

АУТЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЯ ПО ДИНАМИКЕ ПОДПИСИ НА ОСНОВЕ НЕЧЁТКОГО КЛАССИФИКАТОРА

И.А. Ходашинский¹, Е.Ю. Костюченко¹, К.С. Сарин¹, А.Е. Анфилофьев¹,
М.Б. Бардамова¹, С.С. Самсонов¹, И.В. Филимоненко¹

¹Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

Аннотация

Анализ динамики подписи является одним из наиболее быстрых, интуитивно понятных и экономичных инструментов аутентификации пользователей. Динамическое распознавание подписи основано на анализе нескольких характеристик почерка индивидуума, таких как параметры движения, давление, азимут и угол наклона пера в определенные моменты времени, а также скорости и ускорения вышеперечисленных величин. В нашей работе в качестве признаков были использованы постоянная составляющая и первые семь гармоник разложения данных сигналов в ряд Фурье. Создание систем подтверждения подлинности подписи включает следующие этапы: предобработка, отбор информативных признаков, классификация. Для отбора признаков использованы бинарные метаэвристические и детерминированные алгоритмы. Классификация выполнялась с помощью нечеткого классификатора. Параметры нечетких классификаторов настраивались непрерывными метаэвристическими алгоритмами. Работоспособность системы аутентификации проверена на авторской базе данных. База данных содержит 280 оригинальных вариантов подписи одного автора и 1281 вариант фальсификаций (поддельных подписей) семи авторов. Для оценки статистической значимости различий в точности и ошибках нечетких классификаторов, сформированных метаэвристическими алгоритмами, использованы критерий Манна–Уитни (Уилкоксона) и тест Крускала–Уоллиса.

Ключевые слова: распознавание образов, обработка информации, алгоритмы, отбор признаков, нечеткий классификатор, распознавание подписи.

Цитирование: 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. Филимоненко // Компьютерная оптика. – 2018. – Т. 42, №4. – С. 657-666. – DOI: 10.18287/2412-6179-2018-42-4-657-666.

Введение

Понятие «биометрия» связано с измеряемыми физиологическими (отпечатки пальцев, радужка глаза, лицо и др.) и поведенческими (речь, походка, нажатие клавиш, рукописная подпись и др.) особенностями человека, которые могут быть использованы для распознавания индивида. Достоинство биометрических систем определяется двумя основными факторами [1]: 1) пользователям не нужно запоминать пароли, PIN-коды или носить ключи доступа, 2) биометрические данные трудно украсть или подделать.

Биометрическая аутентификация, основанная на поведенческих особенностях человека, сложнее, чем проверка на основе физиологических, но она и более надежна в силу того, что в динамических признаках, характеризующих поведенческие особенности, проявляется уникальность человека, которую гораздо сложнее фальсифицировать.

Рукописная подпись является исторически подтвержденным и наиболее часто используемым средством аутентификации человека, выполняется она путем перемещения пера по поверхности графического планшета и имеет три важнейших атрибута: форму, динамику и вариативность. Динамика является самой важной характеристикой подписи [2]. Форма – осознанный аспект, динамика и вариативность – бессознательный.

Построение систем подтверждения подлинности подписи включает следующие этапы: предобработка, отбор информативных признаков, классификация.

Целью нашей работы является описание нового подхода к биометрической аутентификации на основе динамики подписи с использованием нечетких классификаторов и отбора информативных биометрических признаков.

1. Близкие работы

1.1. Подпись

Рукописная подпись – это общепринятый биометрический атрибут, используемый для аутентификации личности. Можно выделить два класса атрибутов рукописной подписи. Первый использует так называемую статическую (off-line) подпись и основан на анализе геометрических признаков подписи, таких как отношения формы и размера и т. д. Другой подход основан на анализе динамики процесса (on-line) подписи. Аутентификация пользователя по динамике подписи намного эффективнее статической [3], потому что: а) динамика подписи является очень индивидуальной характерной чертой подписывающего лица, б) трудно подделать, в) восстановление подписи по форме сигналов, описывающих динамику, довольно трудная задача. Существуют различные подходы к анализу динамики подписи [3].

Первый подход основан на использовании глобальных и локальных признаков для обучения классификатора. Здесь глобальные признаки формируются из подписи как единого образа, а локальные признаки выделяются из отдельных частей подписи. В [4] рассматривается метод, в основу которого положены при-

знаки подписи, определенные с применением нечеткой логики. Нечеткая система применяется для проверки динамической подписи с использованием глобальных признаков в [3]; отбор признаков выполняется здесь с использованием генетического алгоритма, причем для каждого подписанта определен индивидуальный набор признаков; признакам заданы веса, которые учитываются в процессе классификации.

Функциональный подход, в котором атрибуты подписи (скорость, ускорение, направление движения пера, давление) представляют в виде временных рядов. Сравнение выполняется с использованием мер эластичного расстояния, например, динамической трансформации временной шкалы [5–7].

Разделение подписи на регионы, которые используются на этапе обучения и проверки. В [8] предложен алгоритм, разбивающий скорость на три диапазона; авторы показывают, что только диапазон средних скоростей может успешно использоваться для распознавания. В [9] подпись разбивалась на несколько разделов, каждый раздел имел свою собственную кодовую книгу и значение веса; окончательный результат основан на слиянии результатов оценки каждой кодовой книги. В [10] предлагается разделение подписи на вертикальные и горизонтальные участки; вертикальные сегменты соответствуют начальным, средним и последним моментам времени выполнения подписи; горизонтальные участки соответствуют областям подписи, связанным с высокой и низкой скоростью пера и высоким и низким давлением пера.

Гибридный подход основан на сочетании различных методов, упомянутых выше. В [11] предлагается подход для проверки подписи на основе временного ряда (x, y) , коэффициентов подобия и статистики Хотеллинга, позволяющей сократить объем данных, необходимых для классификации; признаки и ассоциированные с ними коэффициенты подобия создают новые составные признаки, применение которых повышает точность классификации. В [12] представлена система аутентификации на основе ансамбля локальных, региональных и глобальных сопоставлений; здесь учитываются следующие подходы: слияние двух локальных методов, использующих метод динамической трансформации временной шкалы, подход на основе скрытой марковской модели, где каждая подпись описывается с помощью её локальных свойств.

1.2. Отбор признаков

Увеличение количества признаков приводит к ухудшению эффективности работы алгоритмов обучения, вопреки интуитивному представлению о том, что большее количество признаков позволит получить больше информации и с большей точностью выполнить задачу классификации. Причина в том, что по мере увеличения количества признаков алгоритмы нуждаются в большем количестве данных для обучения, необходимых для построения правил, определяющих соотношение между этими признаками и меткой класса. Кроме того, признаки, не содержащие информации о метке класса, могут способствовать

неправильной классификации и замедлять процесс обучения. Поэтому разработка методов отбора подмножества признаков, которые классифицируют данные более точно, чем весь набор признаков, является актуальной задачей [13].

Отбор признаков – это процедура выделения из исходного множества такого подмножества признаков, которое в полной мере соответствует решаемой задаче или задаче обучения. Цель отбора признаков: 1) избежать переобучения, 2) уменьшить объем анализируемых данных, 3) улучшить эффективность классификации, 4) устранить нерелевантные и «шумовые» признаки, 5) повысить интерпретируемость полученных результатов [14].

Методы отбора признаков делятся на три категории: фильтры (*filters*), обертки (*wrappers*) и встроенные (*embedded*) методы [15]. Метод фильтров основан на обобщенных свойствах обучающих данных и не включает в процесс отбора признаков собственно алгоритм построения классификатора. Достоинством данного метода является сравнительно невысокая вычислительная сложность, неплохая обобщающая способность и независимость от классификатора; основным недостатком данного метода заключается в том, что признаки, как правило, отбираются независимо друг от друга [15]. Метод обертки включает построение классификатора в процессе отбора признаков и использует прогностическую точность классификатора для оценки эффективности отобранного подмножества признаков. Такое взаимодействие с классификатором, как правило, дает результаты лучшие, чем метод фильтра, однако при этом возрастает вычислительная сложность метода и существует риск переобучения [15]. Встроенные методы выполняют отбор признаков в процессе обучения и включают алгоритм отбора признаков в алгоритм построения классификатора.

В работе [15] отмечается, что не существует лучшего метода отбора признаков, и усилия специалистов должны быть сосредоточены на поиске хорошего метода для каждой конкретной проблемы.

2. Биометрические признаки

Онлайн-система анализа рукописной подписи основана на динамическом методе, который использует графический планшет в качестве устройства ввода. Планшет позволяет формировать следующие динамические последовательности дискретного времени: 1) положение пера по осям x, y, z ; 2) P – давление, прикладываемое пером; 3) азимут α ; 4) угол высоты пера относительно планшета θ . Указанные последовательности раскладываются в ряды с использованием преобразования Фурье, Уолша, Хаара или вейвлет-преобразования [16].

В нашей работе в качестве признаков были использованы постоянная составляющая и первые семь гармоник разложения подписи в ряд Фурье [17]. Параметры и сформированные на их основе пронумерованные признаки приведены в табл. 1.

Для решения задач аутентификации использовался графический планшет WACOM. Подготовка обу-

чающих данных для построения классификаторов, реализующих процедуру аутентификации, проходила следующим образом. Всего было задействовано восемь пользователей. Первый многократно вводил на графическом планшете свою подпись. Остальные фальсифицировали подпись первого пользователя. Варианты подписей каждого пользователя варьировались в диапазоне от 119 до 280. Каждая полученная подпись, прежде чем из нее были извлечены вышеописанные признаки, подверглась предобработке, которая заключается в исправлении технических ошибок съема, приведении ориентации подписи к стандартному виду, приведении к единому масштабу и обнаружении точного момента начала и завершения простреления подписи.

Табл. 1. Классифицирующие признаки рукописной подписи

Параметр x	1	2	3	4	5	6	7	8
Скорость x	49	50	51	52	53	54	55	56
Ускорение x	97	98	99	100	101	102	103	104
Параметр y	9	10	11	12	13	14	15	16
Скорость y	57	58	59	60	61	62	63	64
Ускорение y	105	106	107	108	109	110	111	112
Параметр z	17	18	19	20	21	22	23	24
Скорость z	65	66	67	68	69	70	71	72
Ускорение z	113	114	115	116	117	118	119	120
Параметр P	25	26	27	28	29	30	31	32
Скорость P	73	74	75	76	77	78	79	80
Ускорение P	121	122	123	124	125	126	127	128
Параметр α	33	34	35	36	37	38	39	40
Скорость α	81	82	83	84	85	86	87	88
Ускорение α	129	130	131	132	133	134	135	136
Параметр θ	41	42	43	44	45	46	47	48
Скорость θ	89	90	91	92	93	94	95	96
Ускорение θ	137	138	139	140	141	142	143	144

3. Постановка задачи

Пусть имеется универсум $U=(A, C)$, где $A = \{x_1, x_2, \dots, x_n\}$ – множество входных признаков, $C = \{1, 2, \dots, m\}$ – множество классов. Пусть $\mathbf{X} = x_1 \times x_2 \times \dots \times x_n \in \mathcal{R}^n$ – n -мерное пространство признаков. Объект в заданном универсуме характеризуется своим вектором значений признаков. Задача классификации заключается в предсказании класса объекта по его вектору значений признаков [18].

Задача отбора признаков заключается в поиске на заданном множестве признаков \mathbf{X} такого их подмножества, которое при уменьшении числа признаков не приводило бы к уменьшению точности классификации; решение представляется в виде бинарного вектора $\mathbf{S} = (s_1, s_2, \dots, s_n)^T$, где $s_i = 0$ означает, что i -й признак не участвует в классификации, $s_i = 1$ означает, что i -й признак используется классификатором. Для каждого подмножества признаков оценивается точность классификации.

Нечеткий классификатор может быть представлен в виде функции, которая присваивает точке \mathbf{x} в пространстве входных признаков метку класса с вычисляемой степенью уверенности:

$$f : \mathcal{R}^n \rightarrow [0, g]^m .$$

Основой нечеткого классификатора является продукционное правило следующего вида:

$$R_i : \text{ЕСЛИ } s_1 \wedge x_1 = A_{1i} \text{ И } s_2 \wedge x_2 = A_{2i} \text{ И } \dots \text{ И } s_n \wedge x_n = A_{ni} \\ \text{ТО class} = L_i,$$

где $L_i \in \{1, 2, \dots, m\}$ – выход i -го правила, A_{ki} – нечеткий терм, характеризующий k -й признак в i -м правиле ($i \in [1, R]$, $k = \overline{1, n}$); запись $s_i \wedge x_i$ указывает на наличие ($s_i = 1$) или отсутствие ($s_i = 0$) признака в классификаторе; R – число правил.

Выход нечеткого классификатора представляет собой вектор $(\beta_1, \beta_2, \dots, \beta_m)^T$, где

$$\beta_j = \sum_{i=1, R}^n \prod_{k=1}^n \mu_{ki}(x_k), \quad j = \overline{1, m}, \quad \mu_{ki}(x_k) -$$

значение функции принадлежности нечеткого термина A_{ki} в точке x_k .

В нашей работе класс определяется следующим образом:

$$\text{class} = \arg \max_{1 \leq j \leq m} \beta_j .$$

На таблице наблюдений $\{(x_p, c_p), \quad p = \overline{1, z}\}$ мера точности классификации может быть задана следующим образом:

$$E(\theta, \mathbf{S}) = \frac{\sum_{p=1}^z \begin{cases} 1, & \text{если } c_p = \arg \max_{1 \leq j \leq m} f_j(\mathbf{x}_p; \theta, \mathbf{S}) \\ 0, & \text{иначе} \end{cases}}{z} ,$$

где $f(\mathbf{x}_p; \theta, \mathbf{S})$ – выход нечеткого классификатора с параметрами θ и признаками \mathbf{S} в точке \mathbf{x}_p . Проблема построения нечеткого классификатора сводится к поиску максимума указанной функции в пространстве \mathbf{S} и $\theta = (\theta^1, \theta^2, \dots, \theta^D)$:

$$\begin{cases} E(\theta, \mathbf{S}) \rightarrow \max \\ \theta_{\min}^i \leq \theta^i \leq \theta_{\max}^i, \quad i = \overline{1, D}, \\ s_j \in \{0, 1\}, \quad j = \overline{1, n} \end{cases}$$

где $\theta_{\min}^i, \theta_{\max}^i$ – нижняя и верхняя границы каждого параметра соответственно. Проблема относится к классу NP-трудных. В нашей работе предлагается решать указанную проблему в три этапа: на первом этапе проводится отбор признаков с помощью бинарных метаэвристических методов, на втором этапе генерируется структура нечеткого классификатора, на третьем этапе ведется обучение нечетких классификаторов на созданной базе правил и отобранных признаках с использованием непрерывных метаэвристических методов.

4. Метаэвристические методы

4.1. Гравитационный алгоритм оптимизации

Гравитационный алгоритм основан на использовании законов тяготения [19]. На вход алгоритму подаются следующие параметры: вектор признаков \mathbf{S} , количество векторов P , максимальное количество

итераций T , начальное значение гравитационной постоянной G_0 , коэффициенты α и малая константа ε . Популяция векторов $\Theta = \{\theta_1, \theta_2, \dots, \theta_P\}$ создается на основе входного вектора θ_1 , полученного алгоритмом генерации структуры классификатора.

Далее итеративно рассчитывается значение массы и ускорение каждой частицы:

$$m_i(t) = \frac{(1 - E[\mathbf{S}, \theta_i(t)] - E[\mathbf{S}, \theta_{worst}(t)])}{(E[\mathbf{S}, \theta_{best}(t)] - E[\mathbf{S}, \theta_{worst}(t)])},$$

$$a_i^d(t) = \sum_{j=1, j \neq i}^P rand(0; 1) \cdot G(t) \cdot \frac{M_j(t) \cdot (\theta_j^d(t) - \theta_i^d(t))}{(\|\theta_j(t) - \theta_i(t)\| + \varepsilon)},$$

где $\theta_{worst}(t)$ и $\theta_{best}(t)$ – векторы, имеющие на текущей итерации наименьшую и наибольшую точность соответственно; $d = 1, |\theta_i|$ – порядковый номер элемента вектора; $rand(0; 1)$ – случайное число, принадлежащее интервалу $[0; 1]$; $M_j(t) = m_j(t) / \sum_{k=1}^P m_k(t)$ – нормированное значение массы j -й частицы; $i = \overline{1, P}$; $G(t) = G_0 \cdot (t/T)^\alpha$ – значение гравитационной постоянной. Элементы вектора обновляются следующим образом:

$$\theta_i^d(t+1) := \theta_i^d(t) + V_i^d(t+1),$$

где $V_i^d(t+1) = rand(0; 1) \cdot V_i^d(t) + a_i^d(t)$. Алгоритм завершается, когда число итераций t станет равным T , при этом в θ_{best} будет искомое решение.

Бинарный вариант данного алгоритма используется для поиска бинарного вектора \mathbf{S}_{best} , показывающего наличие признаков в классификаторе, при котором достигается максимальная точность классификации. Начальная популяция $\mathbf{S} = \{\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_P\}$ задается случайным образом. Элементы вектора обновляются путем преобразования численного значения скорости в бинарный эквивалент с помощью функции трансформации, которая определяет вероятность изменения значения элемента вектора на противоположное:

$$\left\{ \begin{array}{l} \text{ЕСЛИ } (rand(0;1) < F(V_i^d(t+1))), \text{ ТО } p = 1 \\ \text{ИНАЧЕ } p = 0 \\ S_i^d(t+1) = S_i^d(t) \oplus p \end{array} \right. ,$$

где $F(V_i^d(t+1)) = \left| (2/\pi) \arctan((\pi/2)V_i^d(t+1)) \right|$.

4.2. Сорняковый алгоритм оптимизации

Сорняковый алгоритм относится к семейству алгоритмов, основанных на поведении объектов живой природы. Сорняковый алгоритм является популяционным алгоритмом, отражающим ограниченный по времени жизненный цикл сорняков при их распространении и выживании на ограниченной территории [20].

Непрерывный сорняковый алгоритм оптимизации имеет следующие входные параметры: N – максимальное число итераций; V – максимальное число векторов; n_{min} , n_{max} – минимальное и максимальное число дочерних векторов; σ – параметр нормального распределения.

На начальной популяции векторов $\Theta = \{\theta_1, \theta_2, \dots, \theta_P\}$ определяется приспособленность каждого вектора, а также количество дочерних векторов, которое может породить данный вектор:

$$n^s = \frac{n_{max} - n_{min}}{E_{best} - E_{worst}} E(\theta^s) + \frac{E_{best} n_{max} - E_{worst} n_{min}}{E_{best} - E_{worst}},$$

где E_{worst} , E_{best} – худшее и лучшее значение точности классификации.

Для каждого родительского вектора порождаются дочерние векторы:

$$\sigma_N = \sigma(N-t)/N; a = rand(0; 1); b = rand(0; 1),$$

$$u \rightarrow N(0, \sigma_N) = \sigma_N \cos(b) \sqrt{-2 \ln(a)}, \theta_i^s(t+1) = \theta_i^s(t) + u.$$

Порождённые дочерние векторы объединяются с родительскими в одну популяцию. Из полученной популяции всегда остается только V наиболее приспособленных векторов. Итерации алгоритма повторяются до тех пор, пока не выполнено условие останова алгоритма.

Бинарный вариант данного алгоритма используется для поиска бинарного вектора признаков \mathbf{S}_{best} , здесь трансформация непрерывных значений в бинарные происходит на каждой итерации следующим образом:

$$\left\{ \begin{array}{l} \text{ЕСЛИ } u < \sigma_N, \text{ ТО } S_i^d(t+1) = NOT(S_i^d(t)) \\ \text{ИНАЧЕ } S_i^d(t+1) = S_i^d(t) \end{array} \right. .$$

4.3. Алгоритм «кукушкин поиск»

Метаэвристика «кукушкин поиск» имитирует поведение кукушки в период размножения. Кукушка находит недавно построенные гнезда и подкладывает в них свои яйца (заменяет своими), которые в итоге могут быть выкинуты хозяином гнезда [21, 22].

Ниже кратко представлен алгоритм «кукушкин поиск» применительно к оптимизации параметров θ нечеткого классификатора.

На начальном этапе происходит генерация исходной популяции решений $\Theta = \{\theta_1, \theta_2, \dots, \theta_P\}$, где θ_1 принимает значения параметров, определяемых алгоритмом генерации структуры, а $\theta_2, \dots, \theta_P$ определяются случайным образом.

Далее осуществляется итерационный процесс поиска оптимума, который состоит из генерации новых и удаления худших решений популяции. Генерация новых решений осуществляется путем изменения всех текущих элементов векторов решений на случайную величину полета Леви:

$$\theta^{new} = \theta + Levi ,$$

где $Levi = \gamma u / |v|^{1/\beta}$, γ – коэффициент прыжка полета Леви (рекомендуемое значение 0,01), $\beta = 1,5$; $[1, 2]u, v$ – нормально распределенные случайные величины $v \sim N(0; \sigma_v^2)$, $u \sim N(0; \sigma_u^2)$, $\sigma_v^2 = 1$,

$$\sigma_u^2 = \left\{ \frac{\Gamma(1+\beta) \cdot \sin(\pi \cdot \beta/2)}{\Gamma((1+\beta)/2) \cdot 2^{(\beta-1)/2}} \right\}^{1/\beta} ,$$

здесь $v \sim N(0; \sigma - v - 2)\Gamma(x)\Gamma(\cdot)$ – гамма-функция.

Если новый вектор решения показывает лучшую точность классификации, чем соответствующий вектор популяции, то происходит замена вектора популяции на новое решение. Удаление худшего решения в популяции – это удаление решения k , такого, что

$$k = \arg \min_{1 \leq i \leq P} (E(\theta_i, S)).$$

Решение удаляется с вероятностью p , которая задается в начале алгоритма, и вместо удаленного решения генерируется новое случайным образом.

Процесс продолжается заданное число итераций T . Алгоритм возвращает вектор решения, показавший лучшее значение точности классификации среди векторов популяции Θ .

4.4. Алгоритм обезьян

Алгоритм обезьян основан на наблюдениях за передвижением обезьян в горной местности [23]. В процессе выполнения алгоритма выполняются три оператора: движение вверх, локальный прыжок, глобальный прыжок.

Шаги алгоритма описаны ниже:

Шаг 1. Задание параметров алгоритма: размер популяции M , шаг a , интервал локального прыжка b , интервалы глобального прыжка c и d .

Шаг 2. Инициализация популяции $\theta_i = (\theta_{i1}, \theta_{i2}, \dots, \theta_{iD})$, $i = 1, \dots, M$.

Шаг 3. Движение агента вверх: $\theta_{i\text{new}} = \theta_{i\text{old}} + \Delta\theta_i$, где

$$\Delta\theta_{ij} = a \cdot \text{sign} \left(\frac{E(\theta_i + \Delta\mu_i) - (E(\theta_i - \Delta\mu_i))}{2\Delta\mu_{ij}} \right),$$

$$\Delta\mu_i = \begin{cases} a, & \text{если } t = \text{rand}(0,1) > 1/2 \\ -a, & \text{иначе} \end{cases}$$

Шаг 4. Локальный прыжок: $\theta_{i\text{new}} = \theta_{i\text{old}} + \mathbf{L}_i$, где \mathbf{L}_i – вектор длиной D равномерно распределённых чисел, каждое из которых находится в диапазоне $[-b, b]$.

Шаг 5. Глобальный прыжок: $\theta_{i\text{new}} = \theta_{i\text{old}} + \mathbf{G}_i$, где $G_{ij} = \gamma(\rho_j - \theta_{ij})$, $\rho_j = (\sum_{i=1}^M \theta_{ij})/M$, γ – действительное число из интервала $[c, d]$.

4.5. Алгоритм «стадо криля»

Алгоритм «стадо криля» имитирует жизненные повадки антарктических крилей, их перемещение в процессе поиска пищи с учётом взаимосвязанных факторов [24]. Одним из таких факторов является наличие хищника, которое заставляет крилей держаться вместе и увеличивает плотность стада (N); другой фактор – поиск пищи (F) и третий фактор – случайное блуждание криля (RD). На вход алгоритму подаются следующие параметры: вектор признаков S , размер популяции P , максимальное количество итераций t_{max} , коэффициент инерции при движении членов стада $w_n \in [0, 1]$, малое положительное число ε_1 , коэффициент инерции при поиске пищи $w_f \in [0, 1]$, максимальная скорость блуждания RD_{max} , константа $C_t \in [0, 1]$.

Популяция векторов $\Theta = \{\theta_1, \theta_2, \dots, \theta_P\}$ создается на основе входного вектора θ_1 , полученного алгорит-

мом генерации структуры классификатора. Затем на каждой итерации t рассчитываются факторы и изменяются элементы вектора θ :

$$\mathbf{N}_i(t+1) = N_{\text{max}} \alpha_i + w_n \mathbf{N}_i(t),$$

$$\alpha_i = \alpha_i^{\text{local}} + \alpha_i^{\text{target}}, \alpha_i^{\text{local}} = \sum_{j=1}^{NN} \hat{K}_{i,j} \hat{\theta}_{i,j},$$

$$\hat{\theta}_{i,j} = \frac{\theta_j - \theta_i}{\|\theta_j - \theta_i\| + \varepsilon_1}, r_i = \frac{1}{5P} \sum_{j=1}^P \|\theta_i - \theta_j\|,$$

$$\hat{K}_{i,j} = \frac{E_j - E_i}{E_{\text{worst}} - E_{\text{best}}}, \alpha_i^{\text{target}} = C_{\text{best}} \hat{K}_{i,\text{best}} \hat{\theta}_{i,\text{best}},$$

где N_{max} – максимальная скорость крилей в стаде; α_i^{local} – вектор движения, учитывающий влияние соседей i -го криля; α_i^{target} – вектор движения, учитывающий влияние лучшего криля в стаде, E_i – точность классификации i -го решения; E_{worst} , E_{best} – худшее и лучшее значение точности классификации соответственно; NN – число соседей, находящихся от i -го криля на расстоянии меньше, чем r_i ; $\text{rand} \in [0, 1]$ – случайное число.

$$\mathbf{F}_i(t+1) = V_f \beta_i + w_f \mathbf{F}_i(t), \beta_i = \beta_i^{\text{food}} + \beta_i^{\text{best}},$$

$$\beta_i^{\text{food}} = C_{\text{food}} \hat{K}_{i,\text{food}} \hat{\theta}_{i,\text{food}}, C_{\text{food}} = 2(1 - t/t_{\text{max}}),$$

$$\theta_{i,\text{food}} = \sum_{i=1}^N \frac{1}{K_i} \theta_i / \sum_{i=1}^N \frac{1}{K_i},$$

где V_f – скорость перемещения крилей при поиске пищи; N – число крилей в популяции; $\theta_{i,\text{best}}$ – лучшая из всех предыдущих позиция i -го криля.

$$\Delta\theta_i = \mathbf{N}_i + \mathbf{F}_i + \mathbf{RD}_i, \mathbf{RD}_i = RD_{\text{max}} (1 - t/t_{\text{max}}) \delta,$$

$$\theta_i(t+1) = \theta_i(t) + \Delta\theta_i \cdot \Delta t, \Delta t = C_t \sum_{j=1}^{NV} (UB_j - LB_j),$$

где RD_{max} – максимальная скорость блуждания; δ – случайный вектор направления движения, элементами которого являются равномерно распределённые величины из диапазона $[-1, 1]$; NV – общее число параметров; LB_j , UB_j – нижняя и верхняя граница j -го параметра соответственно.

4.6. Алгоритм отбора признаков FCBF

Алгоритм FCBF (*Fast Correlation-Based Filter*) [25] относится к алгоритмам-фильтрам, которые в своей работе не используют построение классификатора. Алгоритм основан на вычислении приращения информации признаков экспериментальных данных.

5. Эксперимент

База данных содержит 280 оригинальных вариантов подписи одного автора и 1281 вариант фальсификаций (поддельных подписей) семи авторов. Необходимость в многократном представлении одной и той же подписи обусловлена изменчивостью в написании, вызванной физическими или психологическими причинами.

Варианты подлинной и фальсифицированной подписи приведены на рис. 1 ниже.



Рис. 1. Образцы подписи: подлинной (а), фальсифицированной (б)

Методика эксперимента по отбору признаков метаэвристическими алгоритмами включает следующие этапы:

- 1) разбиение набора данных на десять выборок в соответствии со схемой кросс-валидации;
- 2) на всех выборках запуск алгоритма отбора признаков: бинарного метаэвристического по пять раз, FCBF один раз;
- 3) для каждой из десяти выборок определение набора признаков, на котором получена наиболее высокая точностью на обучающих данных.

В табл. 2 представлены результаты работы гравитационного алгоритма. Здесь E_{tra} , E_{tst} – точность в процентах на обучающей и тестовой выборке соответственно; E_1 , E_2 – ошибки первого и второго рода в процентах; F – число признаков. Число правил в этом эксперименте равно двум.

Табл. 2. Гравитационный алгоритм

N	E_{tra}	E_{tst}	E_1	E_2	F	Признаки
1	99,95	99,55	1,43	0,24	4	5, 9, 17, 33
2	99,77	99,55	1,07	0,32	6	29,53, 101, 33,121,123
3	99,96	99,81	1,08	0	6	1, 4, 33, 53, 86, 101
4	99,91	99,87	0	0,16	5	18, 33, 48, 78, 125
5	99,95	99,68	1,44	0,08	7	33,48,140 4,74,77,85,
6	99,79	99,74	1,07	0,08	4	3,33,36, 58
7	99,89	99,42	2,14	0,24	3	10, 17, 33
8	99,81	99,81	0,72	0,08	3	18, 33, 95
9	99,88	99,74	0,36	0,24	3	5, 33, 137
10	99,65	99,42	1,80	0,32	4	2,33,62,67
ср.	99,86	99,66	1,11	0,18	4,5	

В табл. 3 представлены результаты работы сорнякового алгоритма.

Табл. 3. Сорняковый алгоритм

N	E_{tra}	E_{tst}	E_1	E_2	F	Признаки
1	98,18	98,02	2,5	1,95	1	33
2	99,41	99,42	2,14	0,31	2	33,4
3	98,90	98,59	3,93	0,86	2	33,9
4	90,42	90,32	3,21	0,16	3	33,4,9
5	98,95	98,91	5,71	0,08	4	33,9,1,5
6	99,81	99,68	0,71	0,23	2	33,18
7	99,81	99,81	1,07	0	3	33,4,144
8	99,76	99,68	1,07	0,16	3	33,4,18
9	99,94	99,87	0	0,16	4	33,5,18,48
10	97,28	97,44	14,29	0	8	33,9,1,5,87 121,25,94
ср.	98,25	98,17	3,46	0,39	3,2	

В табл. 4 представлены результаты классификации, выполненные последовательным запуском следующих алгоритмов: FCBF для отбора признаков, субтрактивной кластеризации для генерации структуры нечеткого классификатора [26], «кукушкин поиск»

для оптимизации параметров классификатора. Число правил в этом эксперименте равно трём.

Табл. 4. FCBF+«кукушкин поиск»

N	E_{tra}	E_{tst}	E_1	E_2	F	Признаки
1	99,86	99,36	0	0,78	3	33, 144, 29
2	99,86	100,0	0	0	3	33, 144, 29
3	98,79	99,36	0	0	2	33, 144
4	99,86	100,0	0	0	3	33, 144, 29
5	99,79	100,0	0	0	3	33, 144, 29
6	99,93	98,72	7,14	0	3	33, 144, 29
7	99,86	99,36	3,57	0	3	33, 144, 29
8	99,79	100,0	0	0	3	33, 144, 29
9	99,86	100,0	0	0	3	33, 144, 29
10	99,86	100,0	0	0	3	33, 144, 29
ср.	99,75	99,68	1,07	0,08	2,9	

В табл. 5 представлены результаты работы классификации, выполненные последовательным запуском следующих алгоритмов: алгоритм генерации базы правил по экстремумам классов [18], жадный алгоритм для отбора признаков, алгоритм обезьян для оптимизации параметров классификатора.

Табл. 5. Жадный+Алгоритм обезьян

N	E_{tra}	E_{tst}	E_1	E_2	F	Признаки
1	99,81	99,74	0,71	0,16	2	18, 33
2	99,99	100,00	0	0	4	18, 33, 48, 60
3	99,88	99,68	0,36	0,32	4	18, 33, 47, 60
4	99,99	99,81	0,36	0,16	4	5, 18, 33, 48
5	99,99	99,87	0	0,16	5	18,33,47, 48, 60
6	99,99	99,87	0	0,16	5	3, 18, 33, 48, 75
7	99,79	99,81	0	0,23	6	5,18,33,47,48,60
8	99,91	99,49	1,79	0,23	10	5,14,18,33,47, 50,58,64,77,141
9	99,92	99,81	0,71	0,08	12	11, 14, 18, 25, 33, 47, 48, 50, 53, 64, 75, 77
10	99,96	99,62	1,79	0,08	13	5,14, 18, 25, 33, 47, 48, 50, 58, 62, 64, 77, 141
ср.	99,92	99,77	0,57	0,16	6,5	

В табл. 6 представлены результаты работы алгоритма «стадо криля».

Табл. 6. Алгоритм «стадо криля»

N	E_{tra}	E_{tst}	E_1	E_2	F	Признаки
1	99,67	99,04	3,2	0,5	5	13,26,31, 32,33
2	99,52	99,04	3,6	0,4	5	29,33,40, 41,44
3	98,53	98,27	6,1	0,8	1	33
4	99,69	99,23	2,9	0,3	2	29,33
5	99,50	99,10	3,9	0,2	4	33,57,71, 111
6	99,93	99,36	2,5	0,2	4	4,33,74, 136
7	99,63	99,23	2,9	0,3	4	10,24,33, 46
8	98,93	98,21	5,7	1,0	4	33,60,69, 95
9	99,39	99,10	2,9	0,6	3	13,32,33
10	99,61	98,91	3,9	0,5	4	26,31,33, 40
ср.	99,44	98,95	3,8	0,5	3,6	

Для оценки статистической значимости различий в точности и ошибках нечётких классификаторов, сформированных гравитационным алгоритмом (ГА), «кукушкин поиск» (КП), сорняковым алгоритмом

(СА), алгоритмом обезьян (АО) и алгоритмом крилей (АК), использованы критерий Манна–Уитни (Уилкоксона) и тест Крускала–Уоллиса.

В табл. 7 приведены результаты проверки нулевой гипотезы (медианы выборок статистически неразличимы) критерием Манна–Уитни при попарном сравнении алгоритмов, здесь в ячейке таблицы указаны значимости для точности на тестовой выборке, для ошибки первого и для ошибки второго рода.

Табл. 7. Итоги проверки нулевой гипотезы критерием Манна–Уитни

Алгоритм	ГА	СА	АО	АК	КП
ГА	1,000	0,094 0,139 0, 969	0,146 0,061 0,590	1,8E-4 1,8E-4 0,007	0,465 0,043 0,005
СА	0,094 0,139 0, 969	1,000	0,018 0,008 0,727	0,761 0,198 0,063	0,029 0,019 0,007
АО	0,146 0,061 0,590	0,018 0,008 0,727	1,000	1,8E-4 1,6E-4 0,002	0,614 0,249 0,005
АК	1,8E-4 1,8E-4 0,007	0,761 0,198 0,063	1,8E-4 1,6E-4 0,002	1,000	0,002 0,007 0,001
КП	0,465 0,043 0,005	0,029 0,019 0,007	0,614 0,249 0,005	0,002 0,007 0,001	1,000

Ранжирование алгоритмов с позиций точности классификации представленных рукописных подписей выглядит следующим образом: $ГА \approx АО \approx КП \} СА \approx АК$, здесь « $\}$ » задает отношение «предпочтительнее», « \approx » – «статистически равнозначны». Аналогично может быть представлено ранжирование алгоритмов с позиций ошибки первого рода. Ранжирование алгоритмов с позиций ошибки второго рода выглядит следующим образом: $КП \} ГА \approx СА \approx АО \} АК$, однако необходимо учесть, что классификатор на основе алгоритма «кукушкин поиск» имеет три нечетких правила, а классификаторы, обученные другими алгоритмами, имеют два правила.

Критерий Крускала–Уоллиса проверяет нулевую гипотезу о равенстве медиан в пяти выборках. При проверке различий в точности классификации получено значение значимости $2,5E-4$, что существенно меньше $0,05$, из чего можно сделать вывод: существует статистически значимое различие между медианами точности на уровне значимости $0,05$. При проверке различий в ошибках первого рода получено значение значимости $1,1E-4$, что существенно меньше $0,05$, из чего можно сделать вывод: существует статистически значимое различие между медианами ошибок первого рода на уровне значимости $0,05$. При проверке различий в ошибках второго рода получено значение значимости $3,6E-4$, что существенно меньше $0,05$, из чего можно сделать вывод: существует статистически значимое различие между медианами ошибок второго рода на уровне значимости $0,05$.

Заключение

Для решения задачи аутентификации пользователя по динамике рукописной подписи предложено использовать нечеткие классификаторы. Рассмотрены

метаэвристические алгоритмы для отбора признаков и оптимизации параметров классификаторов. Сравнительный статистический анализ не выявил алгоритма, превосходящего остальные по эффективности построения классификаторов, из чего можно сделать вывод: для решения задач аутентификации должны использоваться различные алгоритмы создания нечетких классификаторов.

Благодарности

Работа выполнена при финансовой поддержке Министерства образования и науки РФ в рамках базовой части государственного задания в сфере научной деятельности (проект № 8.9628.2017/8.9).

Литература

1. **Ortega-Garcia, J.** Authentication gets personal with biometrics / J. Ortega-Garcia, J. Bigun, D. Reynolds, J. Gonzalez-Rodriguez // IEEE Signal Processing Magazine. – 2004. – Vol. 21, Issue 2. – P. 50-62. – DOI: 10.1109/MSP.2004.1276113.
2. **Gupta, G.** A review of dynamic handwritten signature verification / G. Gupta, A. McCabe // Technical report at James Cook University, Australia. – 1997.
3. **Zalasiński, M.** Fuzzy-genetic approach to identity verification using a handwritten signature / M. Zalasiński, K. Spałka, L. Rutkowski. – In Book: Advances in data analysis with computational intelligence / ed. by A. Gawęda, J. Kacprzyk, L. Rutkowski, G. Yen. – Cham: Springer, 2018. – P. 375-394. – DOI: 10.1007/978-3-319-67946-4_17.
4. **Аникин, И.В.** Распознавание динамической рукописной подписи на основе нечеткой логики / И.В. Аникин, Э.С. Анисимова // Вестник Казанского государственного энергетического университета. – 2016. – Т. 31, № 3. – С. 48-64.
5. **Kholmatov, A.** Identity authentication using improved online signature verification method / A. Kholmatov, B. Yanikoglu // Pattern Recognition Letters. – 2005. – Vol. 26, Issue 15. – P. 2400-2408. – DOI: 10.1016/j.patrec.2005.04.017.
6. **Sharma, A.** An enhanced contextual DTW based system for online signature verification using vector quantization / A. Sharma, S. Sundaram // Pattern Recognition Letters. – 2016. – Vol. 84, Issue C. – P. 22-28. – DOI: 10.1016/j.patrec.2016.07.015.
7. **Xia, X.** Signature alignment based on GMM for on-line signature verification / X. Xia, Z. Chen, F. Luan, X. Song // Pattern Recognition. – 2017. – Vol. 65, Issue C. – P. 188-196. – DOI: 10.1016/j.patcog.2016.12.019.
8. **Khan, M.A.U.** Velocity-image model for online signature verification / M.A.U. Khan, M.K. Khan, M.A. Khan // IEEE Transactions on Image Processing. – 2006. – Vol. 15, Issue 11. – P. 3540-3549. – DOI: 10.1109/TIP.2006.877517.
9. **Razzak, M.I.** Multilevel fusion for fast online signature recognition using multi-section VQ and time modelling / M.I. Razzak, B. Alhaqabani // Neural Computing and Applications. – 2015. – Vol. 26, Issue 5. – P. 1117-1127. – DOI: 10.1007/s00521-014-1779-6.
10. **Spałka, K.** A new algorithm for identity verification based on the analysis of a handwritten dynamic signature / K. Spałka, M. Zalasiński, L. Rutkowski // Applied Soft Computing. – 2016. – Vol. 43. – P. 47-56.
11. **Doroz, R.** Dynamic signature verification method based on association of features with similarity measures / R. Doroz,

- P. Porwik, T. Orczyk // *Neurocomputing*. – 2016. – Vol. 171. – P. 921-931. – DOI: 10.1016/j.neucom.2015.07.026.
12. **Nanni, L.** Combining local, regional and global matchers for a template protected on-line signature verification system / L. Nanni, E. Maiorana, A. Lumini, P. Campisi // *Expert Systems with Applications*. – 2010. – Vol. 37, Issue 5. – P. 3676-3684. – DOI: 10.1016/j.eswa.2009.10.023.
 13. **Jalalirad, A.** Using feature-based models with complexity penalization for selecting features / A. Jalalirad, T. Tjalkens // *Journal of Signal Processing Systems*. – 2018. – Vol. 90, Issue 2. – P. 201-210. – DOI: 10.1007/s11265-016-1152-3.
 14. **Alkuhlani, A.** Multistage feature selection approach for high-dimensional cancer data / A. Alkuhlani, M. Nassef, I. Farag // *Soft Computing*. – 2017. – Vol. 21, Issue 22. – P. 6895-6906. – DOI: 10.1007/s00500-016-2439-9.
 15. **Bolón-Canedo, V.** Feature selection for high-dimensional data / V. Bolón-Canedo, N. Sánchez-Marño, A. Alonso-Betanzos. – London: Springer, 2015. – 148 p. – ISBN: 978-3-319-21857-1.
 16. **Иванов, А.И.** Оценка надежности верификации автографа на основе искусственных нейронных сетей, сетей многомерных функционалов Байеса и сетей квадратичных форм / А.И. Иванов, П.С. Ложников, А.Е. Сулавко // *Компьютерная оптика*. – 2017. – Т. 41, № 5. – С. 765-774. – DOI: 10.18287/2412-6179-2017-41-5-765-774.
 17. **Kostyuchenko, E.** Integration of Bayesian classifier and perceptron for problem identification on dynamics signature using a genetic algorithm for the identification threshold selection / E. Kostyuchenko, M. Gurakov, E. Krivonosov, M. Tomyshev, R. Mescheryakov, I. Hodashinsky. – In book: *Advances in Neural Networks – ISNN 2016* / ed. by L. Cheng, Q. Liu, A. Ronzhin. – Cham: Springer, 2016. – P. 620-627. – DOI: 10.1007/978-3-319-40663-3_71.
 18. **Мех, М.А.** Сравнительный анализ применения методов дифференциальной эволюции для оптимизации параметров нечетких классификаторов / М.А. Мех, И.А. Ходашинский // *Известия Российской академии наук. Теория и системы управления*. – 2017. – № 4. – С. 65-75. – DOI: 10.7868/S0002338817040060.
 19. **Rashedi, E.** GSA: A gravitational search algorithm / E. Rashedi, H. Nezamabadi-pour, S. Saryazdi // *Information Sciences*. – 2009. – Vol. 179, Issue 13. – P. 2232-2248. – DOI: 10.1016/j.ins.2009.03.004.
 20. **Mehrabian, A.R.** A novel numerical optimization algorithm inspired from weed colonization / A.R. Mehrabian, C. Lucas // *Ecological Informatics*. – 2006. – Vol. 1, Issue 4. – P. 355-366. – DOI: 10.1016/j.ecoinf.2006.07.003.
 21. **Yang, X.-S.** Engineering optimisation by cuckoo search / X.-S. Yang, S. Deb // *International Journal of Mathematical Modelling and Numerical Optimisation*. – 2010. – Vol. 1, Issue 4. – P. 330-343. – DOI: 10.1504/IJMMNO.2010.035430.
 22. **Yang, X.-S.** Cuckoo search: recent advances and applications / X.-S. Yang, S. Deb // *Neural Computing and Applications*. – 2014. – Vol. 24, Issue 1. – P. 169-174. – DOI: 10.1007/s00521-013-1367-1.
 23. **Zhao, R.** Monkey algorithm for global numerical optimization / R. Zhao, W. Tang // *Journal of Uncertain Systems*. – 2008. – Vol. 2, Issue 3. – P. 165-176.
 24. **Gandomi, A.H.** Krill herd: A new bio-inspired optimization algorithm / A.H. Gandomi, A.H. Alavi // *Communications in Nonlinear Science and Numerical Simulation*. – 2012. – Vol. 17, Issue 12. – P. 4831-4845. – DOI: 10.1016/j.cnsns.2012.05.010.
 25. **Yu, L.** Feature selection for high-dimensional data: A fast correlation-based filter solution / L. Yu, H. Liu // *Proceedings of the 12th International Conference on Machine Learning (ICML-03)*. – 2003. – P. 856-863.
 26. **Sarin, K.S.** Identification of fuzzy classifiers based on the mountain clustering and cuckoo search algorithms / K.S. Sarin, I.A. Hodashinsky // *2017 International Siberian Conference on Control and Communications (SIBCON)*. – 2017. – 6 p. – DOI: 10.1109/SIBCON.2017.7998553.

Сведения об авторах

Ходашинский Илья Александрович, доктор технических наук, профессор, профессор кафедры комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС) Томского государственного университета систем управления и радиоэлектроники (ТУСУР). Область научных интересов: вычислительный интеллект. E-mail: hodashn@rambler.ru.

Костюченко Евгений Юрьевич, кандидат технических наук, доцент, доцент кафедры КИБЭВС ТУСУР. Область научных интересов: вычислительный интеллект, биометрические характеристики, анализ речи. E-mail: key@keva.tusur.ru.

Сарин Константин Сергеевич, кандидат технических наук, доцент кафедры КИБЭВС ТУСУР. Область научных интересов: анализ данных, машинное обучение. E-mail: sks@security.tomsk.ru.

Анфилофьев Александр Евгеньевич, аспирант кафедры КИБЭВС ТУСУР. Область научных интересов: базы данных, вычислительные системы, вычислительный интеллект. E-mail: yowwi00@gmail.com.

Бардамова Марина Борисовна, аспирант кафедры КИБЭВС ТУСУР. Область научных интересов: вычислительный интеллект, биомедицинские технологии, машинное обучение. E-mail: 722bmb@gmail.com.

Самсонов Сергей Сергеевич, студент кафедры КИБЭВС ТУСУР. Область научных интересов: вычислительный интеллект. E-mail: 723_sss@fb.tusur.ru.

Филимоненко Игорь Витальевич, студент кафедры безопасности информационных систем ТУСУР. Область научных интересов: вычислительный интеллект. E-mail: ifilimon96@mail.ru.

ГРНТИ: 28.23.15, 28.19.31, 81.93.29

Поступила в редакцию 21 марта 2018 г. Окончательный вариант – 2 апреля 2018 г.

DYNAMIC-SIGNATURE-BASED USER AUTHENTICATION USING A FUZZY CLASSIFIER

I.A. Hodashinsky¹, E.Yu. Kostyuchenko¹, K.S. Sarin¹, A.E. Anfilofiev¹, M.B. Bardamova¹, S.S. Samsonov¹, I.V. Filimonenko¹

¹*Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia*

Abstract

Dynamic signature verification is one of the most fast, intuitive, and cost effective tools for user authentication. Dynamic signature recognition uses multiple characteristics in the analysis of an individual's handwriting. Dynamic characteristics include the velocity, acceleration, timing, pressure, and direction of the signature strokes, all analyzed in the x , y , and z directions. In this paper, the constant term and the first seven harmonics of the Fourier series expansion of the signature were used as features. The authentication systems development includes the following stages: preprocessing, feature selection, classification. Binary metaheuristic algorithms and deterministic algorithms are used to select attributes. The classification was carried out using a fuzzy classifier. The fuzzy classifiers parameters were tuned using continuous metaheuristic algorithms. The efficiency of the authentication system was verified on the author's database. The database contains 280 original variants of the signature of one author and 1281 variants of counterfeit signatures of seven authors. To assess the statistical significance of differences in the accuracy and error rates of the fuzzy classifiers formed by metaheuristic algorithms, the Mann-Whitney (-Wilcoxon) U-test to compare medians and the Kruskal-Wallis test were used.

Keywords: pattern recognition, information processing, algorithms, feature selection, fuzzy classifier, signature recognition.

Citation: Hodashinsky IA, Kostyuchenko EYu, Sarin KS, Anfilofiev AE, Bardamova MB, Samsonov SS, Filimonenko IV. Dynamic-signature-based user authentication using a fuzzy classifier. *Computer Optics* 2018; 42(4): 657-666. DOI: 10.18287/2412-6179-2018-42-4-657-666.

Acknowledgements: The work was supported by the Ministry of Education and Science of the Russian Federation under project No. 8.9628.2017/8.9 (the basic part of the state research contract).

References

- [1] Ortega-Garcia J, Bigun J, Reynolds D, Gonzalez-Rodriguez J. Authentication gets personal with biometrics. *IEEE Signal Process Mag* 2004; 21(2): 50-62. DOI: 10.1109/MSP.2004.1276113.
- [2] Gupta G, McCabe A. A review of dynamic handwritten signature verification. Technical report at James Cook University, Australia 1997.
- [3] Zalasinski M, Cpałka K, Rutkowski L. Fuzzy-Genetic Approach to Identity Verification Using a Handwritten Signature. In Book: Gawęda A, Kacprzyk J, Rutkowski L, Yen G, eds. *Advances in data analysis with computational intelligence*. Cham: Springer; 2018: 375-394. DOI: 10.1007/978-3-319-67946-4_17.
- [4] Anikin IV, Anisimova ES. Dynamic handwritten signature recognition based on fuzzy logic [In Russian]. *Bulletin of Kazan State Power Engineering University* 2016; 31(3): 48-64.
- [5] Kholmatov A, Yanikoglu B. Identity authentication using improved online signature verification method. *Pattern Recogn Lett* 2005; 26(15): 2400-2408. DOI: 10.1016/j.patrec.2005.04.017.
- [6] Sharma A, Sundaram S. An enhanced contextual DTW based system for online signature verification using vector quantization. *Pattern Recogn Lett* 2016; 84(C): 22-28. DOI: 10.1016/j.patrec.2016.07.015.
- [7] Xia X, Chen Z, Luan F, Song X. Signature alignment based on GMM for on-line signature verification. *Pattern Recogn* 2017; 65(C): 188-196. DOI: 10.1016/j.patcog.2016.12.019.
- [8] Khan MAU, Khan MK, Khan MA. Velocity-image model for online signature verification. *IEEE Trans Image Process* 2006; 15(11): 3540-3549. DOI: 10.1109/TIP.2006.877517.
- [9] Razzak MI, Alhaqbani B. Multilevel fusion for fast online signature recognition using multi-section VQ and time modelling. *Neural Computing and Applications* 2015; 26(5): 1117-1127. DOI: 10.1007/s00521-014-1779-6.
- [10] Cpałka K, Zalasinski M, Rutkowski L. A new algorithm for identity verification based on the analysis of a handwritten dynamic signature. *Appl Soft Comput* 2016; 43: 47-56. DOI: 10.1016/j.asoc.2016.02.017.
- [11] Doroz R, Porwik P, Orczyk T. Dynamic signature verification method based on association of features with similarity measures. *Neurocomputing* 2016; 171: 921-931. DOI: 10.1016/j.neucom.2015.07.026.
- [12] Nanni L, Maiorana E, Lumini A, Campisi P. Combining local, regional and global matchers for a template protected on-line signature verification system. *Expert Systems with Applications* 2010; 37(5): 3676-3684. DOI: 10.1016/j.eswa.2009.10.023.
- [13] Jalalirad A, Tjalkens T. Using Feature-Based Models with Complexity Penalization for Selecting Features. *J Signal Process Syst* 2018; 90(2): 201-210. DOI: 10.1007/s11265-016-1152-3.
- [14] Alkuhlani A, Nassef M, Farag I. Multistage feature selection approach for high-dimensional cancer data. *Soft Computing* 2017; 21(22): 6895-6906. DOI: 10.1007/s00500-016-2439-9.
- [15] Bolón-Canedo V, Sánchez-Marroño N, Alonso-Betanzos A. Feature selection for high-dimensional. London: Springer; 2015. ISBN: 978-3-319-21857-1.
- [16] Ivanov AI, Lozhnikov PS, Sulavko AE. Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms [In Russian]. *Computer Optics* 2017; 41(5): 765-774. DOI: 10.18287/2412-6179-2017-41-5-765-774.
- [17] Kostyuchenko E, Gurakov M, Krivonosov E, Tomyshev M, Mescheryakov R, Hodashinsky I. Integration of Bayes-

- ian classifier and perceptron for problem identification on dynamics signature using a genetic algorithm for the identification threshold selection. In Book: Cheng L, Liu Q, Ronzhin A, eds. *Advances in Neural Networks 2016*: 620-627. DOI: 10.1007/978-3-319-40663-3_71.
- [18] Mekh MA, Hodashinsky IA. Comparative analysis of differential evolution methods to optimize parameters of fuzzy classifiers. *J Comput Syst Sci Int* 2017; 56(4): 616-626. DOI: 10.1134/S1064230717040116.
- [19] Rashedi E, Nezamabadi-pour H, Saryazdi S. GSA: A gravitational search algorithm. *Information Sciences* 2009; 179(13): 2232-2248. DOI: 10.1016/j.ins.2009.03.004.
- [20] Mehrabian AR, Lucas C. A novel numerical optimization algorithm inspired from weed colonization. *Ecological Informatics* 2006; 1(4): 355-366. DOI: 10.1016/j.ecoinf.2006.07.003.
- [21] Yang X-S, Deb S. Engineering optimisation by cuckoo search. *International Journal of Mathematical Modelling and Numerical Optimisation* 2010; 1(4): 330-343. DOI: 10.1504/IJMMNO.2010.035430.
- [22] Yang X-S, Deb S. Cuckoo search: recent advances and applications. *Neural Comput & Applic* 2014; 24(1): 169-174. DOI: 10.1007/s00521-013-1367-1.
- [23] Zhao R, Tang W. Monkey Algorithm for Global Numerical Optimization. *Journal of Uncertain Systems* 2008; 2(3): 165-176.
- [24] Gandomi AH, Alavi AH. Krill herd: A new bio-inspired optimization algorithm. *Communications in Nonlinear Science and Numerical Simulation* 2012; 17(12): 4831-4845. DOI: 10.1016/j.cnsns.2012.05.010.
- [25] Yu L, Liu H. Feature selection for high-dimensional data: A fast correlation-based filter solution. *Proc 12th Int Conf Machine Learn (ICML-03)* 2003: 856-863.
- [26] Sarin KS, Hodashinsky IA. Identification of fuzzy classifiers based on the mountain clustering and cuckoo search algorithms. *2017 International Siberian Conference on Control and Communications (SIBCON) 2017*. DOI: 10.1109/SIBCON.2017.7998553.

Author's information

Ilya Alexandrovich Hodashinsky. Received the Dr. Sc. degree in 2004 from the Tomsk State University of Control Systems and Radioelectronics, Russia. Received the Professor title at the 2011. He is a professor of Tomsk State University of Control Systems and Radioelectronics. His main research interests include the computational intelligence, pattern recognition, and data mining. E-mail: hodashn@rambler.ru.

Evgeny Yurievich Kostyuchenko. Candidate of Engineering Sciences, associate professor, associate professor of Tomsk State University of Control Systems and Radioelectronics. Area of scientific interests: computing intelligence, biometric characteristics, speech analysis. E-mail: key@keva.tusur.ru.

Konstantin Sergeevich Sarin. Candidate of Engineering Sciences department, associate professor of Complex Information Security of Computer Systems in Tomsk State University of Control Systems and Radioelectronics. Main research interests include data mining and machine learning. E-mail: sks@security.tomsk.ru.

Aleksandr Evgenevich Anfilofev, postgraduate student, department of Tomsk State University of Control Systems and Radioelectronics. Scientific interests: databases, computer system and computational intelligence. E-mail: yowwi00@gmail.com.

Marina Borisovna Bardamova. She is a postgraduate student of Tomsk State University of Control Systems and Radioelectronics. Research interests are computational intelligence, biomedical technology, and machine learning. E-mail: 722bmb@gmail.com.

Sergey Sergeevich Samsonov, student of the Tomsk State University of Control Systems and Radioelectronics. Area of scientific interests: computing intelligence. E-mail: 723_sss@fb.tusur.ru.

Igor Vitalyevich Filimonenko. He is a student of Tomsk State University of Control Systems and Radioelectronics. His main research interests include the computational intelligence, pattern recognition. E-mail: ifilimon96@mail.ru.

Received March 21, 2018. The final version – April 2, 2018.
