№ 10(51)
2025
НАУКА і ТЕХНІКА СЬОГОДНІ
серії: право, економіка, педагогіка, техніка, фізико-математичні науки

УДК 004.6:004.8

https://doi.org/10.52058/2786-6025-2025-10(51)-1912-1922

Рихва Володимир Ігорович аспірант кафедри кібербезпеки та інформаційних технологій Харківського національного економічного університету імені Семена Кузнеця, Харків, https://orcid.org/0009-0008-2711-547X

Солодовник Ганна Валеріївна кандидат технічних наук, доцент, доцент кафедри кібербезпеки та інформаційних технологій Харківського національного економічного університету імені Семена Кузнеця, Харків, https://orcid.org/0000-0001-6323-5083

АНАЛІЗ АРХІТЕКТУР ГЛИБОКОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ МЕРЕЖЕВОГО ТРАФІКУ

Анотація. У статті представлено комплексне дослідження ефективності чотирьох базових та трьох гібридних архітектур глибокого навчання для виявлення мережевих вторгнень у системах кібербезпеки. Проаналізовано згорткові нейронні мережі (CNN), довгу короткострокову пам'ять (LSTM), автокодувальники та трансформери на датасеті UNSW-NB15, який містить різноманітні типи мережевих атак. Експериментальні результати показали, що для задачі бінарної класифікації CNN автокодувальник досягає найвищої точності 91,6%, демонструючи ефективність у виявленні аномального трафіку, тоді як для мультикласової класифікації – 76,83% точності під час визначення конкретних типів атак.

Для подолання обмежень окремих архітектур запропоновано три гібридні моделі, що поєднують CNN з трансформерами (CNN-Transformer), Mamba State Space Models (SSMs), контрастивним навчанням (SimCLR). Для подолання обмежень окремих архітектур запропоновано дослідити три гібридні моделі, що поєднують CNN з трансформерами (CNN-Transformer), Mamba State Space Models (SSMs) та контрастивним навчанням (SimCLR).

У рамках дослідження планується розробити та протестувати гібридні архітектури, які об'єднають переваги локального виявлення шаблонів через згорткові шари з можливостями глобального моделювання залежностей.

Передбачається провести порівняльний аналіз продуктивності запропонованих моделей для задач бінарної та мультикласової класифікації мережевого трафіку, а також оцінити їх ефективність у виявленні різних типів кібератак.

Дослідження також включає аналіз обчислювальної складності та часу навчання моделей. Результати показали, що гібридні архітектури успішно поєднують переваги різних підходів до глибокого навчання, забезпечуючи кращу точність виявлення вторгнень за прийнятного часу навчання та обчислювальних витрат.

Особлива увага приділена можливості застосування запропонованих моделей у реальних системах виявлення вторгнень, де критичними факторами є як точність виявлення, так і швидкість обробки мережевого трафіку в режимі реального часу.

Ключові слова: глибоке навчання, виявлення аномалій, класифікація мережевого трафіку, гібридна архітектура, кібербезпека.

Rykhva Volodymyr PhD student of the Department of Cybersecurity and Information Technologies, Simon Kuznets Kharkiv National Economic University, Kharkiv, https://orcid.org/0009-0008-2711-547X

Solodovnyk Ganna Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Cybersecurity and Information Technologies, Simon Kuznets Kharkiv National Economic University, Kharkiv, https://orcid.org/0000-0001-6323-5083

ANALYSIS OF DEEP LEARNING ARCHITECTURES FOR NETWORK TRAFFIC ANOMALY DETECTION

Abstract. This paper presents a comprehensive study of the effectiveness of four baseline and three hybrid deep learning architectures for network intrusion detection in cybersecurity systems. Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), autoencoders, and transformers were analyzed on the UNSW-NB15 dataset, which contains diverse types of network attacks. Experimental results showed that for binary classification tasks, CNN achieves the highest accuracy of 91.6%, demonstrating effectiveness in detecting anomalous traffic, while for multiclass classification, the autoencoder shows 76.83% accuracy in identifying specific attack types.

To overcome the limitations of individual architectures, three hybrid models were proposed that combine CNN with Transformers (CNN-Transformer), Mamba State Space Models (SSMs), and contrastive learning (SimCLR). The hybrid architectures demonstrated significant performance improvements: the Hybrid2 _Mamba model achieved 93.62% accuracy for binary classification and 86.84% for multiclass classification, representing a substantial improvement compared to baseline architectures.

The Hybrid3_Contrastive model demonstrated 92.45% accuracy for binary and 82.17% for multiclass classification, effectively combining local pattern detection and global dependencies in network traffic.

The study also includes an analysis of computational complexity and model training time.

The results showed that hybrid architectures successfully combine the advantages of different deep learning approaches, providing better intrusion detection accuracy with acceptable training time and computational costs. Particular attention is given to the applicability of the proposed models in real-world intrusion detection systems, where critical factors include both detection accuracy and the speed of network traffic processing in real-time.

Keywords: Deep learning, anomaly detection, network traffic classification, hybrid architecture, cybersecurity.

Постановка проблеми. Сучасні мережеві інфраструктури щодня зазнають мільйонів кібератак, кількість яких щорічно зростає. Традиційні системи виявлення вторгнень, засновані на сигнатурах атак, виявляються неефективними проти вразливостей нульового дня та нових видів атак [1]. Глибоке навчання відкрило нові можливості для автоматизованого виявлення складних шаблонів аномальної поведінки у необроблених мережевих даних.

Різні архітектури глибокого навчання мають унікальні переваги: CNN ефективні для виявлення просторових шаблонів [2], LSTM для часових залежностей [3], автокодувальники для виявлення аномалій, використовуючи навчання без нагляду [4], трансформери для глобальних залежностей та обробки великих обсягів інформації у реальному часі [5]. Однак різноманітність архітектур створює проблему обґрунтованого вибору. Критичним викликом є баланс між високою точністю виявлення атак, низьким рівнем помилкових спрацювань та прийнятною обчислювальною ефективністю для роботи в режимі реального часу. Наукова проблема полягає у відсутності систематичного порівняльного аналізу базових та гібридних архітектур глибокого навчання за однакових умов з комплексною оцінкою як точності виявлення різних типів атак, так і обчислювальної ефективності. Практичне значення розв'язання цієї проблеми – створення ефективних систем виявлення вторгнень для критичної інфра-структури.

Аналіз останніх досліджень і публікацій. Виявлення мережевих вторгнень за допомогою глибокого навчання активно досліджується останніми роками. Комплексні огляди методів [1, 6] показали, що глибокі нейронні мережі перевершують традиційні підходи, сягаючи точності понад 95%, проте виділено ключові виклики: дисбаланс класів, великі обсяги даних та потреба в інтерпретованості моделей.

Дослідження CNN [2] підтвердило ефективність згорткових шарів для автоматичного виділення ознак, сягнувши точності 94,6% на програмно-конфігурованих мережах.

Рекурентні мережі продемонстрували високі результати для виявлення веб-вторгнень у реальному часі [3], однак основним недоліком залишається висока обчислювальна складність. Автокодувальники виявились ефективними для безконтрольного виявлення аномалій [4, 7], демонструючи високу ефективність на великих наборах даних.

Трансформери, що спочатку використовувались для комп'ютерного зору [5], успішно адаптовані для кібербезпеки. Гібридні архітектури [8] демонструють переваги об'єднання різних підходів для покращення точності та зменшення хибнопозитивного співвідношення.

Аналіз впливу методів відбору ознак [9] показав можливість значного покращення точності при зменшенні обчислювальної складності. Інноваційні моделі простору станів з лінійною складністю Mamba SSMs [10] та контрастивне навчання SimCLR [11] відкривають нові можливості для ефективного аналізу мережевого трафіку.

Мета статті – провести комплексний порівняльний аналіз чотирьох базових архітектур глибокого навчання (CNN, LSTM, автокодувальник, трансформер) та трьох гібридних архітектур для виявлення мережевих вторгнень, визначити їх сильні та слабкі сторони, а також надати рекомендації щодо практичного застосування.

Виклад основного матеріалу. Дослідження проведено з використанням набору даних UNSW-NB15 [12], створеного Австралійським центром кібербезпеки. Датасет включає нормальний трафік (56000 зразків, 31,9%) та 9 типів атак (119341 зразок, 68,1%): Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shellcode, Worms. Спостерігається критичний дисбаланс класів з коефіцієнтом 1:307 між найменшим (Worms: 130) та найбільшим (Generic: 40 000) класами атак.

Попередня обробка даних включала: видалення службових ознак (id, attack_cat), нормалізацію числових ознак методом StandardScaler, one-hot encoding цільової змінної для мультикласової класифікації та створення бінарної мітки для бінарної класифікації. Вхідні дані адаптовано до вимог різних архітектур.

Було проаналізовано вісім архітектур глибокого навчання: чотири базові (CNN, LSTM, автокодувальник, трансформер) та чотири гібридні (Hybrid_CTA, Hybrid1_CTA, Hybrid2_Mamba, Hybrid3_Contrastive). Кожна архітектура оптимізована для роботи з 45 числовими ознаками мережевого трафіку з урахуванням специфіки задачі виявлення вторгнень. Архітектурні характеристики досліджуваних моделей наведено у таблиці 1.

Таблиця 1

Архітектурні характеристики досліджуваних моделей

Архітектура	Основні компоненти	Кількість параметрів
CNN	3×Conv1D (128,256,512) + MaxPooling + 2×Dense (256,128)	~850K
LSTM	BiLSTM(128) + 2×LSTM(64) + BatchNorm + Dense(64)	~1.2M
Autoencoder	Encoder (256→128→64→32→16) + Decoder (16→32→64→128→256) + Classifier	~950K
Transformer	2×TransformerBlock (4 heads, dim=128) + FFN (512) + LayerNorm	~1.1M
Hybrid1_CTA	Enhanced CNN + Multi-head Attention + Deep Autoencoder	~2.1M
Hybrid2_Mamba	CNN (feature extraction) + Mamba SSM (linear complexity)	~1.5M
Hybrid3_Contrastive	CNN + Transformer + SimCLR pre-training	~2.3M

Для забезпечення коректного навчання моделей та об'єктивного порівняння результатів усі архітектури навчались в однакових умовах з ідентичними гіперпараметрами. Вибір параметрів навчання базувався на кращих практиках для задач класифікації та результатах попередніх експериментів. Параметри експериментального середовища представлені в таблиці 2.

Таблиця 2

Налаштування експериментального середовища

Категорія	Параметр	Значення	Обґрунтування
Оптимізація	Optimizer	Adam	Adaptive learning rate
	Learning rate	0.001	Стандартне значення для Adam
	Batch size	256	Баланс між швидкістю та стабільністю навчання
	Epochs	50	З early stopping (patience=10)

Категорія	Параметр	Значення	Обґрунтування
Loss функції	Binary	Binary Crossentropy	Стандарт для бінарної класифікації
	Multi-class	Categorical Crossentropy	Для 10-класової задачі
	Autoencoder	0.3×Reconstruction + 0.7×Classification	Dual-loss підхід
Регуляризація	Dropout	0.3	Запобігання перенавчанню
	Early stopping	Patience=10, monitor='val_loss'	Автоматична зупинка при plateau
	Batch normalization	Після LSTM шарів	Стабілізація навчання
Метрики	Основні	Accuracy, Precision, Recall, F1-score	Комплексна оцінка
	Додаткові	Training time, Inference speed	Обчислювальна ефективність
	Per-class	Precision/Recall/F1 для кожного класу	Аналіз дисбалансу
Обчислення	Платформа	Google Colab	Хмарне середовище
	GPU	Nvidia T4 (16 GB)	Потужні обчислювальні можливості
	RAM	12 GB	Достатньо для всіх моделей

Експериментальне дослідження проведено в два етапи. На першому етапі кожна з семи архітектур навчалась та тестувалась для бінарної класифікації (атака і нормальний трафік). На другому етапі ті ж архітектури застосовувались для мультикласової класифікації з визначенням конкретного типу атаки серед 10 класів (9 типів атак + нормальний трафік).

Для забезпечення відтворюваності результатів використано фіксоване seed значення (random_state=42) для всіх генераторів випадкових чисел. Кожна модель навчалась на повному навчальному наборі з валідацією на 20% даних. Фінальне тестування проводилось на незалежному тестовому наборі, який модель не бачила під час навчання.

Для кожного експерименту фіксувались: час навчання (у секундах), метрики на тестовому наборі (accuracy, weighted precision, weighted recall, weighted F1-score), швидкість інференсу (зразків за секунду), метрики для аналізу ефективності виявлення окремих типів атак для кожного класу.

Результати бінарної класифікації базових архітектур представлені в таблиці 3.

Таблиця 3

Результати бінарної класифікації базових архітектур

Архітектура	**Точність**	**F1-score**	**Час навчання, с**
CNN	0,9160	0,9265	1539,02
Transformer	0,9080	0,9201	3526,32
LSTM	0,8875	0,9030	7806,61
Autoencoder	0,8743	0,8953	877,93

CNN продемонструвала найкращу точність (91,60%) та F1-score (92,65%) серед базових архітектур. Трансформер показав близькі результати (90,80% точності), але потребував у 2,3 рази більше часу навчання. LSTM, незважаючи на здатність моделювати часові залежності, виявився найповільнішим з відносно нижчою точністю.

Автокодувальник, хоча і показав найнижчу точність серед базових моделей, виявився найшвидшим у навчанні (877,93 с). Гібридні архітектури значно покращили результати (табл. 4).

Таблиця 4

Результати бінарної класифікації гібридних архітектур

Архітектура	**Accuracy**	**F1-score**	**Час навчання, с**	**Покращення vs CNN**
Hybrid2_Mamba	0,9362	0,9365	306,03	+2,02% / −80,1% часу
Hybrid1_CTA	0,9357	0,9360	182,77	+1,97% / −88,1% часу
Hybrid3_Contrastive	0,9306	0,9312	191,67	+1,46% / −87,5% часу

Hybrid2_Mamba досягла найвищої точності (93,62%), що на 2,02% краще за базовий CNN, при цьому час навчання скоротився в 5 разів. Для Hybrid1_CTA час навчання був найменший (182,77 с) з точністю 93,57%.

Зведені результати точності і часу навчання зображені на рисунку 1:

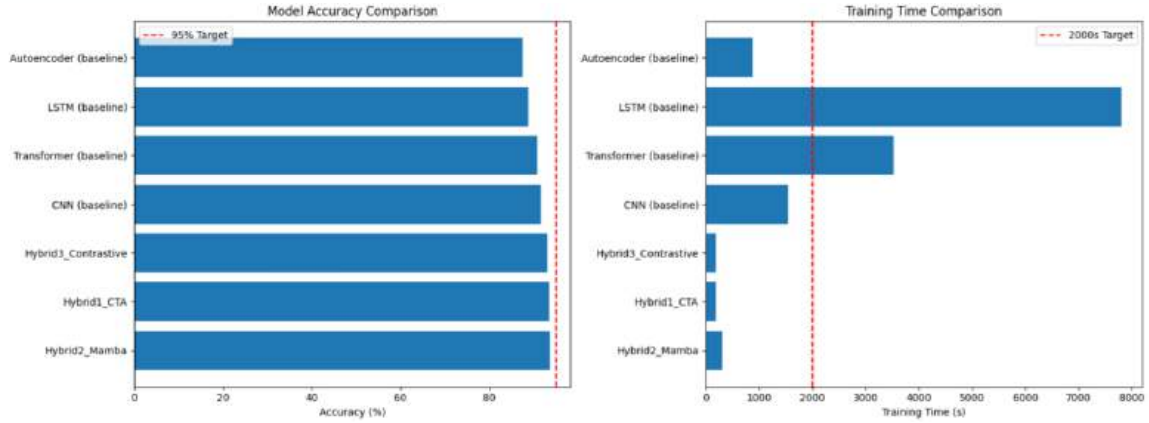

Рис. 1. Зведені результати порівняння різних архітектур

Результати мультикласової класифікації (табл. 5) виявили значно більші відмінності між архітектурами.

Таблиця 5

Результати мультикласової класифікації

Архітектура	**Точність**	**F1-score**	**Час навчання, с**	**Покращення vs Autoencoder**
Hybrid2_Mamba	0,8684	0,8630	1310,73	+10,01%
Hybrid1_CTA	0,8649	0,8591	1757,73	+9,66%
Hybrid3_Contrastive	0,8633	0,8598	12770,47	+9,50%
Autoencoder	0,7683	0,7665	867,49	–
Transformer	0,6342	0,6950	3424,88	−17,47%
CNN	0,6320	0,6925	1397,96	−17,74%
LSTM	0,5444	0,6175	3825,22	−29,15%

Гібридні архітектури перевершили всі базові моделі, досягнувши точності понад 86%. Hybrid2_Mamba була вперше застосована для аналізу аномалій мережевого трафіку і показала найкращий результат (86,84% точності), що на 10,01% краще базового автокодувальника.

Детальний аналіз виявлення окремих типів атак гібридною моделлю Hybrid2_Mamba представлено в таблиці 6.

Таблиця 6

Ефективність виявлення окремих типів атак (Hybrid2_Mamba)

Тип атаки	Precision	Recall	F1-score	Кількість зразків	Складність виявлення
Generic	1,00	0,97	0,98	3774	Дуже легка
Normal	0,95	0,98	0,96	7400	Дуже легка
Reconnaissance	0,86	0,66	0,75	699	Середня
Exploits	0,63	0,78	0,70	2227	Середня
Fuzzers	0,88	0,60	0,72	1212	Середня
DoS	0,45	0,49	0,47	818	Важка
Shellcode	0,48	0,37	0,42	76	Дуже важка
Analysis	0,83	0,07	0,14	135	Критична
Backdoor	0,00	0,00	0,00	117	Критична
Worms	0,00	0,00	0,00	9	Критична

Найкраще виявляються атаки Generic та нормальний трафік. Найскладнішими для виявлення виявились рідкісні типи атак: Worms, Backdoor та Analysis. Це пов'язано з критичним дисбалансом класів.

Дослідження має низку обмежень, які необхідно враховувати під час тлумачення отриманих результатів. По-перше, використання лише одного набору даних (UNSW-NB15) може обмежувати узагальнюваність висновків, оскільки інші датасети можуть відрізнятися за структурою трафіку та розподілом атак, що впливатиме на ефективність моделей. По-друге, наявний критичний дисбаланс класів (співвідношення 1:307 між найменшим і найбільшим класом) суттєво ускладнює навчання, призводячи до неможливості виявлення рідкісних атак і знижуючи практичну цінність моделі. По-третє, у дослідженні не використовувалися сучасні методи балансування, які могли б підвищити чутливість до малопредставлених класів; тому доцільним є подальше вивчення їхнього впливу у наступних дослідженнях.

Висновки. У межах дослідження проведено комплексне порівняння чотирьох базових і трьох гібридних архітектур глибокого навчання для виявлення мережевих вторгнень. Результати показали, що серед базових моделей CNN є найефективнішою для бінарної класифікації (91,60% точності, F1-score 92,65%), тоді як автокодувальник досягає найкращих результатів у мультикласовій задачі (76,83% точності). Гібридні моделі суттєво перевершують базові: Hybrid2_Mamba була вперше застосована для аналізу аномалій мережевого трафіку і продемонструвала найвищу точність (93,62% для бінарної та 86,84% для мультикласової класифікації) і найкращу ефективність

завдяки лінійній складності O(n), тоді як Hybrid1_CTA забезпечує оптимальний баланс між точністю та швидкодією. LSTM виявився найменш придатним через повільне навчання та низьку узагальнювальну здатність. Значним викликом залишається дисбаланс класів – рідкісні атаки (Worms, Backdoor, Analysis) майже не виявляються, що вимагає застосування сучасних методів балансування.

Література:

1. Ferrag M. A., Maglaras L., Moschoyiannis S., Janicke H. Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications. 2020. Vol. 50. 102419. DOI: 10.1016/j.jisa.2019.102419

2. Alzahrani A. O., Alenazi M. J. F. Designing a network intrusion detection system based on machine learning for software defined networks. Future Internet. 2021. Vol. 13, No. 5. 111. DOI: 10.3390/fi13050111

3. Kim A., Park M., Lee D. H. AI-IDS: Application of deep learning to real-time Web intrusion detection. IEEE Access. 2020. Vol. 8. P. 70245–70261. DOI: 10.1109/ACCESS.2020.2986882

4. Yin C., Zhang S., Wang J., Xiong N. N. Anomaly detection based on convolutional recurrent autoencoder for IoT time series. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2022. Vol. 52, No. 1. P. 112–122. DOI: 10.1109/TSMC.2020.2968516

5. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations (ICLR). 2021. DOI: 10.48550/arXiv.2010.11929

6. Hassan M. M., Gumaei A., Alsanad A., Alrubaian M., Fortino G. A hybrid deep learning model for efficient intrusion detection in big data environment. Information Sciences. 2020. Vol. 513. P. 386–396. DOI: 10.1016/j.ins.2019.10.069

7. Mighan S. N., Kahani M. A novel scalable intrusion detection system based on deep learning. International Journal of Information Security. 2021. Vol. 20. P. 387–403. DOI: 10.1007/s10207-020-00508-5

8. Ahmad Z., Shahid Khan A., Wai Shiang C., Abdullah J., Ahmad F. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies. 2021. Vol. 32, No. 1. e4150. DOI: 10.1002/ett.4150

9. Kasongo S. M., Sun Y. Performance analysis of intrusion detection systems using a feature selection method on the UNSW-NB15 dataset. Journal of Big Data. 2020. Vol. 7, No. 1. P. 1–20. DOI: 10.1186/s40537-020-00379-6

10. Gu A., Dao T. Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752. 2023. DOI: 10.48550/arXiv.2312.00752

11. Chen T., Kornblith S., Norouzi M., Hinton G. A simple framework for contrastive learning of visual representations. International Conference on Machine Learning. 2020. P. 1597–1607.

12. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Military Communications and Information Systems Conference (MilCIS). 2015. P. 1–6. DOI: 10.1109/MilCIS.2015.7348942

References:

1. Ferrag M. A., Maglaras L., Moschoyiannis S., Janicke H. Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications. 2020. Vol. 50. 102419. DOI: 10.1016/j.jisa.2019.102419

2. Alzahrani A. O., Alenazi M. J. F. Designing a network intrusion detection system based on machine learning for software defined networks. Future Internet. 2021. Vol. 13, No. 5. 111. DOI: 10.3390/fi13050111

3. Kim A., Park M., Lee D. H. AI-IDS: Application of deep learning to real-time Web intrusion detection. IEEE Access. 2020. Vol. 8. P. 70245–70261. DOI: 10.1109/ACCESS. 2020.2986882

4. Yin C., Zhang S., Wang J., Xiong N. N. Anomaly detection based on convolutional recurrent autoencoder for IoT time series. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2022. Vol. 52, No. 1. P. 112–122. DOI: 10.1109/TSMC.2020.2968516

5. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations (ICLR). 2021. DOI: 10.48550/arXiv.2010.11929

6. Hassan M. M., Gumaei A., Alsanad A., Alrubaian M., Fortino G. A hybrid deep learning model for efficient intrusion detection in big data environment. Information Sciences. 2020. Vol. 513. P. 386–396. DOI: 10.1016/j.ins.2019.10.069

7. Mighan S. N., Kahani M. A novel scalable intrusion detection system based on deep learning. International Journal of Information Security. 2021. Vol. 20. P. 387–403. DOI: 10.1007/s10207-020-00508-5

8. Ahmad Z., Shahid Khan A., Wai Shiang C., Abdullah J., Ahmad F. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies. 2021. Vol. 32, No. 1. e4150. DOI: 10.1002/ett.4150

9. Kasongo S. M., Sun Y. Performance analysis of intrusion detection systems using a feature selection method on the UNSW-NB15 dataset. Journal of Big Data. 2020. Vol. 7, No. 1. P. 1–20. DOI: 10.1186/s40537-020-00379-6

10. Gu A., Dao T. Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752. 2023. DOI: 10.48550/arXiv.2312.00752

11. Chen T., Kornblith S., Norouzi M., Hinton G. A simple framework for contrastive learning of visual representations. International Conference on Machine Learning. 2020. P. 1597–1607.

12. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Military Communications and Information Systems Conference (MilCIS). 2015. P. 1–6. DOI: 10.1109/MilCIS.2015.7348942