

Л. В. Жукова, ст. препод.¹, вед. аналитик-математик⁴, lvzhukova@hse.ru,

И. М. Ковальчук, аспирант², ст. аналитик⁴, ikovalchuk@ec-leasing.ru,

А. А. Кочнев, аспирант³, ст. аналитик⁴, akochnev@ec-leasing.ru,

В. Р. Чугунов⁴, ген. директор, vchugunov@ec-leasing.ru

¹Национальный исследовательский университет "Высшая школа экономики", Москва

²ФГБОУ ВО "МИРЭА — Российский технологический университет"

³РЭУ имени Г. В. Плеханова, Москва

⁴ЗАО "ЕС-лизинг", Москва

Построение шкалы выявления мошеннической деятельности в сети Интернет с помощью машинного обучения¹

Повсеместная цифровизация общества и развитие информационных технологий способствуют увеличению методов взаимодействия между финансовыми организациями и потенциальными потребителями финансовых услуг. В то же время появление новых финансовых продуктов неизбежно ведет к росту угроз, а использование информационных технологий содействуют постоянному "совершенствованию" мошеннических схем и недобросовестному оказанию услуг, негативно влияющих как на финансовый рынок в целом, так и на его отдельных участников, таких как финансовые организации и их клиенты.

В связи с развитием современного общества большая часть финансовых операций, в том числе мошеннических, перешла в Интернет. При удаленном оказании услуг сложнее отследить и привлечь бенефициара к ответственности, однако способы пресечь мошенническую деятельность все равно есть, но они обусловлены высокими трудозатратами на мониторинг и анализ, потому что в сети Интернет расположены огромные объемы неструктурированной информации (BigData). В основе решения выявления нелегальной деятельности на финансовом рынке лежит разведка на основании открытых данных, применение методов машинного обучения и методов системного анализа.

В статье рассмотрены виды финансовых услуг, предоставляемых в сети Интернет, среди которых наиболее распространена мошенническая деятельность. Для выявления нелегальных финансовых услуг выделены и сгруппированы критерии в зависимости от вклада в принятие решения. Основным результатом исследования является построение шкалы комплексного индикатора, с помощью которого разработана математическая модель, основанная на выделенных критериях и методах машинного обучения, для выявления степени нелегальности финансовых услуг, оказываемых в сети Интернет.

Ключевые слова: противодействие недобросовестным практикам, финансовые услуги, интернет-мошенничество, машинное обучение, системный анализ, математическая модель, большие данные, комплексный индикатор, анализ открытых данных, мониторинг в сети Интернет

Введение

Главный надзорный орган РФ в сфере финансовых услуг — Центральный банк Российской Федерации, согласно Федеральному закону

"О Центральном банке Российской Федерации (Банке России)" от 10.07.2002 N 86-ФЗ (ред. от 11.06.2021), уполномочен осуществлять мониторинг за деятельностью организаций, предоставляющих финансовые услуги, и вести деятельность по противодействию недобросовестным практикам. Основными объектами мониторинга недобросовестных практик Центрального банка РФ являются:

¹ Статья подготовлена по материалам доклада на Седьмой Международной конференции "Актуальные проблемы системной и программной инженерии" АПСПИ 2021.

- 1) микрофинансовые организации;
- 2) форекс-брокеры;
- 3) финансовые пирамиды;
- 4) платежные системы;
- 5) операторы электронных денежных средств (электронные кошельки).

Всего за 2020 г. выявлено 1549 субъектов нелегальной финансовой деятельности, по аналитическим данным [1], на каждые два легальных субъекта приходится один нелегальный.

В настоящее время Центральным Банком РФ ведется активная работа по противодействию недобросовестным практикам. Однако развитие сферы интернет-услуг и продолжающаяся цифровизация экономики требуют модификации существующих методов выявления субъектов нелегальной деятельности: автоматизации процесса мониторинга интернет-пространства с помощью методов обработки неструктурированных данных в целях выявления субъектов финансовой деятельности, отличающихся наличием признаков недобросовестной деятельности. Для усовершенствования системы мониторинга необходимо автоматизировать все этапы — не только поиска информации, но также анализа и оценки вероятности нелегальности интернет-ресурсов для оптимизации затрат времени специалистов на принятие решений.

Современная практика такова, что необходимо исследование самих интернет-ресурсов, на которых в той или иной форме предлагаются финансовые услуги, так как не всегда есть возможность установления их связи с бенефициарами. Мошенники не указывают информацию о юридическом лице или указывают ложную. Это затрудняет выявление субъекта, оказывающего недобросовестные услуги, и привлечение его к ответственности. Перейти от поиска мошенников к поиску интернет-ресурсов без необходимости установления бенефициаров, ведущих нелегальную деятельность, возможно благодаря современным технологиям обработки неструктурированных данных, с помощью которых можно проводить автоматизированный мониторинг и информировать о факте оказания незаконных финансовых услуг. Этот подход позволит расширить область поиска недобросовестных финансовых практик, снизить риски вовлечения граждан и организаций в подобную деятельность и оперативно пресекать ее. В свою очередь собранная информация об интернет-ресурсе поможет в поиске и привлечении к ответственности нелегальных участников финансового рынка.

Для последующей информатизации мониторинга интернет-ресурсов поставлена цель ис-

следования — разработать модель шкалы оценки вероятности мошеннической или нелегальной деятельности субъектов, оказывающих финансовые услуги в сети Интернет.

В отличие от существующего процесса проверки деятельности юридических лиц в данной работе предложено следующее.

1. Систематически проводить автоматизированный мониторинг интернет-ресурсов в целях выявления фактов безлицензионного оказания финансовых услуг.

2. В качестве объекта исследования перейти с проверки деятельности юридических лиц, оказывающих финансовые услуги, на проверку нелегальности интернет-ресурсов, на которых оказываются данные услуги, что позволит расширить область поиска нелегальных участников финансового рынка.

3. Для оценки нелегальности финансовой деятельности использовать методы интеллектуальной обработки неструктурированных и слабоструктурированных данных.

Анализ предметной области и нормативных документов регулятора позволяет выявить критерии нелегальности, присущие каждому рассматриваемому виду финансовой деятельности. Исходя из анализа информации регулятора и научных исследований, сформирован перечень критериев нелегальности. С помощью методов машинного обучения оценена их важность в зависимости от вклада в принятие решения о нелегальности интернет-ресурса. На основе полученной системы критериев построена математическая модель комплексного индикатора для оценки вероятности нелегальной финансовой деятельности.

1. Анализ предметной области

В настоящее время финансовое мошенничество приобретает все больший размах в сети Интернет. Являясь принципиально новым явлением, финансовое интернет-мошенничество имеет характерные признаки, которые отличают его от других видов преступных действий:

- высокая степень латентности;
- многообразие способов совершения преступлений;
- транснациональность и глобальный характер деятельности;
- сложность привлечения к ответственности с юридической точки зрения [2].

Помимо известных и изученных видов интернет-мошенничества, таких как фишинг, краудфан-

динг, мошеннические письма на электронную почту и т. д. [3, 4], в последнее время набирает популярность направление оказания нелегальных финансовых услуг в сети Интернет. При таком мошенничестве создаются интернет-ресурсы, которые оказывают финансовые услуги без лицензии, что открывает им возможности для мошеннической деятельности. В отличие от прямого интернет-мошенничества, нелегальная финансовая деятельность не обязательно является мошеннической, однако отсутствие контроля со стороны надзорных органов предоставляет такую потенциальную возможность [5]. Несмотря на высокую степень освещенности вопроса интернет-мошенничества [2—4, 6, 7], рассматриваемое в данном исследовании направление, такое как оказание нелегальных финансовых услуг, остается малоизученным и слабо формализованным. Проблемой в устранении интернет-ресурсов, оказывающих подобные услуги, со стороны контрольных органов является отсутствие эффективного механизма автоматизированного мониторинга и выявления актов нелегального оказания финансовых услуг в сети Интернет.

Для анализа видов мошеннических финансовых услуг в сети Интернет в данном исследовании предлагается отталкиваться от классификации ресурсов, предложенной регулятором [5]. Авторами рассмотрены классы интернет-ресурсов в зависимости от видов финансовых услуг, в большей степени представляющих интерес для Центрального Банка РФ с точки зрения распространенности нелегальной деятельности в сети Интернет.

1. Микрофинансовая организация (МФО) — организация, занимающаяся выдачей микрозаймов физическим и юридическим лицам на короткий срок под высокий процент [8]. В сети Интернет размещены услуги большого числа МФО, однако большинство из них ведет свою деятельность без лицензии [9]. Они привлекают клиентов заниженными (по сравнению с легальными организациями) требованиями для выдачи займов. Помимо проблемы теневой экономики, не облагаемой налоговыми обязательствами, риски взаимодействия с подобными организациями несут и клиенты (непрозрачные условия, скрытые комиссии, невозврат залоговых ценностей).

2. Форекс-дилер (Forex) — это профессиональный участник финансового рынка, который заключает сделки от своего имени и за собственный счет с физическими лицами [10]. В настоящее время на территории Российской Федерации легально оказывают свою деятельность только четыре форекс-дилера, в то время как у тысяч других

форекс-дилеров отсутствуют лицензии на ведение такой деятельности [5, 11].

3. Финансовая пирамида (*High Yield Investment Program*, HYIP-проект) — это система, построенная на постоянном притоке денежных средств, поступающих извне от новых участников [12]. Подобие инвестиционных фондов, которые обещают высокую окупаемость, выплачивая прежним участникам денежные средства за счет поступлений от новых участников [13]. Финансовые пирамиды или HYIP-проекты запрещены на территории Российской Федерации [14]. Однако в сети Интернет данный вид сервисов маскируется под инвестиционные фонды и даже игры (ввод и вывод денежных средств происходит через виртуальные предметы), что не делает их легальными. Запрещенность данного вида финансовой деятельности, к сожалению, не отталкивает граждан, которые верят в высокую доходность HYIP-проектов.

4. Платежная система (ПС) — это связанная договорными отношениями общность юридических лиц, которые объединились в целях осуществления перевода денежных средств в электронной или физической форме [15, 16]. Платежные системы, предоставляющие свои услуги без лицензии Центрального банка РФ, зачастую привлекают клиентов более низкими комиссиями за переводы (или их полным отсутствием) по сравнению с легальными конкурентами [17].

5. Оператор электронных денежных средств (электронный кошелек, e-Wallet, ЭДС) — юридическое лицо, обеспечивающее перевод денежных средств в электронные эквиваленты для последующих операций над ними [15, 18]. Данный вид финансовых интернет-ресурсов тесно связан с платежными системами, зачастую они комбинируются (платежная система с виртуальным счетом).

Для обеспечения возможности оценки нелегальности деятельности необходимо определить критерии, указывающие на признаки мошеннической деятельности, применительно к каждому из рассмотренных видов финансовых услуг. Основным источником информации при определении перечня критериев нелегальности финансовой деятельности являются открытые материалы Центрального банка РФ, такие как [1, 19, 20] описывающие факторы развития мошеннической деятельности, классификацию ресурсов и критерии мошеннической деятельности. При определении критериев стоит отталкиваться от требований, предъявляемых Центральным банком РФ к содержанию информации на финансовых интернет-ресурсах, описанных в письме от 23 октября 2009 г. № 128-Т "О Рекомендациях по информационному

содержанию и организации Web-сайтов кредитных организаций в сети Интернет" [21].

Помимо официальных данных надзорного органа, проблема определения признаков мошенничества в финансовых услугах освещена в научных работах [13, 16, 22, 23]. В работе [16] авторами исследования рассматривается роль Центрального банка РФ в процессе предотвращения мошенничества на финансовом рынке, приводятся классификация видов нелегальной деятельности и их отличительные особенности. В работе [22] авторами рассматриваются виды нелегальных финансовых услуг, особое внимание уделяется критериям нелегальности микрофинансовых организаций и финансовых пирамид. В работах [13, 23] детально исследованы финансовые пирамиды, их разновидности и способы выявления.

В результате анализа предметной области выявлены виды финансовых услуг, в которых распространена мошенническая деятельность. Исследованы источники информации о видах финансовой деятельности, на основе которых можно сформировать перечень критериев, указывающих на нелегальность оказываемых услуг. Однако не все критерии имеют одинаковую значимость при принятии решения о нелегальности ресурса. В связи с этим далее в работе с помощью методов машинного обучения оценивается важность критериев в зависимости от вклада в принятие решения о недобросовестности оказываемых услуг и формируется математическая модель, основанная на системе критериев для оценки вероятности нелегальности деятельности на интернет-ресурсе.

2. Анализ критериев нелегальности с помощью методов машинного обучения

На основе проведенного анализа сформирован общий перечень из 37 критериев нелегальности финансовой деятельности. Однако *набор критериев*, применяемых к объекту исследования, которым является интернет-ресурс, зависит от вида оказываемых финансовых услуг.

Все критерии проверки нелегальности финансовой деятельности ресурсов в сети Интернет можно разделить на три типа [24]:

- 1) *прямые* — 8 критериев;
- 2) *косвенные* — 27 критериев;
- 3) *транзитивные* — 2 критерия.

Критерии первого типа однозначно и формально устанавливают факт нелегальности проверяемого юридического лица или индивидуально-предпринимателя (ИП). Например, к таким критериям относится проверка регистрации

юридического лица по ЕГРЮЛ/ЕГРИП¹ и наличие у организации лицензии на осуществление финансовой деятельности. Однако не всегда выполняются требования по размещению необходимой информации для установления связи между интернет-ресурсом и бенефициаром, соответственно, невозможно применить критерии первого типа.

Мошенники могут разместить на сайте заимствованные сведения о легальном участнике финансового рынка, что делает результат проверок недостоверным. В таких ситуациях целесообразно воспользоваться критериями, *косвенно* указывающими на недобросовестность ресурса. К ним может относиться, например, зарубежный хостинг сайта. К такому хостингу часто прибегают мошенники для затруднения процесса блокировки их деятельности.

Транзитивными являются критерии, напрямую не указывающие на нелегальность деятельности, однако от них зависит возможность срабатывания критериев другого типа. Например, при отсутствии реквизитов невозможна проверка соответствия деятельности юридического лица заявленным финансовым услугам.

Фрагмент перечня критериев нелегальности финансовой деятельности представлен в табл. 1.

Используя группы критериев, характеризующих интернет-ресурс, можно с определенной долей вероятности определить нелегальность его финансовой деятельности. Однозначное заключение по этому вопросу может дать только эксперт в предметной области, который на заключительном этапе проверяет наиболее приоритетные ресурсы из ограниченного списка. Поэтому можно считать, что рассматриваемая система критериев оценивания нелегальности оказываемых услуг выступает в качестве системы поддержки принятия решения экспертом. Для применения критериев необходима математическая модель, учитывающая типы и вес каждого критерия.

Для ранжирования критериев по их влиянию на вероятность нелегальности ресурса могут быть применены различные методы машинного обучения, классифицирующие ресурсы по их *набору критериев*: деревья решений, случайный лес. С помощью решающих правил, построенных для набора независимых переменных, вся выборка разбивается на узлы в виде древовидной иерархической структуры. При построении дерева решений независимые переменные ранжируются по

¹ ЕГРЮЛ — Единый государственный реестр юридических лиц; ЕГРИП — Единый государственный реестр индивидуальных предпринимателей.

Фрагмент выделенных критериев

Прямые	Проверка лицензии ЦБ РФ на наличие, подлинность и срок действия; проверка сведений о компании в ЕГРЮЛ; обещание сверхдоходности инвестиций; проверка вида деятельности организации по ОКВЭД*; предоставление займа по ставкам, отличающимся от нормали по рынку
Косвенные	Высокая доля негативных отзывов; учредитель/руководитель массовый регистратор; офшорная регистрация организации; наличие предварительных (специальных) взносов; указание на выгодные условия или условия, значительно отличающиеся от рыночных
Транзитивные	Отсутствие публикации учредительных документов; отсутствие контактной информации
* ОКВЭД — Общероссийский классификатор видов экономической деятельности.	

значимости независимых правил. С помощью этих методов предлагается отобрать важные критерии, значимые для определения целевых и нецелевых интернет-ресурсов.

В качестве зависимой переменной выступает разметка интернет-ресурсов на целевые и нецелевые. Параметрами в модели выступают значения критериев. С учетом значений 37 критериев нелегальности ресурса выбраны алгоритмы машинного обучения метода деревьев решений QUEST, CHAID. Эти алгоритмы были выбраны для решения, так как они удовлетворяют следующим условиям: в качестве зависимой можно использовать как качественные, так и количественные переменные, также преимущество методов в быстром вычислении для большой размерности данных.

Суть этих алгоритмов — разбиение, основанное на некоторой метрике (например, хи-квадрат), позволяющей для категориальной зависимой переменной (целевой/нецелевой интернет-ресурс) и нескольких независимых переменных (предикторов) выявить факторы, наилучшим образом объясняющие различия между категориями зависимой переменной (например, выделяет группы с наибольшим и наименьшим процентом целевых интернет-ресурсов). Наилучшее решение находят с помощью максимизации различий между группами путем перебора всех предикторов. Результатом является дерево классификации — набор последовательно выделенных подгрупп с наибольшими различиями целевой переменной (например, группы с максимальным и минимальным процентом целевых интернет-ресурсов). Это позволяет определить, сочетание каких признаков сильнее всего влияет на целевую переменную, а также найти наиболее перспективные целевые группы.

Математически можно сформулировать задачу следующим образом: решающее дерево $a(x)$ разбивает все пространство признаков на некоторое число непересекающихся подмножеств $\{J_1, \dots, J_n\}$, и в каждом подмножестве J_j выдает константный прогноз w_j :

$$a(x) = \sum_{j=1}^n w_j [x \in J_j].$$

В рамках данного исследования проверены прогностические возможности выбранных методов классификации для каждого типа ресурсов. В результате были определены оптимальные алгоритмы и оптимальные деревья решений для каждого типа ресурса.

Для решения поставленной задачи, в отличие от рассмотренных ранее работ, в данной работе предложен подход к построению комплексного индикатора оценки вероятности нелегальной деятельности интернет-ресурса. Для построения математической модели используют открытые сведения об интернет-ресурсе — контент и метаданные сайта. Контентом является текстовое содержимое, картинки и документы, представленные на интернет-ресурсе. Метаданные представляют собой информацию об интернет-ресурсе, например, доменное имя, IP-адрес, дату регистрации домена. Для оценки вероятности нелегальной деятельности собран и проанализирован большой объем данных из сети Интернет. Для применения критериев к найденным интернет-ресурсам использовали алгоритмические методы и методы машинного обучения. На основе полученной информации построена математическая модель шкалы комплексного индикатора.

3. Оценка вероятности нелегальной финансовой деятельности с помощью математической модели

Для расчета шкалы комплексного индикатора [25] оценки вероятности нелегальной деятельности предлагается математическая модель, объектом управления которой является интернет-ресурс. У каждого объекта выделяются четыре компоненты, основанные на трех группах критериев и результате прогноза нелегальности ресурса, полученного с помощью дерева решений.

Обозначим q_j^r — значение критерия r для объекта управления j . Если критерий сработал, то значение критерия равно 1, если нет — 0. Для каждого класса объекта, описанного в разд. 1, определен свой набор критериев. Например, если у объекта класс определен как МФО, то для него не рассчитываются критерии других классов. Всего таких классов объекта ($class_j$) пять:

- 1) МФО;
- 2) ЭДС;
- 3) Форекс;
- 4) Пирамида;
- 5) ПС.

Для каждого объекта управления формируется множество значений его компонент следующим образом:

Компонента 1 (k_1) показывает, что сработал любой критерий, явно указывающий на нелегальность объекта управления. Компонента k_1 принимает значение 1 при условии, если хотя бы одно из $q_j^r = 1$, где r входит во множество прямых критериев. В ином случае $k_1 = 0$.

Компонента 2 (k_2) указывает на уровень нелегальности объекта, основанный только на косвенных критериях, поэтому носит вспомогательный характер. Компонента k_2 определяется суммой сработавших значений критериев, входящих во множество косвенных в соответствии с классом объекта. Полученное значение нормализуется с помощью деления на максимальное значение суммы сработавших критериев:

$$k_2 = \frac{\sum_{r=1}^m q_j^r}{\max \sum_r q^r | class_j},$$

где m — общее число проверяемых критериев ($1 < m < 37$).

Следовательно, $0 \leq k_2 \leq 1$. Большое значение компоненты k_2 показывает, что данный объект нельзя оставлять без внимания эксперта. Применение этой компоненты необходимо, так как

в случае отсутствия прямых критериев нужно дать оценку объекту или подкрепить уверенность в нелегальности объекта, если сработали прямые критерии.

Компонента 3 (k_3) основана на срабатывании транзитивных критериев. Компонента также является вспомогательной и показывает, что отсутствует необходимая информация для применения прямых и косвенных критериев. Важная информация об объекте, такая как ИНН, ОГРН, наименование юр. лица и т. п. может быть не указана или специально скрываться, что делает объект подозрительным, так как невозможна точная автоматизированная проверка нелегальности и необходима экспертная оценка деятельности. Компонента k_3 принимает значение 1 при условии, если хотя бы одно из $q_j^r = 1$, где r входит во множество транзитивных критериев. Если не сработал ни один транзитивный критерий, то $k_3 = 0$.

Компонента 4 (k_4) представляет значение шкалы вероятности нелегальности объекта, основанное на методе деревьев решений, описанном в разд. 2. Компонента включает в себя все группы критериев в зависимости от определенного класса объекта. Шкала принимает два значения (0, 1) в зависимости от прогнозируемой вероятности (P) нелегальности и порогового значения, установленного для каждого класса объектов:

- низкая (0): $P(Y=1) \leq S$;
- высокая (1): $P(Y=1) > S$,

где S — пороговое значение для отсекающей вероятности нелегальности, определяемое на основе минимизации ошибок первого и второго рода; Y — признак нелегальности ($Y=1$) или легальности ($Y=0$) ресурса.

На основе полученных значений компонент определяется Z_j — значение комплексного индикатора вероятности нелегальности объекта управления:

$$Z_j = F(k_{1j}, k_{2j}, k_{3j}, k_{4j} | class_j = m), \text{ где } m = \{1:5\},$$

F — это функция, объединяющая разработанные компоненты, представленная линейной функцией с весами ω_i ,

$$Z_j = F(k_{ji}) = \sum_{i=1}^4 k_{ji} * \omega_i. \quad (1)$$

Предложенная концепция апробирована для оценки вероятности нелегальности 5619 интернет-ресурсов, найденных посредством тематических поисковых запросов. Интернет-ресурсы разделены на целевые (имеющие признаки нелегальной деятельности) и нецелевые (имеющие контент, от-

носящийся к данной тематике, но не имеющие признаков незаконной деятельности). Число таких интернет-ресурсов в зависимости от типа представлено в табл. 2.

Таблица 2

Число интернет-ресурсов

Тип ресурса	Нецелевые	Целевые
ЭДС	682	227
Форекс	214	168
МФО	1746	531
ПС	367	65
Пирамида	587	1032
ИТОГО	3596	2023

На основе проведенного анализа интернет-ресурсов рассчитаны четыре компонента, определены веса, отражающие соотношение важности компонент между собой, а также получена шкала комплексного индикатора, отражающая вероятность нелегальности ресурса.

Веса компонент в формуле (1):

- перед k_{1j} , $\omega_1 = 0,4$;
- перед k_{2j} , $\omega_2 = 0,2$;
- перед k_{3j} , $\omega_3 = 0,2$;
- перед k_{4j} , $\omega_4 = 0,2$.

Шкала комплексного индикатора вероятности нелегальности ресурса:

- 1) низкая: $Z_j < 0,2$;
- 2) средняя: $0,2 \leq Z_j < 0,6$;
- 3) высокая: $Z_j \geq 0,6$.

Высокое значение комплексного индикатора говорит о том, что на данные ресурсы следует обратить внимание эксперта в первую очередь. Среднее значение комплексного индикатора указывает на невозможность однозначного автоматизированного определения нелегальности ресурса, поэтому принятие решения остается за экспертом. При низком значении можно утверждать, что ресурс легален.

Общая ошибка классификации (ошибка первого рода) не превысила 20 % для всех типов ресурсов.

Примеры интернет-ресурсов с высокой вероятностью нелегальной деятельности приведены в табл. 3.

В табл. 4 представлено заключение специалиста в предметной области относительно нелегальности отобранных интернет-ресурсов по состоянию на июнь 2021 г. Как видно из данных таблицы, результаты апробации разработанной модели на выборке из трех интернет-ресурсов подтверждаются экспертной оценкой.

В зависимости от типа интернет-ресурса (МФО, Пирамида, Форекс, ПС и ЭДС) вероятность неле-

Таблица 3

Интернет-ресурсы с высокой вероятностью нелегальности

Интернет-ресурс, домен	Компонента 1	Компонента 2	Компонента 3	Компонента 4	Значение комплексного индикатора	Тип ресурса
wforex.com	1	0,333	1	1	3	Форекс
vk.com/dengi_sumom	1	0,111	1	1	3	Пирамида
instagram.com/kredit_online_kaz	1	0,0741	1	1	3	МФО

Таблица 4

Экспертное мнение специалиста относительно нелегальности выборки интернет-ресурсов

Интернет-ресурс, домен	Тип ресурса	Экспертное мнение
wforex.com	Форекс	На сайте представлен сервис по проведению конверсионных (обменных) операций на валютном рынке Forex. Деятельность осуществляется без лицензии Центрального банка РФ
vk.com/dengi_sumom	Пирамида	На странице представлены услуги компании Finiko, которая внесена Центральным банком РФ в список компаний с признаками финансовой пирамиды
instagram.com/kredit_online_kaz	МФО	На странице представлена микрофинансовая организация, осуществляющая выдачу кредитов физическим лицам без лицензии Банка России

гальности рассчитана различными алгоритмами. Для МФО используется алгоритм QUEST, для всех остальных типов - CHAID.

Выявлены различные критерии наиболее важные для классификации интернет-ресурсов:

- для ресурсов, относящихся к МФО, - предоставление займа по минимальному набору документов;
- для ресурсов, относящихся к Пирамидам, - обещание высокой доходности;
- для ресурсов, относящихся к Форекс, — отсутствие уведомления по рискам;
- для интернет-ресурсов, относящихся к платежным системам и электронным денежным средствам, - короткий период деятельности организации.

Одним из результатов моделирования является список проанализированных интернет-ресурсов, отсортированный по убыванию значений комплексного индикатора и характеризующий уменьшение вероятности нелегальности.

Заключение

Проблема интернет-мошенничества в сфере финансовых услуг в последние годы привлекает все больше внимания со стороны контролирующих органов. Реализуются различные подходы и методы для выявления таких ресурсов: контекстный поиск, отслеживание рекламных объявлений и др.

В работе предложена модель для выявления признаков нелегальности интернет-ресурса с помощью алгоритма построения комплексного индикатора. Предлагается рассчитывать комплексный индикатор как сумму четырех компонент. Компоненты индикатора основаны на критериях, полученных в процессе исследования предметной области.

Разработанная модель апробирована на 5619 интернет-ресурсах, выявленных в результате автоматизированного мониторинга, получены количественные оценки ошибок классификации, проанализированы отдельные результаты.

Преимуществом предлагаемого подхода является возможность осуществления автоматизированной проверки нелегальности интернет-ресурса в целях ускорения принятия решения специалистом в предметной области.

Предложенная на основе анализа открытых структурированных и неструктурированных данных концепция оценки вероятности нелегальности интернет-ресурса позволяет классифицировать ресурсы по степени признаков нелегальности

и впоследствии применить превентивные меры со стороны надзорного органа для снижения рисков вовлечения граждан и организаций в незаконную деятельность.

Список литературы

1. **Противодействие** недобросовестным практикам ЦБ РФ. Аналитика. URL: <https://www.cbr.ru/inside/analitics/>
2. **Никитина И. А.** Финансовое мошенничество в сети Интернет // Вестник Томского государственного университета. 2010. № 337. С. 122—124.
3. **Магомедов Ш. М.** Финансовое мошенничество в сети "Интернет" // Сборник материалов международной научно-практической конференции "Эффективность бизнеса в условиях международной нестабильности". 2017. С. 60—74.
4. **Harjot Kaur, Er. Prince Verma.** K-MLP Based Classifier for Discernment of Gratuitous Mails using N-Gram Filtration // International Journal of Computer Network and Information Security (IJCNIS). 2017. Vol. 9, N. 7. P. 45—58.
5. **Противодействие** недобросовестным практикам ЦБ РФ. URL: <https://cbr.ru/inside/>
6. **Багаутдинов Ф. Н., Хафизова Л. С.** Финансовое мошенничество (уголовно-правовой и криминологический аспекты противодействия). М.: Юрлитинформ, 2008. 280 с.
7. **Безверхов А.** Развитие понятия мошенничества в отечественном праве // Уголовное право. 2016. № 4. С. 9—12.
8. **О микрофинансовой деятельности и микрофинансовых организациях** // Федеральный закон от 2 июля 2010 года № 151-ФЗ 440. URL: <http://consultant.ru>
9. **Замалева Л. Р.** Развитие микрофинансовых организаций в кредитной системе России // Материалы IV Всероссийской научно-практической (заочной) конференции "Современные проблемы и перспективы развития банковского сектора". 2019. С. 32—38.
10. **О рынке ценных бумаг**: Федеральный закон Российской Федерации от 22.04.1996 № 39-ФЗ. URL: <http://consultant.ru>
11. **Гончаров И. Н.** Анализ деятельности форекс-дилеров как профессиональных участников рынка ценных бумаг (финансового рынка) // Материалы Седьмой международной научно-практической конференции "Развитие теории и практики управления социальными и экономическими системами". 2018. С. 15—19.
12. **Махова А. В., Нелипа А. В.** Финансовые пирамиды в современной российской экономике // Экономика и бизнес: теория и практика. 2020. № 4-2 (62). С. 147—151.
13. **Магомедова С. З., Казимагомедова А. С.** Сущность и социальная опасность финансовых пирамид // Экономические исследования и разработки. 2020. № 4. С. 171—174.
14. **УК РФ** Статья 172.2. Организация деятельности по привлечению денежных средств и (или) иного имущества (введена Федеральным законом от 30.03.2016 N 78-ФЗ). URL: <http://consultant.ru>
15. **О национальной** платежной системе: Федеральный закон от 27.06.2011 № 161-ФЗ. URL: <http://consultant.ru>
16. **Телюкова Ю. М.** Роль Центрального банка в профилактике финансовых пирамид // Сборник статей XIII Международного научно-исследовательского конкурса "Студент года 2020". 2020. С. 132—136.
17. **Хоменко Е. Г.** Платежные системы как элементы национальной платежной системы России и их классификация // Вестник университета имени О. Е. Кутафина (МГЮА). 2017. № 1 (29). С. 122—134.
18. **Дадаев Я. Э.** Электронные деньги в России: проблемы и перспективы развития // Вестник Чеченского государственного университета. 2018. № 4 (32). С. 34—38.
19. **Концепция** противодействия недобросовестным действиям на финансовом рынке. URL: https://www.cbr.ru/Content/Document/File/48603/concept_countersing_unfair_actions.pdf

20. **Микрофинансирование.** ЦБ РФ. URL: <https://www.cbr.ru/microfinance/>

21. **Письмо ЦБР** от 23 октября 2009 г. № 128-Т "О Рекомендациях по информационному содержанию и организации Web-сайтов кредитных организаций в сети Интернет" URL: <https://www.garant.ru/products/ipo/prime/doc/489901/>

22. **Путинцева Е. Э., Широкова О. В.** Пресечение нелегальной деятельности на финансовом рынке // Устойчивое развитие науки и образования. 2018. № 12. С. 29—32.

23. **Хорунин А. Ю., Фильчакова Н. Ю., Рудченко Н. П.** "Финансовые пирамиды": к вопросу об актуальности и со-

циальной опасности явления в современности // Финансовые исследования. 2018. № 4 (61). С. 88—93.

24. **Кочнев А. А.** Построение алгоритма системы мониторинга и выявления нелегальных финансовых услуг в сети Интернет // Информационные технологии в государственном управлении. Цифровая трансформация и человеческий капитал: сб. науч. тр. 19-й науч.-практ. конф. М.: Проспект, 2021. С. 74—81.

25. **Богданова Т. К., Жукова Л. В.** Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных // Бизнес-информатика. 2021. Т. 15, № 2. С. 21—33.

Building the Scale for Fraud Detection on the Internet Using ML

L. V. Zhukova^{1,4}, lvzhukova@hse.ru, **I. M. Kovalchuk**^{2,4}, ikovalchuk@ec-leasing.ru, **A. A. Kochnev**^{3,4}, akochnev@ec-leasing.ru, **V. R. Chugunov**⁴, vchugunov@ec-leasing.ru,

¹ National Research University Higher School of Economics, Moscow, 101000, Russian Federation

² MIREA — Russian Technological University, Moscow, Russian Federation

³ PLEKHANOV Russian University of Economics, Moscow, Russian Federation

⁴ JSC "EC-LEASING" CO., Moscow, 117587, Russian Federation

Corresponding author:

Zhukova Ludmila V., Senior Lecturer, National Research University Higher School of Economics, Moscow, 101000, Russian Federation, Leading Analyst, JSC "EC-LEASING" Co., Moscow, 117587, Russian Federation
E-mail: lvzhukova@hse.ru

Received on November 08, 2021

Accepted on March 31, 2022

The widespread digitalization of the modern society and the development of information technology have increased the number of methods of interaction between financial institutions and potential consumers of financial services. At the same time, the advent of new financial products inevitably leads to an increase in threats, and the use of information technology facilitates the continuous "improvement" of fraudulent schemes and unfair services that negatively impact both the financial market as a whole and its individual participants, such as financial institutions and their clients.

Due to the development of modern society, most financial transactions have moved to the Internet, including the fraudulent ones. When services are provided remotely, it is more difficult to trace and prosecute the beneficiary, but there are still ways to stop fraudulent activity. They can be characterized as labour-consuming, as the monitoring and analysis of huge amounts of unstructured information (BigData) located on the Internet take great amount of time and effort. The solution to detecting illegal activity in the financial market is based on open data intelligence, the application of machine learning methods and systems analysis techniques.

The article examines the types of financial services provided on the Internet, among which fraudulent activities are most common. In order to identify illegal financial services, criteria are identified and grouped according to their contribution to the decision-making process. The main result of the study is the construction of the scale of a complex indicator, which is used to develop a mathematical model based on the selected criteria and machine learning methods to identify the extent of illegality of financial services provided on the Internet.

Keywords: counteraction to unfair practices, financial services, Internet fraud, machine learning, system analysis, mathematical model, big data, complex indicator, open data analysis, Internet monitoring

For citation:

Zhukova L. V., Kovalchuk I. M., Kochnev A. A., Chugunov V. R. Building the Scale for Fraud Detection on the Internet Using ML, *Programmnaya Ingeneriya*, 2022, vol. 13, no. 5, pp. 247—256.

DOI: 10.17587/prin.13.247-256

References

1. **Protivodejstvie** nedobrosovestny'm praktikam CзB RF. Analitika, available at: <https://www.cbr.ru/inside/analitics/> (in Russian).
2. **Nikitina I. A.** Finansovoe moshennichestvo v seti Internet, *Vestnik Tomskogo gosudarstvennogo universiteta*, 2010, no. 337, pp. 122–124 (in Russian).
3. **Magomedov Sh. M.** Finansovoe moshennichestvo v seti "Internet", *Sbornik materialov mezhdunarodnoj nauchno-prakticheskoy konferencii "E'ffektivnost' biznesa v usloviyax mezhdunarodnoj nestabil'nosti"*, 2017, pp. 60–74 (in Russian).
4. **Harjot Kaur, Er. Prince Verma.** K-MLP Based Classifier for Discernment of Gratuitous Mails using N-Gram Filtration, *International Journal of Computer Network and Information Security (IJCNIS)*, 2017, vol. 9, no. 7, pp. 45–58.
5. **Protivodejstvie** nedobrosovestny'm praktikam CзB RF, available at: <https://cbr.ru/inside/> (in Russian).
6. **Bagautdinov F. N., Hafizova L. S.** *Finansovoe moshennichestvo (ugolovno-pravovoj i kriminologicheskij aspekt) protivodejstviya*, Moscow, YurLitinform, 2008, 280 p. (in Russian).
7. **Bezverhov A.** Razvitie ponyatiya moshennichestva v otechestvennom prave, *Ugolovnoe parvo*, 2016, no. 4, pp. 9–12. (in Russian).
8. **O mikrofinansovoj** deyatel'nosti i mikrofinansovy'x organizacijax" Federal'ny'j zakon ot 2 iyulya 2010 goda № 151-FZ 440, available at: <http://consultant.ru> (in Russian).
9. **Zamaleeva L. R.** Razvitie mikrofinansovy'x organizacij v kreditnoj sisteme Rossii, *Materialy IV Vserossijskoj nauchno-prakticheskoy (zaочноj) konferencii "Sovremennyye problemy i perspektivy razvitiya bankovskogo sektora"*, 2019, pp. 32–38 (in Russian).
10. **O ry'nke cenny'x bumag:** Federal'ny'j zakon Rossijskoj Federacii ot 22.04.1996 № 39-FZ, available at: <http://consultant.ru> (in Russian).
11. **Goncharov I. N.** Analiz deyatel'nosti foreks-dilerov kak professional'ny'x uchastnikov ry'nka cenny'x bumag (finansovogo ry'nka), *Materialy Sed'moj mezhdunarodnoj nauchno-prakticheskoy konferencii "Razvitie teorii i praktiki upravleniya social'ny'mi i e'konomicheskimi sistemami"*, 2018, pp. 15–19 (in Russian).
12. **Maxova A. V., Nelipa A. V.** Finansovy'e piramidy v sovremennoj rossijskoj e'konomie, *E'konomika i biznes: teoriya i praktika*, 2020, no. 4-2 (62), pp. 147–151 (in Russian).
13. **Magomedova S. Z., Kazimagomedova A. S.** Sushhnost' i social'naya opasnost' finansovy'x pyramid, *E'konomicheskie issledovaniya i razrabotki*, 2020, no. 4, pp. 171–174 (in Russian).
14. **UK RF Stat'ya 172.2.** Organizaciya deyatel'nosti po privlecheniyu denezhny'x sredstv i (ili) inogo imushhestva (vvedena Federal'ny'm zakonom ot 30.03.2016 N 78-FZ), available at: <http://consultant.ru> (in Russian).
15. **O nacional'noj** platelyzhnoj sisteme: Feder. zakon ot 27.06.2011 № 161-FZ, available at: <http://consultant.ru> (in Russian).
16. **Telyukova Yu. M.** Rol' Central'nogo banka v profilaktike finansovy'x pyramid, *Sbornik statej XIII Mezhdunarodnogo nauchno-issledovatel'skogo konkursa "Student goda 2020"*, 2020, pp. 132–136 (in Russian).
17. **Xomenko E. G.** Platelyzhny'e sistemy kak e'lementy nacional'noj platelyzhnoj sistemy Rossii i ix klassifikaciya, *Vestnik universiteta imeni O. E. Kutafina (MGYuA)*, 2017, no. 1(29), pp. 122–134 (in Russian).
18. **Dadaev Ya. E.** E'lektronny'e den'gi v Rossii: problemy i perspektivy razvitiya, *Vestnik Chechenskogo gosudarstvennogo universiteta*, 2018, no. 4 (32), pp. 34–38 (in Russian).
19. **Koncepciya** protivodejstviya nedobrosovestny'm dejstviyam na finansovom ry'nke, available at: https://www.cbr.ru/Content/Document/File/48603/concept_countersing_unfair_actions.pdf (in Russian).
20. **Mikrofinansirovanie.** CзB RF, available at: <https://www.cbr.ru/microfinance/> (in Russian).
21. **Pis'mo CзBR** ot 23 oktyabrya 2009 g. No. 128-T "O Rekomendaciyax po informacionnomu soderzhaniyu i organizacii Websajtov kreditny'x organizacij v seti Internet", available at: <https://www.garant.ru/products/ipo/prime/doc/489901/> (in Russian).
22. **Putinceva E. E., Shirokova O. V.** Presechenie nelegal'noj deyatel'nosti na finansovom ry'nke, *Ustojchivoe razvitie nauki i obrazovaniya*, 2018, no. 12, pp. 29–32 (in Russian).
23. **Horunin A. Yu., Fil'chakova N. Yu., Rudchenko N. P.** "Finansovy'e piramidy": k voprosu ob aktual'nosti i social'noj opasnosti yavleniya v sovremennosti, *Finansovy'e issledovaniya*, 2018, no. 4 (61), pp. 88–93 (in Russian).
24. **Kochnev A. A.** Postroenie algoritma sistemy monitoringa i vy'yavleniya nelegal'ny'x finansovy'x uslug v seti Internet. *Informacionny'e tekhnologii v gosudarstvennom upravlenii. Cifrovaya transformaciya i chelovecheskij kapital: sbornik nauchny'x trudov 19-j nauch.-prakt. konf.*, Moscow, Prospekt, 2021, pp. 74–81 (in Russian).
25. **Bogdanova T. K., Zhukova L. V.** Ocenka sostoyaniya ob'ekta upravleniya na osnove universal'nogo kompleksnogo indikatora s ispol'zovaniem strukturirovanny'x i nestrukturirovanny'x danny'x, *Biznes-informatika*, 2021, vol. 15, no. 2, pp. 21–33 (in Russian).

ООО "Издательство "Новые технологии". 107076, Москва, ул. Матросская Тишина, д. 23, стр. 2
Технический редактор *Е. М. Патрушева*. Корректор *А. В. Чугунова*.

Сдано в набор 31.03.2022 г. Подписано в печать 28.04.2022 г. Формат 60×88 1/8. Заказ PI521
Цена свободная.

Оригинал-макет ООО "Авансд солишнз". Отпечатано в ООО "Авансд солишнз".
119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: www.aov.ru