**К вопросу о совершенствовании системы статистических показателей безопасности применения цифровых технологий**

*Использование информационно-коммуникационных технологий в условиях современной реальности ведет не только к получению известных положительных социально-экономических эффектов, но и предполагает неизбежное возникновение связанных с этим негативных последствий.**С вопросами защиты интересов личности, общества и государства от внутренних и внешних информационных угроз связано понятие информационной безопасности. Расширение ее границ – от низкоуровневых технических мероприятий по защите информационных систем (компьютерной безопасности и кибербезопасности) до стратегического управления экономическим производством в данном контексте – привело к появлению более релевантной категории: цифровой безопасности* *– и диктует целесообразность объединения множества этих терминов в обобщающее понятие «безопасность применения цифровых технологий».*

*Вопросы безопасности применения цифровых технологий все чаще становятся предметом статистических исследований, научная и практическая ценность которых напрямую зависит от степени актуальности используемой системы показателей. Цель данной статьи состоит в том, чтобы проанализировать существующую систему статистических показателей и обозначить перспективы по ее совершенствованию. Для достижения поставленной цели в процессе анализа были использованы методы дескриптивной статистики, корреляционно-регрессионный анализ, а также методы машинного обучения (в частности, различные варианты алгоритмов классификации на основе деревьев решений).*

*Результаты обзора современных методологических разработок, применяемых на международном и национальном уровнях, позволили сделать вывод о том, что исчисляемые статистические показатели способны количественно оценивать самые разнообразные аспекты безопасности применения цифровых технологий. Вместе с тем, вполне объяснимая «автономность» этих методологий существенно затрудняет создание комплексного представления об измеряемых процессах, например, их позиции в общей структуре и роли в возможных причинно-следственных связях. С целью заполнения этого методологического проблема в работе предложен подход к динамическому формированию системы статистических показателей путем включения в нее новых индикаторов, фиксирующих состояние этапов жизненного цикла процесса нарушения безопасности применения цифровых технологий. Сам процесс нарушения безопасности рассматривается при этом как последовательность «субъект – мотив – объект – метод – последствия» с подразделением по направлениям (технологическое, социальное, экономическое) и масштабу охвата (микро-, мезо-, макро-).*

*Поиск информационных и методических источников, имеющих опосредованное отношение к статистическим (инспектировалась широко известная веб-платформа Kaggle, объединяющая исследователей в области науки о данных и машинного обучения), предоставил возможность провести анализ результатов тематического опроса компаний. В итоге моделирования влияния безопасности применения цифровых технологий на эффективность экономической деятельности было установлено, что вариация фактора безопасности примерно на треть определяет уровень экономической эффективности компании в рамках представленной совокупности. Использование методов машинного обучения позволило получить предсказательные модели с приемлемым уровнем качества классификации компаний на эффективные и неэффективные в том же факторном контексте.*

*С практической точки зрения предложенный в работе подход к формированию системы статистических показателей будет полезен при анализе и администрировании процессов нарушения безопасности применения цифровых технологий в масштабе предприятий, регионов и страны в целом.*

*Ключевые слова*: информационно-коммуникационные технологии, информационная безопасность, безопасность применения цифровых технологий, система статистических показателей

**Обзор существующих официальных методологий**

Цифровая трансформация, при всех ее уже претворенных в жизнь или еще только ожидаемых позитивных социально-экономических эффектах, неизбежно создает информационно-технологические неопределенности и уязвимости, способные представлять собой потенциальную угрозу интересам общества, бизнеса и государства, ставить под удар их информационную безопасность.

Система официальных взглядов и основные положения в этой сфере отражены в Доктрине информационной безопасности РФ, Федеральном законе «Об информации, информационных технологиях и о защите информации», семействе стандартов (например, ГОСТ Р ИСО/МЭК 27000-2012 «Информационная технология. Методы и средства обеспечения безопасности. Системы менеджмента информационной безопасности») и других тематических источниках. На международном уровне изучением проблематики информационной безопасности занимается ряд международных организаций, каждая из которых действует в определенном профильном секторе: так, Организация экономического сотрудничества и развития (ОЭСР, OECD) акцентирует внимание на социальных и экономических аспектах; разработка технических стандартов возложена, в частности, на Международную организацию по стандартизации (ИСО, ISO) и Международную электротехническую комиссию (МЭК, IEC); вопросами киберпреступности занимаются Совет Европы, Управление Организации Объединенных Наций по наркотикам и преступности (UNODC) и Интерпол.

Сегодня в мире отмечается четкая тенденция к инверсии – если ранее информационный контур безопасности ведения бизнеса очерчивали технические специалисты, то теперь политику информационной безопасности определяют представители высшего менеджмента исходя из оценки рисков, прежде всего, экономического характера. В этом контексте привычная категория «информационная безопасность» фактически трансформировалась в безопасность цифровую, что вполне явно прослеживается в тематических публикациях Организации экономического сотрудничества и развития: “OECD Guidelines for information security” (2002), “OECD Digital Security Risk Management” (2015), “OECD Policy Framework on Digital Security” (2022). Последняя из перечисленных публикаций содержит свод рекомендаций по управлению цифровой безопасностью применительно к разным уровням детализации: от базисного, интуитивно понятного, отражающего социально-экономические эффекты и до самого продвинутого, затрагивающего технические аспекты цифровых технологий. Структура политики такого управления пластична и в последней версии (2022 г.) представлена на рис. 1.



Рис. 1. **Уровни иерархической структуры политики цифровой безопасности OECD**

Источник: [1, p. 7].

Рекомендации ОЭСР, если рассматривать их в качестве методологических основ для количественного анализа процессов цифровой безопасности, обладают весьма заметным недостатком: они лишь декларируют направления деятельности, но не предлагают механизм практической реализации (в отличие, например, от разработок этой международной организации в области методологии измерения информационного общества).

Международный союз электросвязи (МСЭ, ITU) – одна из старейших в мире организаций, действующих сегодня под эгидой Организации объединенных наций – в некоторой степени исправила этот пробел, разработав (совместно с компанией ABI Research) Глобальный индекс кибербезопасности (Global Cybersecurity Index, GCI) и впервые опубликовав его в 2015 году [3]. Этот интегральный показатель отражает успехи стран мира на пути к достижению целей по основным направлениям кибербезопасности (см. таблицу 1).

Таблица 1

**Компоненты Глобального индекса кибербезопасности**

|  |  |
| --- | --- |
| Наименование | Содержание |
| Правовые меры | Оценка наличия правовых институтов и структур, занимающихся вопросами кибербезопасности и киберпреступности |
| Технические меры | Оценка наличия технических институтов и платформ, занимающихся кибербезопасностью |
| Организационные меры | Оценка наличия институтов координации политики и стратегий развития кибербезопасности на национальном уровне |
| Меры по наращиванию потенциала | Оценка наличия программ исследований и разработок, образования и обучения, сертифицированных специалистов и агентств государственного сектора, способствующих наращиванию потенциала кибербезопасности |
| Меры сотрудничества | Оценка наличия партнерств, механизмов сотрудничества и сетей обмена информацией в области кибербезопасности |

В 2020 году (последняя публикация Индекса пришлась на 2021 г. [4], а следующая ожидается в 2023 г.) в состав проиндексированной совокупности вошло 169 государств, где Российская Федерация, набрав 98,06 балла из 100, разделила 5-е место с Объединенными Арабскими Эмиратами и Малайзией в общем рейтинге и единолично заняла 1-е место в рейтинге стран Содружества Независимых Государств. Специализированная анкета для построения Индекса включает ряд вопросов по каждой из сформулированных мер и предполагает получение ответов с использованием профильных компетенций в части законов и нормативных актов, деятельности общественных организаций, научных школ и центров разработок, а также других аспектов из области кибербезопасности. Полученные данные обобщаются в показатели, которым на основе оценок специально привлекаемых экспертов присваиваются определенные веса в зависимости от важности их вклада, после чего исчисляется сам Индекс путем осреднения показателей на арифметической или геометрической основе. К сожалению для статистики, Глобальный индекс кибербезопасности не несет в себе каких-либо данных о количественной стороне деятельности участников процесса создания киберугроз и их отражения, их мотивах и последствиях.

Очевидно, что приведенные выше методологические разработки в области информационной (цифровой, кибер-) безопасности носят более концептуальный, чем практический характер, и в поисках конкретики следует обратиться к международным и национальным органам статистики. Учитывая существенный уровень гармонизации существующей статистической методологии и во избежание повторяемости показателей, логично будет в первом случае рассмотреть результаты обследований организаций (например, стран Евросоюза), а во втором – населения Российской Федерации.

Статистическая служба Европейского союза (Евростат, Eurostat) ежегодно (с 2002 г.) проводит и публикует результаты опроса сообщества по теме использования информационно-коммуникационных технологий (ИКТ) и электронной коммерции на предприятиях. Наиболее актуальные статистические данные были получены в результате обследования предприятий, проведенных Национальными статистическими органами стран-членов Евросоюза в первые месяцы 2022 года: было опрошено около 150400 предприятий с 10 и более сотрудниками или самозанятыми лицами из 1,47 млн предприятий в ЕС.

Некоторые дескриптивные статистики совокупности стран ЕС по ряду показателей информационной безопасности предприятий [4] приведены в таблице 2. Как отмечают евростатистики, в 2022 году 92 % предприятий ЕС с 10 и более сотрудниками или самозанятыми лицами использовали хотя бы одну меру для обеспечения целостности, доступности и конфиденциальности данных и систем ИКТ. Более чем каждое третье предприятие (37 %) сообщило о наличии документов, устанавливающих меры, практику или процедуры по обеспечению безопасности ИКТ. На каждом четвертом предприятии (24 %) эти документы были определены или рассмотрены в течение последних 12 месяцев. Каждое четвертое предприятие (25 %) было застраховано от инцидентов безопасности ИКТ. Наконец, в 2022 году более чем каждое пятое предприятие (22 %) столкнулось с последствиями инцидентов безопасности, связанных с ИКТ. Следует отметить, что значения коэффициента вариации (*KV*) с его пороговым значением в 33 процента характеризуют совокупность стран Евросоюза как «пограничную» в оценке ее однородности по большинству показателей.

Таблица 2

**Информационная безопасность предприятий в странах Европейского союза**

(в процентах от обследуемых предприятий)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Показатели | *X min* |  | *X max* | *Медиана* | *KV, %* |
| Используют как минимум одну меру безопасности ИКТ | 62 | 88,0 | 98 | 90 | 8,6 |
| Информируют работников об их обязанностях в области безопасности ИКТ | 32 | 57,0 | 75 | 60 | 18,6 |
| Имеют разработанные инструкции по мерам, практикам или процедурам по безопасности ИКТ | 15 | 37,8 | 68 | 36 | 34,1 |
| Имеют страховку от инцидентов, связанных с ИКТ | 4 | 20,8 | 71 | 14 | 75,0 |
| Составили или проверили существующие инструкции по безопасности ИКТ предприятия в течение последних 12 месяцев | 7 | 24,9 | 43 | 21 | 38,2 |
| Пережили инциденты безопасности, связанные с ИКТ, имевшие некоторые последствия в отчетном (2021) году | 11 | 20,5 | 44 | 19 | 35,6 |

Министерство цифрового развития, связи и массовых коммуникаций РФ (Минцифры России) в рамках реализации федерального проекта «Информационная безопасность» национальной программы «Цифровая экономика Российской Федерации» осуществляет разработку методик расчета тематических показателей, находящих отражение в публикациях об использовании информационных технологий населением и организациями. Федеральная служба государственной статистики (Росстат) в сотрудничестве с Минцифры России и Национальным исследовательским университетом «Высшая школа экономики» (НИУ ВШЭ) регулярно публикует статистические сборники, посвященные информационному обществу и цифровой экономике. Результаты проведенного Росстатом в 2022 г. Выборочного федерального статистического наблюдения по вопросам использования населением информационных технологий и информационно-телекоммуникационных сетей содержат, в частности, данные об инцидентах информационной безопасности и средствах защиты, применяемых для их предотвращения (табл. 3 и 4).

Таблица 3

**Население, столкнувшееся с проблемами информационной безопасности, по субъектам Российской Федерации**

(в процентах от обследуемых лиц в возрасте от 15 до 74 лет)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Показатели | *X min* |  | *X max* | *Медиана* | *KV, %* |
| *Сталкивались с такими проблемами:* |  |  |  |  |  |
| Заражение вирусом, что привело к потере информации и/или времени на их удаление | 0,7 | 5,1 | 22,0 | 4,6 | 76,3 |
| Несанкционированный доступ к компьютеру (информационным ресурсам, информационным системам) | 0,1 | 1,7 | 19,2 | 1,2 | 142,8 |
| Несанкционированная рассылка (спам) | 3,8 | 23,5 | 81,9 | 21,5 | 55,0 |
| Получение по электронной почте мошеннических писем с просьбой выслать персональные данные | 0,1 | 2,6 | 9,2 | 2,0 | 84,6 |
| Перенаправление на фальшивые сайты с просьбой указать персональные данные | 0,0 | 1,9 | 8,0 | 1,4 | 88,1 |
| Посещение детьми нежелательных сайтов, контакты детей с потенциально опасными людьми через сеть Интернет | 0,0 | 0,6 | 3,3 | 0,4 | 98,4 |
| Хищение денежных средств или персональных данных | 0,0 | 0,5 | 5,7 | 0,3 | 162,3 |
| Использование мобильного телефона неизвестными лицами | 0,0 | 1,7 | 39,2 | 0,4 | 320,3 |
| Использование электронной почты неизвестными лицами | 0,0 | 1,0 | 28,6 | 0,3 | 369,1 |
| Другие проблемы информационной безопасности | 0,2 | 3,5 | 2,4 | 20,8 | 111,4 |
| *Не сталкивались с проблемами информационной безопасности* | 12,8 | 69,4 | 92,2 | 70,9 | 20,5 |

Таблица 4

**Население, использовавшее средства защиты информации, по субъектам Российской Федерации**

(в процентах от обследуемых лиц в возрасте от 15 до 74 лет)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Показатели | *X min* |  | *X max* | *Медиана* | *KV, %* |
| *Использовали – всего:* | 36,1 | 71,7 | 98,8 | 73,0 | 16,1 |
| *из них:* |  |  |  |  |  |
| Антивирусные средства | 29,2 | 68,5 | 98,8 | 69,9 | 18,1 |
| Антиспамовые фильтры | 1,7 | 15,2 | 13,6 | 53,9 | 58,4 |
| Средства родительского контроля или фильтрации Интернет-ресурсов | 0, | 2,1 | 1,7 | 9,9 | 87,8 |
| Другие средства защиты | 0,1 | 2,1 | 1,1 | 12,4 | 121,4 |
| *Не используют средства защиты* | 0,4 | 18,3 | 16,9 | 60,0 | 52,0 |
| *Затруднились ответить* | 0,6 | 10,1 | 10,2 | 28,4 | 61,2 |

Явно прослеживаемая в таблицах высокая степень неоднородности совокупности регионов России по большинству показателей указывает на наличие типических региональных групп по признакам интенсивности и номенклатуры угроз информационной безопасности: где-то население постоянно сталкивается с такого рода инцидентами, где-то – лишь изредка.

Проведенный обзор используемых в настоящее время официальных методологий профильных и статистических организаций выявил ряд проблем, среди которых множественность понятия информационной безопасности и фрагментарность структуры ее индикаторов. Это обстоятельство в рамках исследования требует выполнения ряда действий. Во-первых, во избежание несогласованности в дефинициях и для упрощения понимания предметной области будет логичным считать вышеуказанный набор понятий безопасности (информационная, информационно-коммуникационная, компьютерная, кибер-, цифровая) нестрогими синонимами и объединить в общий термин «безопасность применения цифровых технологий». Во-вторых, в целях повышения качества и развития существующей системы статистических показателей представляется верным четко определить ее структуру, а также увеличить «мощность» этой системы путем включения в нее новых индикаторов.

**Формирование динамической системы показателей**

Необходимым условием получения научно обоснованных результатов оценки и анализа объекта статистического исследования служит наличие соответствующей системы показателей, представляющей собой «комплекс взаимосвязанных и расположенных в логической последовательности показателей, всесторонне характеризующих состояние и развитие массовых явлений общественной жизни» [10, с. 21]. Большое значение теории статистических показателей придавала еще советская школа статистики: «Система, совокупность всегда более значима, чем сумма отдельных частей, так как помимо информации о частях она несет информацию о том новом, что появляется в результате взаимодействия частей, информацию о развитии системы в целом. Рассматриваемое положение математиками в общем виде формулируется так: функция системы больше суммы функций составляющих ее частей, т. е. *f(x, y) > f (x) + f(y)»* [11, с. 226]. При этом, согласно положениям диалектического материализма, справедливо отмечалось, что «системы статистических показателей нельзя представлять стабильными, они изменяются вслед за изменениями отражаемой ими действительности» [там же, с. 228].

Современными отечественными учеными-статистиками предлагается (например, в [12, c. 80]) осуществлять формирование системы статистических показателей, руководствуясь рядом принципов, среди которых фигурируют такие как: *принцип системного подхода* (в совокупности разнообразные статистические показатели полностью описывают объект исследования), *принцип информативности при минимизации числа статистических показателей* (показатели должны быть максимально информативны, при этом их количество должно быть минимально) и *принцип количественной определенности оценки* (для показателей должна быть определена количественная оценка, они должны иметь эталонное или нормативное значение, должен быть известен диапазон принимаемых значений); однако же, разумеется, далеко не всегда эти принципы могут быть практически выполнимы.

В данном теоретическом контексте очевидно, что система статистических показателей безопасности применения цифровых технологий – в силу высокой степени динамики процессов в этой сфере – не может являться косной структурой, а должна, в первую очередь, отражать логику особенностей этого развития. Такой подход выражен на рис. 2, где концептуальная схема формирования системы показателей содержит два компонента, отображающих аспекты оценки и анализа: характерный, учитывающий особенности предметной области – «Процесс», и общий («Направление» и «Масштаб»), свойственный для логики построения систем показателей в целом.



Рис. 2. **Концептуальная схема формирования системы статистических показателей безопасности применения цифровых технологий**

В результате состав показателей системы формируется динамически при сохранении общей структуры, основанной на логике взаимосвязи аспектов (пример практической реализации схемы представлен в таблице 5).

Таблица 5

**Примеры индикаторов-представителей системы статистических показателей безопасности применения цифровых технологий**

|  |  |  |
| --- | --- | --- |
| Направление | Показатели по звеньям процесса | Масштаб |
|  | ***Субъект*** |  |
| Технологич. | Соотношение числа инцидентов, связанных с внутренними и внешними для организации источниками киберугроз | Микро |
| Экономич. | Число сотрудников по информационной безопасности в штате организации | Микро |
| Социологич. | Доля домохозяйств, пострадавших от злонамеренных действий неизвестных лиц через социальные сети и онлайн-сервисы | Макро |
|  | ***Мотив*** |  |
| Технологич. | Доля инцидентов, связанных с кибермошенничеством | Микро |
| Технологич. | Число инцидентов, связанных с кибертерроризмом | Макро |
| Экономич. | Доля организаций, столкнувшихся с вымогательством посредством использования цифровых технологий | Мезо  Макро |
| Социологич. | Доля лиц в домохозяйстве, сталкивавшихся с угрозами разглашения личной информации с целью вымогательства | Микро |
|  | ***Объект*** |  |
| Технологич. | Число инцидентов, связанных с попытками проникновения внутрь периметра информационной системы организации | Микро |
| Технологич. | Доля инцидентов в центрах хранения данных, связанных с кибербезопасностью | Макро |
| Экономич. | Структура внутренних информационных ресурсов организации | Микро |
| Экономич. | Доля организаций, имеющих полнофункциональный сайт с административной частью | Мезо  Макро |
| Социологич. | Число учетных записей в социальных сетях в среднем на одного члена домохозяйства | Микро |
|  | ***Метод*** |  |
| Технологич. | Число инцидентов, связанных с отказом в обслуживании запросов клиентов сайта организации в связи с кибератаками | Микро |
| Технологич. | Число зарегистрированных вредоносных программ по видам | Макро |
| Экономич. | Доля организаций, подвергшихся кибератакам по видам | Макро |
| Социологич. | Доля домохозяйств, столкнувшихся с кибермошенничеством с использованием методов социальной инженерии | Макро |
|  | ***Последствия*** |  |
| Технологич. | Объем утраченных данных организации вследствие кибератаки | Микро |
| Экономич. | Величина ущерба по причине кибератаки на сайт организации | Микро |
| Экономич. | Величина финансовых потерь в результате кибератак | Макро |
| Социологич. | Периодичность потери контроля над учетными записями | Микро |
| Социологич. | Доля домохозяйств, отказавшихся от использования цифровых услуг правительства и бизнеса в целях защиты информации | Мезо  Макро |

Примечательным моментом перечня является ряд показателей, весьма репрезентативно представляющих технологический уровень предметной области и оперирующих термином «инцидент». Под инцидентом информационной безопасности (ИБ) понимается появление одного или нескольких нежелательных или неожиданных событий ИБ, с которыми связана значительная вероятность компрометации бизнес-операций и создания угроз ИБ. Инциденты имеют широкую классификацию, но наиболее актуально выделять две их группы: преднамеренные, включающие в себя весь набор способов и методов создания угроз ИБ (фишинг, брутфорс, программы-вымогатели, черви, трояны и т.п.), и случайные, вызванные ошибками пользователей, нелицензированным программным обеспечением и пр.

В этой связи оценку и анализ безопасности применения цифровых технологий весьма удобно представлять как изучение звеньев некоторого процесса отслеживания такого инцидента (субъект-мотив-объект-методы-последствия). Например, от его инициатора (злоумышленника) через мотив (хулиганство, мошенничество, терроризм) к объекту (информационный ресурс гражданина, корпорации, правительства) посредством всего спектра доступных методов и до результата в виде финансовых и репутационных потерь. При этом возможен и позитивный вариант такого процесса: например, субъект (инженер службы ИБ) – мотив (защита информационного ресурса) – объект (информационные системы и сети) – методы (алгоритм действий со стороны службы ИБ) – последствия (угроза конфиденциальности, целостности и доступности ресурса предотвращена).

Вполне ясно, что исчисление приведенных в перечне показателей, относящихся к мезо- (регион страны, вид экономической деятельности) и макро- (страна) уровням, требует проведения специальных выборочных обследований и на практике трудно реализуемо в связи с массой ожидаемых проблем в решении программно-методологических и организационных вопросов.

**Анализ данных опроса о готовности компаний к киберугрозам**

Сегодня в поисках данных для анализа не обязательно проводить обследование самому: можно обратиться к открытым источникам, наиболее известным из которых является веб-платформа Kaggle – созданная в 2010 году и действующая под характерным лозунгом «Home of Data Science» система организации конкурсов по исследованию данных: здесь пользователи и организации могут публиковать наборы данных, исследовать их и создавать предсказательные модели.

В рамках настоящей работы и с целью проведения анализа взаимосвязи между безопасностью применения цифровых технологий и ее социально-экономическими эффектами на Kaggle был обнаружен набор данных, полученных по результатам исследования [13] «Готовность российских компаний к киберугрозам. Cyber risks readiness. Russia 2018-2020». Данные, представленные в панельном виде, включают 1146 наблюдений за 3 года с 2018 по 2020 гг. для 382 российских компаний различных видов деятельности (информационные технологии и телекоммуникации, финансы, строительство, производство, энергетика, медицина и др.). Этот массив характеризуется как финансовыми показателями (среди которых, например, показатель собственного капитала (ROE), известный по модели Дюпон (DuPont model)), так и результатами экспертных оценок готовности компаний к киберугрозам.

Для проведения анализа из исходного пула индикаторов был отобран ряд показателей – их характеристики представлены в таблице 6. Во избежание разночтений и, что, несомненно, окажет влияние на особенности моделирования, типы этих величин определялись на основе следующей классификации: числовые (непрерывные и дискретные) и категориальные (номинальные, т.е. неупорядоченные и, собственно, порядковые или ординальные). Наряду с этим, в соответствии с целью анализа, отобранные показатели в зависимости от своей роли были поделены на факторные и результативные.

Таблица 6

**Набор показателей результатов исследования готовности российских компаний к киберугрозам**

|  |  |  |  |
| --- | --- | --- | --- |
| Метка | Роль | Тип | Определение |
| ROE | результат | непрерывный | *Показатель рентабельности собственного капитала компании* |
| ROA | результат | непрерывный | *Показатель рентабельности активов компании* |
| IND | фактор | номинальный | *Показатель принадлежности компании к определенному виду экономической деятельности.* Принимает значения от 1 (‘IT и Телеком компании’) до 7 (‘Другие’) |
| INFR | фактор | ординальный | *Показатель уровня готовности организации к киберугрозам с точки зрения инфраструктуры*.  Принимает полученные на основе экспертных оценок значения от 1 (‘низкий’) до 5 (‘высокий’) |
| PEOPLE | фактор | ординальный | *Показатель уровня готовности организации к киберугрозам с точки зрения менеджмента организации и уровня подготовленности сотрудников*.  Методика и диапазон оценок те же |
| PARTNERS | фактор | ординальный | *Показатель уровня готовности организации к киберугрозам с точки зрения взаимодействия организации с партнерами и поставщиками*. Методика и диапазон оценок те же |

Анализ значений матрицы коэффициентов парной корреляции (таблица 7) показал наличие тесных связей между некоторыми показателями, в итоге из их числа были исключены две величины – Показатель собственного капитала (ROA) и Показатель готовности персонала к киберугрозам (PEOPLE).

Таблица 7

**Коэффициенты парной корреляции набора показателей**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Метки | ROE | ROA | INFR | PEOPLE | PARTNERS |
| ROE | 1,000 | 0,008 | 0,117 | 0,0783 | 0,001 |
| ROA | 0,008 | 1,000 | 0,521 | 0,357 | 0,256 |
| INFR | 0,117 | 0,521 | 1,000 | 0,706 | 0,295 |
| PEOPLE | 0,078 | 0,357 | 0,706 | 1,000 | 0,229 |
| PARTNERS | 0,001 | 0,256 | 0,295 | 0,229 | 1,000 |

Моделирование путем последовательного включения в уравнение множественной регрессии оставшихся факторов (8 показателей) привело к следующим результатам. Доля общей дисперсии определена на треть (значение скорректированного коэффициента детерминации составило 0,323). По величине коэффициентов при вошедших в модель регрессорах оказалось, что наибольшее влияние на эффективность экономической деятельности компаний оказывает «инфраструктурный» фактор (INFR: 11,497), далее следует сотрудничество с деловыми партнерами (PARTNERS: 1,316). Оба эти показателя, учитывая их ординальный характер, были введены в модель как дискретные числовые переменные. Исходный фактор (IND) в процессе моделирования был преобразован в шесть фиктивных переменных (седьмая, отражающая отраслевое значение «Другие», была исключена во избежание проявления мультиколлинеарности). Единственным из вошедших в уравнение регрессии «отраслевых» факторов стал индикатор принадлежности компании к сфере медицины (Medicine: -1,271). Показатель F-статистики, характеризующий качество модели, составил 140,8. Все коэффициенты при регрессорах значимы при 1%-ом уровне. В целом, можно утверждать, что, учитывая все же опосредованное влияние безопасности применения цифровых технологий на экономическую деятельность, модель выполнила свою функцию – четко показала вклад технологий в процесс экономического производства в размере 32,3 процента.

Однако, настоящий набор данных предоставляет возможность расширить границы анализа, проводимого традиционными статистическими методами, в направлении алгоритмов машинного обучения, что вполне приемлемо: «В контексте предсказательного моделирования какова разница между машинным обучением и статистикой? Четкой разграничительной линии, которая разделяет эти две дисциплины, нет. Машинное обучение тяготеет к большему вниманию к разработке эффективных алгоритмов, которые масштабируются до больших данных в целях оптимизации предсказательной модели. Статистика обычно больше сосредоточена на теории вероятностей и опорной структуре модели» [14, с. 252]. В этой связи было решено осуществить построение и обучение модели классификации набора данных, где целевой переменной (выходом, output) является принадлежность компании к определенному классу, объединившему ряд видов деятельности, а факторами (предикторами, features) – полученные экспертным путем оценки степени готовности компании к киберугрозам. Для формирования таких классов, прежде всего, из совокупности в силу понятных причин были исключены компании информационно-технологической и телекоммуникационной сферы, после чего оставшиеся компании были распределены на два класса: «сервисные» (финансовые и медицинские) и «производственные» (промышленность, строительство, энергетика).

Классифицирование проводилось несколькими широко известными методами: от самого простого в вычислительном отношении «K ближайших соседей» (K Nearest Neighbors) и до основанных на алгоритме «Дерево решений» (собственно, сам Decision Tree, а также Random Forest и один из вариантов семейства методов Boosting). В качестве оценки использовалась метрика ROC AUC (aria under the curve), т. е. площадь под кривой, характеризующей скорость обучения модели, основанной на том или ином алгоритме классификации: чем больше площадь, чем выше качество обученной модели (результаты представлены на рис. 3).

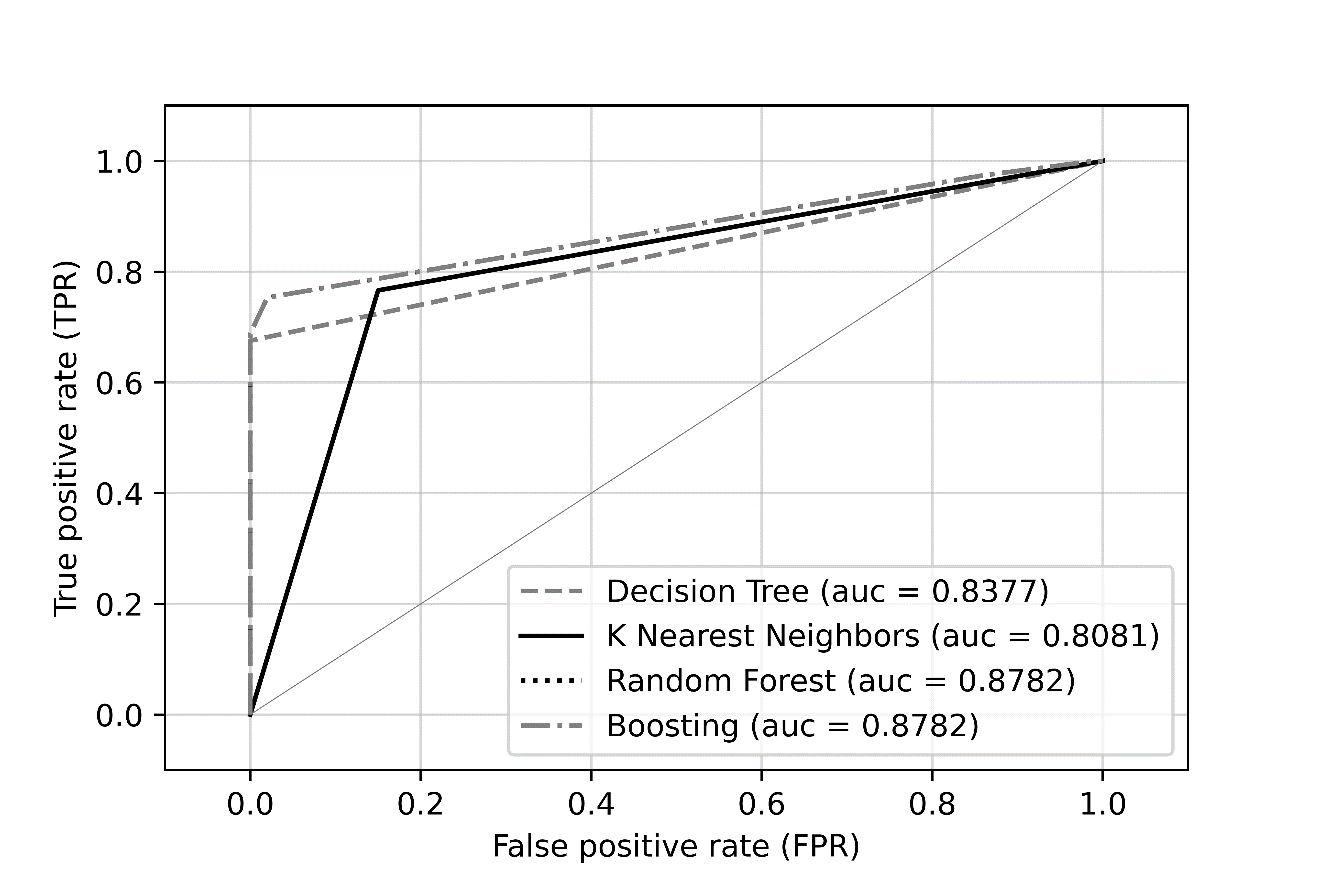


Рис. 3. **Графики ROC выбранного пула классификаторов**

Очевидно, что все три «древесных» варианты начинают старт одинаково эффективно (их графики буквально сливаются в одну линию), после чего, те из них, что используют не одно «дерево», а множество (Random Forest и Boosting), показывают лучший результат. Таким образом, с помощью методов машинного обучения удалось обобщить особенности отраслевой дифференциации компаний по степени их готовности к киберугрозам.

**Основные выводы**

Безопасность применения цифровых технологий, в силу их важности для жизни современного общества и функционирования экономики, является актуальной предметной областью для статистических исследований. Современные методологические разработки международных организаций способствуют решению задач по количественному измерению процессов безопасности применения цифровых технологий, но лишь фрагментарно.

Комплексный охват предмета исследования требует формирования профильной системы показателей, для чего целесообразно использовать представление об исследуемом процессе как звеньях логической цепи: субъект – мотив – объект – метод – последствия. Подобный подход позволяет структурировать анализ процессов сферы безопасности применения цифровых технологий, проецировать их развитие по направлениям (социальное, экономическое, технологическое) и по масштабу (микро-, мезо-, макроуровни).

Проведенный анализ результатов тематического опроса компаний показал, что экономическая эффективность организаций-респондентов на треть зависела от их уверенности в способности противостоять киберугрозам. Выявлена отраслевая дифференциация уверенности респондентов в этом вопросе.

Материалы статьи, в т. ч. данные и программные скрипты, размещены в репозитории автора по адресу: https://github.com/karyshev63rus/it\_security.

**Литература**

1. OECD Policy Framework on Digital Security: Cybersecurity for Prosperity. P.: OECD Publishing, 2022. 38 p. doi: https://doi.org/10.1787/a69df866-en.
2. OECD Measuring digital security risk management practices in business. P.: OECD Publishing, 2019. 63 p. doi: https://doi.org/10.1787/7b93c1f1-en.
3. ITU Publications. Global Cybersecurity Index 2015. URL: https://www.itu.int/pub/D-STR-SECU-2015.
4. ITU Publications. Global Cybersecurity Index 2020. URL: https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E.
5. Eurostat Statistics Explained. IT security in enterprises. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=ICT\_security\_in\_enterprises.
6. **Schatz D., Bashroush R.** Economic valuation for information security investment: a systematic literature review // Inf Syst Front. 2017. 19. P. 1205–1228 (2017). doi: https://doi.org/10.1007/s10796-016-9648-8.
7. **Хочуева Ф.А., Шугунов Т.Л., Жуков А.З., Ингушев Ч.Х.** Информационная безопасность сквозь призму цифровой экономики // Современные наукоемкие технологии. 2018. № 11 (часть 1) С. 65–71.
8. **Yerina A. M., Honchar I. A., Zaiets S. V.** Statistical indicators of cybersecurity Development in the context of Digital transformation of economy and Society // Sci. in nov. 2021. V. 17, no. 3. P. 3—13. doi: https:// doi.org/10.15407/sci ne17.03.003.
9. **Bernik I., Prislan K.** (2016) Measuring Information Security Performance with 10 by 10 Model for Holistic State Evaluation // PLoS ONE. 11(9): e0163050. doi: https://doi.org/10.1371/journal.pone.0163050.
10. **Зарова Е. В., Проскурина Н. В.** Теоретические основы региональной статистики. Самара: Изд-во СГЭА, 2004. 62 с.
11. **Суслов И. П.** Теория статистических показателей. М.: Статистика, 1975. 264 c.
12. **Алетдинова А. А.** Формирование системы статистических показателей инновационного потенциала организации // Экономика, статистика и информатика. Вестник УМО. №6(2), 2011 C. 78-81.
13. Готовность российских компаний к киберугрозам. URL: https://www.kaggle.com/datasets/stanislavkurovskiy/cybersecurity-russia2018-2020.
14. **Брюс П., Брюс Э., Гедек П.** Практическая статистика для специалистов Data Science. СПб.: БВХ-Петербург, 2021. 352 с.