Зайвый Илья Александрович
Студент 4 курса
Факультет Информационная безопасность
ФГАОУ ВО НИЯУ МИФИ
г. Москва

Бехтин Артем Владимирович
Студент 4 курса
Факультет Информационная безопасность
ФГАОУ ВО НИЯУ МИФИ
г. Москва

ПРИМЕНЕНИЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ В РЕШЕНИИ ЗАДАЧИ ОПТИМИЗАЦИИ ПОДБОРА ПОРОГОВОГО ЗНАЧЕНИЯ ДЛЯ МОДУЛЯ АНАЛИЗА КЛАВИАТУРНОГО ПОЧЕРКА

Аннотация: *В работе рассматривается задача автоматической калибровки порогового значения метрики риска в модуле анализа клавиатурного почерка системы непрерывной аутентификации пользователей ОС Windows. Специфика задачи определяется одноклассовым характером обучающей выборки, несбалансированностью классов и недифференцируемостью целевой функции: метрики TPR и FPR являются кусочно-постоянными функциями порога, что исключает применение градиентных методов. Для поиска оптимального порога предложен генетический алгоритм с вещественным кодированием (RCGA), в котором ограничение на вероятность ложного срабатывания (FPR ≤ 5%) учитывается методом штрафных функций. Метрика риска вычисляется как ошибка реконструкции автоэнкодера с последующей Z-score нормализацией. Экспериментальная проверка на датасете KeyRecs показала быструю и воспроизводимую сходимость (3–5 поколений, < 2 с) и позволила получить*

порог θ = 2.02σ при TPR = 60,09% и FPR = 4,93%. Показано, что RCGA обеспечивает контролируемый компромисс между безопасностью и удобством пользователя.*

Ключевые слова: *Генетические алгоритмы, подбор порогов, RCGA, обнаружение аномалий, биометрия, информационная безопасность.*

Zayvy Ilya Alexandrovich
Student 4 term
Faculty of Information security
NRNU MEPhI
Moscow

Bekhtin Artem Vladimirovich
Student 4 term
Faculty of Information security
NRNU MEPhI
Moscow

APPLICATION OF GENETIC ALGORITHMS IN SOLVING THE PROBLEM OF OPTIMIZING THE THRESHOLD VALUE FOR THE KEYBOARD HANDWRITING ANALYSIS MODULE

Annotation: The paper addresses the automatic calibration of the risk-metric threshold in a keystroke-dynamics module of a continuous authentication system for Windows operating systems. The problem is characterised by a one-class training set, class imbalance, and a non-differentiable objective function: the TPR and FPR metrics are piecewise-constant functions of the threshold, which precludes gradient-based optimisation. A real-coded genetic algorithm (RCGA) is proposed to search for the optimal threshold, with the false-positive-rate constraint (FPR ≤ 5%) incorporated via a penalty method. The risk metric is computed as the autoencoder

reconstruction error followed by Z-score normalisation. Experimental evaluation on the KeyRecs dataset showed fast and reproducible convergence (3–5 generations, < 2 s) and yielded an optimal threshold $\theta^* = 2.02\sigma$ with TPR = 60.09% and FPR = 4.93%. The results confirm that RCGA provides a controllable trade-off between security and usability.

Key words: Genetic algorithms, threshold selection, RCGA, anomaly detection, biometrics, and information security.

Цель работы

В рамках разработки системы непрерывной аутентификации пользователей [4, с. 125] возникает задача динамического определения порогового значения метрики риска, разделяющего легитимное поведение пользователя и потенциальные попытки несанкционированного доступа.

Система функционирует в условиях:

- одноклассовой обучающей выборки: на этапе регистрации доступны только данные легитимного пользователя;
- несбалансированности данных: объем легитимной выборки существенно отличается от совокупного объема данных потенциальных атакующих;
- недифференцируемости целевой функции: метрики качества классификации (TPR, FPR) являются кусочно-постоянными функциями порогового параметра, что исключает применение градиентных методов оптимизации;
- высоких требований к скорости сходимости: процедура калибровки порога должна завершаться в приемлемые временные рамки для обеспечения практической применимости.

Для решения задачи поиска оптимального порога θ^* был выбран генетический алгоритм с вещественным кодированием (Real-Coded Genetic Algorithm, RCGA) [3, с. 266] по следующим основаниям:

- отсутствие требований к дифференцируемости: метод относится к классу методов нулевого порядка и не требует вычисления производных целевой функции;

- глобальная оптимизация: стохастическая природа алгоритма снижает риск застревания в локальных экстремумах, характерных для ступенчатых ландшафтов приспособленности;

- гибкость ограничений: возможность включения жестких ограничений на параметры решения (FPR ≤ 5%) непосредственно в функцию приспособленности [1, с. 314];

- масштабируемость: архитектура алгоритма допускает расширение пространства поиска на многомерный случай (оптимизация вектора гиперпараметров) без принципиальной модификации ядра;

- вычислительная эффективность: при оптимизации скалярного параметра время оценки одного поколения составляет единицы миллисекунд, что обеспечивает практическую применимость.

Теоретическое описание алгоритма

Математическая модель системы

Пусть задана выборка признаков клавиатурного почерка легитимного пользователя $L = \{(\tau_i, \lambda_i)\}_{i=1}^{N}$, где τ_i – время удержания клавиши, λ_i – латентность между нажатиями.

Автоэнкодер $f_\phi: R^2 \to R^2$с параметрами ϕ обучается минимизировать функцию потерь [5, с. 5] по формуле (1):

$$L(\phi) = \frac{1}{N}\sum_{i=1}^{N} \quad \| x_i - f_\phi(x_i) \|^2, x_i = [\tau_i, \lambda_i]^T \qquad (1)$$

Для произвольного окна из W последовательных нажатий метрика риска определяется как среднеквадратичная ошибка реконструкции [5, с. 6] по формуле (2):

$$\rho = \frac{1}{W}\sum_{j=1}^{W} \quad \| x_j - f_\phi(x_j) \|^2 \qquad (2)$$

При программной реализации алгоритма и исследованиях на тестовом

датасете выяснилось, что значения метрики риска для части атакующих пользователей находятся в диапазоне, характерном для легитимного пользователя, что ограничивает максимально достижимое значение TPR. Для обеспечения интерпретируемости и персонализации порогового значения применяется процедура Z-score нормализации согласно формуле (3):

$$z = \frac{\rho - \mu_L}{\sigma_L + \varepsilon} \quad (3)$$

где $\mu_L = E[\rho \mid L]$, $\sigma_L = \sqrt{Var[\rho \mid L]}$ – эмпирические моменты распределения риска легитимного пользователя, $\varepsilon = 10^{-6}$ – константа для обеспечения численной устойчивости.

Формализация задачи оптимизации

Требуется найти скалярный порог θ^*, максимизирующий вероятность обнаружения атак (True Positive Rate, TPR) при ограничении на вероятность ложного срабатывания (False Positive Rate, FPR) [1, с. 313], описываемого формулой (4):

$$\theta^* = arg\, max_{\theta \in [\theta_{min}, \theta_{max}]} TPR(\theta) \quad (4)$$

при условии:

$$FPR(\theta) \leq \alpha, \alpha = 0.05$$

где TPR и FPR вычисляются по формулам (5) и (6) соответственно:

$$TPR(\theta) = \frac{1}{|A|} \sum_{a \in A} I[z_a > \theta], \quad (5)$$

$$FPR(\theta) = \frac{1}{|L|} \sum_{l \in L} I[z_l > \theta] \quad (6)$$

A – выборка данных атакующих, $I[\cdot]$ – индикаторная функция.

Архитектура генетического алгоритма

Применяется вещественное кодирование: каждая особь популяции представляет собой скалярное значение $\theta \in R$ [3, с. 267]. Диапазон поиска $[\theta_{min}, \theta_{max}] = [0.5, 4.0]$ определяется на основе эмпирических перцентилей распределения -оценок.

Фитнес-функция реализует метод штрафных функций для учета ограничения на FPR [1, с. 314]:

где $\beta = 10.0$— коэффициент штрафа, обеспечивающий приоритет соблюдения ограничения.

Используемые операторы эволюции:

Инициализация: популяция размера $P = 50$ формируется равномерным случайным распределением в допустимом диапазоне.

Отбор: применяется турнирный отбор с размером турнира $k = 3$.

Кроссовер: оператор расширенного промежуточного рекомбинирования (BLX-α) с параметром $\alpha = 0.3$[3, с. 275].

Мутация: адаптивная гауссова мутация с затухающей дисперсией:

$$\theta' = \theta + N(0, \sigma_g), \sigma_g = \sigma_0 \cdot \left(\frac{g}{G}\right) + \sigma_{min}$$

где g— номер текущего поколения, $G = 50$— максимальное число поколений, $\sigma_0 = 0.05$, $\sigma_{min} = 0.01$.

Элитизм: пять особей с наивысшими значениями фитнес-функции переносятся в следующее поколение без изменений.

Критерий остановки:

Алгоритм завершает работу после исчерпания максимального числа поколений G. Дополнительный критерий ранней остановки по сходимости фитнес-функции не применяется ввиду малого времени вычисления.

После определения оптимального порога θ^*классификация каждого окна признаков осуществляется по правилу:

$$decision(z) = \{ACCEPT, z \leq \theta^* \; REJECT, z > \theta^*$$

Программная реализация

Среда разработки и особенности реализации

Разработка выполнена в среде операционной системы Windows 10 с использованием интерпретатора Python 3.9. Применены следующие библиотеки:

Таблица 1 – Используемый стек технологий

Библиотека	Версия	Назначение

TensorFlow	2.13.0	Инференс автоэнкодера
scikit-learn	1.3.0	Предобработка данных, масштабирование
pandas	2.0.3	Загрузка и обработка табличных данных
NumPy	1.24.3	Векторизованные вычисления
Matplotlib	3.7.2	Визуализация результатов

Легитимная выборка: данные пользователя «bob», собранные в ходе регистрации системы, объемом 893 окна по 60 нажатий. Признаки: hold_time, latency (единицы измерения — миллисекунды).

Выборка атак: объединенные данные 102 пользователей из датасета KeyRecs free-text [2, с. 3], нормализованные и пропущенные через модель пользователя «bob». Объем: 448 524 окна.

Предобработка:

- фильтрация некорректных значений (hold_time ≤ 0, latency ≤ 0, hold_time > 2000 мс, latency > 3000 мс);
- масштабирование признаков с использованием StandardScaler, обученного на данных легитимного пользователя;
- формирование скользящих окон фиксированного размера W=60;
- Z-score нормализация метрики риска относительно статистик легитимной выборки.

Результаты экспериментов

После выполнения генетического алгоритма получены следующие значения:

Таблица 2 – Результаты выполнения алгоритма

Параметр	Значение
Оптимальный порог $\theta*$	2.0229σ
TPR (обнаружение атак)	60.09%
FPR (ложные срабатывания)	4.93%
Число поколений до сходимости	3–5
Время выполнения оптимизации	< 2 с

График сходимости алгоритма демонстрирует монотонный рост функции приспособленности с выходом на плато в пределах первых пятнадцати поколений, что свидетельствует о корректной работе операторов эволюции и адекватности выбранного диапазона поиска.

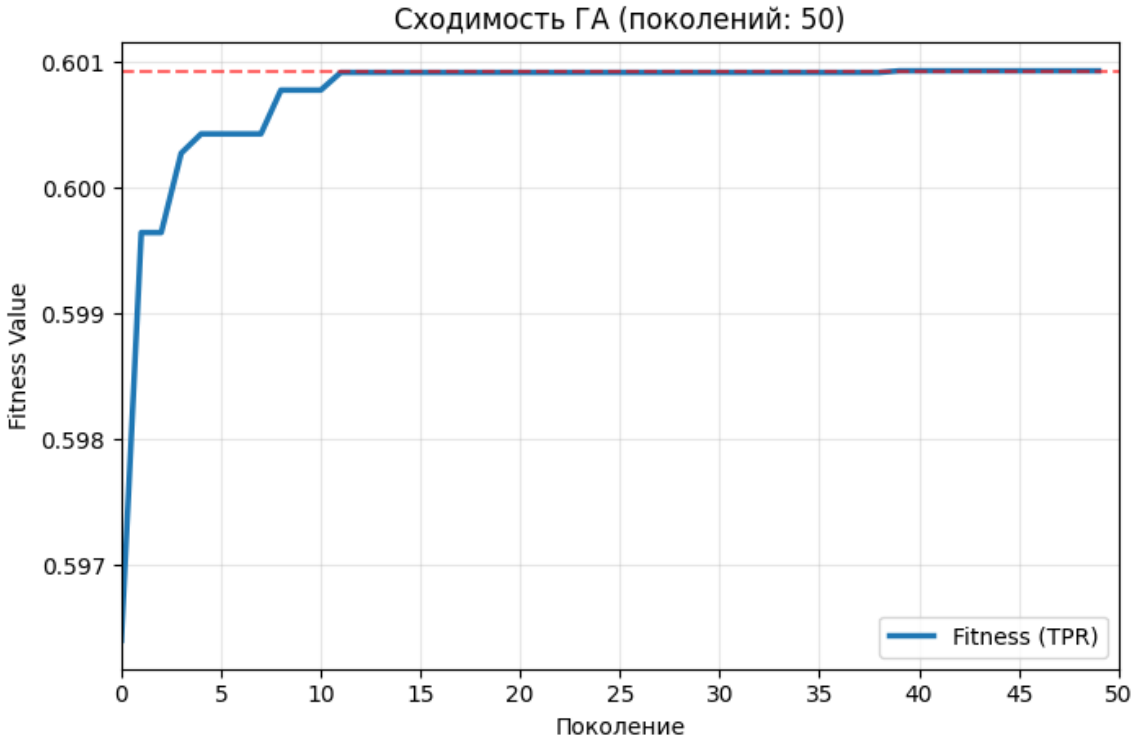

Рисунок 1 – Сходимость реализованного алгоритма

На рисунке 2 заметно, что основной пик даже Z-score нормализованных

метрик риска приходится на пик для легитимного пользователя. В данном случае приходится постоянно балансировать между TPR и FPR. Для системы по умолчанию было установлено значение FPR не выше 5%, таким образом, максимального TPR удалось добиться лишь в 60%, однако данное ограничение связано непосредственно с реализацией автоэнкодера и фактической схожести обученной модели с тестовыми нелегитимными датасетами.

При долгом использовании системы (около 7 часов) с определенным алгоритмом порогом 2.02, триггер на превышение порога сработал лишь один раз.

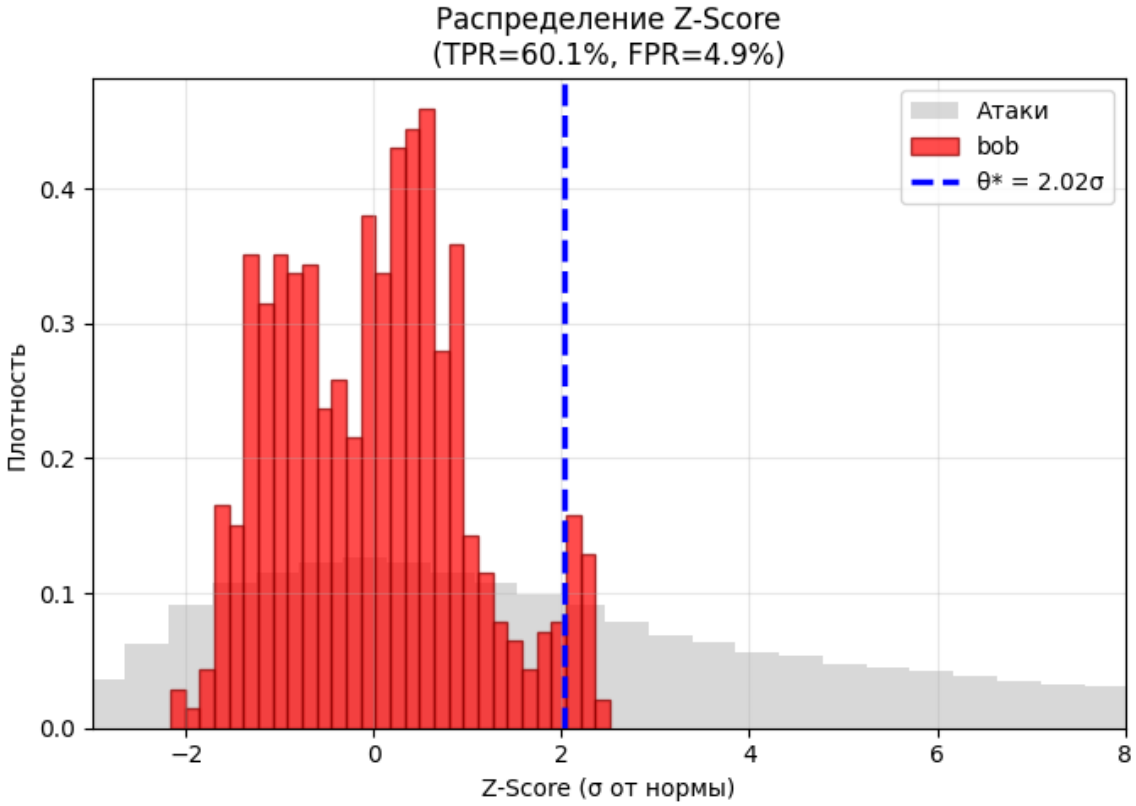

Рисунок 2 – Распределение Z-score нормализованных метрик риска

Заключение

Применение RCGA позволило автоматизировать процесс калибровки порогового значения с соблюдением заданного ограничения на вероятность ложного срабатывания (FPR ≤ 5%). Алгоритм продемонстрировал быструю сходимость (3–5 поколений) и воспроизводимость результатов.

Полученное значение порога $\theta^* = 2.0229$ вместо используемого до этого статического 2.5 обеспечивает баланс между безопасностью (TPR =

60.09%) и удобством пользователя (FPR = 4.93%), что соответствует требованиям к системам непрерывной аутентификации уровня MVP.

Значение TPR ≈ 60% отражает системное ограничение автоэнкодеров, оптимизирующих ошибку реконструкции, а не разделяющую способность между классами [5, с. 9]. Пользователи со статистически близкими паттернами клавиатурного почерка не могут быть надежно отделены на основе исключительно метрики MSE [4, с. 130].

Генетический алгоритм является обоснованным выбором для задачи оптимизации порога в условиях недифференцируемой целевой функции и жестких ограничений. Для одномерного пространства поиска вычислительные затраты минимальны, а гибкость настройки функции приспособленности обеспечивает контроль над компромиссом безопасности и удобства.

Разработанный подход к оптимизации порога обнаружения аномалий на основе генетического алгоритма с вещественным кодированием и Z-score нормализацией метрики риска является работоспособным решением для системы непрерывной аутентификации. Полученные результаты могут быть использованы в качестве базовой конфигурации с последующей донастройкой под конкретные условия эксплуатации.

Список использованных источников:

1. Deb K. An Efficient Constraint Handling Method for Genetic Algorithms // Computer Methods in Applied Mechanics and Engineering. — 2000. — Vol. 186, № 2–4. — P. 311–338.

2. Dias T., Vitorino J., Maia E., Sousa O., Praça I. KeyRecs: A Keystroke Dynamics and Typing Pattern Recognition Dataset // Data in Brief. — 2023. — Vol. 50. — Article ID 109509. — DOI: 10.1016/j.dib.2023.109509.

3. Herrera F., Lozano M., Verdegay J. L. Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis // Artificial Intelligence Review. — 1998. — Vol. 12, № 4. — P. 265–319.

4. Killourhy K. S., Maxion R. A. Comparing Anomaly-Detection

Algorithms for Keystroke Dynamics // Proceedings of the IEEE/IFIP International Conference on Dependable Systems & Networks (DSN). — 2009. — P. 125–134.

5. Sakurada M., Yairi T. Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction // Proceedings of the MLSDA 2014 Workshop. — 2014. — P. 4–11.