

Статические алгоритмы выделения уникальных признаков из рукописной подписи человека

Р.Н. Дзямко-Гамулец

Московский технический университет связи и информатики

Аннотация: Одними из самых надёжных способов подтверждения личности являются биометрические методы аутентификации. Существует два вида методов: статические и динамические. К статическим методам относятся сканирование отпечатков пальцев, 3D лица, рисунков вен, глазной сетчатки и т.д. К динамическим методам относятся верификация голоса, клавиатурного почерка и рукописной подписи. Цель данной работы - разобрать один из динамических методов биометрической аутентификации, который может быть использован в большинстве внутренних и внешних информационных систем, как один из инструментов авторизации или подтверждения намерений со стороны пользователя. В данной работе рассматриваются алгоритмы, выделяющие уникальные признаки из статических характеристик подписи, так как большинство особенностей подписи выделяются именно из зависимостей от написания отдельных фрагментов изображения подписи.

Ключевые слова: статические алгоритмы, метрики, длина подписи, масштабирование, угол подписи.

Введение

Среди способов авторизации пользователей в информационных системах большими перспективами обладают биометрические методы аутентификации, основанные на поведенческих характеристиках человека, учитывающих особенности подсознательных движений пользователя в процессе воспроизведения определенных действий (аутентификация по рукописному/клавиатурному почерку, голосу и др.) [1].

В данной статье рассматривается один из таких методов – аутентификация человека по его рукописной подписи в режиме онлайн.

Общий алгоритм работы метода заключается в сборе экземпляров подписей, выделении уникальных признаков, нахождении усреднённой подписи и сравнении тестируемых экземпляров с усреднённой и её признаками.

Существуют различные способы выделения признаков, они могут быть классифицированы по-разному в зависимости от того, на чем делается

акцент в классификации. В частности, алгоритмы выделения признаков можно разделить на две группы статические и динамические [2].

Статические работают с подписью как со множеством точек, без учёта временного распределения написания этих точек, даже если само устройство фиксирует временные метки их появления.

Динамические анализируют всю информацию, полученную от устройства в процессе ввода подписи, в этом случае подпись представлена в виде двух функций координат, зависящих от времени.

Считается, что динамические признаки дают больше информации, однако статические позволяют выявить особенности структуры подписи в качестве геометрической кривой, что даёт возможность исследовать и сравнивать именно геометрические особенности имеющихся экземпляров и выявлять закономерности, которые могли бы оставаться незамеченными при динамическом анализе. В данной работе рассматриваются именно статические признаки и алгоритмы для их выделения.

Способы сравнения подписей

Существует большое количество подходов, которые по исходным данным получают тензорные величины (они в частных случаях могут быть обычными числами, векторами, матрицами), называемые признаками.

Именно эти признаки и используются для того, чтобы производить сравнение подписей. На множестве таких объектов можно ввести понятие метрики. Данная величина позволит определить, насколько близко или далеко два объекта расположены друг от друга.

Введение метрики превращает имеющееся множество в метрическое пространство, однако при условии, если сама метрика введена корректно [3]. Для этого требуется, чтобы функция ρ , называемая метрикой, удовлетворяла следующим аксиомам (1-3):

$$\rho(x, y) = 0 \Leftrightarrow x = y, \quad (1)$$

$$\rho(x, y) = \rho(y, x), \quad (2)$$

$$\rho(x, z) \leq \rho(x, y) + \rho(y, z), \quad (3)$$

для $\forall x, y, z \in M$, где M – множество рассматриваемых элементов.

Функция метрики определена следующим образом (4):

$$\rho: M \times M \rightarrow \mathbb{R}, \quad (4)$$

то есть это отображение, на котором любым двум элементам из множества M ставится в соответствие действительное число.

Таким образом, определив понятие метрики, можно сравнить два экземпляра подписи на похожесть. Однако необходимо определить, каким образом выделяются статические признаки в виде тензорных величин.

В данной работе рассматриваются следующие признаки:

1. Угол наклона подписи;
2. Отношение высоты к ширине;
3. Длина подписи, как геометрической кривой;
4. Матрица расстояний между точками подписи;
5. Массив эллипсоидов инерции подписи.

Здесь и далее считается, что вся необходимая предобработка для конкретного признака была выполнена ранее.

Угол наклона подписи

Во время написания подписи устройство не всегда находится ровно в горизонтальном или ровно в вертикальном положении. Для решения данной проблемы проводится предобработка данных, один из этапов которой

заключается в повороте подписи на некоторый угол, который определяется, исходя из её структуры и позволяет значительно уменьшить проблемы, возникающие из-за положения устройства [4].

Однако угол наклона устройства также может дать некоторую информацию о пользователе. Кроме того, поскольку наклон определяется, исходя из геометрии подписи, а поддельная подпись может нарушить её, это закономерно приведёт к изменению угла наклона.

Алгоритм нахождения угла наклона состоит из следующих этапов:

1. Нахождение на временной шкале центральной точки во время создания подписи;
2. Разбиение всего множества точек на два подмножества: точки, временные метки которых меньше найденной середины, и те, которые больше;
3. Нахождение геометрического центра каждого из этих двух подмножеств;
4. Проведение прямой через центры каждого из подмножеств. Полученный угол между прямой и осью абсцисс и будет углом наклона подписи к оси абсцисс.

Формула расчёта угла наклона подписи имеет следующий вид (5):

$$\alpha = \tan^{-1} \left(\frac{y_{2c} - y_{1c}}{x_{2c} - x_{1c}} \right) (\pm \pi), \quad (5)$$

где (x_{1c}, y_{1c}) – координаты геометрического центра первого множества точек, а (x_{2c}, y_{2c}) – второго, число π прибавляется в зависимости от необходимой четверти.

В итоге сравнение подписей по углу наклона является сравнением двух чисел, что выражается следующей формулой (6):

$$\rho(\alpha_1, \alpha_2) = |\alpha_1 - \alpha_2|, \quad (6)$$

Отношение высоты к ширине

Данный признак позволяет выделить соотношение сторон подписи. Для этого необходимо определить следующие понятия: ширина подписи, высота подписи. Предположим, что подпись представляет собой множество, состоящее из N точек (7):

$$P = \{P_i, i = \overline{1..N}\}, \quad (7)$$

Тогда каждая точка имеет вид (8):

$$P_i = (x_i, y_i), i = \overline{1..N}, \quad (8)$$

Рассмотрим множество из координат x всех точек P_i , а также аналогично для y (9):

$$X = \{x_i, i = \overline{1..N}\}, Y = \{y_i, i = \overline{1..N}\}, \quad (9)$$

Тогда ширина подписи (10):

$$w = \max\{x_i\} - \min\{x_i\}, \quad (10)$$

Высота подписи равна (11):

$$h = \max\{y_i\} - \min\{y_i\}, \quad (11)$$

Таким образом, получается коэффициент (12):

$$k = \frac{h}{w}, \quad (12)$$

Сравнение происходит по правилам обычных чисел, что приводит к итоговому выражению сравнения подписей по признаку (13):

$$\rho(k_1, k_2) = |k_1 - k_2|, \quad (13)$$

Длина подписи как геометрической кривой

Длина подписи зависит от размера области, которую использует пользователь для ввода подписи на устройстве. На различных устройствах размер экрана может отличаться, что приводит к тому, что размеры подписи зависят от устройства ввода.

Однако для конкретного устройства пользователь обычно имеет более-менее одинаковый способ ввода, что делает возможным сравнение экземпляров по рассматриваемому параметру.

Кроме того, неправильно выполненная подпись приведёт к искажению длины подписи, что позволит улучшить защиту от фальсификации [5].

Пусть N – количество точек подписи, тогда формула расчёта длины выглядит следующим образом (14):

$$l = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}, \quad (14)$$

В данном случае кривая аппроксимируется ломаной, и её длина вычисляется как сумма длин отрезков линий, соединяющих соседние точки.

Стоит отметить, что данные вычисления следует проводить до масштабирования, при этом их можно проводить до поворота подписи или же после, так как поворот геометрической кривой не влияет на её длину [6].

Вычисление расстояния между длинами различных экземпляров подписей происходит по правилам сравнения чисел, аналогично предыдущим случаям.

Матрица расстояния между точками подписи

Матрица расстояний – ещё один признак, который произвольному экземпляру кривой подписи ставит в соответствие матрицу, составленную из расстояний между каждой парой точек этой кривой. Это позволяет анализировать структурную составляющую, которая зависит от расположения точек друг относительно друга.

Пусть текущий экземпляр подписи имеет N точек. Матрица расстояний представляет числовую матрицу с числами r_{ij} (15), равными расстоянию между точками P_i и P_j (16, 17):

$$r_{ij} = \rho(P_i, P_j), \quad (15)$$

$$P_i = (x_i, y_i), P_j = (x_j, y_j), \quad (16)$$

$$\rho(P_i, P_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (17)$$

Обозначим эту матрицу (18):

$$M = \{m_{ij}\}, i = \overline{1..N}, j = \overline{1..N}, \quad (18)$$

Для сравнения двух матриц и задания метрики на этом множестве можно взять следующую функцию (19):

$$\rho(M_1, M_2) = |M_1 - M_2|_2, \quad (19)$$

то есть берётся 2-векторная норма от разности матриц.

Разность вычисляется поэлементно. В итоге векторная 2-норма для матрицы будет иметь вид (20):

$$|M|_2 = \sqrt{\sum_{i=1}^N \sum_{j=1}^N m_{ij}^2}, \quad (20)$$

Массив эллипсоидов инерции подписи

Ещё один признак, который позволяет определить структуру подписи – набор эллипсоидов инерции точек подписи относительно различных осей и центров. Данный подход позволяет проанализировать разброс точек в различных областях, отличающихся размерами и положением. Эллипсоид инерции определяется моментами инерции относительно трёх осей, поэтому для его определения необходимо вычислить сами моменты инерции.

Для выделения признаков по этому методу рассматривается множество точек как некоторая масса и вычисляются её моменты инерции относительно каждой из трёх осей [7]. Оси x и y расположены в плоскости ввода устройства, а ось z – перпендикулярно ему. Считается, что каждая точка имеет единичную массу.

Алгоритм расчёта следующий – сперва рассматривается вся подпись, ищется центр масс точек. Далее найденный центр и будет центром

новых координат, полученных параллельным переносом исходных осей [8].
Относительно этих новых осей вычисляются три момента инерции (21-23):

$$I_x = \sum_{i=1}^N r_{xi}, \text{ где } r_{xi} = |y_i|, \quad (21)$$

$$I_y = \sum_{i=1}^N r_{yi}, \text{ где } r_{yi} = |x_i|, \quad (22)$$

$$I_z = \sum_{i=1}^N r_{zi}, \text{ где } r_{zi} = \sqrt{x_i^2 + y_i^2}, \quad (23)$$

Здесь используются координаты точек в новых осях. Далее исходная область делится на восемь подобластей: по две области при делении одной горизонтальной или вертикальной линией, проходящих через центр, а также по четыре области при одновременном использовании горизонтальной и вертикальной линий [9]. В каждой из этих подобластей снова находятся три момента инерции, но учитываются только те точки, которые попадают в эту область.

Затем каждая из подобластей опять разбивается. Останавливать этот процесс следует после заранее заданного числа разбиений или же пока полученная область не будет слишком мала.

Под k подразумевается количество получившихся областей при таком разбиении. Необходимо выбрать способ нумерации этих областей, тогда в качестве выделенного признака получается следующий вектор (24):

$$I = (I_{x1}, I_{y1}, I_{z1}, \dots, I_{xk}, I_{yk}, I_{zk}), \quad (24)$$

Сравнение двух таких векторов можно делать по евклидовой норме разности (25, 26):

$$\rho(I_1, I_2) = |I_1 - I_2|_2, \quad (25)$$

$$\text{где } |I|_2 = \sqrt{I_{x1}^2 + \dots + I_{zk}^2}, \quad (26)$$

Данный признак выделяется после всех этапов предобработки и анализирует структуру распределения точек на различных участках подписи как геометрической кривой.

Заключение

В нашей работе были рассмотрены алгоритмы выделения признаков, использующие данные о расположении точек подписи без учёта их временной зависимости. Рассмотрены как простейшие способы выделения признаков, такие, как угол поворота подписи или её длина, как геометрическая кривая, так и более сложные в виде вектора моментов инерции различных областей, а также матрицы расстояний между точками подписи.

Было определено понятие метрики для возможности сравнения двух экземпляров подписи, на основе выделенных из них признаков.

Сами признаки используются в качестве способа сравнения двух подписей для определения подлинности полученного экземпляра подписи. Качественный выбор признаков является непростой задачей, где главный критерий удачного выбора - проведение эксперимента и анализ результатов [10].

Однако статические методы анализа подписи имеют ограничения, например, они не учитывают временную зависимость написания геометрических точек [11]. Поэтому рекомендуется дополнительно применять и другие алгоритмы, не имеющие описанных выше ограничений.

Литература

1. Постникова М.С. Современные технологии обеспечения информационной безопасности // Тенденции развития науки и образования. 2021. № 72-1. С. 98-100.
URL: elibrary.ru/download/elibrary_45777209_34845254.pdf

2. Егорова С.С. Актуальные проблемы анализа рукописной подписи: судебно-почерковедческая экспертиза // Проблемы становления гражданского общества. Сборник статей IX Международной научной студенческой конференции. Иркутск, 2021. С. 276-277.
URL: elibrary.ru/download/elibrary_46576420_64350411.pdf
 3. Друки А.А., Милешин М.А. Применение нейросетевых алгоритмов для распознавания рукописных подписей на изображениях // Молодёжь и современные информационные технологии. Сборник трудов XII Международной научно-практической конференции студентов, аспирантов и молодых учёных. Национальный исследовательский Томский политехнический университет. 2014. С. 116-117.
URL: elibrary.ru/download/elibrary_23291859_60897958.pdf
 4. Аникин И.В., Анисимова Э.С. Распознавание рукописных подписей на основе нечётких признаков и метода потенциалов // Информация и безопасность. 2016. Т. 19. № 4. С. 567-570.
URL: elibrary.ru/download/elibrary_27194996_10853860.pdf
 5. Поздняков М.А. Применение алгоритма скелетизации изображения рукописной подписи // Решетневские чтения. Материалы XXV Международной научно-практической конференции, посвящённой памяти генерального конструктора ракетно-космических систем академика М.Ф. Решетнева. В 2-х частях. Под общей редакцией Ю.Ю. Логинова. Красноярск, 2021. С. 387-388.
URL: elibrary.ru/download/elibrary_47711945_39680611.pdf
 6. Ходашинский И.А., Бардамова М.Б. Модификации алгоритма прыгающих лягушек для отбора признаков в нечётком классификаторе при аутентификации пользователя по рукописной подписи
-

- // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 75-83.
URL: elibrary.ru/download/elibrary_44368696_86649911.pdf
7. Анисимова Э.С. Методы идентификации личности на основе рукописной подписи // Математика и её приложения: фундаментальные проблемы науки и техники. Сборник трудов всероссийской конференции. Алтайский государственный университет. 2015. С. 146-150.
URL: elibrary.ru/download/elibrary_24936236_13802316.pdf
8. Ахунджанов У.Ю. Нейросетевой подход для верификации рукописной подписи в режиме OFF-LINE // Компьютерные технологии и анализ данных (СТДА'2022). Материалы III Международной научно-практической конференции. Минск, 2022. С. 18-21.
URL: elibrary.ru/download/elibrary_48504124_14953133.pdf
9. Качайкин Е.И., Куликов С.В. Получение биометрических параметров высокого качества из статического изображения рукописной подписи // Инфокоммуникационные технологии. 2015. Т. 13. № 4. С. 446-450.
URL: elibrary.ru/download/elibrary_25993179_69687020.pdf
10. Заливин А.Н., Балабанова Т.Н., Прохоренко Е.И., Васильева Н.В. Сегментация сканированного рукописного текста на словные фрагменты // Экономика. Информатика. 2021. Т. 48. № 2. С. 383-391.
URL: elibrary.ru/download/elibrary_46437615_97360207.pdf
11. Ахунджанов У.Ю., Старовойтов В.В. Предварительная обработка изображений рукописных подписей для последующего распознавания // Системный анализ и прикладная информатика. 2022. № 2. С. 4-9.
URL: elibrary.ru/download/elibrary_49059362_16539285.pdf
-

References

1. Postnikova M.S. Tendencii razvitiya nauki i obrazovaniya. 2021. № 72-1. pp. 98-100. URL: elibrary.ru/download/elibrary_45777209_34845254.pdf.
2. Egorova S.S. Problemy` stanovleniya grazhdanskogo obshhestva. Sbornik statej IX Mezhdunarodnoj nauchnoj studencheskoj konferencii. Irkutsk, 2021. pp. 276-277. URL: elibrary.ru/download/elibrary_46576420_64350411.pdf
3. Druki A.A., Milesin M.A. Molodyozh` i sovremennyy`e informacionny`e texnologii. Sbornik trudov XII Mezhdunarodnoj nauchno-prakticheskoy konferencii studentov, aspirantov i molody`x uchyony`x. Nacional`ny`j issledovatel`skij Tomskij politexnicheskij universitet. 2014. pp. 116-117. URL: elibrary.ru/download/elibrary_23291859_60897958.pdf
4. Anikin I.V., Anisimova E`.S. Informaciya i bezopasnost`. 2016. T. 19. № 4. pp.567-570. URL: elibrary.ru/download/elibrary_27194996_10853860.pdf
5. Pozdnyakov M.A. Reshetnevskie chteniya. Materialy` XXV Mezhdunarodnoj nauchno-prakticheskoy konferencii, posvyashhyonnoj pamyati general`nogo konstruktora raketno-kosmicheskix sistem akademika M.F. Reshetneva. V 2-x chastyax. Pod obshhej redakciej Yu.Yu. Loginova. Krasnoyarsk, 2021. pp. 387-388. URL: elibrary.ru/download/elibrary_47711945_39680611.pdf
6. Xodashinskij I.A., Bardamova M.B. Informacionny`e i matematicheskie texnologii v nauke i upravlenii. 2020. № 4 (20). pp. 75-83. URL: elibrary.ru/download/elibrary_44368696_86649911.pdf
7. Anisimova E`.S. Matematika i eyo prilozheniya: fundamental`ny`e problemy` nauki i texniki. Sbornik trudov vserossijskoj konferencii.



- Altajskij gosudarstvennyj universitet. 2015. pp. 146-150.
URL: elibrary.ru/download/elibrary_24936236_13802316.pdf
8. Axundzhanov U.Yu. Komp'yuternye texnologii i analiz dannyx (CTDA'2022). Materialy III Mezhdunarodnoj nauchno-prakticheskoy konferencii. Minsk, 2022. pp. 18-21.
URL: elibrary.ru/download/elibrary_48504124_14953133.pdf
9. Kachajkin E.I., Kulikov S.V. Infokommunikacionnye texnologii. 2015. T. 13. № 4. pp. 446-450.
URL: elibrary.ru/download/elibrary_25993179_69687020.pdf
10. Zalivin A.N., Balabanova T.N., Proxorenko E.I., Vasil'eva N.V. E'konomika. Informatika. 2021. T. 48. № 2. pp. 383-391.
URL: elibrary.ru/download/elibrary_46437615_97360207.pdf
11. Axundzhanov U.Yu., Starovojtov V.V. Sistemnyj analiz i prikladnaya informatika. 2022. № 2. pp. 4-9.
URL: elibrary.ru/download/elibrary_49059362_16539285.pdf

Дата поступления: 13.02.2024

Дата публикации: 17.03.2024