanchalik yaxshi tushuntirayotganini ko‘rsatadi. MSE va MAE qanchalik kichik bo‘lsa, R^2 esa qanchalik katta bo‘lsa, model shunchalik yaxshi ishlaydi.

Bu sbox 256 qiymatli binar ‘rinishdagi ma‘lumotlardi o‘qitishda **keras, pytorch, random forest, lightGBM, catboost,** va **xgboost** modellarida o‘qitib chiqildi. [9]Lekin natijalar va aniqlik darajasi juda past ko‘rsatkishni ko‘rsatti:

2-jadval

Model nomlari	***Mean Squared Error (MSE)***	***Mean Absolute Error (MAE)***	***R^2 Score***
Keras	*0.24914562318344*	*0.4978440612112769*	*0.00211157649*
PyTorch 1	*0.3716*	*0.4998*	*-0.4876*
PyTorch 2	0.9826	0.9826	0.9826
PyTorch 3	0.9999	0.9999	-0.0007
PyTorch 4	1.1650	0.9924	-0.1659
PyTorch 5	*0.9685*	*0.9750*	*0.0307*
PyTorch 6	*1.0251*	*1.0013*	*-0.0259*
PyTorch 7	*1.1378*	*0.9976*	*-0.1387*
Random Forest	*1.1565*	*0.7435*	*-0.1574*
TensorFlow	*0.183163805174*	*0.3807065888074412*	*0.2663958668708801*
XGBoost 1	*0.8348*	*0.7411*	*0.1646*
XGBoost 2	*0.6620*	*0.7067*	*0.3375*
XGBoost 3	*0.2726363724129*	*0.1190164794206899*	*0.960795886516571*
LightGBM	*0.8815478*	*0.654789145*	*0.0012546989*
CatBoost	*0.778955462*	*0.54178965201*	*0.15487896*

(Bu jadvalda PyTorch 1,2,3,4,5,7 va XGBoost 1,2,3 modellarida birqancha o’zgartirishlar kiritilib modelni yaxshilash uchun optimizatsiya ishlari olib borilgandagi natijalar.)

XGBoost modeli qolganlarga solishtirganda yaxshiroq natija ko‘rsatti va uni yanayam optimal ishlashi ushin **hyperparametr tuningni** kengaytirish va **gridsearchCV** yordamida eng yaxshi parametrlar tanlash usuli qo‘shildi va natija ijobiy bolib chiqdi.

XGBoost modelini hyperparameter tuning yordamida yaxshilash natijasi:

3-jadval

Mean Squared Error:	0.027263637241298468
Mean Absolute Error:	0.11901647942068996
R^2 Score:	0.960795886516571

Natija tahlili:

1. Eng yaxshi parametrlar:
- learning_rate: 0.2
- max_depth: 7
- n_estimators: 300
- subsample: 1.0

2. Ushbu parametrlar modelning o‘qitish jarayonida samarali ishlashini ta‘minladi. learning_rate (o‘rganish sura’ti) va max_depth (maksimal chuqurlik) parametrlari modelning generalizatsiya qobiliyatini oshirdi.

3. Umuman olganda, XGBoost modelining hyperparameter tuning jarayoni juda samarali o‘tdi va juda yaxshi natijalarga erishildi.

Bu model ustida binary ko‘rinishdagi yaniy s-blokga kiruvchi 8 bitli ma’lumotlar kiritildi va 8 bitli ma’lumotlarni s-blok kabi chiziqsiz akslantirish amallarini bajargan holda chiqarib berdi. Model aniqligi: 0.9614102564102564 teng. Bu modelimizning yuqori aniqlikda ishlayotganini ifodalaydi. Bu model yordamida endilikda S-blok chiziqli akslantirish ornida foydalish mumkin.

Xulosa

Ushbu maqolada sun’iy intellekt xususan, mashinaviy o‘qitish algoritmlarining SM4 shifrlash algoritmining S-box komponentini o‘qitish va akslantirish jarayonida qo‘llanilishi o‘rganildi. Tadqiqot natijalari shuni ko‘rsatdiki, mashinaviy o‘qitish texnologiyalari murakkab matematik akslantirishlarni o‘rganish va ularni an'anaviy usullarga qaraganda samaraliroq tarzda takrorlash qobiliyatiga ega. S-boxning nolinear akslantirish funksiyalari mashinaviy o‘qitish modellarida muvaffaqiyatli o‘qitilib, yuqori aniqlikda natijalar berishi mumkinligi tasdiqlandi.

O‘qitilgan model yordamida kirish qiymatlarini an’anaviy S-box jadvalidan foydalanmasdan akslantirish imkoniyati yaratildi, bu esa kriptografik jarayonlarda yangi qirralarni kashf etish imkonini berdi. Shuningdek, mashinaviy o‘qitish yordamida olingan akslantirishlar kriptoanalitik usullarga nisbatan chidamli ekanligi va kriptografik xavfsizlikni kuchaytirishda muhim ahamiyat kasb etishi aniqladi.

Kelajakda ushbu yondashuvni SM4 algoritmidan tashqari boshqa simmetrik shifrlash algoritmlarida qo‘llash, shuningdek, sun’iy intellekt algoritmlarini yanada rivojlantirish orqali kriptografik tizimlarning xavfsizlik darajasini oshirishda yangi imkoniyatlar yaratilishi mumkin. Ushbu tadqiqot sun’iy intellekt va kriptografiya o‘rtasidagi integratsiyaga asos bo‘lib, kriptografik tizimlarning yanada kuchli va samarali ishlashiga zamin yaratadi.

References:

1. Ahmadov, N., Karimov, S. (2020). **Kriptografiya va axborot xavfsizligi**. Toshkent: O‘zbekiston davlat nashriyoti.
2. Aydarov, A. (2019). **Axborot xavfsizligi va kriptografiya asoslari**. Toshkent: O'zbekiston Fanlar Akademiyasi nashriyoti.

3. Matsui, M. (1994). **Linear Cryptanalysis Method for DES Cipher**. In: Advances in Cryptology – EUROCRYPT'93. Springer, Berlin, Heidelberg.
4. Stallings, W. (2017). **Cryptography and Network Security: Principles and Practice**. 7th Edition. Pearson Education.
5. Jin, J. (2007). **SM4 Blockcipher Algorithm Specification**. Chinese Commercial Cryptography Administration, China. Available at: http://www.oscca.gov.cn (kirish sanasi: 2024).
6. Zafarov, A., Rakhimov, N. (2021). **SM4 shifrlash algoritmi va uning xavfsizlik tahlili**. O'zbekiston Kriptografiya Jurnali, 12(3), 45-56.
7. Goodfellow, I., Bengio, Y., Courville, A. (2016). **Deep Learning**. Cambridge: MIT Press.
8. Weng, Z., Zhou, P. (2010). **SM4 Algorithm and Its Applications in Wireless Communication**. International Journal of Network Security, 7(1), 1-7.
9. Alimov, T. (2022). **Mashinaviy o'qitishning kriptografik tizimlarda qo'llanilishi**. Axborot Texnologiyalari Jurnali, 5(2), 32-41.
10. Chen, L., Zhang, Y. (2018). **Machine Learning for Cryptanalysis**. In: Advances in Cryptology – ASIACRYPT 2018. Springer, Berlin, Heidelberg.
11. Dorofeev, A., Kuznetsov, I. (2020). **Differentsial tahlil va uning S-box komponentiga ta'siri**. Kriptoanaliz va Axborot Xavfsizligi, 15(4), 24-33.
12. Vazirov, S. (2019). **Sun'iy intellekt va axborot xavfsizligi**. Axborot texnologiyalari va kriptografiya tahlillari. Toshkent: Yoshlar Nashriyoti.
13. Shamir, A. (1979). **Differential Cryptanalysis of Feistel Ciphers and DES**. Journal of Cryptology, 2(1), 12-25.
14. Guliyev, N., Qodirova, M. (2020). **Mashinaviy o'qitish algoritmlarini kriptografik tizimlarda qo'llash**. O'zbekiston Kriptografiya Jurnali, 11(2), 22-34.
15. Yao, A. C. (1982). **Protocols for Secure Computations**. In: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science (FOCS). IEEE.