Применение нечеткой логики для приоритизации рисков, связанных с человеческим фактором

Д. М. Курпаченко

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

dkurpachenko@gmail.com

Аннотация. **Разработана нечеткая модель интегральной оценки потенциальной уязвимости сотрудников организации. На основе анализа научной литературы выделены семь критериев, отражающих профессиональный статус, информационную открытость и личностно-поведенческие характеристики персонала. С использованием алгоритма нечеткого вывода Мамдани построена система лингвистических переменных, функций принадлежности и базы правил, формализующая логику экспертного оценивания. Приведен пример работы модели для двух гипотетических сотрудников, демонстрирующий ее работоспособность.**

Ключевые слова: нечеткая логика; человеческий фактор; оценка рисков; информационная безопасность; поддержка принятия решений; алгоритм Мамдани

I. Введение

Проблема человеческого фактора остаётся одной из ключевых в обеспечении информационной безопасности организаций. В литературе часто отмечается, что сотрудники зачастую являются самым слабым звеном в защите, а их ошибки могут привести к серьёзнейшим инцидентам [1]. Это подтверждается и статистическими исследованиями [2]. В то же время научно подтверждается, что возможно выявить потенциальные риски ещё до их реализации [3], для этого активно используются методы социотехнического тестирования и сбор информации из открытых источников [4]. В дальнейшем полученные данные позволяют укрепить защищённость организации.

Для формализации процесса оценки рисков, связанных с человеческим фактором. является целесообразным использовать аппарат нечёткой логики. Данный подход уже зарекомендовал себя, как полезный при работе с оценкой поведения сотрудников [5]. Кроме работы с оценкой вероятности утечки информации на основе данных сотрудников, нечёткую логику можно использовать и для оценки подверженности сотрудников конкретным видам фишинга [1], что можно рассматривать как наиболее близкое к данному докладу исследование.

Между тем, существующие работы либо опираются на статичный набор критериев [1], либо используют сложные для интерпретации методы машинного обучения [3]. В данном докладе развиваются идеи из исследования [1] для интеграции их в процесс сбора информации о сотрудниках. Целью работы является разработка нечеткой модели для интегральной оценки потенциальной уязвимости сотрудника, которая могла бы служить инструментом поддержки принятия решений для специалистов по безопасности.

II. Исходные данные и система оценок

Для разработки модели оценки потенциальной уязвимости сотрудников необходима система критериев, отражающих различные аспекты, влияющие на итоговый показатель. При выборе критериев в данном докладе учитывались 2 аспекта: доступность (возможность сбора данных для оценки) и теоретическая обоснованность использования критерия для оценки рисков. Все критерии предлагается разделить на три группы, характеризующие профессиональный статус сотрудника, его информационную открытость и личностно-поведенческие характеристики. Для каждого критерия вводится лингвистическая переменная с соответствующим терм-множеством, что позволяет перейти от качественных описаний к формализованным оценкам, пригодным для дальнейшей обработки методами нечеткой логики.

A. Профессионально-должностные критерии

Критерий 1. Критичность должности / уровень доступа (K_1). Критерий отражает значимость сотрудника для ключевых бизнес-процессов организации и объём его привилегий в информационном пространстве компании. Очевидна связь между уровнем доступа сотрудника и объёмом потенциальных последствий его действий для системы. Лингвистическая переменная «Критичность» задается терм-множеством {Низкая, Средняя, Высокая}. Низкий уровень соответствует рядовым исполнителям без доступа к конфиденциальной информации, высокий — руководителям подразделений и сотрудникам, имеющим прямые доступы к финансовым, технологическим или управляющим системам.

B. Информационно-цифровые критерии

Критерий 2. Доступность публичной информации (K_2). Критерий характеризует объём и глубину сведений о сотруднике, доступных в открытых источниках. Переменная «Информационная открытость» принимает значения {Низкая, Средняя, Высокая}. Низкий уровень соответствует наличию лишь минимальных сведений (имя, должность) на официальном сайте организации, высокий — наличию развернутых профилей в нескольких социальных сетях, публикаций, интервью, а также фактов утечки учетных данных.

Критерий 3. Внешняя профессиональная активность (K_3). Критерий отражает причастность сотрудника к

профессиональному сообществу – конференции, ведение блогов, участие в круглых столах… Повышенная активность создаёт дополнительные сценарии взаимодействия с человеком. Лингвистическая переменная «Активность» имеет термы {Низкая, Средняя, Высокая}.

C. Личностно-поведенческие критерии

Критерий 4. Отношение к работодателю (K_4). Критерий учитывает отношение сотрудника к организации и руководительскому составу. Негативное отношение воспринимается, как повышенная вероятность преднамеренных и непреднамеренных деструктивных действий. Переменная «Лояльность» принимает значения {Негативная, Нейтральная, Позитивная}. Оценка может производиться на основе данных внутренних опросов, анализа текучести кадров в подразделении, а также (с соблюдением этических норм) на основе публичных высказываний сотрудника.

Критерий 5. Социально-экономическое положение (K_5). Критерий основывается на факторах, характеризующих социально-экономическую уязвимость человека. Финансовые трудности и социальные проблемы могут повышать склонность человека к риску или делать его более податливым к внешнему влиянию. Лингвистическая переменная «Социальная стабильность» задается термами {Стабильное, Нестабильное, Критическое}. В рамках данной работы этот критерий рассматривается исключительно как теоретически значимый фактор; его практическое измерение требует соблюдения соответствующих правовых и этических норм.

Критерий 6. Склонность к риску / психологические особенности (K_6). Критерий характеризует личностные черты человека, влияющие на восприимчивость к методам социальной инженерии. Переменная «Психологический профиль» принимает значения {Низкая, Средняя, Высокая} склонность к риску/доверчивость.

Критерий 7. Цифровая грамотность (K_7). Критерий привязан к уровню понимания сотрудником принципов кибербезопаности. Данный критерий является обратным по отношению к риску: чем выше грамотность, тем ниже потенциальная уязвимость. Переменная «Осведомленность» имеет термы {Низкая, Средняя, Высокая}. Оценка может производиться на основе результатов прохождения тренингов, тестирований на фишинг, данных о нарушениях политик безопасности.

Таким образом, предлагаемая система включает семь критериев, охватывающих три ключевых аспекта: профессиональный статус (K_1), информационную открытость (K_2, K_3) и личностно-поведенческие характеристики (K_4, K_5, K_6, K_7). Совокупность перечисленных критериев образует многомерное пространство признаков, на основе которого в дальнейшем будет построена система нечеткого вывода для интегральной оценки потенциальной уязвимости сотрудника.

III. Построение нечеткой модели оценки

Для решения задачи исследования наиболее подходящим представляется использование алгоритма нечёткого вывода Мамдани. Данный алгоритм широко применяется в задачах оценки и классификации благодаря своей интерпретируемости: база правил формируется на естественном языке и легко может быть скорректирована экспертом, данных подход используется и в наиболее близком исследовании [1]. Процесс нечеткого вывода включает следующие этапы: формирование базы правил, фаззификация входных переменных, агрегирование, активизация подзаключений, аккумуляция и дефаззификация (приведение к четкому значению). Сведём предложенные лингвистические переменные в табл. 1.

ТАБЛИЦА I.

Критерий	**Лингвистическая переменная**	**Терм-множество**
K_1	«Критичность должности»	{Низкая (Н), Средняя (С), Высокая (В)}
K_2	«Информационная открытость»	{Низкая (Н), Средняя (С), Высокая (В)}
K_3	«Профессиональная активность»	{Низкая (Н), Средняя (С), Высокая (В)}
K_4	«Лояльность»	{Негативная (Нег), Нейтральная (Нейтр), Позитивная (Поз)}
K_5	«Социальная стабильность»	{Стабильная (Ст), Нестабильная (Нест), Критическая (Кр)}
K_6	«Склонность к риску»	{Низкая (Н), Средняя (С), Высокая (В)}
K_7	«Цифровая грамотность»	{Низкая (Н), Средняя (С), Высокая (В)}

В качестве выходной переменной ожидается «Индекс потенциальной уязвимости» P со значениями {Низкий (Н), Средний (С), Высокий (В)}.

Для лингвистических переменных будем использовать трапециевидные и треугольные функции принадлежности, как наиболее простые при экспертной настройке. Универсум для всех переменных примем в интервале [0, 1]. Ввиду ограниченного объема приведем описание функций принадлежности для выходной переменной «Индекс потенциальной уязвимости»:

- Терм «Низкий»: трапециевидная функция с параметрами [0; 0; 0,2; 0,35];
- Терм «Средний»: треугольная функция с параметрами [0,2; 0,5; 0,8];
- Терм «Высокий»: трапециевидная функция с параметрами [0,65; 0,8; 1; 1].

Иллюстрация функций для выходной принадлежности приведена на рис. 1.

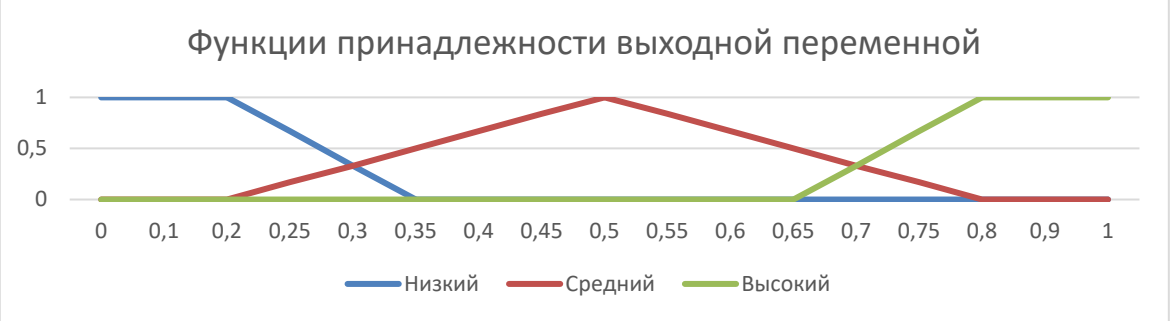

Рис. 1. Иллюстрация функций для выходной переменной

Аналогичным образом задаются функции принадлежности для входных переменных. Параметры функций могут уточняться в процессе экспертного оценивания и настройки модели в зависимости от условий организации и внешних факторов.

База правил является ядром нечёткой модели. Для 7 входных критериев полная база могла бы содержать 2187 правил, однако на практике используется сокращенный набор, охватывающий наиболее значимые комбинации критериев. Приведём несколько типовых правил, которые охватывают часть типовых сценариев:

1) ЕСЛИ (Критичность = Высокая) И (Информационная открытость = Высокая) И (Лояльность = Негативная) И (Социальная стабильность = Критическая), ТО (Индекс уязвимости = Высокий), так как сотрудник с высокими привилегиями, открытой информацией о себе, негативно настроенный к работодателю и находящийся в сложных жизненных обстоятельствах представляет наибольший риск;

2) ЕСЛИ (Критичность = Высокая) И (Лояльность = Позитивная) И (Цифровая грамотность = Высокая), ТО (Индекс уязвимости = Средний), высокий уровень доступа может частично компенсироваться лояльностью и высокой осведомленностью, снижая итоговую оценку до среднего уровня.

Как уже указывалось ранее, полный набор правил формируется экспертом, исходя из условий организации.

Для получения четкого числового значения «Индекса потенциальной уязвимости» используется следующая процедура:

1) Фаззификация: для каждого сотрудника по каждому критерию определяется степень принадлежности к соответствующим термам на основе заданных функций принадлежности;

2) Агрегирование и активизация: для каждого правила вычисляется степень истинности посылки (как минимум конъюнкций) и производится "усечение" функции принадлежности заключения;

3) Аккумуляция: все усеченные функции принадлежности, полученные для выходной переменной, объединяются (обычно с использованием операции максимума);

4) Дефаззификация: полученное нечеткое множество преобразуется в четкое число. Наиболее часто используется метод центра тяжести, при котором итоговое значение вычисляется как абсцисса центра тяжести результирующей фигуры.

Итоговое значение $P \in [0, 1]$ интерпретируется следующим образом:

- $0 \le P < 0{,}35$ — низкий уровень потенциальной уязвимости;
- $0{,}35 \le P < 0{,}65$ — средний уровень;
- $0{,}65 \le P \le 1$ — высокий уровень.

IV. Пример работы модели

Смоделируем двух сотрудников организации и проверим их по упрощённому набору из 6 правил, чтобы доказать практическую применимость предложенного метода.

Сотрудник А (рядовой бухгалтер): Критичность должности – низкая; Информационная открытость – низкая; Профессиональная активность – низкая; Лояльность – позитивная; Социальная стабильность – стабильная; Склонность к риску – низкая; Цифровая грамотность – средняя.

Сотрудник Б (руководитель отдела закупок): Критичность должности – высокая; Информационная открытость: высокая; Профессиональная активность – высокая; Лояльность – нейтральная; Социальная стабильность – нестабильная; Склонность к риску – средняя; Цифровая грамотность – средняя.

Смоделируем эти же характеристики в числах. Для сотрудника А:

- K_1=0.2 → Низкая: 0.8, Средняя: 0.2, Высокая: 0;
- K_2=0.1 → Низкая: 1.0;
- K_3=0.1 → Низкая: 1.0;
- K_4=0.9 → Поз: 1.0;
- K_5=0.8 → Стаб: 1.0;
- K_6=0.2 → Низкая: 0.8, Средняя: 0.2;
- K_7=0.5 → Средняя: 1.0.

Для сотрудника Б:

- K_1=0.9 → Высокая: 1.0
- K_2=0.85 → Высокая: 1.0
- K_3=0.9 → Высокая: 1.0
- K_4=0.5 → Нейтр: 0.8, Нег: 0.2
- K_5=0.4 → Нестаб: 1.0
- K_6=0.5 → Средняя: 1.0
- K_7=0.5 → Средняя: 1.0

В табл. 2 приведены правила, по которым мы сравним представленных выше сотрудников.

ТАБЛИЦА II.

№	$К_1$	$К_2$	$К_3$	$К_4$	$К_5$	$К_6$	$К_7$	**Р**
1	В	С	—	Нег	Кр	—	—	В
2	—	—	—	Поз	—	—	В	Н
3	С	С	—	—	—	В	Н	С
4	В	—	—	Поз	—	—	В	С
5	В	В	—	—	—	—	—	В
6	Н	—	—	Поз	—	—	—	Н

Примечание: В — высокий, С — средний, Н — низкий, Нег — негативная, Поз — позитивная, Кр — критическая, «—» — любое значение.

Рассчитаем правила для каждого из сотрудников.

Для сотрудника А сработает (не в 0) только правило 6: $К_1$=Н: 0.8; $К_4$=Поз: 1.0; → min(0.8, 1.0) = 0.8 (Низкий риск по правилу).

Для сотрудника Б сработает только правило 5: $К_1$=В: 1.0; $К_2$=В: 1.0 → min(1.0, 1.0) = 1.0 (Высокий риск по правилу).

Далее возьмём предложенные в докладе функции принадлежности:

Для А: Усекаем «Низкий» на уровне 0.8 → центр тяжести будет около 0.16, что попадает под низкий риск.

Для Б: Усекаем «Высокий» на уровне 1.0 (полностью) → центр тяжести около 0.86, что попадает под высокий риск.

Итого получаем для сотрудника А Р ≈ 0.16 (низкий риск), и для сотрудника Б Р ≈ 0.86 (высокий риск). Полученные значения соответствуют экспертным ожиданиям: сотрудник с низкими привилегиями и закрытым профилем попадает в зону низкого риска, тогда как руководитель с высоким доступом, открытой информацией и признаками нестабильности — в зону высокого.

V. Заключение

В работе разработана нечеткая модель интегральной оценки потенциальной уязвимости сотрудников организации. На основе анализа научной литературы выделены семь критериев, отражающих профессиональные, информационные и личностно-поведенческие характеристики персонала. С использованием аппарата нечеткой логики построена система вывода Мамдани, включающая лингвистические переменные, функции принадлежности и базу правил, формализующую логику экспертного оценивания.

Практическая значимость работы заключается в возможности использования предложенной модели службами организаций для приоритизации внимания к сотрудникам, представляющим повышенный риск, и при планирования профилактических мероприятий.

Дальнейшие исследования могут быть направлены как на валидацию модели на реальных данных и сравнении с экспертными оценками, так и на разработку программного прототипа.

Список литературы

[1] Lambat Y., Ayres N., Maglaras L., & Ferrag M. A. (2021). A Mamdani Type Fuzzy Inference System to Calculate Employee Susceptibility to Phishing Attacks. Applied Sciences, 11(19), 9083. https://doi.org/10.3390/app11199083

[2] Авезова Я. Актуальные киберугрозы для организаций в СНГ: H2 2024 — Q3 2025 / Я. Авезова, В. Беседина // PT Expert Security. 2025. — URL: https://ptsecurity.com/research/analytics/cis-cyberthreat-landscape-h2-2024-q3-2025/ (дата обращения: 15.03.2026).

[3] Федин Ф.О. Применение самоорганизующихся карт в целях совершенствования системы защиты информации автоматизированной системы / Ф.О. Федин, И.К. Андреев // Вестник компьютерных и информационных технологий. 2025. № 5. С. 41-47. – DOI: 10.14489/vkit.2025.05.pp.041-047.

[4] Dacko S. Discover Your Vulnerabilities before Hackers Do! [Электронный ресурс] // Social Engineer. 2024. – URL: https://www.social-engineer.com/discover-your-vulnerabilities-before-hackers-do/ (дата обращения: 15.03.2026).

[5] Váczi D. Fuzzy-based Cybersecurity Risk Analysis of the Human Factor from the Perspective of Classified Information Leakage / D. Váczi, E. Tóth-Laufer, T. Szádeczky // 18th IEEE International Symposium on Intelligent Systems and Informatics (SISY). – Subotica, Serbia: IEEE, 2020. P. 113-118. – DOI: 10.1109/SISY50555.2020.9217053.