

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ

INFORMATION TECHNOLOGIES AND SYSTEMS

УДК 621.317+681.849

<https://doi.org/10.29235/1561-8358-2020-65-4-506-512>

Поступила в редакцию 18.11.2019

Received 18.11.2019

В. И. Соловьев¹, О. В. Рыбальский², В. В. Журавель³, В. К. Железняк⁴

¹Восточноукраинский национальный университет им. В. Даля, Северодонецк, Украина

²Национальная академия внутренних дел, Киев, Украина

³Киевский научно-исследовательский экспертно-криминалистический центр МВД Украины, Киев, Украина

⁴Полоцкий государственный университет, Новополоцк, Беларусь

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ВЫЯВЛЕНИЯ МОНТАЖА ЦИФРОВЫХ ФОНОГРАММ

Аннотация. Экспериментально доказана возможность создания построенной на основе нейронных сетей глубокого обучения эффективной системы, которая предназначена для выявления следов монтажа в цифровых фонограммах. Смысл эксперимента состоял в исследовании способности систем на основе таких сетей выявлять паузы со следами монтажа. Создан экспериментальный массив данных в звуковом редакторе из фонограмм, записанных на различной аппаратуре цифровой звукозаписи (при частоте дискретизации 44,1 кГц). Из него производился предварительный отбор пауз длительностью от 100 мс до нескольких секунд. Из отобранных 1000 пауз в автоматическом (компьютерном) режиме сформирован массив фрагментов пауз, из которого сгенерированы массивы фрагментов пауз разной длительности размерностью около 100 000. Для формирования массива фрагментов пауз с монтажом случайным образом выбранные паузы делились на части в произвольном соотношении. После этого из полученных частей создавались новые паузы с зафиксированным местом монтажа. Общий массив всех фрагментов пауз был разбит на тренировочный и тестовый массивы. Определялась наибольшая эффективность, достигаемая на тестовом массиве в процессе обучения. В общем случае эта эффективность определяется наибольшей величиной вероятности правильной классификации фрагментов с монтажом и фрагментов без монтажа. Предложена научно обоснованная методология выявления признаков монтажа в цифровых фонограммах на основе нейронных сетей глубокого обучения. Проведенные эксперименты показали, что возможно построение эффективной системы для выявления таких следов. Дальнейшее развитие методологии должно быть направлено на поиск путей повышения вероятности правильной бинарной классификации исследуемых пауз.

Ключевые слова: нейронная сеть глубокого обучения, точки монтажа, цифровая фонограмма, цифровой монтаж, экспертиза

Для цитирования: Применение нейронных сетей глубокого обучения для выявления монтажа цифровых фонограмм / В. И. Соловьев [и др.] // Вес. Нац. акад. навук Беларусі. Сер. фіз.-тэхн. навук. – 2020. – Т. 65, № 4. – С. 506–512. <https://doi.org/10.29235/1561-8358-2020-65-4-506-512>

Victor I. Solovyov¹, Oleg V. Rybalskiy², Vadym V. Zhuravel³, Vladimir K. Zheleznyak⁴

¹V. Dal Eastukrainian National University, Severodonetsk, Ukraine

²National Academy of Internal Affairs, Kyiv, Ukraine

³Kyiv Scientifically-Research Expertly-Criminalistics Center of the Ministry of Internal Affairs of Ukraine, Kyiv, Ukraine

⁴Polotsk State University, Novopolotsk, Belarus

APPLICATION OF NEURON NETWORKS OF DEEP LEARNING FOR EXPOSURES EDITING OF DIGITAL PHONOGRAMS

Abstract. Possibility of creation of effective system, which is intended for exposure of tracks of editing in digital phonograms and is built on the basis of neuron networks of the deep learning, is experimentally proven. Sense of experiment consisted in research of ability of the systems on the basis of such networks to expose pauses with tracks of editing. The experimental array of data is created in a voice editor from phonograms written on the different apparatus of the digital audio

recording (at frequency of discretisation 44,1 kHz). A preselection of pauses was produced from it, having duration from 100 ms to a few seconds. From 1000 selected pauses the array of fragments of pauses is formed in the automatic (computer) mode, from which the arrays of fragments of pauses of different duration are generated by a dimension about 100 000. For forming of array of fragments of pauses with editing, the chosen pauses were divided into casual character parts in arbitrary correlation. Afterwards, the new pauses were created from it with the fixed place of editing. The general array of all fragments of pauses was broken into training and test arrays. The maximum efficiency, achieved on a test array in the process of educating, was determined. In general case this efficiency is determined by the maximum size of probability of correct classification of fragments with editing and fragments without editing. Scientifically reasonable methodology of exposure of signs of editing in digital phonograms is offered on the basis of neuron networks of the deep learning. The conducted experiments showed that the construction of the effective system is possible for the exposure of such tracks. Further development of methodology must be directed to find the ways to increase the probability of correct binary classification of investigated pauses.

Keywords: neuron network of deep learning, points of editing, digital phonogram, digital editing, examination

For citation: Solovyov V. I., Rybalskiy O. V., Zhuravel V. V., Zheleznyak V. K. Application of neuron networks of deep learning for exposures editing of digital phonograms. *Vesti Natsyyanal'nai akademii nauk Belarusi. Seryya fizika-technichnykh nauk = Proceedings of the National Academy of Sciences of Belarus. Physical-technical series*, 2020, vol. 65, no. 4, pp. 506–512 (in Russian). <https://doi.org/10.29235/1561-8358-2020-65-4-506-512>

Введение. Проверка отсутствия (или наличия) следов монтажа в цифровых фонограммах является одной из наиболее важных проблем экспертизы материалов и аппаратуры звукозаписи. В ряде работ отечественных и зарубежных авторов рассматривались вопросы ее решения [1–5]. В них предложены разные подходы – от спектрального анализа шумов до проверки метаданных. Однако до настоящего времени не найден метод, обеспечивающий высокую степень вероятности выявления монтажа.

Отметим, что в большинстве предложенных решений обязательно проведение сравнительных исследований параметров образцовых (экспериментальных) и спорных (представленных на экспертизу) фонограмм. Поэтому для записи образцовых фонограмм требуется предъявление на экспертизу аппаратуры, на которой записывались спорные фонограммы, что часто вообще невыполнимо из-за ряда специфических обстоятельств [6].

Давно разработанный в СССР и ныне широко используемый в странах ЕС метод, который основан на проверке частоты (или фазы) сигнала сетевой наводки, зафиксированного в фонограмме [3, 7], имеет ряд существенных недостатков. Из них следует отметить, например, необходимость постоянной фиксации частоты электросети во всех точках страны и, что наиболее существенно, невозможность выявления монтажа, выполненного методом вырезания и перестановки (используя паузы в речевых сигналах) фрагментов одной фонограммы [8]. Такой монтаж наиболее сложен для выявления экспертизой.

В наших предыдущих работах, например [9], были рассмотрены модели проявления в сигналах следов цифрового монтажа фонограмм, выполненного способом вырезания и перестановки фрагментов. Далее было экспериментально доказано, что следы такого монтажа проявляются в виде изменений фрактального состава монтируемых сигналов [10, 11]. Вместе с тем было показано, что, несмотря на наличие таких проявлений, их выделение из сигналов в виде конкретных признаков невозможно без проведения обширных исследований. Для этого потребовалось создать специальное программное обеспечение [10]. На сегодняшний день не предложено научно-обоснованного подхода, который бы позволил применять единую методологию выявления монтажа. Мы полагаем, что такая методология должна базироваться на большом объеме экспериментальных данных, поскольку апробация на массовом статистическом материале любого предлагаемого подхода в режиме «ручного исследования» практически нереализуема.

Последующий поиск направления создания таких программ показал, что наиболее перспективным представляется использование нейронных сетей глубокого обучения [11].

Цель данной работы – экспериментально доказать возможность создания эффективной системы, предназначенной для выявления следов монтажа в цифровых фонограммах и построенной на основе такой сети.

Постановка и проведение эксперимента. Эксперимент необходим для выяснения принципиальной пригодности выбранного направления. Его смысл состоит в исследовании способности систем на основе нейронных сетей глубокого обучения выявлять паузы со следами монтажа. Для создания экспериментального массива данных в звуковом редакторе из фонограмм, записанных на различной аппаратуре цифровой звукозаписи (при частоте дискретизации 44,1 кГц),

производился предварительный отбор пауз длительностью от 100 мс до нескольких секунд. Отбор производился в «ручном режиме». Таких пауз было отобрано примерно 1000, и они составили исходную базу эксперимента. Из отобранных пауз в автоматическом (компьютерном) режиме был сформирован массив фрагментов пауз. Для этого паузы сканировались окном длительностью 15–30 мс с интервалом 1 мс и из них вырезались фрагменты с длительностью окна сканирования. Таким образом, были сгенерированы массивы фрагментов пауз без монтажа разной длительности размерностью около 100 000. Для создания массива фрагментов пауз с монтажом случайным образом выбранные паузы делились на части в произвольном соотношении. После чего из полученных частей формировались новые паузы с зафиксированным местом монтажа. Затем сканированием этих пауз временным окном из зоны локализации монтажа вырезались фрагменты пауз с монтажом одинаковой длительности. Размерность массива фрагментов пауз с монтажом также составляла примерно 100 000.

Эти массивы фрагментов пауз без монтажа и с монтажом явились исходными массивами (dataset) для обучения нейронной сети.

Разработанная программная система базировалась на платформе PITON и библиотеке глубокого обучения нейронных сетей keras (backend tensorflow) со следующими основными параметрами нейронной сети: полносвязная нейронная сеть до 50 нейронных слоев с прореживанием (dropout) и пакетной нормализацией. Решалась задача бинарной классификации на массиве dataset. В далее проиллюстрированных экспериментах фрагменты массивов пауз длительностью 20 мс имели размерность 882.

Общий массив всех фрагментов пауз был разбит на тренировочный и тестовый массивы. Устанавливалась наибольшая эффективность, достигаемая на тестовом массиве в процессе обучения. В общем случае эта эффективность определяется наибольшей величиной вероятности правильной классификации фрагментов с монтажом и фрагментов без монтажа.

Пример формы настройки параметров нейронной сети, использованной при проведении эксперимента, показан на рис. 1.

Рис. 1. Окно параметров экспериментальной нейронной сети глубокого обучения

Fig. 1. Window of parameters of experimental neuron network of the deep learning

Процесс проведения эксперимента с использованием нейронной сети глубокого обучения проиллюстрирован на рис. 2 и 3. На рис. 2 показан результат, достигнутый сетью на обучающем массиве. Как видно, эффективность, достигнутая на базе обучения за 575 эпох, составляет 0,85. Такой результат получен в процессе обучения на компьютере с графическим адаптером (GPU).

На этом же рисунке указано, что на тестовом массиве максимальная эффективность составляет 0,76. На рис. 3 показаны совместные графики для тренировочного и тестового массивов фрагментов пауз.

В рамках проведенной оптимизации сети достигнута эффективность классификации 0,76. Безусловно, эта вероятность с точки зрения многообразия вариантов оптимизации не может являться предельной. Как показывают предварительные эксперименты, существенное повышение размерности массивов фрагментов пауз и, как следствие, возможности увеличения сложности нейронной сети позволяет увеличить эффективность.

Однако в задачах подобного рода вряд ли можно ожидать столь же высокой эффективности классификации, как, например, в задачах классификации изображений (0,99). В то же время структура задачи позволяет предложить рациональные решения, которые можно применить в задачах выявления монтажа. Ключевым моментом в решении данной задачи является введение понятия «пороговой» вероятности.

В задачах классификации обученная модель всегда позволяет получить вероятности правильной классификации. Как показывает анализ, около 30 % фрагментов пауз имеют вероятность такой классификации ниже значения 0,55. При превышении этого порога вероятности правильной классификации резко возрастают и достигают значения 0,99 и выше. На рис. 4 и 5 приведены графики зависимостей вероятности правильной классификации и прогнозируемых вероятностей ошибок классификации от «пороговой» вероятности. Отметим, что значения вероятности 0 и 1, приведенные на графиках рис. 4 и 5, в реальности являются лишь приближенными к ним значениями, но разрешающая способность выбранного масштаба изображения не позволяет отобразить это явление.

Также отметим, что в весьма специфических задачах практического выявления цифрового монтажа в фонограммах можно установить высокую «пороговую» вероятность, например равную 0,9. При этом вероятность правильной классификации будет не ниже 0,99. Но для решаемой задачи с практической точки зрения такая вероятность не очень высока. Так, на фрагментах пауз без монтажа общей длительностью 1 с при такой вероятности правильной классификации возможны 10 ложных результатов.

Введение «пороговой» вероятности при решении практических задач экспертизы

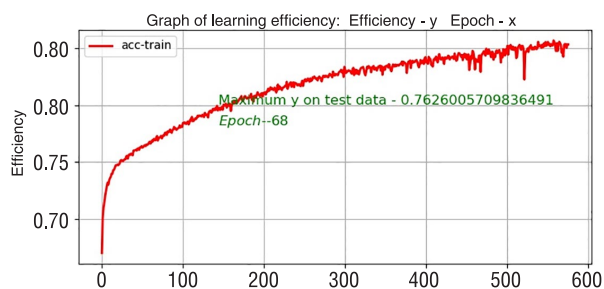


Рис. 2. График эффективности обучения
Fig. 2. Chart of efficiency of educating

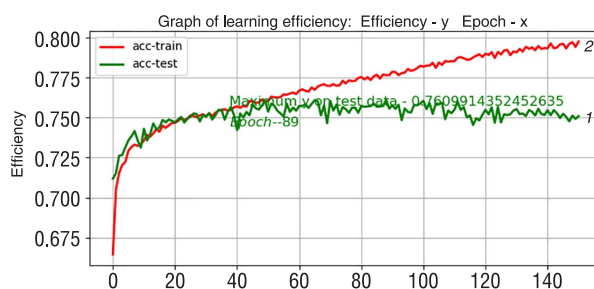


Рис. 3. Графики эффективности обучения: тестовый (1) и тренировочный (2) массивы
Fig. 3. Charts of efficiency of educating: test (1) and training (2) arrays

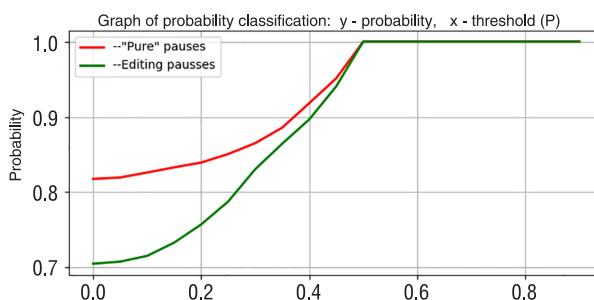


Рис. 4. Графики эффективности классификации
Fig. 4. Charts of efficiency of classification

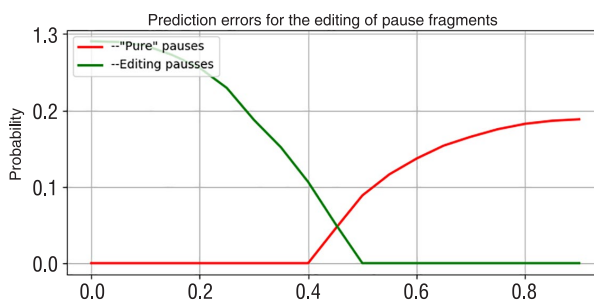


Рис. 5. Графики прогнозируемых ошибок классификации
Fig. 5. Charts of the forecast errors of classification

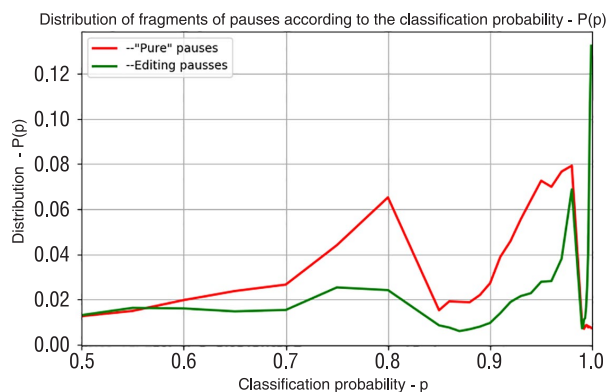


Рис. 6. Распределения фрагментов пауз с монтажом и без монтажа по вероятности правильной классификации (дискретный аналог плотностей вероятности)

Fig. 6. Distributions of fragments of pauses with editing and without editing on probability of correct classification (discrete analogue of probability density)

правильной классификации для этих фрагментов. Эти графики, полученные на тестовом массиве фрагментов, представлены рис. 6.

Графики показывают, что весовые доли фрагментов пауз с монтажом с высокой степенью вероятности правильной классификации значительно превышают весовые доли таких пауз без монтажа. Это значит, что среди пауз с монтажом преобладают паузы с высокой степенью вероятности правильной классификации, что свидетельствует о возможности и целесообразности применения нейронных сетей глубокого обучения для построения системы выявления монтажа в цифровых фонограммах.

При этом следует отметить, что для проведения экспериментов мы использовали лишь один из вариантов конфигурации таких сетей – полносвязную нейронную сеть. Однако существуют еще несколько вариантов их конфигурации, что говорит о возможности построения еще более эффективной системы.

Выводы. Проведенные экспериментальные исследования показали принципиальную пригодность нейронных сетей глубокого обучения для построения системы выявления следов цифровой обработки в фонограммах. Перспективность применения таких сетей состоит в возможности дальнейшего совершенствования как их конфигурации, так и методов предварительной обработки информации, что позволяет рассчитывать на достижение высокой эффективности такой системы.

Список использованных источников

1. Рыбальский, О.В. Основные положения теории выявления следов цифровой обработки фонограмм и особенности ее программной и методической реализации в экспертизе материалов и средств видеозвукозаписи. Ч. 1 / О.В. Рыбальский // *Захист інформації*. – 2006. – № 1. – С. 71–76.
2. Nicolalde, D.P. Evaluating digital audio authenticity with spectral distances and ENF phase change / D.P. Nicolalde, J.A. Apolinario // *IEEE International Conference on Acoustics, Speech and Signal Processing*. – 2009. – Vol. 1. – P. 1417–1420. <https://doi.ieeecomputersociety.org/10.1109/ICASSP.2009.4959859>
3. Nicolalde, D.P. Audio authenticity: Detecting ENF discontinuity with high precision phase analysis / D.P. Nicolalde, J.A. Apolinario, L. W. P Biscainho // *IEEE Transactions on Information Forensics and Security*. – 2010. – Vol. 5. – P. 534–543. <https://doi.org/10.1109/TIFS.2010.2051270>
4. Korycky, R. Methods of Time-Frequency Analysis in Authentication of Digital Audio Recordings / R. Korycky // *INTL Journal of Electronics and Telecommunications*. – 2010. – Vol. 56, № 3. – P. 257–262. <https://doi.org/10.2478/v10177-010-0033-0>
5. Бобрицкий, С.М. Дослідження ознак монтажу записів, виконаних цифровими записувачами пристроями / С.М. Бобрицкий, С.В. Стороженко // *Теорія та практика судової експертизи і криміналістики: зб. наук. праць* / редкол.: М.Л. Цимбал [та ін.]. – Харків: Право, 2011. – Вип. 11. – С. 353–361.

соответствует пропуску фрагментов пауз с признаками монтажа и ложных срабатываний с низкой вероятностью. В то же время это соответствует природе исследуемого объекта (пауз). Это явление обусловлено тем, что и в паузах без монтажа, как правило, могут возникать выбросы малого уровня.

Чтобы прийти к окончательному выводу о возможности использования таких сетей для выявления монтажа, следовало установить, насколько велики будут потери в эффективности выявления пауз с монтажом при введении «пороговой» вероятности. Для этого были проведены дополнительные исследования отдельно для пауз с монтажом и для пауз без монтажа («чистых» пауз). Определялись распределения весовых долей фрагментов пауз по вероятности их правильной классификации (дискрет-

6. Журавель, В. В. Особенности формирования фонограмм, записанных с телефонных каналов / В. В. Журавель // Сучас. спец. техника. – 2015. – №4 (43). – С. 26–31.
7. Основы экспертного криминалистического исследования магнитных фонограмм / А. А. Ложкевич [и др.]. – М.: ВНИИ МВД СССР, 1977. – 172 с.
8. Рыбальский, О. В. Модели нестандартной подделки цифровых фонограмм / О. В. Рыбальский // Реєстрація, зберігання і обробка даних. – 2003. – Т. 5, №4. – С. 25–32.
9. Рыбальский, О. В. Следы монтажа в цифровых фонограммах, выполненного способом вырезания и перестановки фрагментов / О. В. Рыбальский, В. И. Соловьев, В. В. Журавель // Реєстрація, зберігання і обробка даних. – 2016. – Т. 18, №1. – С. 32–41.
10. Рыбальский, О. В. Экспериментальная проверка эффекта изменения фрактального состава сигналов при монтаже фонограммы способом вырезания и перестановки фрагментов / О. В. Рыбальский, В. И. Соловьев, В. В. Журавель // Сучас. спец. техника. – 2016. – №3. – С. 75–85.
11. Рыбальский, О. В. Основные требования к системе выявления точек цифрового монтажа в фонограммах и методология ее создания / О. В. Рыбальский, В. И. Соловьев, В. В. Журавель // Інформатика та математичні методи в моделюванні. – 2018. – Т. 8, №3. – С. 232–237.

References

1. Rybal'skii O. V. Substantive provisions of theory of exposure of tracks of digital treatment of phonograms and feature of her programmatic and methodical realization in examination of materials and facilities of the audio videorecording. Part 1. *Zakhyist informatsii = Ukrainian Information Security Research Journal*, 2006, no. 1, pp. 71–76 (in Russian).
2. Nicolalde D. P., Apolinario J. A. Evaluating digital audio authenticity with spectral distances and ENF phase change. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, vol. 1, pp. 1417–1420. <https://doi.ieeecomputersociety.org/10.1109/ICASSP.2009.4959859>
3. Nicolalde D. P., Apolinario J. A., Biscainho L. W. P. Audio authenticity: Detecting ENF discontinuity with high precision phase analysis. *IEEE Transactions on Information Forensics and Security*, 2010, vol. 5, pp. 534–543. <https://doi.org/10.1109/TIFS.2010.2051270>
4. Korycky R. Methods of Time-Frequency Analysis in Authentication of Digital Audio Recordings. *INTL Journal of Electronics and Telecommunications*, 2010, vol. 56, no. 3, pp. 257–262. <https://doi.org/10.2478/v10177-010-0033-0>
5. Bobritskii S. M., Storozhenko S. V. Research of signs of editing of the records executed by digital record devices. *Teoriya ta praktyka sudovoyi ecpertyzy i kriminalistiki: zbirnik naukovikh prats* [Theory and Practice of Forensic Science and Criminology: Collection of Scientific Papers]. Kharkov, Pravo Publ., 2011, Issue 11, pp. 353–361 (in Ukrainian).
6. Zhuravel' V. V. Features of forming of phonograms written in from telephone channels. *Suchasna spetsialna tekhnika* [Modern Special Technoque], 2015, no. 4, pp. 26–31 (in Russian).
7. Lozhkevich A. A., Snetkov V. A., Chivanov V. A., Sharshunskii V. L. *Bases of Expert Criminalistics Research of Magnetic Phonograms*. Moscow, All-Union Scientific Research Institute of the USSR Ministry of Internal Affairs, 1972. 172 p. (in Russian).
8. Rybal'skii O. V. Models of non-standard imitation of digital phonograms. *Regestratsiya, zberigannya i obrobka danykh = Data Recording, Storage and Processing*, 2003, no. 4, pp. 25–32 (in Russian).
9. Rybal'skii O. V., Solov'ev V. I., Zhuravel' V. V. Tracks of editing are in digital phonograms, excision and transposition of fragments executed by a method. *Regestratsiya, zberigannya i obrobka danykh = Data Recording, Storage and Processing*, 2016, no. 1, pp. 32–41 (in Russian).
10. Rybal'skii O. V., Solov'ev V. I., Zhuravel' V. V. Experimental verification of effect of change of fractal composition of signals at editing of phonogram the method of excision and transposition of fragments. *Suchasna spetsialna tekhnika* [Modern Special Technoque], 2016, no. 3, pp. 75–85 (in Russian).
11. Rybal'skii O. V., Solov'ev V. I., Zhuravel' V. V. Basic requirements to the system of exposure of points of the digital editing in phonograms and methodology of her creation. *Informatica ta matematichi metody v modeluvanni = Informatics and Mathematical Methods in Simulation*, 2018, vol. 8, no. 3, pp. 232–237 (in Russian).

Інфармацыя аб аўтарах

Соловьев Виктор Иванович – кандидат технических наук, доцент, заместитель заведующего кафедрой компьютерных систем и сетей, Восточноукраинский национальный университет им. В. Даля (пр. Центральный, 59а, 93400, Северодонецк, Украина). <https://orcid.org/0000-0003-3541-4740>. E-mail: edemsvi@gmail.com

Рыбальский Олег Владимирович – доктор технических наук, профессор, профессор кафедры информационных технологий и кибербезопасности, Национальная академия внутренних дел (пл. Соломенская, 1, 03035, Киев, Украина). <https://orcid.org/0000-0002-1440-8344>. E-mail: rov_1946@ukr.net

Information about the authors

Victor I. Solovyov – Ph. D. (Engineering), Associate Professor, Deputy of Manager of the Department of the Computer Systems and Networks, V. Dal East Ukrainian National University (59a, Tsentralnyi Ave., 93400, Severodonetsk, Ukraine). <https://orcid.org/0000-0003-3541-4740>. E-mail: edemsvi@gmail.com

Oleg V. Rybalskiy – D. Sc. (Engineering), Professor, Professor of the Department of Information Technologies and Cybernetic Safety, National Academy of Internal Affairs (1, Solom'yanska Are., 03035, Kyiv, Ukraine). <https://orcid.org/0000-0002-1440-8344>. E-mail: rov_1946@ukr.net

Журавель Вадим Васильевич – кандидат технических наук, заведующий лабораторией исследований в сфере информационных технологий, Киевский научно-исследовательский экспертно-криминалистический центр Министерства внутренних дел Украины (ул. Владимирская, 15, 01001, Киев, Украина). E-mail: fonoscopia@ukr.net

Железняк Владимир Кириллович – доктор технических наук, профессор, заведующий кафедрой информационной безопасности, Полоцкий государственный университет (ул. Блохина, 29, 211440, Новополоцк, Республика Беларусь). E-mail: v.zheleznyak@psu.by

Vadym V. Zhuravel – Ph. D. (Engineering), Head of the Laboratory of Researches in the Field of Information Technologies, Kyiv Scientifically-Research Expertly-Criminalistics Center of the Ministry of Internal Affairs of Ukraine (15, Vladimirska Str., 01001, Kyiv, Ukraine). E-mail: fonoscopia@ukr.net

Vladimir K. Zheleznyak – D. Sc. (Engineering), Professor, Head of the Department of Informative Safety, Polotsk State University (29, Blokhin Str., 211440, Novopolotsk, Republic of Belarus). E-mail: v.zheleznyak@psu.by