УДК 004.85

Д.А. Андреевских, С.Д. Разбойников, М.П. Силич, К.С. Сарин

Сравнение применения локальных и глобальных признаков рукописной подписи для аутентификации пользователя

Проводится сравнение двух методов аутентификации пользователя на основе рукописной подписи. Первый основан на извлечении локальных признаков подписи как функций, зависящих от времени ее нанесения. Второй использует глобальные признаки подписи, характеризующие ее целиком без привязки к моментам нанесения. Для каждого из методов представлены алгоритмы аутентификации, проведены эксперименты на реальных рукописных подписях, полученных в ходе исследования.

Ключевые слова: рукописная подпись, аутентификация, машинное обучение, распознавание образов.

doi: 10.21293/1818-0442-2020-23-4-30-34

Поскольку в настоящее время электронные носители информации вытесняют бумажные, актуальной является проблема подтверждения личности в пространстве электронных документов. Для решения этой проблемы разработаны различные средства аутентификации пользователя, в том числе и с использованием динамических биометрических данных. Аутентификация по рукописной подписи является наиболее распространенным и проверенным временем методом подтверждения личности [1, 2]. Рукописная подпись представляется в виде дискретных сигналов, формируемых графическим планшетом. Сигналы образованы перемещением пера по поверхности планшета и описывают три характеристики подписи: форму, динамику и вариативность. Форма представляет графическое изображение подписи. Вариативность и динамика определяют положение, скорость, давление пера в разные моменты времени ее нанесения. Динамическое распознавание подписи предполагает наличие сигналов, описывающих динамику. Если доступно только изображение подписи, то применяют статическое распознавание.

Проверка рукописной подписи является проблемой распознавания образов [3–6]. Важным этапом в решении этой проблемы является извлечение признаков из исходных данных, представленных в виде сигналов. Признаки рукописной подписи – это величины, извлеченные из оцифрованных образцов подписей и описывающие подпись в виде вектора значений. Признаки должны позволять описывать каждую рукописную подпись так, чтобы степень различия между подписями у разных пользователей была максимальной, при этом допуская изменчивость подписей одного и того же пользователя [7, 8].

Признаки разделяют на две основных группы: локальные и глобальные [9–13]. К локальным признакам относятся признаки, характеризующие подпись в определенные отрезки времени. Иными словами, признаки, представимые в виде функции, зависящей от времени нанесения. В числе данных признаков выступают координаты X и Y , величина давления пера на экран планшета, скорость нанесения подписи и др. К глобальным признакам относятся те признаки, которые характеризуют подпись

как единое целое без привязки к отдельным отрезкам времени нанесения. В числе данных признаков можно выделить общее время подписи, количество поднятий пера во время нанесения подписи, средние значения скорости и ускорения ведения пера и др.

Целью настоящей работы является выявление группы признаков, с помощью которой проведение аутентификации пользователей было бы наиболее предпочтительным с точки зрения точности определения субъекта. Соответствующий круг задач заключается в необходимости изучить алгоритмы аутентификации на основе локальных и глобальных признаков, разработать и реализовать их в программной среде MatLab, после чего провести тестирование реализованных алгоритмов и оценить точность распознавания для каждого из алгоритмов.

Методика проведения экспериментов по снятию экземпляров подписи

Для проведения исследования была собрана база данных оригинальных и поддельных подписей по разработанной нами методике. Сигналы рукописных подписей снимались с помощью планшета Wacom STU-540, который фиксирует сигналы координат X и Y , величина давления пера P с частотой 200 Гц.

Участникам нужно было придумать подпись. В течение недели до снятия показаний они тренировались в ее нанесении. В день, когда проходило снятие сигналов подписи, фиксировались десять экземпляров. Процесс нанесения подписи записывался на видеоизображение. Через одну и две недели данная процедура повторялась. Далее было выбрано шесть фальсификаторов среди участников. В течение недели по изображению оригинальной подписи они должны были научиться ее воспроизводить. Спустя неделю каждый фальсификатор вносил в базу пять экземпляров подделанной подписи. После фальсификаторам выдавалось видеоизображение процесса нанесения подписи. Через неделю снималось еще пять экземпляров поддельной подписи. Со следующей недели цикл повторялся.

Таким образом, для каждого пользователя в базе присутствуют 30 оригинальных экземпляров подписей, 30 квалифицированных подделок, полученных на основе изображения оригинальной подписи,

30 квалифицированных подделок на основе видео-изображения процесса нанесения.

Описание наборов данных

Файл с подписью имеет структуру таблицы. Каждая строка представляет значение сигналов подписи в один момент времени. В первых двух столбцах фиксируются координаты X и Y кончика пера, в третьем столбце фиксируется давление пера на экран планшета P .

На рис. 1 продемонстрированы изображения трех подписей и соответствующие им сигналы в каждый момент времени t .

Аутентификация на основе локальных признаков

Алгоритм аутентификации с применением локальных признаков основан на алгоритме динамической трансформации временной шкалы (DTW) [14, 15].

Рис. 1. Сигналы рукописной подписи

Для аутентификации пользователя при помощи алгоритма DTW в среде MatLab были написаны несколько вариантов программы, общий принцип действия которых можно описать следующим алгоритмом:

1. На основе множества оригинальных подписей $G = \{g_1, \dots, g_n\}$ проверяемого пользователя сформировать сигналы координат и величины давления пера подписи-эталона g_{ref} . Сигналы подписи-эталона вычислить как средние значения оригинальных подписей в каждый момент времени.

2. Вычислить пороговое значение DTW-расстояния $d_{threshold}$. Данное значение позволит определить, является ли проверяемая подпись оригинальной.

3. Вычислить DTW-расстояние между проверяемой подписью и подписью-эталон.

4. Если DTW-расстояние меньше $d_{threshold}$, то проверяемая подпись принадлежит пользователю, иначе проверяемая подпись – подделка.

Было проведено шесть экспериментов с различными видами сигналов и определяемым пороговым значением. Ниже представлено их описание.

Вариант 1. Эталон состоит из средних значений координат X , Y и величины давления пера на экран планшета P . Пороговое значение вычисляется по формуле (1):

$$d_{threshold} = \frac{\sum_{i=1}^n d(g_i, g_{ref})}{n} + 3 \cdot \sigma, \quad (1)$$

где $d(g_i, g_j)$ – DTW-расстояние между сигналами подписи g_i и g_j , σ – среднее квадратичное отклонение DTW-расстояний оригинальных подписей:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n \left(d(g_i, g_{ref}) - \frac{\sum_{j=1}^n d(g_j, g_{ref})}{n} \right)^2}{n}}.$$

Вариант 2. Эталон состоит из средних значений координат X , Y и величины давления пера на экран планшета P . Пороговое значение вычисляется по формуле (2):

$$d_{threshold} = \frac{\max_{i=1 \dots n} (d(g_i, g_{ref})) + \min_{i=1 \dots m} (d(s_i, g_{ref}))}{2}, \quad (2)$$

где $S = \{s_1, \dots, s_m\}$ – множество поддельных подписей, используемых для определения порога.

Вариант 3. Эталон состоит из средних значений сигналов абсолютной скорости и абсолютного ускорения, которые были вычислены с помощью сигналов X и Y . DTW-расстояние вычисляется для сигналов скорости V и ускорения A . Пороговое значение $d_{threshold}$ определяется формулой (1).

Вариант 4. Эталон состоит из средних значений сигналов V и A оригинальной подписи. Пороговое значение рассчитывается по формуле (2).

Вариант 5. Эталон формируется на основе значений сигналов X , Y , P , V , A . Пороговое значение определяется формулой (1).

Вариант 6. Эталон формируется на основе значений сигналов X , Y , P , V , A . Пороговое значение определяется формулой (2).

На основе представленных вариантов нахождения порога были проведены эксперименты по аутентификации, сформирована таблица точности (табл. 1) и построен график зависимости точности от числа оригинальных подписей, необходимых для формирования подписи-эталона (рис. 2).

Таблица 1

Значения средней точности для локальных признаков

№ варианта	Количество оригинальных подписей				
	3	4	5	6	7
1	0,724	0,738	0,729	0,723	0,709
2	0,802	0,802	0,783	0,778	0,780
3	0,716	0,744	0,703	0,683	0,676
4	0,768	0,753	0,732	0,736	0,710
5	0,709	0,721	0,734	0,708	0,717
6	0,798	0,787	0,786	0,779	0,767

Анализ полученных результатов выявил, что наилучшая точность аутентификации достигается при использовании варианта 2, т.е. с применением сигналов X , Y , P и пороговым значением, вычисленным по формуле (2). Используемое при этом количество оригинальных подписей n равно 3. Число поддельных подписей для вычисления порога соответствует числу оригинальных подписей, т.е. $m = n$.

Рис. 2. Зависимость средней точности от количества оригинальных подписей для локальных признаков

Аутентификация на основе глобальных признаков

Алгоритм аутентификации пользователя на основе глобальных признаков использует высокоуровневые признаки, представленные в работе [10]. В качестве инструмента принятия решения используются алгоритмы машинного обучения, а именно классификатор K -ближайших соседей (KNN), деревья решений (TW) и метод опорных векторов (SVM). В обучающих данных оригинальные и поддельные подписи помечаются разными метками и, по сути, проблема аутентификации сводится к проблеме бинарной классификации машинного обучения.

Шаги алгоритма аутентификации представлены ниже:

1. Провести извлечение глобальных признаков из обучающих и тестируемых подписей.
2. Сформировать модель машинного обучения на основе оригинальных и поддельных вариантов подписи пользователя.
3. Осуществить прогноз тестируемой подписи с помощью модели.

Для эксперимента была использована база рукописных подписей, содержащая как оригинальные, так и поддельные подписи. Модели строились на различном количестве данных, а именно: количество оригинальных подписей варьировалось от 3 до 7, а количество поддельных – от 5 до 15.

По результатам экспериментов можно сделать следующий вывод. Наилучшие результаты показаны при 15 поддельных подписях. Сравнение точностей моделей представлено в табл. 2 и на рис. 3.

Проанализировав полученные данные экспериментов, выявлено, что алгоритм аутентификации пользователя с использованием модели SVM является наиболее точным. Число оригинальных и поддельных подписей для создания модели использовались 5 и 15 соответственно.

Таблица 2
Значения средней точности для глобальных признаков

Модель	Количество оригинальных подписей				
	3	4	5	6	7
KNN	0,640	0,644	0,644	0,638	0,643
TW	0,807	0,815	0,827	0,820	0,817
SVM	0,872	0,875	0,880	0,867	0,881

Рис. 3. Зависимость средней точности от количества оригинальных подписей для глобальных признаков

Сравнение алгоритмов аутентификации

Чтобы сравнить применение локальных и глобальных признаков, сравним результаты точностей алгоритма на основе DTW и алгоритма на основе SVM. Результаты значений точностей проведенных экспериментов для каждого пользователя показаны в табл. 3 и на рис. 4.

Таблица 3
Сравнение точностей алгоритмов

№ пользователя	Алгоритм	
	DTW	SVM
1	0,778	0,878
2	0,818	0,898
3	0,626	0,868
4	0,696	0,836
5	0,728	0,852
6	0,972	0,95
7	0,676	0,796
8	0,994	0,896
9	0,926	0,948

Рис. 4. График сравнения средней точности алгоритмов с глобальными и локальными признаками

Средняя точность аутентификации по всем пользователям для алгоритма, использующего глобальные признаки, с моделью классификации SVM оказалась выше и составила 0,880. Применение локальных признаков подписи в алгоритме на основе DTW позволило добиться точности 0,802.

Заключение

В результате проделанной работы проведено сравнение применимости локальных и глобальных признаков рукописных подписей для проблемы аутентификации пользователя. Признаки формировались на основе динамических сигналов подписи, снимаемых с поверхности графического планшета. Применение глобальных признаков в совокупности с моделями машинного обучения позволило получить более высокую точность аутентификации по сравнению с применением локальных признаков. Средняя точность в первом случае составила 0,880, во втором – 0,802.

Несмотря на более высокую точность, алгоритмы с применением моделей машинного обучения имеют один существенный недостаток. Для построения данных моделей необходимы экземпляры как оригинальных, так и поддельных подписей. При практическом применении это вызывает трудности, поскольку создать квалифицированную подделку весьма затруднительно, и точность аутентификации данных моделей, таким образом, будет зависеть от качества предоставленных подделок.

Работа выполнена при финансовой поддержке Министерства образования и науки РФ в рамках научных проектов, выполняемых коллективами научных лабораторий образовательных организаций высшего образования, подведомственных Министерству науки и высшего образования Российской Федерации (проект № FEWM-2020-0042).

Литература

1. Ходашинский И.А. Аутентификация пользователя по динамике подписи на основе нечеткого классификатора / И.А. Ходашинский, Е.Ю. Костюченко, К.С. Сарин, А.Е. Анфилофьев, М.Б. Бардамова, С.С. Самсонов, И.В. Филимонов // Компьютерная оптика. – 2018. – Т. 42, № 4. – С. 657–666.
2. Сарин К.С. Метод баггинга и отбор признаков в построении нечетких классификаторов для распознавания рукописной подписи / К.С. Сарин, И.А. Ходашинский // Компьютерная оптика. – 2019. – Т. 43, № 5. – С. 833–845.
3. Encyclopedia of Biometrics / ed. S.Z. Li, A.K. Jain – Boston: Springer, 2015. – 1626 p.
4. Vorugunti C.S. OSVFuseNet: Online Signature Verification by feature fusion and depth-wise separable convolution based deep learning / C.S. Vorugunti, V. Pulabaigari, R.K. Gorthi, P. Mukherjee // Neurocomputing. – 2020. – Vol. 409. – P. 157–172.
5. Song X.Y. Online Signature Verification Based on Stable Features Extracted Dynamically / X.Y. Song, X.H. Xia, F.J. Luan // IEEE Transactions on systems man cybernetics-systems. – 2017. – Vol. 47, No. 10. – P. 2663–2676.
6. Lai S.X. Recurrent Adaptation Networks for Online Signature Verification / S.X. Lai, L.W. Jin // IEEE Transactions on information forensics and security. – Vol. 14, No. 6. – P. 1624–1637.
7. Kholmatov A. Identity authentication using improved online signature verification method / A. Kholmatov, B. Yanikoglu // Pattern Recognition Letters. – 2005. – Vol. 26. – P. 2400–2408.
8. Yang L. Online handwritten signature verification based on the most stable feature and partition / L. Yang, X. Jin,

Q. Jiang // Cluster Computing. – 2019. – Vol. 22. – P. 1691–1701.

9. Wu G. Continuous Identity Authentication Scheme Based on Physiological and Behavioral Characteristics / G. Wu, J. Wang, Y. Zhang, S. Jiang // Sensors. – 2018. – Vol. 22. – P. 179.

10. Fierrez-Aguilar J. An On-line signature verification system based on fusion of local and global information / J. Fierrez-Aguilar, L. Nanni, J. Lopez-Penalba, J. Ortega-Garcia, D. Maltoni // Lecture Notes in Computer Science. – 2005. – Vol. 3546. – P. 523–532.

11. Parodi M. Online Signature Verification: Improving Performance through Pre-classification Based on Global Features / M. Parodi, J.C. Gomez // Lecture Notes in Computer Science. – 2013. – Vol. 8158. – P. 69–76.

12. Jain A. Signature verification using geometrical features and artificial neural network classifier / A. Jain, S.K. Singh, K.P. Singh // Multimedia tools and applications. – 2020. – Vol. 79, No. 27–28. – P. 19993–20018.

13. Ahrabian K. Usage of autoencoders and Siamese networks for online handwritten signature verification / K. Ahrabian, B. BabaAli // Neural computing and applications. – 2019. – Vol. 31, No. 12. – P. 9321–9334.

14. Sakoe H. Dynamic programming algorithm optimization for spoken word recognition / H. Sakoe, S. Chiba // IEEE Transactions on Acoustics, Speech and Signal Processing. – 1978. – Vol. 26, No. 1. – P. 43–49.

15. Salvador S. Toward accurate dynamic time warping in linear time and space / S. Salvador, P. Chan // Intelligent Data Analysis. – 2007. – Vol. 11, No. 5. – P. 561–580.

Андреевских Дарья Алексеевна

Студентка каф. комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС) Томского государственного университета систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7-962-781-73-18
Эл. почта: andreevskih.dar@yandex.ru

Разбойников Сергей Дмитриевич

Студент каф. КИБЭВС ТУСУРа
Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7-923-608-22-08
Эл. почта: leodaf36@mail.ru

Силич Мария Петровна

Д-р техн. наук, профессор каф. автоматизации обработки информации ТУСУРа
Ленина пр-т, д. 40, г. Томск, Россия, 634050
ORCID: 0000-0002-2555-8145
Тел.: +7 (382-2) 70-15-91
Эл. почта: smp@muma.tusur.ru

Сарин Константин Сергеевич

Канд. техн. наук, доцент каф. КИБЭВС ТУСУРа
Ленина пр-т, д. 40, г. Томск, Россия, 634050
ORCID: 0000-0003-3264-7768
Тел.: +7 (382-2) 70-15-29
Эл. почта: sks@security.tomsk.ru

Andreevskih D.A., Razboynikov S.D., Silich M.P., Sarin K.S.
Comparing the use of local and global features of handwritten signatures for user authentication

The paper compares two methods of user authentication based on a handwritten signature. The first is based on the extraction of local signature features as functions that depend on the time of its application. The second extracts global signature features that characterize it entirely without reference to the moments of application. Authentication algorithms are presented for each of the methods; experiments are carried out on real handwritten signatures obtained during the study.

Keywords: handwritten signature, authentication, machine learning, pattern recognition.

doi: 10.21293/1818-0442-2020-23-4-30-34

References

1. Hodashinsky I.A., Kostyuchenko E.Yu., Sarin K.S., Anfilofiev A.E., Bardamova M.B., Samsonov S.S., Filimonenko I.V. Dynamic-signature-based user authentication using a fuzzy classifier. *Computer Optics*, 2018, vol. 42, no. 4, pp. 657–666 (in Russ.).
2. Hodashinsky I.A., Sarin K.S. Bagged ensemble of fuzzy classifier and feature selection for handwritten signature verification. *Computer Optics*, 2019, vol. 43, no. 5, pp. 833–845 (in Russ.).
3. Encyclopedia of Biometrics, ed. S.Z. Li, A.K. Jain, Boston, Springer, 2015. 1626 p.
4. Voruganti C.S., Pulabaigari V., Gorthi R.K., Mukherjee P. OSVFuseNet: Online Signature Verification by feature fusion and depth-wise separable convolution based deep learning. *Neurocomputing*, 2020, vol. 409, pp. 157–172.
5. Song X.Y., Xia X.H., Luan F.J. Online Signature Verification Based on Stable Features Extracted Dynamically. *IEEE Transactions on systems man cybernetics-systems*, 2017, vol. 47, no. 10, pp. 2663–2676.
6. Lai S.X., Jin L.W. Recurrent Adaptation Networks for Online Signature Verification. *IEEE Transactions on information forensics and security*. vol. 14, no. 6, pp. 1624–1637.
7. Kholmatov A., Yanikoglu B. Identity authentication using improved online signature verification method. *Pattern Recognition Letters*. 2005, vol. 26, pp. 2400–2408.
8. Yang L., Jin X., Jiang Q. Online handwritten signature verification based on the most stable feature and partition. *Cluster Computing*. 2019, vol. 22, pp. 1691–1701.
9. Wu G., Wang J, Zhang Y., S. Jiang Continuous Identity Authentication Scheme Based on Physiological and Behavioral Characteristics. *Sensors*. 2018, vol. 22, pp. 179.
10. Fierrez-Aguilar J., Nanni L., Lopez-Penalba J., Ortega-Garcia J., Maltoni D. An On-line signature verification system based on fusion of local and global information. *Lecture Notes in Computer Science*, 2005, vol. 3546, pp. 523–532.
11. Parodi M., Gomez J.C. Online Signature Verification: Improving Performance through Pre-classification Based on

Global Features. *Lecture Notes in Computer Science*, 2013, vol. 8158, pp. 69–76.

12. Jain A., Singh S.K., Singh K.P. Signature verification using geometrical features and artificial neural network classifier. *Multimedia tools and applications*, 2020, vol. 79, no. 27–28, pp. 19993–20018.

13. Ahrabian K., BabaAli B. Usage of autoencoders and Siamese networks for online handwritten signature verification. *Neural computing and applications*, 2019, vol. 31, no. 12, pp. 9321–9334.

14. Sakoe H., Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1978, vol. 26, no. 1, pp. 43–49.

15. Salvador S., Chan P. Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 2007, vol. 11, no. 5, pp. 561–580.

Daria A. Andreevskih

Student, Department of Complex Information Security of Computer Systems
 Tomsk State University of Control Systems and Radioelectronics (TUSUR)
 40, Lenin pr., Tomsk, Russia, 634050
 Phone: +7-962-781-73-18
 Email: andreevskih.dar@yandex.ru

Sergey D. Razboynikov

Student, Department of Complex Information Security of Computer Systems TUSUR
 40, Lenin pr., Tomsk, Russia, 634050
 Phone: +7-923-608-22-08
 Email: leodaf36@mail.ru

Maria P. Silich

Doctor of Engineering Sciences, Professor,
 Department of Information Processing Automation, TUSUR
 40, Lenin pr., Tomsk, Russia, 634050
 ORCID 0000-0002-2555-8145
 Phone: +7(382-2) 70-15-91
 Email: smp@muma.tusur.ru

Konstantin S. Sarin

Candidate of Engineering Sciences, Assistant Professor,
 Department of Complex Information Security of Computer Systems TUSUR
 40, Lenin pr., Tomsk, Russia, 634050
 ORCID 0000-0003-3264-7768
 Phone: + 7 (382-2) 70-15-29
 Email: sks@security.tomsk.ru