DOI 10.37882/2223–2966.2021.04.35

МЕТОДИКА КОСВЕННОЙ ДИНАМИЧЕСКОЙ БИОМЕТРИЧЕСКОЙ ВЕРИФИКАЦИИ

METHOD OF INDIRECT DYNAMIC BIOMETRIC VERIFICATION

D. Trokoz

Summary. The paper proposes a method of continuous biometric verification based on a heuristic algorithm using the mathematical apparatus of wide neural networks, genetic algorithms, and the algebra of hyperdimensional binary vectors. This technique is based on replacing the target verification problem with the task of predicting the time series of actions of the verified operator, providing continuous verification of his personality by analyzing behavioral characteristics. At the same time, for the synthesis of the verification model, only the data of the operator whose verification is necessary are used, which is a significant advantage compared to the methods of binary classification, the use of which requires the presence of alternative biometric images.

Keywords: neural network, dynamic verification, behavioral biometrics, time series forecasting, hyperdimensional vector algebra.

Трокоз Дмитрий Анатольевич
К.т.н., доцент, ФГБОУ ВО «Пензенский государственный технологический университет»
trokoz@penzgtu.ru

Аннотация. В работе предлагается методика непрерывной биометрической верификации, основанная на эвристическом алгоритме, использующем математический аппарат широких нейронных сетей, генетические алгоритмы и алгебру гиперразмерных двоичных векторов. Данная методика основана на замене целевой задачи верификации задачей прогнозирования временного ряда действий верифицируемого оператора, обеспечивая непрерывную проверку его личности путем анализа поведенческих характеристик. При этом для синтеза верифицирующей модели используются данные только того оператора, верификация которого необходима, что является существенным преимуществом по сравнению с методами бинарной классификации, применение которых требует наличие альтернативных биометрических образов.

Ключевые слова: нейронная сеть, динамическая верификация, поведенческая биометрия, прогнозирование временных рядов, алгебра гиперразмерных векторов.

Введение

В настоящее время очень часто возникает потребность верификации личности при доступе к каким-либо аппаратно-программным комплексам или информационным ресурсам. Основной недостаток большинства существующих методов верификации личности заключается в единовременности процесса верификации, который часто совмещается с процессом аутентификации. При этом ни методы парольной защиты, ни методы статической биометрии, такие как биометрическая верификация по изображению лиц [1], дактилоскопических изображений [2] и рисунков ладоней рук [3][4], не могут гарантировать, что после успешного прохождения процедуры верификации личности при авторизации не будет произведена передача доступа третьему лицу.

Решение этой проблемы требует использования методик динамической биометрической верификации [5]. Они способны обеспечить непрерывный анализ поведенческих биометрических характеристик человека с целью перманентной верификации его личности. Главный недостаток существующих методик динамической биометрической верификации — это низкая точность и отсутствие унификации, которая бы позволила применять эти методики без привлечения экспертов [6]. Обе проблемы связаны между собой, поскольку унификация позволит исключить так называемый «человеческий фактор» [7], который часто не позволяет учесть ряд факторов, которые ведут к снижению точности. Предлагаемая в данной работе методика направлена на решение проблемы унификации применения методов машинного обучения для решения задачи динамической биометрической верификации.

Биометрическая верификация, как задача прогнозирования

Предлагаемая методика использует алгоритм машинного обучения, объединяющий генетические алгоритмы для параметрической оптимизации и алгебру гиперразмерных двоичных векторов [8], подробно рассмотренный в работе [9]. Этот алгоритм включает большое количество шагов, многие из которых унифицированы, то есть выполняется единообразно для широкого класса задач. Среди всех шагов, предложенного алгоритма, наибольший интерес представляет два шага: шаг 1 «Анализ прикладной задачи» и шаг 4 «Отображение параметров в гиперразмерное представление», поскольку они будет заметно отличаться для разных задач, поскольку они требуют экспертное

Таблица 1. Управляющие команды от одного оператора

	Маркер времени 1	Маркер времени 2	…	Маркер времени M
Параметр 1	Значение 1 / 1	Значение 1 / 2	…	Значение 1 / M
Параметр 2	Значение 2 / 1	Значение 2 / 2	…	Значение 2 / M
…	…	…	…	…
Параметр N	Значение N / 1	Значение N / 2	…	Значение N / M

вмешательство, то есть в данных шагах велико влияние человеческого фактора, которое будет распространено на качество решения задачи в целом. Рассмотрим подробнее в чем особенность указанных шагов.

Первый шаг предлагаемого алгоритма (анализ прикладной задачи) интересен прежде всего тем, что именно на данном шаге требуется исходную задачу заменить эквивалентной задачей, характерной для нейронных сетей (обычно задача классификации или прогнозирования) [10] [11]. При этом если такая замена будет выполнена недостаточно корректно (именно здесь проявляется человеческий фактор), то успешное решение результирующей задачи не сможет гарантировать требуемое качество решения исходной задачи.

Четвертый шаг (отображение параметров в гиперразмерное представление) играет существенную роль в предложенном алгоритме, поскольку для нейросетевых моделей корректное представление входных и выходных для обучения является определяющим фактором, который оказывает влияние, как на скорость обучения нейронной сети, так и на его качество, что в свою очередь, по сути, определяет, может ли быть найдено решение прикладной задачи за требуемое время с заданной точностью [12].

Рассмотрим методику, которая бы позволила унифицировать описанные шаги на примере решения задачи биометрической верификации операторов малогабаритных летательных аппаратов, путем анализа управляющих команд, информация о типах и времени которых будет интерпретироваться, как биометрические данные оператора. Начнем с первого шага, а именно с перехода от задачи биометрической верификации к задаче характерной для нейросетевых моделей. Еще раз подробно рассмотрим указанную задачу, чтобы лучше понять. Для анализа имеем множество зарегистрированных команд управления летательными аппаратами, маркированные временными метками, для нескольких операторов. При этом известно, какие команды, какими операторами были даны. Требуется по имеющимся данным построить модель, которая позволяла бы, по данным аналогичной структуры в будущем, проверить соответствие между множеством команд и оператором.

Целевая задача может быть сведена к задаче прогнозирования временных рядов для управляющих команд оператора [13–15]. При этом, нужно учитывать, что по одной управляющей команде невозможно предсказать следующую команду, необходим анализ последовательности событий во времени. Такие исследования отражены в ряде работ, посвященным анализу с использованием глубоких нейронных сетей [16],[17]. Во многих работах упоминается проблема ограничения размера временного окна, которая возникает из-за отсутствия технической возможности обеспечить достаточный объем вычислительных ресурсов для работы с глубокими нейронными сетями с большим числом параметров, а недостаточный размер временного окна существенно снижает точность решения задачи. Широкие нейронные сети [18], которые предлагается использовать в данной работе, лишены указанного недостатка, поскольку требуют существенно меньший объем вычислительных ресурсов для обучения.

Предлагаемая нейросетевая модель позволяет предсказывать управляющие команды от оператора, зная несколько предыдущих. Однако, это на первый взгляд не решает начальную проблему, а именно модель не верифицирует оператора. Действительно, но возможна косвенная верификация (отсюда и название данной методики). Если нейросетевая модель способна предсказывать управляющие команды, которые последуют от оператора, с определенной точностью, то этот факт можно использовать для верификации оператора [19]. Если прогнозирование происходит с приблизительно с той же точностью, что и для тестовых данных, то логично предположить, что управляет летательным аппаратом тот самый оператор, на зарегистрированных управляющих командах которого обучалась модель. В противном случае оператор не пройдет верификацию.

Еще один важный этап — это отображение параметров задачи в гиперразмерное представление. Схематично зарегистрированные управляющие команды от одного оператора представлены в таблице. Каждый столбец этой таблицы представляет собой временную метку, соответствующую времени возникновения события (команды оператора). Необходимо отметить, что

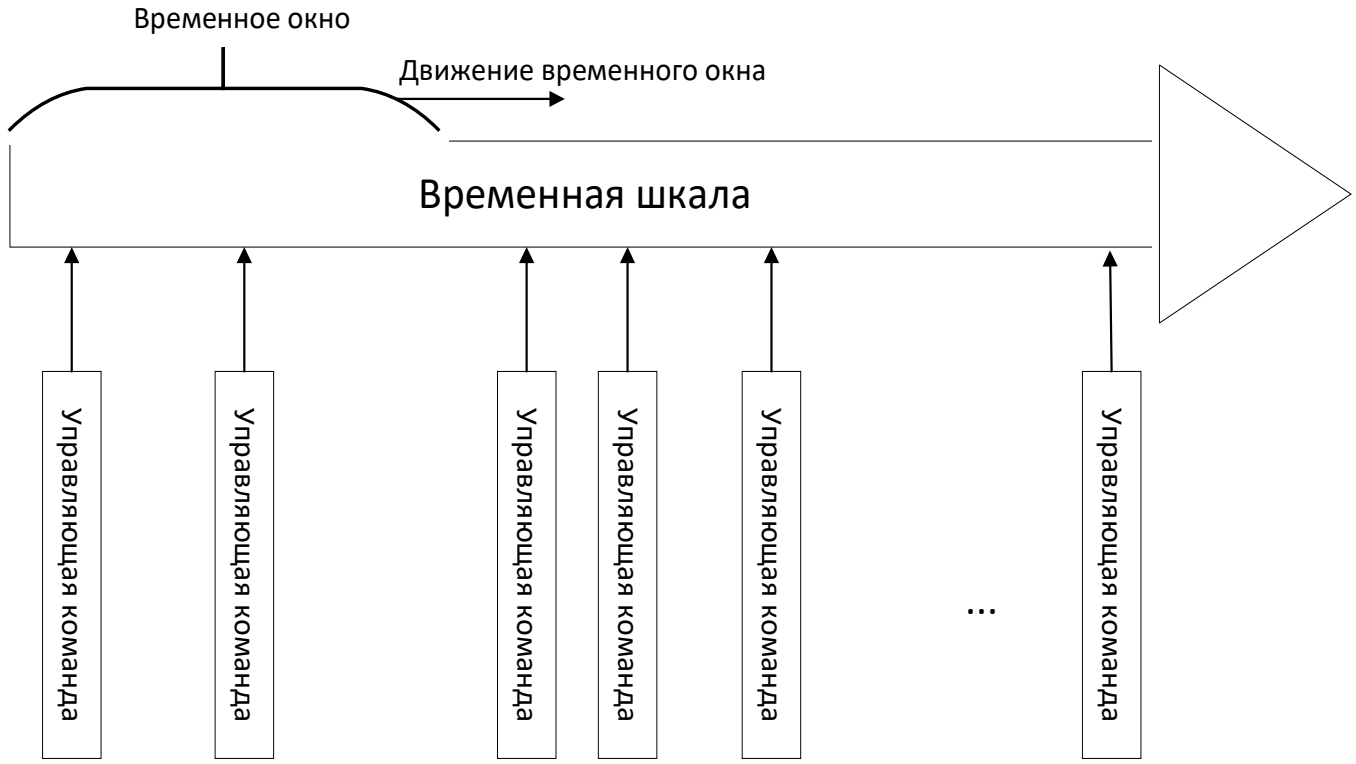

Рис. 1. Концепция временного окна

количество временных меток за один промежуток времени для разных операторов в общем случае различно.

Рассмотрим подробнее какие входные и выходные параметры будет иметь нейросетевая модель для задачи биометрической верификации операторов летательных аппаратов. Как и любая другая нейросетевая модель данные для обучения данной нейронной сети должны представлять множество пар вида «значения входных параметров модели — значения выходных параметров модели». При этом для модели указанные значения параметров должны быть представлены в гиперразмерном виде, то есть в пары вида «гиперразмерный вектор значений входных параметров — гиперразмерный вектор значений выходных параметров».

Количество входных параметров далеко от одного и каждый параметр способен принимать очень большое число значений, как следствие просто сгенерировать необходимое количество (для покрытия всех значений каждого параметра) гиперразмерных двоичных векторов не представляется возможным, поскольку это потребует очень много временных затрат, а сгенерированные векторы попросту не поместятся в оперативную память современных ЭВМ. Кроме того, как было сказано ранее, при работе с временными рядами требуется использовать временное окно.

Временные окна бывают двух типов в зависимости от того, какая характеристика константа для данного окна: количество событий или временная протяженность. Первый тип более прост в реализации, однако, временные окна этого типа неспособны, должны образом учитывать временные задержки между событиями [20]. Поэтому в данной работе предлагается использовать второй тип временных окон. Чтобы лучше понять концепцию временных окон этого типа и проблему, из нее вытекающую рассмотрим рисунок 1. На временной шкале располагаются управляющие команды, которые возникают, когда оператор дает соответствующую команду, используя пульт управления летательным аппаратом, при этом время между отдельными управляющими командами различно.

Временное окно имеет постоянную длину в секундах и перемещается с постоянным шагом, величина которого также равна постоянному количеству секунд. При этом каждая управляющая команда в общем случае содержит значения ряда параметров. В свою очередь входной гиперразмерный вектор нейронной сети должен содержать сразу значения всех параметров всех команд, попадающих во временное окно. Как было сказано ранее, сгенерировать сразу все случайные гиперразмерные векторы для всех возможных значений всех параметров не представляется возможным. Поэтому

Номера разрядов								
4	3	2	1	5	4	3	2	1
Параметр 1				Параметр 2				
1	0	1	1	0	1	0	0	1
Значения разрядов								

Рис. 2. Пример двух входных параметров и их значений

необходимо разработать правила преобразования значений параметров в гиперразмерное представление.

Прежде чем описать предлагаемый подход к синтезу гиперразмерных векторов из классического параметрического представления, рассмотрим две операции, которые могут быть выполнены над гиперразмерными двоичными векторами.

Первой рассматриваемой операцией является произведение двух гиперразмерных векторов. Это бинарная операция, которая представляет собой побитовое «исключающие или» операндов. Результат произведения гиперразмерных двоичных векторов обладает следующим важным свойством: расстояние Хэмминга между результирующим вектором и любым другим синтезированным случайно гиперразмерным вектором, включая векторы, участвующие в произведении, приблизительно равно 0,5. То есть полученный вектор проявляет те же свойства, что и любой другой вектор, синтезированный случайно. Это свойство важнее, чем может показаться на первый взгляд. Оно позволяет использовать результат произведения в операции сложения при формировании входного гиперразмерного вектора нейронной сети.

Чтоб понять, как именно может быть использована указанная операция для уменьшения генерации необходимого числа гиперразмерных двоичных векторов, рассмотрим рисунок 2. На нем представлена пара параметров длиной 4 и 5 бит соответственно.

При использовании подхода, при котором для каждого значения каждого параметра генерируется свой случайный гиперразмерный двоичный вектор потребуется в общей сложности $2^4+2^5=48$ двоичных гиперразмерный векторов. Представим эти векторы как два множества $A^{'} = \{a^{'}\}$ — нумерованное множество гиперразмерных векторов для значений первого параметра (16 векторов) и $A^{"} = \{a^{"}\}$ — нумерованное множество гиперразмерных векторов для значений второго параметра (32 вектора). Нумерация указанных множеств начинается с нуля. Тогда для значений параметров из рисунка 2 результирующий гиперразмерный вектор, включающий оба параметра, будет равен сумме векторов: $a^{'}_{11} + a^{"}_{9}$, поскольку $1011_{(2)} = 11_{(10)}$ и $1001_{(2)} = 9_{(10)}$.

Рассмотрим подход, использующий операцию произведения гиперразмерных векторов для синтеза гиперразмерных векторов любого значения каждого из параметров. Для этого сгенерируем случайные гиперразмерные векторы для каждого параметра в количестве равном на единицу большем, чем количество разрядов соответствующего параметра. Итого для параметров, представленных на рисунке 2 потребуется сгенерировать 5+6=11 гиперразмерных двоичных векторов. Представим эти векторы как два множества $B^{'} = \{b^{'}\}$ — нумерованное множество гиперразмерных векторов для синтеза значений первого параметра (5 векторов) и $B^{"} = \{b^{"}\}$ — нумерованное множество гиперразмерных векторов для синтеза значений второго параметра (6 векторов). Тогда синтезированное значение для каждого параметра будет представлять собой произведение гиперразмерных двоичных векторов, соответствующих разрядам, имеющим единичное значение. Нумерация разрядов каждого параметра начинается с единицы, а нумерация множеств с нуля. При этом нулевой элемент каждого множества всегда включается в произведение, это позволяет избежать получения «пустого» вектора в результате произведения в случае, когда все разряды какого-либо параметра содержат нулевые значения. Результирующий гиперразмерный вектор, включающий оба параметра, будет равен сумме векторов полученных произведений: $b^{'}_{0} \times b^{'}_{1} \times b^{'}_{2} \times b^{'}_{4} + b^{"}_{0} \times b^{"}_{1} \times b^{"}_{4}$. Такой способ синтеза гиперразмерного представления значений параметров позволяет существенно сократить количество необходимых для генерации гиперразмерных двоичных векторов.

Вторая операция, которая будет нужна для синтеза результирующего вектора — операция циклического сдвига. Это унарная операция, в результате которой производится циклический сдвиг значений каждого разряда вектора вправо или влево. Основным свойством данной операции, как и в случае с операцией произведения векторов, является тот факт, что расстояние Хэмминга между результирующим вектором и любым другим синтезированным случайно гиперраз-

мерным вектором, включая вектор, к которому применяется операция сдвига, приблизительно равно 0,5.

Необходимость в этой операции возникает, когда встает вопрос о представлении в виде гиперразмерного двоичного вектора нескольких наборов значений одних и тех же параметров с различными временными метками. Для решения этой проблемы предлагается использовать следующий подход для модификации синтезированного вектора параметров управляющей команды оператора (способ синтеза такого вектора описан выше), связанной с определенной временной меткой. Вычисляется расстояние по временной шкале между значением этой временной метки и началом временного окна. Затем это расстояние дискретизируется с заданным шагом. Полученная величина определяет количества раз, которое применяется операция циклического сдвига к вектору параметров. Затем все модифицированные таким образом векторы параметров складываются, образуя входной вектор нейронной сети.

Теперь рассмотрим, как формируется гиперразмерный вектор значений выходных параметров. Как было сказано, ранее предлагаемая нейросетевая модель должна предсказывать следующую управляющую команду и время ее появления. Соответственно, гиперразмерный вектор значений выходных параметров должен включать управляющую команду и временную метку. Для этого можно использовать способ кодирования команды с использованием операций умножения и сложения гиперразмерных векторов, описанный в предыдущем разделе. А чтобы указать временную метку следует использовать операцию циклического сдвига, как это было показано ранее. Таким образом, выходной вектор будет формироваться по тем же правилам, что и входной с той лишь разницей, что выходной будет включать только одну команду, которая будет идти следом за рядом команд, закодированных во входном векторе.

Описанные выше подходы к формированию пар вида «гиперразмерный вектор значений входных параметров — гиперразмерный вектор значений выходных параметров» позволяют осуществлять эффективное преобразование данных для обучения к гиперразмерному виду. При этом часть данных используется для обучения (порядка 99%), а часть для определения качества достигнутого решения. Указанные подходы вместе с обоснованием задачи прогнозирования управляющей команды по ряду предшествующих зарегистрированных команд, как задачи косвенно решающей задачу верификации, составляют методику косвенной биометрической верификации операторов летательных аппаратов.

Заключение

Преимуществом предложенной методики является возможность обучения нейросетевой модели исключительно на данных от одного оператора, верификация которого необходима. Вторым преимуществом, вытекающим из первого, является существенно меньший объем данных, необходимых для обучения, чем в задачах классификации типа «свой» / «чужой», поскольку требуется обработка только целевых биометрических данных, соответствующих классу «свой». Модель стремится не отличать манеру управления летательным аппаратом одного оператора от множества других, а напротив учится проверять является ли манера управления верифицируемого оператора схожей с той, для которой происходило обучение модели.

Несмотря на перечисленные преимущества, методика косвенной верификации не лишена недостатков. Ключевым недостатком данной методики, является достаточно высокая вероятность ложно положительной верификации для схожих манер управления. Для решения этой проблемы можно использовать большее количество данных для обучения, получаемое от одного оператора. Таким образом, предложенная методика способна обеспечивать непрерывную биометрическую верификацию личности, путем анализа поведенческих характеристик человека, сводя указанную задачу к задаче прогнозирования временных рядов и успешно решая ее.

ЛИТЕРАТУРА

1. Андреева О.В. и др. К вопросу о проведении верификации изображений лиц на основе нейронных сетей //Датчики и системы. — 2014. — № . 5. — С. 56–58.
2. Маркелов К.С. Идентификация и верификация личности–комплексная биометрическая информационная технология //International Journal of Open Information Technologies. — 2015. — Т. 3. — № . 5.
3. Головешко Е.А. Биометрические средства защиты информации //Per aspera ad astra= Через тернии к звёздам. — 2018. — С. 325–330.
4. Грижебовская А.Г., Михалев А.В. Биометрический метод идентификации человека по сосудистому рисунку пальца //Вопросы кибербезопасности. — 2019. — № . 5 (33).
5. Мишин Д.С., Костин С.В., Кузнецов А.А. Перспективы развития методов биометрической идентификации //Научный вестник Орловского юридического института МВД России имени В.В. Лукьянова. — 2018. — № . 2. — С. 152–155.

6. Брюхомицкий Ю.А. Биометрическая идентификация личности методами искусственного интеллекта //Евразийское Научное Объединение.— 2018.— Т. 1.— № . 5 (39).— С. 31.
7. Михайлов А.А., Колосков А.А., Дронов Ю.И. Основные параметры биометрических систем //Алгоритм безопасности.— 2015.— № . 5.— С. 58–61.
8. Пащенко Д.В. и др. Использование алгебры гиперразмерных векторов для эвристического представления данных при обучении широких нейронных сетей // Проблемы управления и моделирования в сложных системах.— 2019.— С. 301–305.
9. Трокоз Д.А. Алгоритм машинного обучения широких нейронных сетей с использованием алгебры гиперразмерных двоичных векторов и генетических алгоритмов / Д.А. Трокоз // Южно-Сибирский научный вестник.— 2020.— № 6.— с. 148–154.
10. Солдатова О.П., Семенов В.В. Применение нейронных сетей для решения задач прогнозирования //Электронный научный журнал «Исследовано в России» http://zhurnal. gpi. ru/articles/2006/136. pdf.— 2006.
11. Талалаев А.А. и др. Анализ эффективности применения искусственных нейронных сетей для решения задач распознавания, сжатия и прогнозирования //Искусственный интеллект и принятие решений.— 2008.— Т. 2.— С. 24–33.
12. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей //Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика.— 2017.— Т. 6.— № . 3.
13. Садовникова Н.А., Шмойлова Р.А. Анализ временных рядов и прогнозирование.— 2016.
14. Дуброва Т.А. Статистические методы прогнозирования.— 2003.
15. Татьянкин В.М. Использование многослойных нейронных сетей в прогнозирование временных рядов //Приоритетные направления развития науки и образования.— 2014.— № . 3.— С. 195–197.
16. Макаренко А.В. Глубокие нейронные сети: зарождение, становление, современное состояние //Проблемы управления.— 2020.— № . 2.— С. 3–19.
17. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей //Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика.— 2017.— Т. 6.— № . 3.
18. Гальцов В.П. Особенности современных нейронных сетей //Вестник науки.— 2019.— Т. 5.— № . 6.— С. 183–186.
19. Катасёв А.С. и др. Нейросетевая модель распознавания пользователей в системах дистанционного обучения //Вестник Казанского технологического университета.— 2015.— Т. 18.— № . 13.
20. Пучков Е.В. Разработка среды моделирования искусственных нейронных сетей. Решение задачи прогнозирования временного ряда //Вестник Ростовского государственного университета путей сообщения.— 2009.— № . 2.— С. 44–50.

Журнал «Современная наука: актуальные проблемы теории и практики»

г. Пенза