DOI 10.37882/2223-2966.2024.05.36

ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЯ НА ОСНОВЕ СИАМСКИХ НЕЙРОННЫХ СЕТЕЙ КАК МОДУЛЬ МНОГОФАКТОРНОЙ АУТЕНТИФИКАЦИИ ДЛЯ СЛОЖНЫХ ЭКОНОМИЧЕСКИХ СИСТЕМ

SIAMESE NEURAL NETWORK-BASED USER IDENTIFICATION AS A MULTIFACTOR AUTHENTICATION MODULE FOR COMPLEX ECONOMIC SYSTEMS

P. Kharlamov
O. Kharlamova
E. Lavrova

Summary. *The object of the study* is the information security of business processes of transferring data subject to trade secrets in complex economic systems, such as territorial scientific and industrial clusters. *The subject of the study* is to develop a method of user identification based on Siamese neural networks as a module (factor) of multifactor authentication to ensure information security of business processes of data transmission, with the regime of trade secrets, in complex economic systems. *The relevance of the problems* stems from the need to improve the existing mechanisms and tools for information security in the transfer of data containing trade secrets between organizations within a complex economic system, such as a territorial scientific and industrial cluster, as the existing methods have a number of significant vulnerabilities, such as a «dictionary attack», are characterized by low accuracy in carrying out user identification. *The aim of the study* is to develop a method of user identification based on Siamese neural networks as a module of multifactor authentication for complex economic systems. A methodological framework for the application of the method of user identification based on Siamese neural networks for complex economic systems is proposed. A method for user identification based on Siamese neural networks as a module for multifactor authentication of complex economic systems such as territorial scientific and industrial clusters is developed. *The conclusion* about the effectiveness of the developed method of user identification based on Siamese neural networks as a module of multifactor authentication in providing information security of data transfer processes, constituting trade secrets in complex economic systems is made.

Keywords: information security, multifactor authentication, complex socio-economic systems, territorial science and industry clusters, data validation, artificial neural networks, Siamese neural networks, business secrets data.

Харламов Павел Сергеевич
младший научный сотрудник, преподаватель, Смоленский филиал РАНХиГС;
магистрант, филиал ФГБОУ ВО «НИУ «МЭИ», Смоленск
pavel_kharlamov.mp67@mail.ru
Харламова Ольга Евгеньевна
Преподаватель, Смоленский филиал РАНХиГС
o.e.kharlamova@mail.ru
Лаврова Елена Викторовна
кандидат экономических наук, доцент, Смоленский филиал РАНХиГС
e.v.lavrova@list.ru

Аннотация. *Объектом исследования* выступает информационная безопасность бизнес-процессов передачи данных, составляющих коммерческую тайну, в сложных экономических системах, таких как территориальные научно-промышленные кластеры. *Предметом исследования* является разработка метода идентификации пользователя на основе сиамских нейронных сетей как модуля (фактора) многофакторной аутентификации для обеспечения информационной безопасности бизнес-процессов передачи данных, имеющих режим коммерческой тайны, в сложных экономических системах. *Актуальность проблематики* обусловлена необходимостью совершенствования существующих механизмов и инструментов обеспечения информационной безопасности при передаче данных, содержащих коммерческую тайну, между организациями, входящими в сложную экономическую систему, например, в территориальный научно-промышленный кластер, поскольку существующие методы имеют ряд значительных уязвимостей, например, к «атаке по словарю», характеризуются низкой точностью при осуществлении идентификации пользователя. *Цель исследования* заключается в разработке метода идентификации пользователя на основе сиамских нейронных сетей как модуля многофакторной аутентификации для сложных экономических систем. Предложена методологическая основа применения метода идентификации пользователя на основе сиамских нейронных сетей для сложных экономических систем. Разработан метод идентификации пользователя на основе сиамских нейронных сетей как модуль многофакторной аутентификации для сложных экономических систем, таких как территориальные научно-промышленные кластеры. Сделан вывод об эффективности внедрения разработанного метода идентификации пользователя на основе сиамских нейронных сетей как модуля многофакторной аутентификации при обеспечении информационной безопасности процессов передачи данных, составляющих коммерческую тайну, в сложных экономических системах.

Ключевые слова: информационная безопасность, многофакторная аутентификация, сложные социально-экономические системы, территориальные научно-промышленные кластеры, валидация данных, искусственные нейронные сети, сиамские нейронные сети, данные, составляющие коммерческую тайну.

Введение

Режим коммерческой тайны в настоящее время установлен во многих организациях, прежде всего в организациях промышленного сектора. Введение ограничения доступа к информации обусловлено ее существенным значением в получении коммерческой выгоды, например, улучшения финансового результата организации, в том числе чистой прибыли организации, улучшения положения на рынке и завоевания большей доли рынка. Большинство организаций, действуя в соответствии с законодательством Российской Федерации, ограничивают лишь количество работников, допущенных к информации, имеющей гриф коммерческой тайны. Однако простое количественное ограничение доступа обеспечивает защищенность указанных данных только в организациях малого бизнеса. Средний и крупный бизнес, осуществляя различные взаимодействия со стейкхолдерами по коммуникационным каналам, подвержен совершенно иным рискам утечки данных, содержащих коммерческую тайну. Прежде всего, рассматривая данную проблему, можно отметить утечку информации такого вида по каналам коммуникации (телекоммуникации) во время реализации бизнес-процесса передачи данных с помощью различных информационных технологий. Вышесказанное справедливо не только для среднего и большого бизнеса, но и для сложных экономических систем, представляющих собой взаимосвязь социально-экономических субъектов.

Главная особенность сложных экономических систем, представленных совокупностью организаций и социально-экономических субъектов, состоит в большом количестве каналов взаимосвязи, созданных с помощью различных телекоммуникационных технологий, например, с помощью многоканальных телекоммуникационных сетей, сетей Intranet и Extranet, а также с помощью почтовых серверов (например, электронной почты), веб-приложений, корпоративных информационных систем. В связи с этим можно констатировать наличие выделенного самостоятельного вспомогательного бизнес-процесса передачи данных в условиях установленного режима коммерческой тайны. Одним из примеров сложной экономической системы являются территориальные научно-промышленные кластеры. Так, в рамках данной системы строится сеть каналов взаимосвязи и взаимодействия на относительно не удаленном расстоянии в пределах одного или нескольких смежных субъектов Российской Федерации между организациями промышленного сектора экономики и научно-исследовательскими организациями. Указанное предполагает, в том числе, и разработку инновационных технологий производства, способных обеспечить конкурентоспособность организаций промышленного сектора, а также конкурентоспособность отдельных видов производимой продукции. Отметим, что вышеуказанная технология производства в большинстве случаев составляет коммерческую тайну и требует обеспечения высокой степени информационной безопасности при ее передаче от научно-исследовательской организации к организации промышленного сектора, входящих в один территориальный научно-промышленный кластер.

Соответственно, рассматривая сеть каналов взаимосвязи и взаимодействия по передаче данных между элементами сложных экономических систем (организаций и их подразделений, а также иных социально-экономических субъектов, входящих в сложную экономическую систему), можно говорить об широком разнообразии данных каналов и необходимости универсального метода обеспечения информационной безопасности бизнеса-процесса передачи данных, составляющих коммерческую тайну. В настоящее время большинство подобных организаций используют уже существующие методы, связанные с разверткой частных виртуальных сетей VPN и обеспечением изоляции основного соединения — использование прикладных протоколов Point-to-Point Protocol (PPTP), Layer 2 Tunneling Protocol (L2TP), Secure Socket Tunneling Protocol (SSTP), Internet Key Exchange (IKEv2). Однако они имеют ряд серьезных уязвимостей (например, прикладной протокол PPTP уязвим к атакам по словарю и атакам типа Bit-flipping), что не позволяет их использовать в сложных экономических системах для обеспечения соответствующего уровня информационной безопасности.

Одним из вариантов решения вышеуказанной проблемы является разработка метода многофакторной аутентификации. Каналы аутентификации при использовании данного метода должны исходить из элементов идентификации сотрудника, принятых в организации. Например, если организация использует для идентификации личную подпись сотрудника, то канал аутентификации должен включать модуль аутентификации по личной подписи сотрудника. При этом в большинстве случаев идентификация сотрудников, связана с анализом исходных изображений или данных, эффективным инструментом для реализации которой являются нейронные сети, прежде всего, сиамские нейронные сети [1–2].

Сиамские нейронные сети как инструмент идентификации

Как отмечено ранее, в настоящее время нейронные сети стали эффективным инструментом идентификации пользователя, работающего с той или иной информационной системой. На вход в нейронную сеть могут подаваться различные данные, однако для задачи идентификации наилучшим образом подходят следующие данные [3–4]:

— числовые последовательности, характеризующие идентификационный номер пользователя;

— изображения, содержащие различную информацию о пользователе (например, его биометрические данные или личную подпись).

Следует отметить, что в большинстве случаев числовые последовательности, передающиеся на входной слой нейронной сети, представляют собой изображения рукописного ввода чисел.

Однако для работы нейронной сети, например, сверточных нейронных сетей, с изображениями необходимы внушительные наборы обучающих данных, что не представляется возможным при решении задачи идентификации пользователя в организации: зачастую в нейронную сеть может быть подан лишь один учебный предмет (исходное изображение или образец), характеризующий конкретно взятого пользователя. В связи с этим в настоящее разработан специальный тип нейронной сети — сиамская нейронная сеть, использующаяся в случаях, когда в каждом классе объектов для идентификации мало единиц данных. Более того, сиамская нейронная сеть для введенного рукописного текста (подписи или числовой последовательности) обеспечивает возможность динамического и корректного распознавания текста, идентифицируя пользователя независимо от скорости письма, угла наклона и правильности написания символьных структур [5–7]. На рисунке 1 представлена структура нейронной сети, которая будет использоваться в разрабатываемом методе идентификации пользователя в условиях режима коммерческой тайны на основе сиамских нейронных сетей для сложных экономических систем, таких как территориальные научно-промышленные кластеры. Сиамская нейронная сеть будет иметь две идентичных подсети с одинаковыми параметрами и весами.

Разработка метода идентификации пользователя в условиях режима коммерческой тайны на основе сиамских нейронных сетей для сложных экономических систем

Для решения поставленной задачи обеспечения высокого уровня информационной безопасности процессов передачи данных, содержащих коммерческую тайну, между элементами (организациями) сложных экономических систем предлагается использование разрабо-

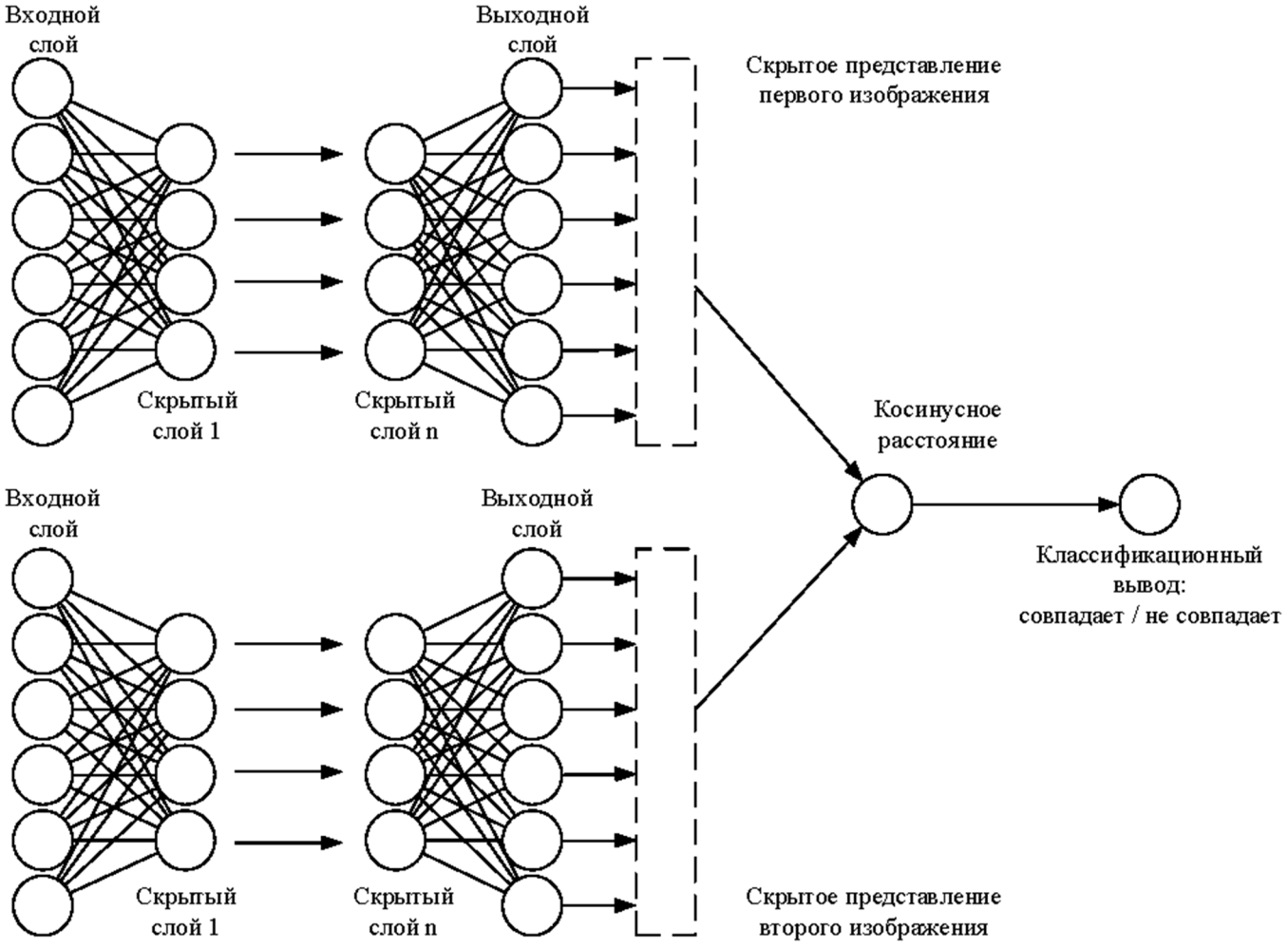

Рис. 1. Структура сиамской нейронной сети, использующейся в разработанном методе идентификации пользователя

танного авторами метода идентификации пользователя в условиях режима коммерческой тайны на основе сиамских нейронных сетей.

В основу метода положены сверточные сиамские нейронные сети, ориентированные на использование изображений в качестве входных данных. Используемая в разработанном методе сиамская нейронная сеть содержит две идентичные подсети, имеющие одну и ту же конфигурацию с одинаковыми параметрами и весовыми коэффициентами, при этом обновление параметров отражается в обеих подсетях [9-10]. При получении входных данных в виде изображения сиамская нейронная сеть отображает их как n-мерный массив данных. Основным принципом ее работы при обучении является ориентированность на демонстрации такого сопоставления, чтобы данные точек из разных классов (образцов изображений, характеризующих личную информацию о сотрудниках организаций, являющихся элементами сложной экономической системы) были расположены как можно дальше при расчете косинусного расстояния, в то время как данные точек из одного класса находились как можно ближе в полученном n-мерном массиве.

В качестве метода обучения был выбран метод обратного распространения ошибки, реализованный в виде последовательности следующих действий:

— прямое распространение сигнала по сиамской нейронной сети, вычисление состояния нейронов сети;
— вычисление значения ошибки нейрона для выходного слоя используемой сиамской нейронной сети;
— обратное распространение ошибки нейрона, последовательный расчет ошибки нейрона скрытого слоя сиамской нейронной сети от конца к началу для всех скрытых слоев;
— обновление весовых коэффициентов сиамской нейронной сети на вычисленную на предыдущем этапе ошибку нейрона скрытого уровня.

В качестве функции активации использована сигмоидальная функция, позволяющая усиливать слабые (нечеткие или размытые фрагменты изображений) входные сигналы сиамской нейронной сети. Используемая в методе функция ошибки — binary cross entropy (BCE) [11–13] с повышением ответственности и чувствительности к выборке сиамской нейронной сети, рассчитывается по формулам (1–3), в которой $y_{исх.}$ — n-массив с бинарными значениями «Истинное (исходное) множество», $y_{пред.}$ — n-массив с бинарными значениями «Множество предсказанных значений».

$$BCE = — (p_1 \cdot (1 — recall) + recall \cdot p_2 \quad (1)$$

$$p_1 = y_{исх.} \cdot \log(y_{пред.}) \quad (2)$$

$$p_2 = (1 — y_{исх.}) \cdot \log(1 — y_{пред.}) \quad (3)$$

Формулы (1–3) являются модификацией классического метода расчета BCE. Даная модификация обусловлена необходимостью дифференциации отклика сиамской нейронной сети и особым акцентом на высокую достоверность и точность распознавания сети при работе с изображениями, в частности, изображениями, содержащими личную подпись сотрудника.

Разработанный вышеприведенный алгоритм применим для метода идентификации пользователя в условиях режима коммерческой тайны на основе сиамских нейронных сетей для сложных экономических систем, таких как территориальные научно-промышленные кластеры в двух вариациях, выделение которых основано на определении типа входных данных сверточной сиамской нейронной сети и на реальных процессах верификации сотрудников в организации-участнике кластера:

— метод идентификации пользователя по биометрических данным (сканирование лица и т.д.) на основе сиамских нейронных сетей;
— метод идентификации пользователя по личной подписи, поставленной в рукописном исполнении на сенсорном экране, на основе сиамских нейронных сетей.

Эффективным и более достоверным является использовании данного метода обеспечения информационной безопасности в комплексе программных решений. В связи с чем необходима интеграция метода в качестве программного модуля в многофакторную аутентификацию пользователя [15–16]. Одним из вариантов многофакторной аутентификации является использование следующих каналов аутентификации, реализованных в виде модулей аутентификации, не требующих активного интернет-соединения (рисунок 2):

— модуль аутентификации по биометрическим данным сотрудника (распознавание отпечатка пальца по ранее занесенным данным) на основе Windows 10 Credential Provider;
— модуль аутентификации по личной подписи сотрудника (в качестве исходной подписи (образца подписи) используется подпись, поставленная сотрудником во время ознакомления его с действующим режимом коммерческой тайны) на основе сиамских нейронных сетей;
— модуль аутентификации посредством применения сотрудником USB-токена.

В данном методе важным является расчет функции потерь для сиамской нейронной сети, реализованной с помощью программного кода, представленного в листинге 1 и написанного на языке высокого уровня Python. Также важным является добавление в метод обратного распространения ошибки оптимизатора для

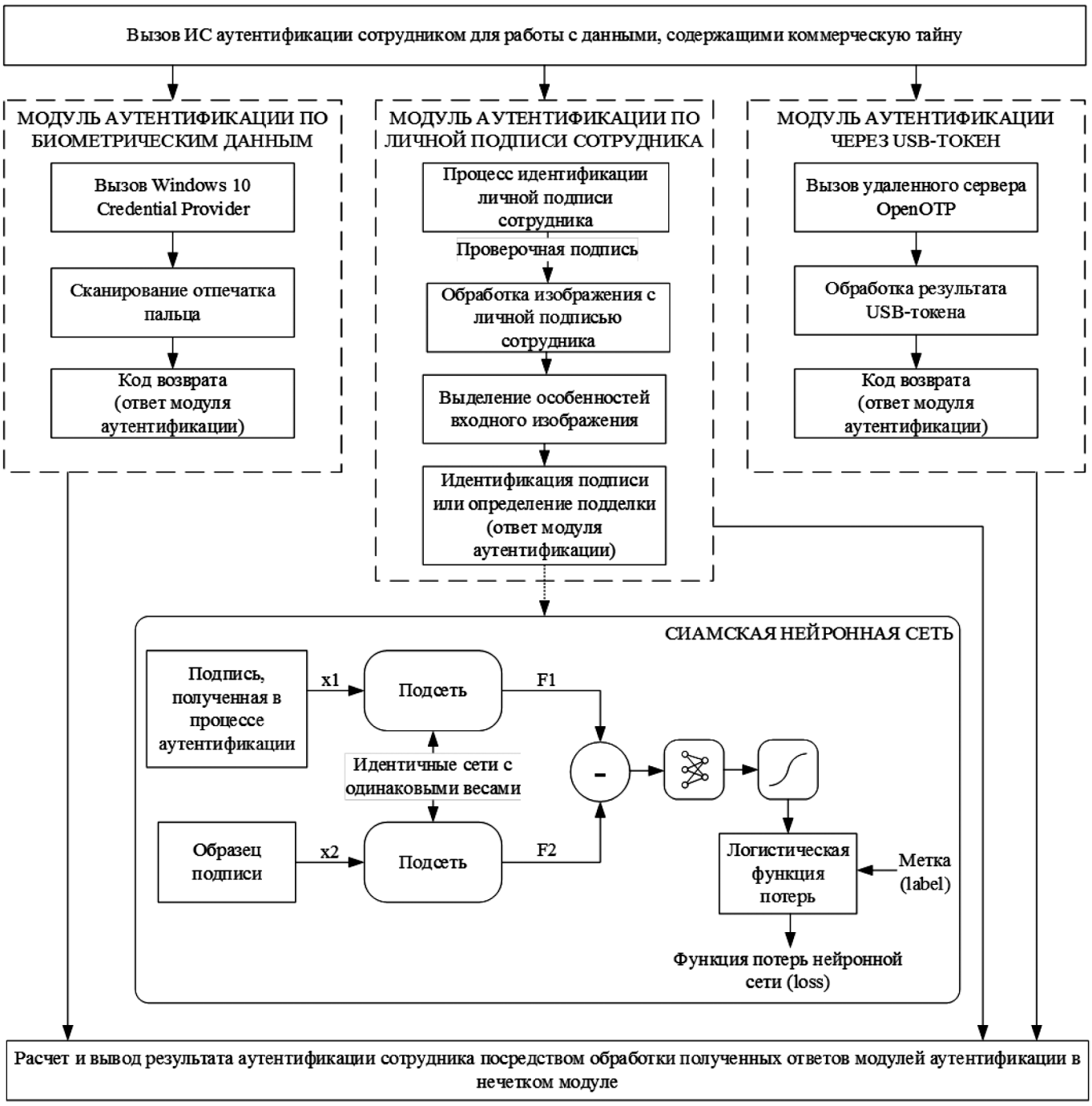

Рис. 2. Схема процесса аутентификации сотрудника с помощью метода многофакторной аутентификации, обеспечивающего информационную безопасность бизнес-процессов передачи данных, составляющих коммерческую тайну, в сложных экономических системах

обновления весовых коэффициентов и расчета показателей для каждой подсети применяемой сверточной сиамской нейронной сети [17–18].

На рисунке 3 представлен оконный интерфейс разработанной информационной системы с результатами работы метода аутентификации по личной подписи сотрудника на основе сиамских нейронных сетей.

Листинг 1

Программный код вычисления функции потерь для сиамской нейронной сети при идентификации пользователя

```
class ContrastiveLoss(torch.nn.Module):
 «»»
 Contrastive loss function.
 Based on:
```

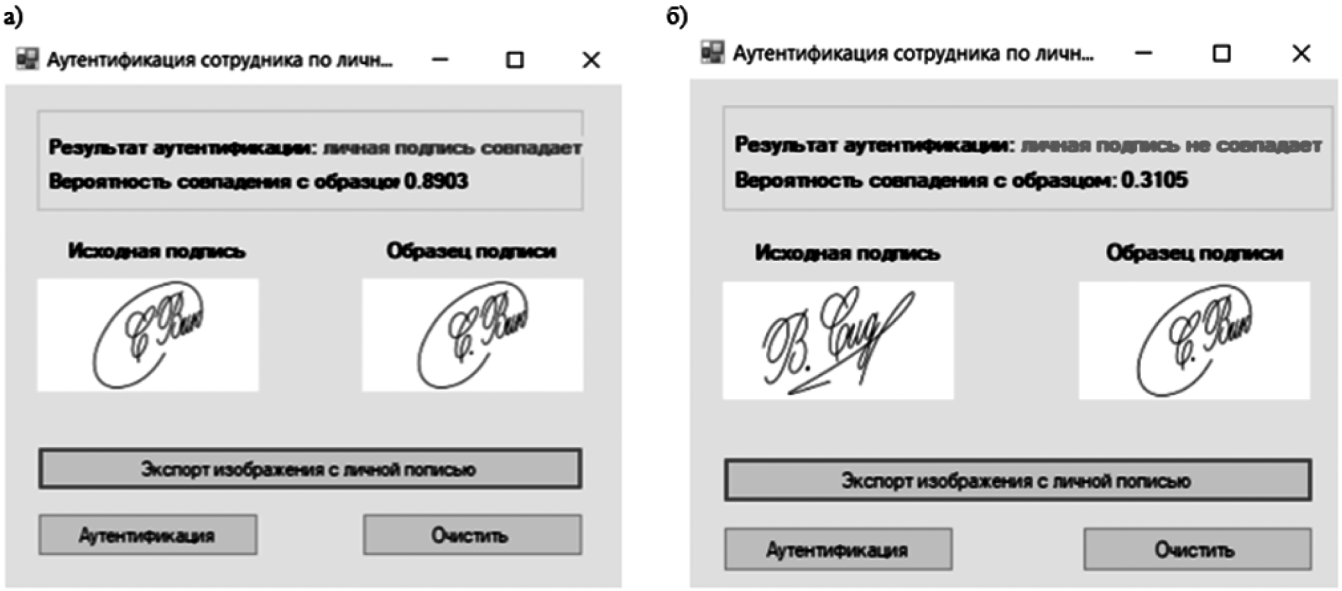

Рис. 3. Аутентификация сотрудника по личной подписи, используя метод аутентификации на основе сиамских нейронных сетей: а) успешной аутентификации подписи: б) неуспешной аутентификации подписи

```
«»»
def __init__(self, margin=1.0):
super(ContrastiveLoss, self).__init__()
self.margin = margin
def forward(self, x0, x1, y):
# euclidian distance
diff = x0 — x1
dist_sq = torch.sum(torch.pow(diff, 2), 1)
dist = torch.sqrt(dist_sq)
mdist = self.margin — dist
dist = torch.clamp(mdist, min=0.0)
loss = y * dist_sq + (1 — y) * torch.pow(dist, 2)
loss = torch.sum(loss) / 2.0 / x0.size()[0]
return loss
```

Результаты работы метода аутентификации по личной подписи сотрудника на основе сиамских нейронных сетей включают результат аутентификации («личная подпись совпадает» или «личная подпись не совпадает»), а также численный показатель вероятности совпадения с образцом подписи (по шкале от 0.0000 до 1.0000), который может быть передан в нечеткий модуль принятия решения о результатах аутентификации.

Заключение

Поставленная проблема обеспечения информационной безопасности при передаче данных, содержащих коммерческую тайну, между организациями-участниками сложных экономических систем, таких как территориальные научно-промышленные кластеры, в настоящее время решается методами, связанными с разверткой частных виртуальных сетей VPN и обеспечением изоляции основного соединения — использование прикладных протоколов PPTP, L2TP, SSTP, IKEv2. Однако по ранее отмеченным причинам они имеют ряд серьезных уязвимостей, что не позволяет их эффективно использовать в сложных экономических системах для обеспечения соответствующего уровня информационной безопасности.

Предложенный вариант решения вышеуказанной системы — многофакторная аутентификация с применением метода идентификации личности сотрудника на основе сиамских нейронных сетей как модуля аутентификации. Предложенная авторами идентификация по личной подписи сотрудника является одним из вариантов данного метода, выбор вариации (например, идентификация по личной подписи или идентификация по биометрическим данным — по изображению лица сотрудника) обосновывается принятым в организации стандартам идентификации. Предложенный способ идентификации подписей сотрудников организации показал свою работоспособность и высокую точность. Более того, предложенный способ возможно использовать в качестве модуля (канала) аутентификации при использовании многофакторной аутентификации, поскольку выходом модуля является показатель вероятности совпадения анализируемой подписи с образцом подписи из набора используемых образцов при обучении сиамской нейронной сети. Соответственно, указанный показатель может использоваться в модуле принятия решения об успешной или не успешной аутентификации сотрудника или пользователя системы, которому предоставлен доступ к данным, содержащим коммерческую тайну, при их передаче между организациями-участниками сложной экономической системы.

ЛИТЕРАТУРА

1. Dli M., Puchkov A., Meshalkin V., Abdeev I., Saitov R., Abdeev R. Energy and Resource Efficiency in Apatite-Nepheline Ore Waste Processing Using the Digital Twin Approach. Energies, 2020, 13, 5829.
2. Puchkov A., Dli M., Lobaneva E., Fedulov Y. Monitoring the Granulometric Composition on the Basis of Deep Neural Networks. In: Zamojski W., Mazurkiewicz J., Sugier J., Walkowiak T., Kacprzyk J. (eds) Theory and Engineering of Dependable Computer Systems and Networks. DepCoS-RELCOMEX 2021. Advances in Intelligent Systems and Computing, 2021, vol. 1389. Springer, Cham.
3. Su-Chang L., Jun-Ho H., Jong-Chan K. Deep Feature Based Siamese Network for Visual Object Tracking // Energies. 2022. № 15(17). P. 6388.
4. Vanita J., Prakhar G., Aditya C., Manas B., Hemanth D. Jude. A Modified Deep Convolution Siamese Network for Writer-Independent Signature Verification // International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems. 2022. № 30(3). P. 479–498.
5. Zhang X., Wu Z., Xie L., Li Y., Li F., Zhang J. Multi-path siamese convolution network for offline handwritten signature verification // ACM International Conference Proceeding Series, 2022, p. 51–58.
6. Vorugunti C.S., Devanur G.S., Mukherjee P., Pulabaigari V. OSVNet: Convolutional siamese network for writer independent online signature verification. Paper presented at the Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2019, p. 1470–1475.
7. Rateria A., Agarwal S. Off-line signature verification through machine learning // The 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON, 2018, 8597090.
8. Zhang H., Piao Y., Huang B., Tan B. SiamMBFAN: Siamese tracker with multi-branch feature aggregation network // Journal of Visual Communication and Image Representation, 2022, 89, 103671.
9. Han S., Shi L., Richie R., Tsui F.R. Building siamese attention-augmented recurrent convolutional neural networks for document similarity scoring // Information Sciences, 2022, 615, 90–102.
10. Sha Y., He Z., Gutierrez H., Du J., Yang W., Lu X. The intelligent detection method for flip chips using CBN-S-net algorithm with SAM images // Journal of Manufacturing Processes, 2022, 83, 60–67.
11. Kalsekar A., Khade R., Jariwala K., Chattopadhyay C. RISC-net: Rotation invariant siamese convolution network for floor plan image retrieval. Multimedia Tools and Applications, 2022, 81(28), 41199–41223.
12. Hong D., Gao L., Yao J., Yokoya N., Chanussot J., Heiden U., Zhang B. Endmember-guided unmixing network (EGU-net): A general deep learning framework for self-supervised hyperspectral unmixing // IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(11), 6518–6531.