ISSN 2708-2032
e-ISSN 2708-2040

INTERNATIONAL
UNIVERSITY

INTERNATIONAL JOURNAL OF INFORMATION & COMMUNICATION TECHNOLOGIES

Volume 2, Issue 4
March 2021

ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН
MINISTRY OF EDUCATION AND SCIENCE OF THE REPUBLIC OF KAZAKHSTAN

ХАЛЫҚАРАЛЫҚ УНИВЕРСИТЕТІ
МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ
INTERNATIONAL UNIVERSITY

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

Том 2, Выпуск 8
December 2021

Издание зарегистрировано Министерством информации и общественного развития Республики Казахстан. Свидетельство о постановке на учет No KZ82VPY00020475 от 20.02.2020 г.

Журнал зарегистрирован в Международном центре по регистрации сериальных изданий ISSN (ЮНЕСКО, г. Париж, Франция)

Выходит 4 раза в год.

УЧРЕДИТЕЛЬ:
АО «Международный университет информационных технологий»

ISSN2708-2032 (print)
ISSN2708-2040 (online)

СОДЕРЖАНИЕ

ИНФОКОММУНИКАЦИОННЫЕ СЕТИ И КИБЕРБЕЗОПАСНОСТЬ

ЦИФРОВЫЕ ТЕХНОЛОГИИ В ЭКОНОМИКЕ И МЕНЕДЖМЕНТЕ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

CONTENTS

INFORMATION AND COMMUNICATION NETWORKS, CYBERSECURITY

DIGITAL TECHNOLOGIES IN ECONOMICS AND MANAGEMENT

INTELLIGENT SYSTEMS

MATHEMATICAL AND COMPUTER MODELING

МАЗМҰНЫ

АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ЖЕЛІЛЕР, КИБЕРҚАУІПСІЗДІК

ЭКОНОМИКАДАҒЫ ЖӘНЕ МЕНЕДЖМЕНТТЕГІ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕЛЕР

МАТЕМАТИКАЛЫҚ ЖӘНЕ КОМПЬЮТЕРЛІК МОДЕЛЬДЕУ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 06–11
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.001
УДК 004.942

Кожахметова Б.А. [1*]**, Губский Д.С.** [1]**, Дайнеко Е.А.**[1]**, Ипалакова М.Т.** [2]

[1]Международный университет информационных технологий, Алматы, Казахстан
[2] Южный федеральный университет, Ростов-на-Дону, Российская Федерация
*b.kozhakhmetova@iitu.edu.kz

ЧИСЛЕННО-МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СОВРЕМЕННЫХ УСТРОЙСТВ СВЧ И КВЧ ДИАПАЗОНОВ НА ПРИМЕРЕ МИКРОПОЛОСКОВОГО РЕЗОНАТОРА

Аннотация. Данная статья посвящена моделированию устройств СВЧ и КВЧ диапазона. В качестве исследуемого объекта выбран микрополосковый резонатор. В статье рассмотрены и выявлены преимущества программных пакетов 3D электромагнитного компьютерного моделирования CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO, используемые для расчета и анализа СВЧ и КВЧ устройств. Для моделирования микрополоскового резонатора выбрана программа компьютерного моделирования CST Microwave Studio, с помощью которого были рассчитаны S-параметры устройства при изменении длины и ширины среднего проводника.

Ключевые слова: микрополосковый резонатор, CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO, математическое моделирование

Введение

На сегодняшний день существует множество программных пакетов 3D электромагнитного компьютерного моделирования и расчёта СВЧ устройств, из которых наиболее популярными являются: CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO и другие. Каждая из этих программ обладает своими преимуществами и позволяет пользователю наиболее эффективно исследовать электромагнитные свойства СВЧ и КВЧ устройств.

EM Software and Systems FEKO

FEKO – программа для моделирования и проектирования СВЧ устройств, позволяющая производить

анализ антенных систем и неоднородных диэлектрических сред разработанная компанией EMSS. Основным отличием данной программы от других программ моделирования является то, что программа дает возможность совмещать численные методы решения трехмерных электродинамических задач с приближенными аналитическими методами, т.е. применение методов моментов в сочетании с методами физической оптики и однородной теорией дифракции. Это позволяет справиться с основным недостатком программ компьютерного моделирования высокочастотных структур, как большие затраты ресурсов при моделировании объектов, размеры которых много больше длины волны [1]. Благодаря данной особенности, можно выделить FEKO среди других программ моделирования СВЧ устройств.

Applied Wave Research Microwave Office

Microwave Office компании AWR - система разработки планарных СВЧ-устройств. Программа предоставляет возможности электромагнитного анализа планарных структур, разработки и моделирования линейных и нелинейных схем, редактирования топологии и 2.5D электромагнитного анализа. Система обладает обширной библиотекой, предназначенной для выполнения анализа частотных характеристик методами гармонического баланса и с помощью рядов Вольтерра.

Вычислительный модуль системы работает в частотной и во временной областях, а также выполняет анализа линейных и нелинейных схем, такие как: анализ схем на основе рядов Вольтерра, анализ переходных процессов, конверсионно-матричный анализ, использование одночастотного и многочастотного методов гармонического баланса для анализа нелинейных схем, метод линейного анализа и др.

Традиционные реализации метода гармонического баланса построены на базе алгоритма анализа низкочастотных аналоговых схем, а система Microwave Office была разработана исключительно для высокочастотных и сверхвысокочастотных приложений, что делает ее значительно быстрее всех существующих продуктов (фактически в реальном времени) [2].

Ansoft High Frequency Structure Simulator

High Frequency Structure Simulator (HFSS) – один из базовых коммерческих пакетов полноволнового трёхмерного электромагнитного моделирования для проектирования СВЧ-структур.

HFSS – стандартно используемое промышленное программное обеспечение для нахождения S-параметров. Программа позволяет создавать модели и рассчитывать трехмерные электромагнитные поля для высокочастотных элементов и узлов. Инженеры полагаются на точность, способности и характеристики HFSS, чтобы проектировать различные устройства (внутренние соединения печатных плат, антенны, биомедицинские устройства). HFSS использует метод конечных элементов (FEM), чтобы вычислить электрическое поведение высокочастотных элементов и узлов. С HFSS можно рассчитать параметры (S, Y, Z), визуализировать трехмерные электромагнитные поля (в ближней и дальней зонах) и создавать эффективные модели, чтобы оценить качество сигнала, включая потери при распространении волны, обратные потери из-за отсутствия согласования импеданса, плохую состыковку и излучение.

Большие возможности открываются при использовании Ansoft HFSS, в частности, для проектирования сложных антенн, включая антенные решетки. В настоящее время высокая точность расчетов методом конечных элементов доказана для сложных волноводных конструкций. Программа HFSS Ansoft, дополнительно к численному методу, содержит аналитические методы, реализованные с помощью макросов, т.е. внутренних функций, рассчитываемых и реализованных в отдельных подпрограммах. Кроме этого, в программе Ansoft HFSS имеется идеальный согласованный слой (Perfect Matched Layer) и периодические граничные условия, а также мощная пост-процессорная обработка [3].

Computer Simulation Technology Microwave Studio

CST Microwave Studio – одна из самых продвинутых и популярных программ моделирования для устройств СВЧ. Программа позволяет производить быстрое и точное численное моделирование СВЧ устройств, также выполнять анализ проблем целостности радиосигналов и электромагнитной совместимости.

С помощью CST Microwave Studio можно проектировать различные модели устройств, такие как: планарные, спиральные, рупорные антенны; делители и сумматоры мощности; микрополосковые, волноводные и диэлектрические фильтры; микрополосковые и волноводные направленные ответвители; соединители; оптические узлы и другие устройства.

Для решения поставленных электродинамических задач в пакете CST Microwave Studio используются следующие методы: метод конечного интегрирования, метод аппроксимации для идеальных граничных условий, метод тонких стенок.

Среда проектирования CST обеспечивает прямой доступ к различным параметрам анализируемых структур и схем, например, геометрическим размерам, характеристикам материалов, номиналам элементов, что делает возможным выполнение быстрой настройки и оптимизацию проектов. Для дополнительной обработки результатов расчета без повторного перезапуска анализа используется метод интеллектуальной интерполяции.

Таким образом, программный пакет компьютерного моделирования CST Microwave Studio является наиболее удобным и оптимальным вариантом для решения поставленной задачи ввиду ряда своих преимуществ перед конкурентными продуктами других фирм производителей [4].

В данной статье рассматривается моделирование измерительного прибора с возможностью удаленного доступа на примере микрополоскового резонатора. Удаленный доступ будет реализован путем разработки клиент-серверного приложения. В созданном приложении серверная компонента включает в себя базу данных собранных лабораторных работ, числовые данные (двоичные файлы) с результатами расчётов и натурных экспериментов для дальнейшей обработки и отображения на экранах измерительных приборов, серверные части программного обеспечения виртуальных моделей. Клиентская компонента отвечает в основном за отображение интерфейса измерительно прибора и вывод необходимой информации. В качестве интерфейса измерительного прибора будет использоваться полностью реалистичная «отрисовка» рабочей панели изучаемого устройства.

Новизна исследования состоит в использовании численно-математического моделирования устройств СВЧ и КВЧ диапазона для создания удаленной виртуальной лаборатории. Теоретическая ценность данной статьи обусловлена тем что, результаты теоретического исследования моделирования на примере микрополоскового резонатора может служить основой для дальнейшей разработки виртуальных моделей устройств данного диапазона. Практическая ценность заключается в том, что данные результаты исследования используются для создания удаленной виртуальной лаборатории для измерения радиотехнических характеристик сигнала.

Модель микрополоскового резонатора

Модель исследуемого микрополоскового резонатора разработана для включения в состав виртуальной лаборатории по изучению устройств СВЧ и КВЧ диапазонов. В соответствии с концепцией виртуальной лаборатории [5] моделируемое устройство должно быть представлено как независимый программный модуль (динамически подключаемая библиотека, dll-файл) с определенным внешним интерфейсом, обеспечивающим взаимодействие с подключаемыми другими приборами и устройствами путем создания необходимых соединений в конфигураторе. Поэтому изучаемое устройство в разработанной концепции должно быть представлено на основе абстрактного «черного ящика», который поддерживает систему входных/выходных сигналов, адекватно реагирует на вызов внутренних функций посредством входных сигналов и возвращает необходимые данные устройствам, которые их запросили. При этом моделируемое устройство действительно может рассматриваться как некий объект, способный реагировать на запрос из вне и возвращать необходимые числовые данные, благодаря переопределению соответствующих виртуальных функций. Например, модель микрополоскового резонатора на запрос из вне (например, от модели векторного анализатора цепей) должна вернуть значение своей характеристики, т.е. параметры S-матрицы. В общем случае не имеет значения откуда и как эти данные были получены. Поэтому есть два оптимальных варианта:

− аналитическое (численное) вычисление каждый раз необходимых данных;

− выбор необходимы данных из базы данных.

В первом случае характеристики устройства должны быть описаны с помощью аналитических (математических) выражений, а в компьютерной модели предусмотрены ограничения численных значений, если они существуют.

В другом случае, необходимо отметить, что выбор из базы данных является общим и легко реализуемым, т.к. все необходимые данные могут быть представлены в двоичном файле, к которому организован быстрый доступ и выбор необходимых значений. Однако, при этом может потребоваться аппроксимация данных, которую можно реализовать методом Ньютона, точности которого вполне достаточно для прорисовки АЧХ.

Поэтому при моделировании простых устройств использовались аналитические вычисления. А при моделировании с помощью различных сторонних пакетов – экспорт значений в базу данных. Рассмотрим его реализацию на примере пакета CST Microwave Studio.

Одним из стандартных форматов экспорта результатов расчета (например, S - параметров) из CST Microwave Studio является текстовый формат файла. Для их преобразования и загрузки в создаваемые модули новых устройств СВЧ были созданы необходимые программные модули, которые осуществляют преобразование и запись данных в специально организованные бинарные файлы (базы данных). Причем необходимо отметить, что каждый раз, создаваемый бинарный файл данных организован так, чтобы обеспечить максимально быстрый доступ к необходимым данным, соответствующим заданным параметрам устройства.

При изменении параметров модели в пользовательском интерфейсе, программное обеспечение модуля автоматически выбирает соответствующий этим параметрам набор значений и использует полученную информацию в выходном сигнале. Следует отметить, что полученные значения являются дискретными и для построения амплитудно-частотной характеристики необходимо применять интерполяцию. Рассчитанные характеристики (S - параметры) микрополоскового резонатора получены при равноотстоящих друг от друга частотах, при этом выбранный шаг изменения частоты достаточно мал. Поэтому, для получения промежуточных значений достаточно построения интерполяционного многочлена Ньютона второй степени. Это справедливо для всех модулей устройств, используемых нами.

Аналогично может быть реализована конвертация результатов экспериментальных исследований новых устройств и обработка стандартных файлов *.s2p, которые можно получить (сохранить) при исследовании устройств на современном высокотехнологичном оборудовании.

Исследуемая модель микрополоскового резонатора представляет собой отрезок волновода с диэлектрической подложкой, на которой расположены три металлических проводника. С помощью пакета проектирования CST Microwave Studio были рассчитаны S-параметры устройства при изменении длины (l) и ширины (s) среднего проводника. Модель микрополоскового резонатора представлена на рисунке 1.

Расчеты S-параметров проведены при изменении длины среднего проводника от 18 мм до 26 мм с шагом 1 мм и его ширины – от 1 мм до 2 мм с шагом 0,1 мм. Полученный набор расчетных данных для различных геометрических размеров был обработан и представлен в бинарном виде.

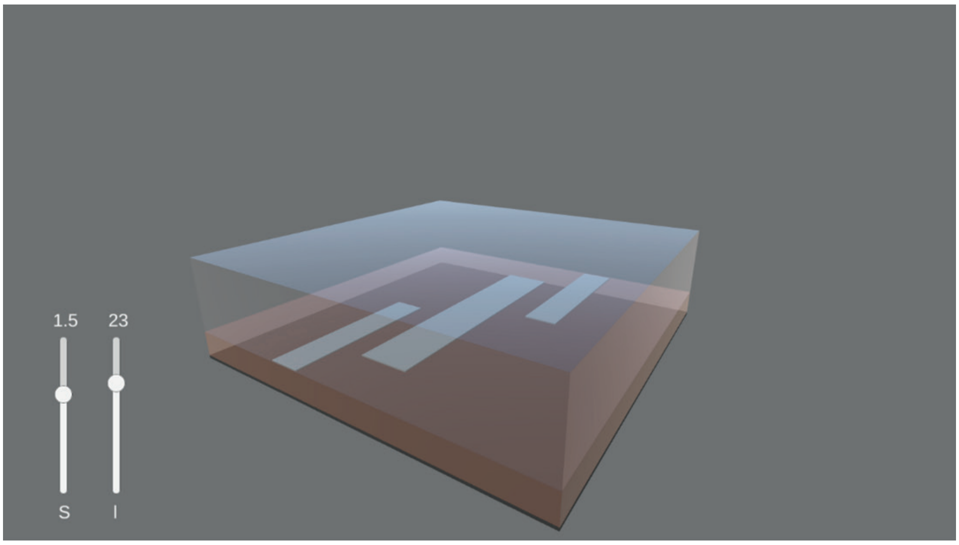

Рисунок 1 – Модель микрополоскового резонатора

Пользовательский интерфейс модели микрополоскового резонатора позволяет изменять длину и ширину среднего проводника. После изменений пользователем размеров резонатора из бинарного файла выбирается нужный набор S-параметров, который обрабатывается соответствующим образом и передается для отображения в измерительный прибор. Процесс обработки происходит в несколько этапов:

− устройство (в нашем случае микрополосковый резонатор) вызывает через свой выходной сигнал соответствующую функцию измерительного прибора (например, векторного анализатора цепей);

− вызванная функция измерительного прибора «понимает», что в устройстве произошли изменения и надо перестроить его АЧХ;

− измерительный прибор вызывает через свою систему сигналов, подключенных в конфигураторе [5], соответствующую функцию устройства и получает необходимые данные для перерисовки АЧХ.

Используя органы управления измерительного прибора, можно проводить необходимые измерения (например, величину затухания, полосу рабочих частот, исследовать их зависимость от геометрических размеров резонатора). На рисунке 2 представлен интерфейс виртуальной лаборатории. Измерительный прибор представлен векторным анализатором цепей R&S®ZVA 40, работающем в диапазоне от 10МГц до 40ГГц.

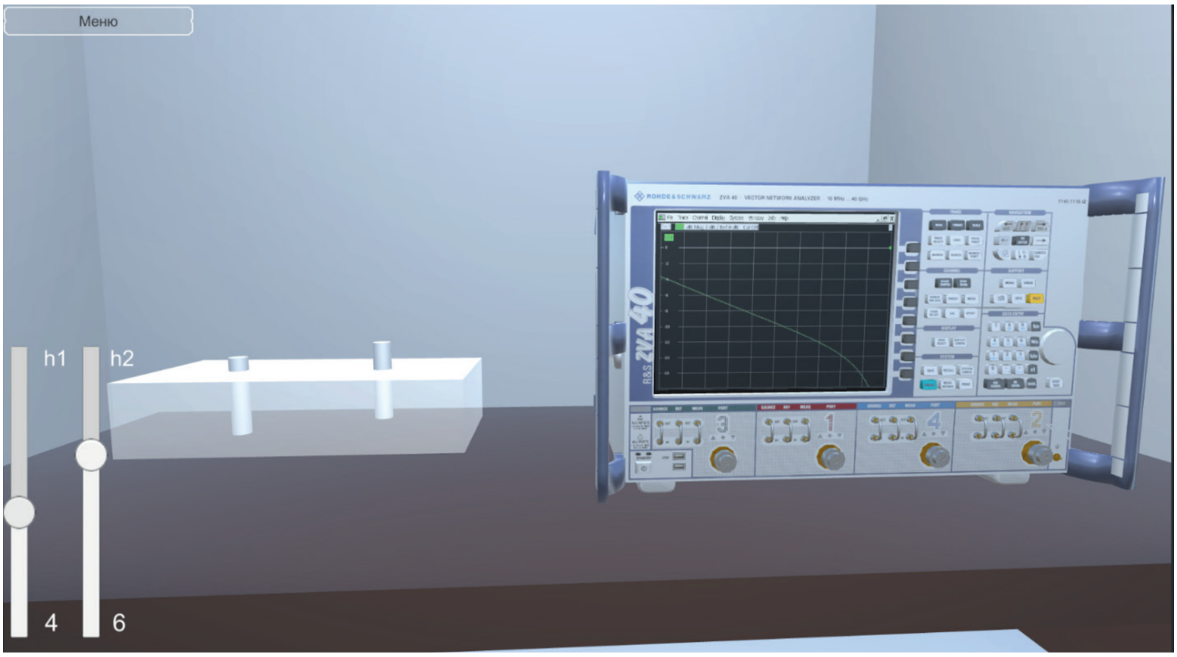

Рисунок 2 – Интерфейс виртуальной лаборатории

Заключение

Вопросам моделирования устройств СВЧ посвящено большое количество работ, в которых авторы показали, что для адекватного описания электродинамических характеристик устройств СВЧ могут быть использованы различные системы автоматизированного проектирования, например, CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO или строгие численно-аналитические методы, обеспечивающие высокую скорость расчета необходимых параметров. При этом авторы отметили преимущества программного пакета компьютерного моделирования CST Microwave Studio при построении моделей различных устройств СВЧ диапазона.

Таким образом, используя в качестве исходных данных результаты расчетов, полученные с помощью программного продукта CST Microwave Studio, можно создать большое количество различных моделей СВЧ устройств и приборов.

СПИСОК ЛИТЕРАТУРЫ

1. FEKO [Электронный ресурс] https://rtf.sfedu.ru/noc1/soft_feko.html / (дата обращения 26.11.2021)

2. Microwave Office, AWR Software [Электронный ресурс] https://www.awr.com/ru/products/microwave-office / (дата обращения 26.11.2021)

3. Ansys HFSS, Best-In-Class 3D High Frequency Electromagnetic Simulation Software [Электронный ресурс] https://www.ansys.com/products/electronics/ansys-hfss / (дата обращения 26.11.2021)

4. CST studio suite electromagnetic field simulation software [Электронный ресурс] CST Studio Suite 3D EM simulation and analysis software (3ds.com) / (дата обращения 26.11.2021)

5. Gubsky, D. S., Kleschenkov, A. B., and Mamay, I. V., «Virtual laboratory for microwave measurements», Computer Applications in Engineering Education, Vol. 27, No. 6, 1496–1505, 2019.

REFERENCES

1. FEKO [Electronic resource] https://rtf.sfedu.ru/noc1/soft_feko.html / (accessed 26.11.2021)

2. Microwave Office, AWR Software [Electronic resource] https://www.awr.com/ru/products/microwave-office / (accessed 26.11.2021)

3. Ansys HFSS, Best-in-Class 3D high Frequency Electromagnetic Simulation Software [Electronic resource] https://www.ansys.com/products/electronics/ansys-hfss / (accessed 26.11.2021)

4. CST studio suite electromagnetic field simulation software [Electronic resource] CST Studio Suite 3D EM simulation and analysis software (3ds.com) /(accessed 26.11.2021)

5. Gubsky, D. S., Kleschenkov, A. B., and Mamay, I. V., «Virtual laboratory for microwave measurements», Computer Applications in Engineering Education, Vol. 27, No. 6, 1496–1505, 2019.

Кожахметова Б.А., Губский Д.С., Дайнеко Е.А., Ипалакова М.Т.

Микрожолақты резонатор мысалында АЖЖ және ЕЖЖ диапазондарының заманауи құрылғыларын сандық-математикалық үлгілеу

Андатпа. Бұл мақала АЖЖ және ЕЖЖ диапазондарының құрылғыларды модельдеуге арналған. Зерттелетін объект ретінде микрожолақты резонатор таңдалды. Мақалада АЖЖ және ЕЖЖ құрылғыларды есептеу және талдау үшін пайдаланылатын CST microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO электромагниттік компьютерлік модельдеудің 3D бағдарламалық пакеттерінің артықшылықтары қарастырылып, анықталды. Микрожолақты резонаторды модельдеу үшін CST Microwave Studio компьютерлік модельдеу бағдарламасы таңдалды, оның көмегімен орташа өткізгіштің ұзындығы мен ені өзгерген кезде құрылғының S-параметрлері есептелді.

Түйін сөздер: микрожолақты резонатор, CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO, математикалық модельдеу.

Kozhakhmetova B.A., Gubsky D.S., Daineko Y.A., Ipalakova M.T..

Numerical and mathematical modeling of modern devices of UHF and EHF bands on the example of a microstrip resonator

Abstract. This article is devoted to modeling UHF and EHF range devices. A microstrip resonator selected as the object under study. The article discusses and identifies the advantages of 3D electromagnetic computer modeling software packages CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO, used for the calculation and analysis of UHF and EHF devices. To simulate a microstrip resonator, the CST

Microwave Studio computer simulation program was selected, with the help of which the S-parameters of the device were calculated when the length and width of the middle conductor changed

Key words: microstrip resonator, CST Microwave Studio, Ansoft HFSS, AWR Microwave Office, EMSS FEKO, mathematical modeling.

Авторлар туралы мәлімет:

Кожахметова Багдат Абдурашидовна, магистр, «Радиотехника, электроника және телекоммуникация» кафедрасының сениор-лекторы, Халықаралық ақпараттық технологиялар университеті.

Губский Дмитрий Семёнович, ф-м.ғ.к, «Қолданбалы электродинамика және компьютерлік модельдеу» кафедрасының доценті, Оңтүстік федералды университет, ORCID: 0000-0001-6651-5953.

Дайнеко Евгения Александровна, PhD, «Радиотехника, электроника және телекоммуникация» кафедрасының ассистент-профессоры Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0001-6581-2622.

Ипалакова Мадина Толегеновна т.ғ.к, «Компьютерлік инженерия» кафедрасының ассистент-профессоры Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0002-8700-1852.

Сведения об авторах:

Кожахметова Багдат Абдурашидовна, магистр, сениор-лектор кафедры «Радиотехника, электроника и телекоммуникации» Международного университета информационных технологий.

Губский Дмитрий Семёнович, к.ф.-м.н., доцент кафедры «Прикладной электродинамики и компьютерного моделирования» Южного федерального университета, ORCID: 0000-0001-6651-5953.

Дайнеко Евгения Александровна, PhD, ассистент-профессор кафедры «Радиотехника, электроника и телекоммуникации» Международного университета информационных технологий, ORCID: 0000-0001-6581-2622.

Ипалакова Мадина Толегеновна, к.т.н., ассистент-профессор кафедры «Компьютерная инженерия» Международного университета информационных технологий, ORCID: 0000-0002-8700-1852.

About authors:

Bagdat A. Kozhakhmetova, master, senior-lecturer of «Radio Engineering, Electronics and Telecommunications» department, International Information Technology University.

Dmitry S. Gubsky, candidate of physical and mathematical sciences, associate professor of «Applied Electrodynamics and Computer Modeling» department, Southern Federal University, ORCID: 0000-0001-6651-5953.

Yevgeniya A. Daineko, PhD, assistant-professor of the «Radio Engineering, Electronics and Telecommunications» department, International Information Technology University, ORCID: 0000-0001-6581-2622.

Madina T. Ipalakova, candidate of technical sciences, assistant-professor of the «Computer Engineering» department, International Information Technology University, ORCID: 0000-0002-8700-1852.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 12–17
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.002

UDC 004.56

Mubarakova S.R. *, Amanzholova S.T., Uskenbayeva R.K.
International Information Technology University, Almaty, Kazakhstan
*E-mail: mubarakova.saltanat@gmail.com

RELEVANCE OF CYBERSECURITY IN THE MODERN WORLD

Abstract. This article discusses one of the most pressing issues of today: relevance of cybersecurity in the modern world. It presents the main facts about cybersecurity, the concept and historical aspect of the development of cybersecurity and analyzes the current state of cybersecurity in Kazakhstan. The relevance of this article lies in the fact that it clearly highlights the role of cybersecurity and its relevance for modern society. In this connection, the article provides up-to-date data on cyberattacks and cybersecurity of our time.

Keywords: cybersecurity, Information Technology, cyberattacks, cybercrime.

Introduction. Today it is difficult to overestimate the importance of cybersecurity in the modern world. This is important because cybersecurity measures are designed to protect against theft and subsequent use of confidential data, personal medical information, intellectual property, government and industry information systems, everything that is stored and managed using information technology. At the moment, the risk of becoming a victim of cybercriminals is quite high. Everyone, from the most popular Internet users to large companies, uses different cloud services to store various personal data. The days when companies could trust antivirus programs and firewalls to protect their information are long gone. In modern reality, the services of cybersecurity specialists are not sufficient for organizing reliable protection. Any office worker, without knowing it, can easily turn into a "tool" of cybercriminals. The role of computer systems and cybersecurity protection measures in modern conditions is enormous [1].

With the development of Internet technology, even electric kettles have "learned" how to receive, process, generate and send digital data over the network. And the more we surround ourselves with high-tech devices, the higher is the risk of becoming a victim of cybercrime. At the same time, cybersecurity is also becoming increasingly important - both as an area of information technology and as a set of tools to ensure the protection of confidential data.

The relevance of the research topic lies in the fact that the number of cybercrimes worldwide has grown enormously over the past few decades, the motives and goals of cybercriminals have changed over time, and the cybercrime rate is rising from year to year. This is evidenced by the huge financial losses of legal entities and structures, as well as the increase in cybercrimes against individuals. This rapidly growing problem requires an effective and speedy solution, as the level of cybercrime and the complexity of crime are increasing, while the processing of cases and the effectiveness of work against criminals in cyberspace are decreasing. This topic is relevant today because the costs of preventing and disposing cybercrimes are growing, more and more legal entities and individuals are trying to protect themselves in advance, but criminals in cyberspace are not stagnating, and methods and types of crimes are getting more and more complicated. Therefore, this area requires constant monitoring and search for solutions.

The purpose of the study is to analyze the relevance of cybersecurity in the modern world.

In accordance with the purpose of the study, we set the following tasks:

1) To study the existing trends in information technology and cybersecurity;

2) To substantiate the relevance of this topic;

3) To study the current status quo in the cybersecurity area;

4) To analyze the cybersecurity situation in Kazakhstan.

The object of the work is the cybersecurity of Kazakhstan at the present time. The subject of the study is cybersecurity in the modern world.

In accordance with the purpose and objectives of the study, the main research methods are:

- study of popular science literature on this problem;
- analysis and synthesis of the collected information;
- analysis of the legal framework;
- analysis of statistical data;
- systematization and generalization of materials, conclusions on this problematic issue.

The scientific significance lies in the fact that the research contributes to the study of the relevance of cybersecurity in the modern world, particularly in Kazakhstan. The practical significance of the work lies in the fact that the results of this work can be used in the efforts to remedy the cybersecurity situation in Kazakhstan and in practical classes on information technology.

The hypothesis of our research: "Currently, cybercrime on the web is growing every year. And despite all the external struggle with this, the problem exists to this day. Moreover, this becomes a global problem for modern man. Cybercrime can grow into a more global problem and become more serious than domestic crimes. Therefore, it is very important to study this phenomenon" [2, p. 63-65].

Methods and materials. The main research methods are: text analysis in the form of analysis of scientific literature related to the topic of information technology and cybersecurity, comparative analysis in the form of studying and summarizing information obtained during the study; statistical methods and synthesis. Comparison and generalization are also used as auxiliary methods of empirical research. To search for existing research, numerous literature search methods were used, including searching in electronic databases and major journals, as well as manually searching for links to identified articles. The main electronic databases were Pubmed, SCOPUS, Web of Science and Science Direct. During the initial search, 51 articles on cybersecurity were found. These articles were further examined using the following three selection criteria: firstly, the studies included in the review should directly address the human factors related to cybersecurity and its consequences; secondly, the studies should be published in peer-reviewed journals; and thirdly, the studies should not focus on the technological dimension of cybersecurity. After applying these three criteria, 50 questions were selected, which were considered in four main areas: cyberattacks, contributing factors and strategies to combat them, cybersecurity in Kazakhstan [3, p. 10-12].

Literature review. Information and cybersecurity is not a new topic for research, as a matter of fact it has been a serious national problem for more than 20 years, which has led to a rapid growth of scientific literature over the past 10 years. A significant contribution to the disclosure of the problems of the formation and development of the information society was made by the socio-philosophical theories of foreign scientists: D. Bell, W. Dayzard, M. Castels, M. McLuhan, J. Masuda, T. Stonier, E. Toffler and others. Given the rapidly growing body of literature on the current cybersecurity issues, the purpose of this review article is to summarize the current literature for researchers, policy makers, practitioners and even the general public. As a result of this review, 51 relevant publications in leading journals on information systems in the period from 2010 to 2020 have been found and analyzed. In particular, there were identified nine main areas of concentration (for example, legal issues, supervision and morality, vulnerabilities, risks and detection), which constituted a substantive basis for the theoretical justification of the basic structures and their interrelations in the study of information systems security. This review has hopefully made an important contribution to cybersecurity research, summarizing the existing literature and providing an exhaustive framework. This review aims to make a new contribution to the synthesis of knowledge in cybersecurity research in three dimensions [4, p. 34-42].

Results and discussions. In the modern world organizations of the commercial, financial, medical, processing and energy sectors, including all government agencies, organize the collection, storage and processing of all information necessary for work, as well as personal data of employees, users, customers and visitors. In principle, all this information should be protected, as it is confidential, and its possible loss or theft can have unpredictable consequences for people and organizations. Organizations that directly provide the infrastructure of entire cities, countries and the global community as a whole are more likely to be subjected to a multi-level complex cyberattack [5].

The very concept of cybersecurity refers to a set of technologies, methods and processes designed to protect the integrity of programs, networks and data from cyberattacks. In other words, cybersecurity is a set of conditions that guarantee protection of all components of information systems from the maximum number of threats and undesirable influences at the physical, financial, emotional, mental, spiritual, educational, political and professional levels. Such influence, or negative consequences in case of errors, accidents, incidents and other damage in cyberspace are considered undesirable [6].

Cybersecurity is security in relation to information technology. This includes all technologies that store, process, or transmit data, such as computers, data networks, and all devices connected or embedded in a network, such as routers and switches. It should be noted that the concept of cybersecurity is a state or process of protecting computer systems aimed at repelling any kind of cyber threats, be it malware, various network attacks such as brute force attacks and DDoS, or even training staff on the methods of protection against social engineering, phishing and other tricks of cybercriminals. As cybersecurity evolves, attackers evolve to exploit weaknesses in the system for profit or simply to prove their case.

Cybersecurity extends to computers, networks, operating systems, applications, and other configurable and programmable components of the IACS system. The concept of cybersecurity was first introduced in 1991 as part of the generalization of digital network technologies, and goes back to the ancient Greek word "cyber". However, this arms race has been going on since the 1950s. For example, launching a cyberattack two decades after the creation of the world's first digital computer in 1943 was not an easy task. Giant electronic machines were not connected to the network, a limited number of people had access to them, and only a few knew how to work with them, so there were practically no threats. So, the history of cybersecurity begins in 1972 with the ARPANET research project, the forerunner of the Internet.

Cybersecurity is rapidly evolving, hackers and security service providers are constantly competing to get around each other, and new threats and innovative ways to combat them are constantly emerging. This review presents the latest trends in cybersecurity. For example, the Covid-19 pandemic has forced most organizations to transfer employees to work from home, often in a very short period of time. Many studies show that after the pandemic, most employees will continue to work remotely. Working from home involves new risks and is one of the most discussed trends in the field of cybersecurity. Home offices tend to be much less secure than centralized offices, which are usually equipped with firewalls and routers, and access control is regulated by the IT security group. The transition to remote work was carried out in a hurry so as not to disrupt work processes, and the security audit could be carried out with less care than usual. Cybercriminals can take advantage of this. Many employees use personal devices for two-factor authentication, and they may well use mobile versions of instant messaging apps such as Microsoft Teams and Zoom. Blurring the boundaries between personal and professional life increases the risk of confidential information falling into the wrong hands. The development of the Internet of Things has also opened up new opportunities for cybercriminals [7]. As a result, the main trend of cybersecurity is to attract the attention of companies to the security problems that have arisen as a result of the transition to remote work: identifying and eliminating new security vulnerabilities, improving systems, implementing security measures and ensuring proper monitoring and documentation.

Currently, cyberattacks are carried out with the aim of extorting money from people or disrupting production or work processes in companies. Cybersecurity is considered as a component of computer security and information technology. Compliance with the basic requirements of information security allows you to keep physical and digital data intact, protect them from disclosure, use, verification or complete destruction, that is, from unauthorized access to them. According to the international security services in the field of cyber threats, about 12 people are attacked every second in the world and about 556 million cybercrimes are committed annually in the world, the damage from which is more than 100 billion US dollars.

According to international cybersecurity experts, in 2019 cyberattacks around the world occurred every 14 seconds. Along with the increase in the number of cyberattacks, the damage they cause is also growing: in 2018 the losses of companies from various sectors of economy amounted to $ 1.5 trillion, while in 2019, according to experts, they already reached 2.5 trillion dollars. In 2022, according to the forecasts of the World Economic Forum, the amount of damage caused to the planet as a result of cyberattacks could grow to $8 trillion. By the way, anyone can be subjected to cyberattacks - both large companies and ordinary users. Many believe that their televisions and other household appliances may not be of interest to hackers. However, few people pay attention to the fact that thanks to these devices it is also possible to access personal data that can be used for various purposes.

Currently, in the Republic of Kazakhstan, the interaction of the IT industry with domestic business is perceived as a promising direction. The message of the President of the Republic of Kazakhstan "Kazakhstan in a new reality: time to act" reflects the fact that large public and private companies spend tens of billions of tenge on the development and attraction of foreign players. The government should establish mutually beneficial cooperation between industry and the IT industry. This will create digital technology platforms that will be able to fuel the digital ecosystem of any industry and make Kazakhstan one of the international centers for processing and storing data.

There are not so many companies engaged in practical provision of information security in Kazakhstan, no more than ten. The main reason: most of the orders were subcontracted in Russia, Israel or European countries. For example, second-tier banks ordered information security not from Kazakhstani, but from foreign suppliers. There are many distributors of software on the Kazakhstan market that sell foreign antiviruses, firewalls and security systems. The main consumer of cybersecurity services is, of course, the banking sector, since the banks' reputation depends on it. In addition, if the information is disclosed, it will be extremely difficult to avoid financial losses. That is, banks must undergo annual inspections, which cost from 20 thousand dollars. If we talk about small businesses, outsourcing of information security will cost from 1 million tenge per month. At the same time, the costs of companies still hardly pay their way, because the market is growing slowly. The ingenuity of hackers and scammers, as well as the emergence of new ways of processing information, stimulate the development of increasingly stringent standards and requirements for information security (IS), which, according to experts, generate new solutions in this area [8].

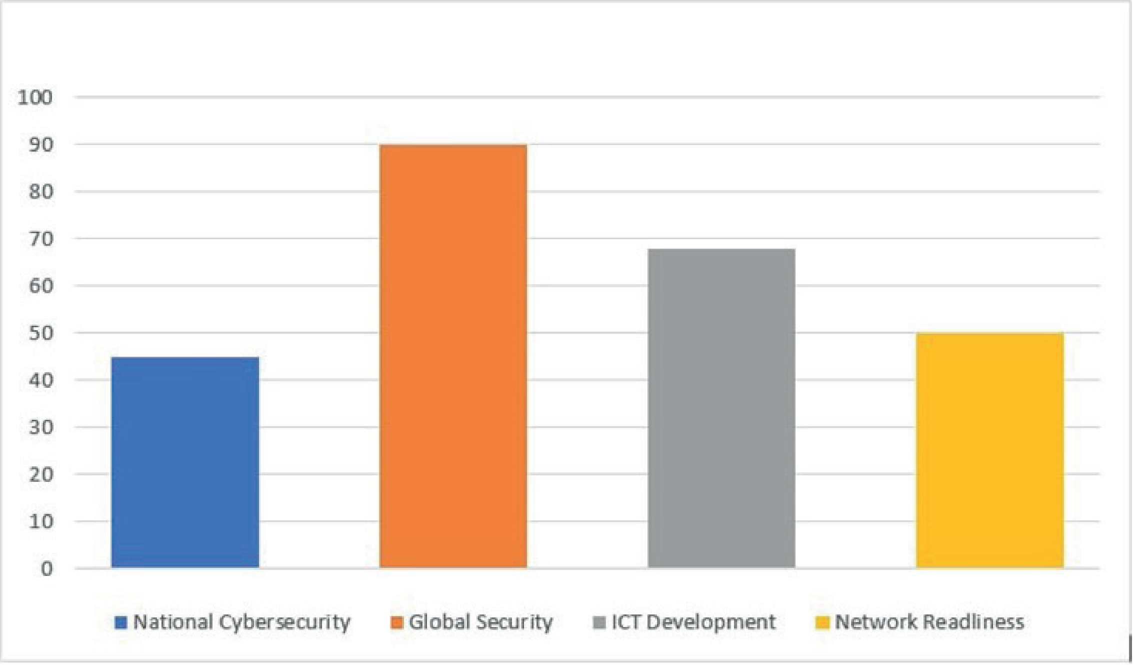

Figure 1 - Percentage of completion of the national Cybersecurity Index

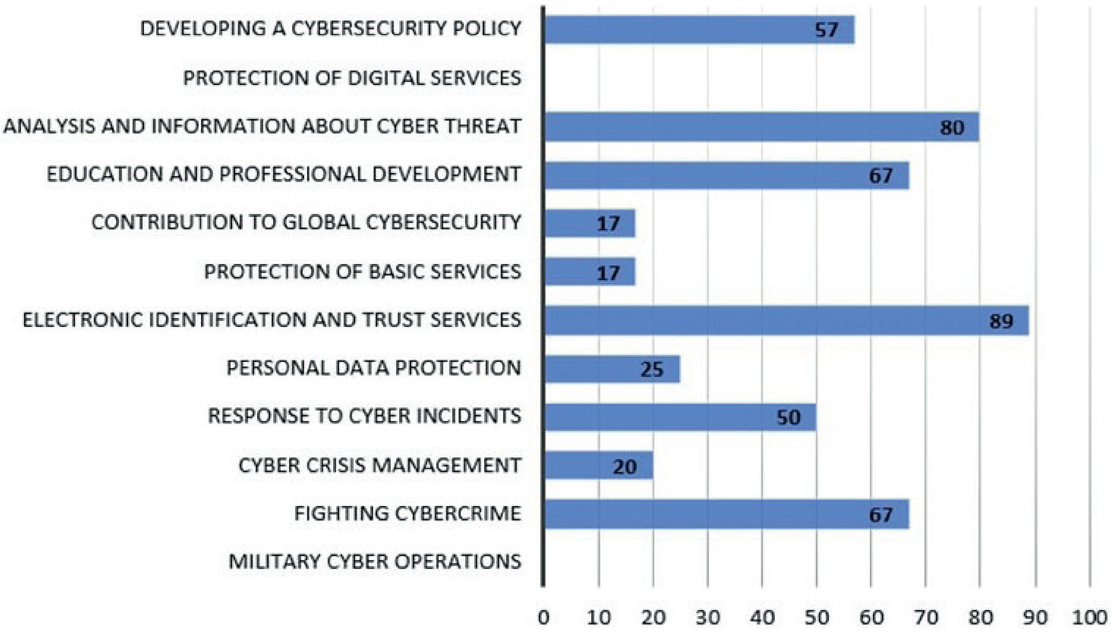

Figure 2 - Percentage of NCSI completion

"Within the framework of the program “Cyber Defense of Kazakhstan”, the state has identified 336 critical cybersecurity facilities, including state institutions, banks and industrial companies, attacks on which may have a national or interstate effect. As of now, testing laboratories have been set up in the country to study malicious code, the National Information Security Coordination Center has been launched. There is also a Private Computer Incident Response Service (CERT) and 7 operational centers for information protection (COMI). The number of grants in this specialty for future specialists has been increased. It is also planned to endow the Information Security Committee with functions to protect personal data, conduct audits and verify the owners of information systems. This will help to improve the situation in the field of information security and personal data protection.

Analysts note the country's successes in the legal sphere. In particular, it is noted that Kazakhstan has uniform requirements in the field of information and communication technologies and information security. The Digitization Initiative attaches increasing importance to an effective cybersecurity strategy. Over the past two years, fundamental conceptual approaches to the development of the country's cybersecurity sphere have been developed in the country. The cybersecurity concept "Cyber Defense of Kazakhstan" has been developed and approved, as well as a number of legislative acts and industry contracts. In addition, testing laboratories have been established to study malicious code, a national information security coordination center has been established, the number of scholarships in this area has been increased, etc. [9, p. 32-34].

Over the past few years, Kazakhstan has developed basic conceptual approaches to the development of the country's cybersecurity sphere. One of the important events is the approval of the Cyber Shield concept of Kazakhstan. The purpose of the concept is to achieve and maintain the level of protection of electronic information resources, information systems and ICT infrastructure from external and internal threats, ensuring the sustainable development of the Republic of Kazakhstan in the conditions of global competition. The implementation of the Cybersecurity Concept "Cyber Shield of Kazakhstan" by 2022 is expected to yield the following results:

− the number of retrained specialists in the field of information security in 2022 will reach 800 people;

− a 50% increase in the share of domestic software products in the field of computerization and communications used in the public and quasi-public sectors in 2022 compared to the base period of 2017;

− the share of IT systems of state bodies, non-state information systems integrated with state IT systems of critical ICT infrastructure facilities associated with information security monitoring centers should be 80% in 2021 and 100% in 2022 [10, p. 16-20].

Conclusion. In general, cybersecurity is the most important area of the IT industry. There are a number of professional certification opportunities for training and gaining experience in the field of cybersecurity. Although billions of dollars are spent on cybersecurity every year, no computer or network is protected from attacks and can be considered completely secure. All devices and IT facilities must be protected from intrusion, unauthorized use and vandalism. In addition, information technology users should be protected from asset theft, extortion, identity theft, loss of privacy and confidentiality of personal information, malicious damage, equipment damage, hacking of business processes and cybercriminals in general. The public should be protected from acts of cyberterrorism, such as compromise or loss of the power grid.

In conclusion, it is shown that strengthening cybersecurity is an essential condition for maintaining peace in the country. And the rapid development of the security market is directly determined by the rapid development and use of technological innovations, as well as the toughening requirements to the protected data security. Thus, it can be assumed that the set goals have been achieved. A lot of new, interesting and useful developments have taken place. The knowledge gained will be useful to all of us in life. The stronger the dependence of society on computer systems becomes, the greater is the vulnerability of Kazakhstan and other countries to all types of cybercriminals. You have to think about security today, tomorrow may be too late.

REFERENCES

1. What is cybersecurity and why is it important? (multipassword.com)
2. T.A. Tereshchenko. Cybersecurity: Problems and Solutions / Natural Sciences and Humanities Research №24(2), 2019. – 63- 65p.
3. Malik, T. N. Cybersecurity: problems and prospects / T. N. Malik. - Text : direct // Young scientist. — 2021. — № 7 (349). — Pp. 10-12.
4. Yablochkin A.S., Koshkin A.P. - Modern directions of research in the field of information security strategies // National security / nota bene. – 2019. – № 5. – 34- 42c.
5. Cybersecurity Trends 2021 | Kaspersky Lab (kaspersky.ru)
6. Cybersecurity and information security (spravochnick.ru)
7. NCSI: Kazakhstan (ega.ee)
8. How Kazakhstan's Cybersecurity is developing | Strategy2050.kz
9. Lim V.B. Development of information security in Kazakhstan//Science, Technology and education, 2020 - 32-34s.
10. Biekenov N. A. Some problems of cybersecurity in the Republic of Kazakhstan // Bulletin of the Institute of Legislation and Legal Information of the Republic of Kazakhstan No. 1 (33) 2014 - 16-20s.

Мубаракова С.Р., Аманжолова С.Т., Ускенбаева Р.К.

Қазіргі әлемдегі кибер қауіпсіздіктің өзектілігі

Андатпа. Бұл мақалада бүгінгі күннің ең өзекті мәселелерінің бірі: қазіргі әлемдегі киберқауіпсіздіктің өзектілігі талқыланады. Мақалада киберқауіпсіздік туралы негізгі фактілер талқыланады. Сондай-ақ киберқауіпсіздікті дамытудың тұжырымдамасы мен тарихи аспектісі қарастырылды. Бұған қоса Қазақстандағы киберқауіпсіздіктің қазіргі жағдайы сараланды. Бұл мақаланың өзектілігі қазіргі қоғам үшін киберқауіпсіздік туралы нақты ақпарат беруінде. Мақаланың негізгі мақсаты – қазіргі әлемдегі киберқауіпсіздіктің өзектілігін талдау. Осыған байланысты мақалада сіз қазіргі заманның кибершабуылдары мен киберқауіпсіздігі туралы өзекті деректерді таба аласыз.

Кілт сөздер: Киберқауіпсіздік, Ақпараттық технологиялар, Кибершабуылдар, Киберқылмыс.

Мубаракова С.Р., Аманжолова С.Т., Ускенбаева Р.К.

Актуальность кибербезопасности в современном мире

Аннотация. В данной статье рассматривается один из наиболее актуальных вопросов сегодняшнего дня: актуальность кибербезопасности в современном мире. В статье рассматриваются основные факты о кибербезопасности. Также были рассмотрены концепция и исторический аспект развития кибербезопасности. После этого было проанализировано текущее состояние кибербезопасности в Казахстане. Актуальность данной статьи заключается в том, что она предоставляет четкую информацию о кибербезопасности для современного общества. Основной целью статьи является анализ актуальности кибербезопасности в современном мире. В связи с этим в статье вы можете ознакомиться с актуальными данными о кибератаках и кибербезопасности нашего времени.

Ключевые слова: Кибербезопасность, Информационные Технологии, Кибератаки, Киберпреступность.

Авторлар туралы мәліметтер:

Мубаракова Салтанат Рахатқызы, «Компьютерлік инженерия және ақпараттық қауіпсіздік» кафедрасының до кторанты, Халықаралық ақ параттық технологиялар ун иверситеті, Манас көш. 8, Алматы, Қазақстан. OR CID: 0000-0002-2394-881X.

Аманжолова Сауле Токсановна, «Компьютерлік инженерия және ақпараттық қауіпсіздік» кафедрасының ассистент-профессоры, техника ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті, Манас көш. 8, Алматы, Қазақстан. ORCID: 0000-0002-6779-9393.

Ускенбаева Раиса Кабиевна, «Компьютерлік инженерия және ақпараттық қауіпсіздік» кафедрасының профессоры, техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университеті, Манас көш. 8, Алматы, Қазақстан. ORCID: 0000-0002-8499-2101.

Сведения об авторах:

Мубаракова Салтанат Рахаткызы, докторант кафедры «Компьютерная инженерия и информационная безопасность», Международный университет информационных технологий, ORCID: 0000-0002-2394-881X.

Аманжолова Сауле Токсановна, кандидат технических наук, ассистент-профессор кафедры «Компьютерная инженерия и информационная безопасность», Международный университет информационных технологий, ORCID: 0000-0002-6779-9393.

Ускенбаева Раиса Кабиевна, доктор технических наук, профессор кафедры «Компьютерная инженерия и информационная безопасность», Международный университет информационных технологий, ORCID: 0000-0002-8499-2101.

About the authors:

Saltanat R. Mubarakova, Doctoral student, Department of Computer Engineering and Information Security, International Information Technology University, ORCID: 0000-0002-2394-881X.

Saule T. Amanzholova, Candidate of Technical Sciences, Assistant–Professor, Department of Computer Engineering and Information Security, International Information Technology University, ORCID: 0000-0002-6779-9393.

Raissa K. Uskenbayeva, Doctor of Technical Sciences, Professor, Department of Computer Engineering and Information Security, International Information Technology University, ORCID: 0000-0002-8499-2101.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 18–37
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.003
УДК 004.056.56

Razaque A.*, Adil A. Zh., Amanzholova S.T., Valiyev B.B.
International Information Technology University, Almaty, Kazakhstan

BLOCKCHAIN TECHNOLOGY-FEATURED NOVEL AIR-CRACKING TOOL FOR WI-FI HACKING DETECTION

Abstract. Wi-Fi plays an important role in promoting several application domains such as business, education, industry, etc. On the other hand, if not handled properly, vulnerabilities of Wi-Fi cause damage to the privacy and confidentiality of the users. Some of the hackers use the Linux tool to exploit the vulnerability of Wi-Fi that allows of the hacking process. In this paper, we introduce a Blockchain Technology-Featured Novel Air-Cracking (BTFAT) method to detect the Linux tool for Wi-Fi security improvement. The proposed BTFAT consists of valuable features (e.g., monitoring, scanning, cracking, and testing) which help detect the Linux tool. The BTFAT is programmed on the C platform. Based on the experimental results, the BTFAT produces higher performance as compared to other existing methods.

Keywords: Wi-Fi, vulnerability, BTFAT, privacy, reliability, testing, blockchain technology

Introduction

Wireless networks are now used everywhere. Wi-Fi is used not only by individual users but also by organizations and companies. Wireless networks are embedded in many areas of our life: social networks, business, work, finance (online payments, banking applications) [1-2].

Wi-Fi can make people's daily lives easier, improve the productivity of many companies, and make it easier for employees to work. But there is also a downside, this is the risk of leakage of confidential information through hacking Wi-Fi [6-7]. Many people, users of various social networks and messengers such as Instagram, Facebook, Twitter, WhatsApp, etc., store their data (photos, correspondence, card data) in their accounts. Hackers can hack Wi-Fi through various attacks and use sniffers to intercept traffic, thus gaining access to personal data. The same situation is possible in large business and financial organizations. This can lead to large financial losses for companies or banks [8-9]. Although networks with blockchain technology have a high level of security, they are also susceptible to hacking by intruders. This may lead to the loss of personal data of users or financial losses of companies and organizations [10-11].

As business depends on data, data acquisition speed and accuracy are crucial. The blockchain is perfect for conveying such information because it offers to authorize members of the network an instant, shared and fully transparent access to information in the register unchanged [3]. The blockchain network allows users to track orders, payments, accounts, products, and more. And since all participants share access to a single source of reliable data, it is possible to view all transaction details at any time to work with greater confidence and gain new benefits and opportunities [4-5].

Recently, many studies have been conducted in the field of hacking Wi-Fi using Linux tools, respectively, there are many solutions to this problem. Wi-Fi hacking using the Wireshark traffic analyzer is based on packet capture (PCAP) [12-13], the WPAclean utility uses a four-way handshake method and a beacon to clear capture files [14-15], there are also Linux tools such as Reaver, which uses a WPS connection as a vulnerability to analyze and hack the network [18-19]. The Wifite tool is designed for hacking a network with various encryption algorithms WEP, WPA, WPA2. Wifite uses a set of attacks on Wi-Fi, including brute-force passwords, handshake capture [16-17]. For network hacking, the Wifite tool has flexible settings [20-21]. Motivated by these challenges, the contributions of this paper are summarized as follows:

- the definition of vulnerability for hacking wireless networks using the technology of the Blockchain-Featured Aircrack-ng.
- hacking a wireless network with blockchain technology in practice.

Billions of users and businesses connect to the global network, use Wi-Fi and networks with blockchain technology. As a result, security becomes the most important issue. The main problem is to investigate the vulnerabilities of blockchain networks and based on the detected vulnerabilities, describe recommendations for protection against hacking, so that users and organizations can be less vulnerable to security attacks [22-23].

The following steps should be considered in investigating security issues against Wi-Fi hacking: (a) investigation of various security mechanisms available for WPA/WPA2 using BTFAT, (b) investigation of vulnerabilities in real-time using the BTFAT, and (c) determining the method of hacking [24-24]. We aim to address these issues and use these solutions in our practical results to make the use of Wi-Fi safer.

The remainder of the paper is organized as follows.

Section II briefly describes the problem and explains its significance. Section III highlights the previous research findings. Section IV describes the state-of-the-art system model. Section V proposes a way of Wi-Fi hacking vulnerabilities using the BTFAT process.

Section VI presents the experimental results and implementation. Section VII gives the discussion of the results. Finally, the conclusions of the paper are presented in Section VIII.

Problem identification

The main problem of this research work is hacking Wi-Fi with Linux using the "Krack" vulnerability (Key Reinstall Attacks). The real problem is researching and finding Wi-Fi vulnerabilities such as incorrectly configured access points, devices with weak encryption keys, impersonating an authorized user. Actions required to resolve this issue:

- to study vulnerabilities and hacking of the Wi-Fi network, then to select the appropriate tool;
- to find the target to attack;
- to check the impact of Pixie dust;
- then to run a full password search. If the PIN code is received but the WPA password is not displayed, to run the commands to get the Wi-Fi password.

Causes for hacking Wi-Fi can be open ports, lack of password protection or weak password protection, lack of data encryption, lack of programs for scanning the network, lack of special services to protect them from attacks. The effects of these causes can be gaining access to the network, interception of network data, commission of various attacks, theft of personal data, interception of passwords, spoofing of the network. The importance of the problem studied in this research paper is that Wi-Fi hacking must be performed as a test of the network and detection of its vulnerabilities to further improve the security of the network perimeter. There are several solutions for hacking Wi-Fi in the form of various attacks such as hacking WPA / WPA2 passwords, attacking WEP, hacking WPS pin, lowering WPA, replacing the true access point with a fake one, fraudulent access point, attacking Wi-Fi access points from global and local networks, denial of service attacks (DoS Wi-Fi), attacks on specific services and functions of routers. An optimistic solution to this problem is to use multiple attacks in combination. This can be implemented using the Linux tool or utility, which includes several or all of the listed kinds of attacks.

Related work

In this section the prominent features of the existing current approaches are summarized.

The main tools for hacking Wi-Fi using Linux are discussed by Sharma [26]. AirSnort uses special algorithms to sort out the password, namely, it analyzes each packet in the network, and when intercepting the required number of data, it decrypts the password from them. AirSnort is available for windows. However, there is one shortcoming - the utility only works with WEP networks.

Bullock & Jeff [27] described the use of packet sniffers Ettercap, Dsniff, and Wireshark for hacking Wi-Fi. Packet sniffers are designed to capture and analyze network traffic. The advantages of traffic analyzers are that they work with the vast majority of known protocols, have a clear and logical graphical interface based on GTK+, and a powerful filter system. Traffic analyzers are also cross-platform and work on such operating systems as Linux, Solaris, FreeBSD, NetBSD, OpenBSD, Mac OS X, and Windows. The disadvantage of these analyzers when hacking Wi-Fi is the need to possess certain skills and abilities in decrypting captured packets. In addition, it is possible to capture packets only in real time.

Li et al. [28] introduces another tool called Reaver for hacking wireless networks that targets certain WPS vulnerabilities. Reaver performs brute force attacks against WPS and registers PIN codes to recover the WPA / WPA2 passphrase. Since many router manufacturers and Internet service providers activate WPS by default, many routers are vulnerable. The disadvantage is that WPS can be disabled.

Wifite is a tool designed to attack multiple wireless networks encrypted using WEP / WPA / WPA2 and WPS. Some parameters are required when WiFite starts working. It records WPA handshakes, automatically disables authentication of connected clients and saves their hacked codes. Hacking Wi-Fi using the Wifite tool is discussed by Sinha [29]. Crunch is a very good and easy-to-use tool for creating custom word lists that can be used in dictionary attacks. Since the success rate of dictionary attacks depends on the quality of the word list, it is impossible to avoid creating your own word lists. The method of hacking the network with the Crunch tool is described by Santo Orcero [30].

MacChanger is a small utility that spoofs a media access control (MAC) address in an arbitrary MAC address. Spoofing the MAC address for Wi-Fi hacking may be necessary to avoid MAC filters or hide the hacker's identity in the wireless network. MacChanger's Wi-Fi hacking approach is discussed by Sinha [31]. After studying these network hacking tools, we have determined that all these tools are essential. However, the above tools have several disadvantages. The disadvantages are that some of these tools can crack only certain encryption algorithms of wireless networks, most of the above tools intercept traffic and hack networks in real time, only at the time of user activity, and have fewer methods for analyzing and hacking a wireless network. But our network hacking tool is not only easy to use, but also has many built-in features for hacking WPA/WPA2/WEP.

System model

The Blockchain technology-featured Aircrack-ng tool is of utmost importance. It successfully detects the Linux tool for Wi-Fi. The BTFAT consists of the features depicted in Figure 1. The features include airdecap-ng, airmon-ng, aireplay-ng, airodump-ng, etc. The airdecap-ng feature decrypts intercepted traffic with a known key, the airmon-ng package puts the network card in the monitoring mode, the airodump-ng feature is a traffic analyzer, it adds traffic to Packet Capture (PCAP) or initialization vectors (IVs) files and shows information about the network. Some of these features are greatly valuable in the Wi-Fi hacking process.

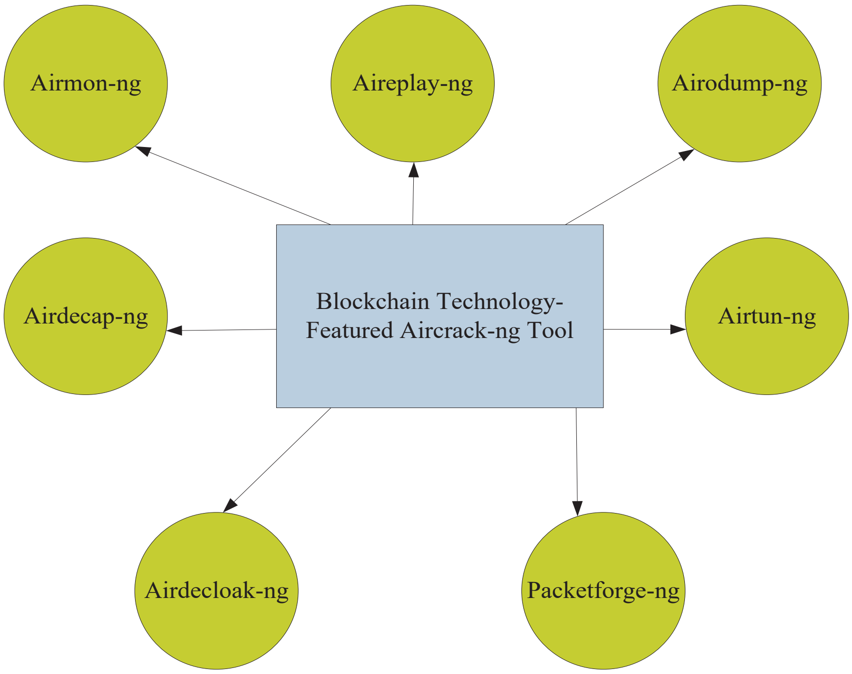

Figure 1 - Components of the Linux BTFAT

To hack Wi-Fi using the BTFAT, the hacker first connects to the Wi-Fi adapter and determines the network interfaces. To do this, the airmon-ng package defines the available network interfaces, as well as the driver. If the network interface driver is detected as a result of the command execution, the network is monitored. Otherwise, the driver is debugged. Network monitoring is performed by the airmon-ng feature as a result of network monitoring, a message should appear indicating that the monitoring mode was successfully enabled on the previously defined interface. Then, using the airodump-ng feature, the listening mode is enabled to determine the available Wi-Fi networks. As a result, the screen displays a list of wireless networks within the range of the Wi-Fi adapter. The screen also displays important characteristics for network hacking, such as the encryption used (WEP, WPA/WPA2), channel, and basic service set id (BSSID). Knowing the necessary information about the network, packets are captured using the airodump-ng package containing the encrypted password. When capturing packets, it is important to capture many IVs packets over 1000. The waiting time depends on the network activity. If no one is connected to the access point, the time may be delayed.

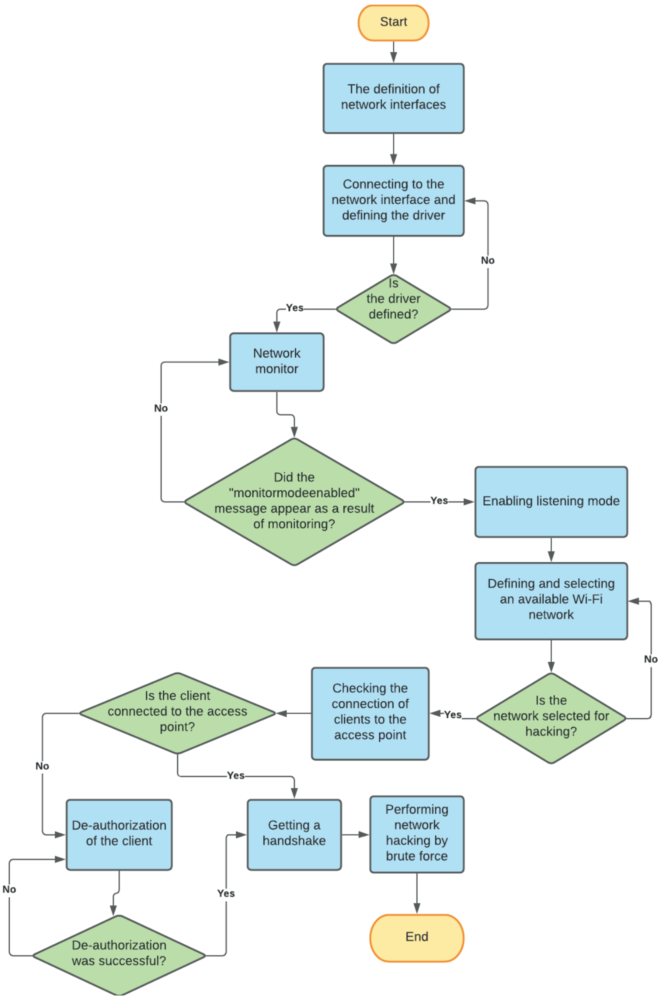
Start
The definition of network interfaces
Connecting to the network interface and defining the driver
Is the driver defined?
No
Yes
Network monitor
No
Did the "monitormodeenabled" message appear as a result of monitoring?
Yes
Enabling listening mode
Defining and selecting an available Wi-Fi network
No
Is the network selected for hacking?
Yes
Checking the connection of clients to the access point
Is the client connected to the access point?
No
Yes
De-authorization of the client
Getting a handshake
Performing network hacking by brute force
No
Yes
De-authorization was successful?
End

Figure 2 - The process of hacking Wi-Fi using the BTFAT

The distance to the access point is not as important as the network activity. To reduce the time for collecting packets, the client logs in. After successful de-authorization of the client, the hacker receives an intercepted handshake. Next, the hacker performs a brute-force hacking using a password dictionary. The process of hacking Wi-Fi using the aircrack-ng tool is depicted in Figure 2.

Figure 3 explains the system model of the Wi-Fi hacking process using the BTFAT and depicts a second (fake) access point created by the hacker during Wi-Fi hacking process. Using this access point, the hacker de-authorizes the user through multiple requests. After reconnecting the user from the real network to the access point created by the hacker, the hacker initiates the handshake. Based on the received handshake, it is possible to hack the network and decrypt the password.

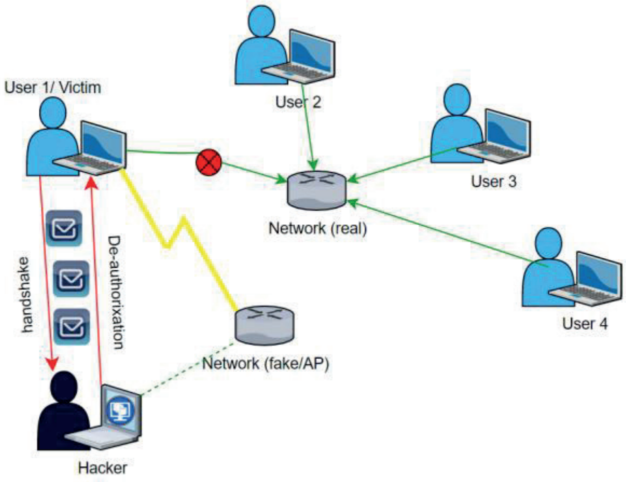

Figure 3 - System model

Proposed Wi-Fi hacking using the BTFAT process

The method proposed for hacking Wi-Fi uses the BTFAT. Before the Wi-Fi is hacked, methods are studied to protect the network. To protect Wi-Fi networks, several well-known methods are used, such as access restriction and authentication methods. This research paper discusses the method of hacking Wi-Fi, which uses the authentication method as a network protection. In turn, authentication methods for network protection are classified: open authentication, Shared Key Authentication (WEP encryption), Mac address authentication, Wi-Fi protected access (WPA), Wisconsin-Internet protected Access2 (WPA2), Cisco Centralized Key Management (CCKM). The BTFAT breaks WEP, WPA, and WPA2 keys. The process of hacking Wi-Fi with the BTFAT consists of three phases:

- packet-capturing and saving processes
- client de-authorization process
- Wi-Fi blockchain-featured hacking process

A. *Packet-capturing and saving processes*

This process is implemented at the beginning and is necessary for collecting IVs data packets. During this process, the network is monitored, as a result of which there are available network interfaces. After that, the hacker connects to them and captures the packets. Then all packages are saved in a single file. IVs packets contain the necessary information to decrypt the password of the required network. Packet-capturing and saving processes are presented in Table 1.

Table 1 - Packet-capturing and saving processes

```
Algorithm-1: Packet-Capturing and Saving Processes
1. Initialization: {$N_c$: Network Channel; $M_{pa}$: MAC address of Access Point; I: Interface; $P_{cf}$ : Packets-
   captured file; $L_t$ : Linux tool; $N_m$: Network monitoring; $N$: Network; $F_0$: Folder; P: Packets}
2. Input: {$N_c$ , $M_{pa}$, I}
3. Output: {$P_{cf}$}
4. Set $N_c$ , $M_{pa}$, I
5. Do Process $N_m \in N \leftarrow L_t$
6. While $N_m \in N<1$
7. Capture P
8. Sum $P_{cf} = P + 1$
9. Do $N_m = 0$
10. Save $P_{cf}$ to $F_0$
11. End while
```

Algorithm-1 explains the packing capturing and saving processes. In step 1, variables are initialized for packet capturing and saving. Steps 2-3 explain the input and output variables respectively. Step 4 uses the components (e.g., network channel, physical address of the access point and interface) for network monitoring process. Step 5 shows the process of using Linux tool on the network for network monitoring process. Steps 6-9 shows the entire network monitoring process and attempts to capture the packets, which are stored into the packet-capturing list. This process continues until the entire network is monitored and all of the packets are stored into the packet-capturing list. In step 10, the packet-capturing list is saved into folder for further process.

There are several properties that define packet capture:

- the total time it takes to capture packets;
- the average interval between adjacent packets;
- the average packet waiting time.

Definition-1: the average value of the interval between adjacent packets τ_a is the average time of packet captures between the previous and subsequent packets and is calculated by the equation (1):

$$\tau_a = \frac{1}{M} \times \sum_{s=0}^{M} (a_{t+1} - a_t) \qquad (1)$$

Where α_t: moments of time when packets arrive; M: number of analyzed intervals.

Theorem-1: The higher the load on the connection channel, the longer is the total time required to capture packets.

Proof: The channel load factor is calculated by the equation (2):

$$L_c = \frac{\sum P_t}{\sum E_p} \qquad (2)$$

Where P_t: capture time of the packet; E_p: end time of processing of the i-th packet.

The number of packets and their size (in bytes) and the time of traffic measurement are known. Then, the total capture time of the packet is equal to:

$$\sum P_t = \frac{(B+N)\times 8}{V} \qquad (3)$$

Where P_t: capture time of the packet; B: number of bytes transmitted; N: number of packets captured; V: packet capture rate.

The total processing time of the i-th packet is equivalent to the time of traffic measurement and is determined by the equation (4):

$$\sum E_p = \varphi \qquad (4)$$

Where φ: the time of traffic measurement.

Based on the previous equations, the channel load factor is equal to:

$$L_c = \frac{(B+N)\times 8}{V\varphi} \qquad (5)$$

Where B: number of bytes transmitted; N: number of packets captured; V: packet capture rate; φ: the time of traffic measurement.

Thus, the higher the channel load factor, the longer the packet capture time.

Hypothesis-1: The higher the packet intensity detected during network monitoring, the shorter is the packet capture time.

Proof: The packet capture time can be determined by the equation (6):

$$T_c = \frac{I_p}{1-M_a\times I_p} \qquad (6)$$

Where I_p: packet intensity (packets/sec); M_a: average network monitoring time.

Let the packet capture time be expressed in terms of traffic intensity T_I and packet length L_p, and channel throughput T_h:

$$I_p = \frac{T_I}{L_P} \qquad (7)$$

Where T_I: traffic intensity; L_p: packet length.

The average network monitoring time is determined by the equation (8):

$$M_a = \frac{L_P}{T_C} \qquad (8)$$

Where L_p: the packet length; T_h: the channel throughput; M_a: the average network monitoring time.

Then, the equations (7) and (8) are substituted in the equation (6):

$$T_h = T_I + \frac{L_P}{T_c} \qquad (9)$$

Where T_h: the channel throughput; T_I: the traffic intensity; L_p: the packet length; T_c: the packet capture time.

Based on the above equations, corollary-1 is derived.

Corollary-1: To reduce packet capture time, the bandwidth of the channel must be high.

B. *Client de-authorization process*

After finding the network interfaces and selecting an access point for hacking, a handshake should be conducted. To receive a handshake, the user must be active on the network. If there is no activity, the activity is created by deactivating the client. During the client deactivation process, the access point (fake) sends requests to the client until the client reconnects to the network. Thus, if deactivation is successful, the hacker receives a handshake. Client de-authorization and handshake recording process are given in Table 2.

Table 2 - Client de-authorization and handshake recording process

```
Algorithm-2: Client de-authorization and handshake recording process
1. Initialization: {$A_{pc}$: client's physical address; $A_{pa}$: physical address of the access point; $H$: handshake; $I$: interface; $C$: client; $A_p$ access point; $S$: client's SSID; $P$: password; $R_s$: recconnect, $P_{cf}$: packets-captured file}
2. Input: { $A_{pc}$, $A_{pa}$, $I$ }
3. Output: {$H$}
4. Set $A_{pc}$, $A_{pa}$, $I$
5. $A_p$ requests $\Rightarrow C \rightarrow R_s$
6. While $A_p = R_s$
7. Do $A_p \leftarrow H \in (P, S)$
8. Record $H$ to $P_{cf}$
9. End while
```

Algorithm-2 explains the client de-authorization and handshake recording processes. In step 1, the variables are initialized for the process of client de-authorization and handshake recording. Steps 2-3 give the input and output processes, respectively. Step 4 uses the network components (e.g., client's physical address, physical address of the access point, interface) for implementing requests. In step 5 requests are sent from the access point to the client to reconnect to the network. Steps 6-7 explain passing the handshake, which includes the password and client ID number to the access point. This process continues while the client is reconnecting to the network. In step 8 the received handshake is written into the captured packets that were received during network monitoring in the previous algorithm for further use in the Wi-Fi hacking process.

The time of de-authorization is characterized by the following properties:

- the total time of sending requests to the user;
- the total intensity of answers received by the hacker;
- processing of responses received from the user and establishing a handshake.

Definition-2: The total intensity of responses received by the hacker β_T is the sum of the intensity of the flow of requests sent to the user $\beta_H = (1 - C) \times \beta$ and the intensity of processed responses sent by the user $\beta_U = (1 - C) \times \beta$ and is calculated by the equation (10):

$$\beta_T = \beta + (1 - C) \times \beta \qquad (10)$$

Where β: the intensity of the elementary stream requests; C: probability of self-classification of a new request stream by a second access point.

Theorem-2: The intensity of sending requests by the hacker affects the performance of processing requests by the user and the average delay in sending requests.

Proof: The performance of processing requests by the user *(P_R)* is determined by the equation (11):

$$P_R = \frac{\beta + (1 - C) \times \beta}{\omega} \qquad (11)$$

Where P_R the performance of processing requests by the user ω: the intensity of the query processing; β: intensity of sending requests; C: probability of self-classification of a new request stream by a second access point.

The probability that the communication channel for sending the request is free (P_c) can be obtained by the equation (12):

$$P_C = \frac{1}{\frac{P_R^{m+1}}{m! \times (m - P_R)} + \sum_{m=0}^{m} \frac{P_R^m}{m!}} \qquad (12)$$

Where m: the number of processors.

The average delay in sending requests (D_A) can be obtained based on the number of requests sent (S_R), depending on the average number of requests in the queue (Q_A):

$$Q_A = \frac{P_R^{m+1} * P_C}{mm!\,(1 - \frac{P_R}{m})^2}, \qquad (13)$$

$$S_R = Q_A + P_C \;, \qquad (14)$$

$$D_A = \frac{S_R}{\beta + (1 - C) \times \beta} \qquad (15)$$

Where *m:* the number of processors; P_c: the probability that the communication channel for sending the request is free; P_R: the performance of request processing by the user; *β:* the intensity of the elementary stream requests; *C:* the probability of self-classification of a new request stream by a second access point.

Hypothesis-2: The smaller the volume of transmitted requests, the longer it takes for a hacker to get a handshake.

Proof: Each request has the same length and requires a transmission time (τ_T). The time of transmission of the message about the client's acceptance of the request is assumed to be equal to τ_R. The time for sending a request (τ_S) is calculated using the equation (16):

$$\tau_S = N \times \tau_T + \tau_R + N \times \tau_A + \tau_W \qquad (16)$$

Where τ_T: the transmission time of the request; N: the number of requests; τ_A: the average processing time of the response received by the hacker; τ_W: the average waiting time for a request in the queue until the communication line is free; τ_R: the time of transmission of the message about the client's acceptance of the request.

Since the bandwidth of the communication channel and the average length of each request are known, the average time for its transmission can be determined by the equation (17):

$$\tau_T = \frac{R_v}{C_h} \qquad (17)$$

Where R_v: a known volume of the request in bits; C_h: channel capacity bit/sec.

The time of transmission of the message about the client's acceptance of the request is calculated similarly by the equation (18):

$$\tau_R = \frac{R_A}{C_h} \qquad (18)$$

Where R_A: known volume of the request acceptance message; C_h: channel capacity bit/sec.

To calculate the average waiting time τ_W and the average message processing time τ_A, it is assumed that the input stream of packets from the user forms a simple stream with an average intensity μ, and the average service time $A_S$calculated by the equation: (19):

$$A_s = \frac{1}{\gamma} \qquad (19)$$

The request received in the buffer will wait until the communication line is released, i.e. until the processing of the message about the acceptance of the previous request is completed. Probabilities of finding a packet in a buffer queue of infinite length is calculated by the equation (20):

$$Q(N,L) = \frac{\frac{L^N}{N!} * \frac{1}{1 - L/N}}{\sum_{k=0}^{N-1} \frac{L^k}{k!} + \frac{L^N}{N!} * \frac{N}{N-L}} \qquad (20)$$

Where $L = \frac{\mu}{\gamma}$: full input load.

The average number of requests can be found by the equation (21):

$$A_N = \frac{L}{N-L} * Q(N,L) \qquad (21)$$

Where A_N: the average number of requests; Q*(N,L)*: the probabilities of finding a packet in a buffer queue of infinite length; *L*: the full input load.

Based on the previous equations, the average waiting time for a request in the queue is calculated by the equation (22):

$$\tau_W = \frac{A_N}{\mu} = \frac{Q(N,L)}{\gamma(N-L)} \qquad (22)$$

Where A_N: average number of requests; μ: average intensity.

The average processing time of a single request is determined by the equation (23):

$$\tau_A = \frac{L}{\mu} = \frac{1}{\gamma} \qquad (23)$$

Where *L:* full input load; μ average intensity.

Thus, if the parameters τ_W and τ_A are unchanged, the time of sending the request τ_S is determined by the equation (24):

$$\tau_S = \frac{N * R_v}{C_h} + \frac{R_A}{C_h} + \frac{N}{\gamma} + \frac{Q(N,L)}{\gamma(N-L)} \qquad (24)$$

Where R_v: the known volume of the request in bits; C_h: the channel capacity bit/sec; R_A: the known volume of the request acceptance message; Q*(N,L):* the probabilities of finding a packet in a buffer queue of infinite length; *L:* the full input load, N: the number of requests.

Corollary-2: To reduce the volume of transmitted requests, it is necessary to transmit requests of greater length to speed up the time of receiving the handshake. This corollary was based on the analysis of equation (24).

Considering that all requests have equal length and average transmission time, the duration of the communication channel occupation when transmitting one request after establishing a connection between the hacker and the user is determined by the equation (25):

$$\tau_H = N \times \tau_T + N \times \tau_A + \tau_M \qquad (25)$$

Where $R_M = \tau_M \times C_h$: the volume of transmitted requests.

Thus, the total time for sending requests is determined by the equation (26):

$$\tau = \tau_S + \tau_H = N \times \tau_S + N \times \tau_H + \tau_M + \tau_R + \tau_W \qquad (26):$$
$$= \frac{N * R_v}{C_h} + \frac{N}{\gamma} + \frac{R_A}{C_h} + \frac{R_M}{C_h} + \frac{Q(N,L)}{\gamma(N-L)}$$

Where τ_S: the time for sending a request; R_v: the known volume of the request in bits; C_h: the channel capacity bit/sec; R_A: the known volume of the request acceptance message; Q*(N,L):* the probabilities of finding a packet in a buffer queue of infinite length; *L:* the full input load, N: the number of requests.

C. *Wi-Fi Blockchain-Featured Hacking Process*

The last phase is hacking the wireless network using BTFAT. Blocks in the blockchain system can only create a certain number of bitcoins, transactions must have a certain format and correct signatures for spent bitcoins, a transaction cannot be performed twice within the same blockchain, etc. The blockchain cannot be hacked by attacking the encrypted traffic of an individual node: if the consensus rules are violated in the block, the blockchain system denies the operation of an individual node, even if other nodes believe that an intrusion into the chain of records did not occur.

To hack network with blockchain technology, ARP spoofing is performed after capturing the network traffic. This is an attack committed when sending ARP messages to the local network. The purpose of this attack is to link the hacker's MAC address to the IP address of another host, such as the default gateway. Thus, any traffic directed to a specific IP address is sent to the hacker. After making an attack on the network, the hacker inserts a malicious script into the HTML pages that the user views the command "to call the miner" and deploys an HTTP server on its computer to serve the miner. Figure 2.1 depicts the process of hacking Wi-Fi with BTFAT. The goal of the third phase is to carry out an autonomous attack on the Wi-fi network to introduce a malicious code. With the help of the built-in BATTAT tools, it is possible to analyze and edit traffic. For the purity of the hacking process, only one line of code is embedded in the HTML page, which calls the miner. After the traffic is captured, the JavaScript code is embedded in it, and an injector is created. The created injector adds a string to the HTML with a call to the JavaScript miner. The packet-capturing and saving processes are shown in Table 3.

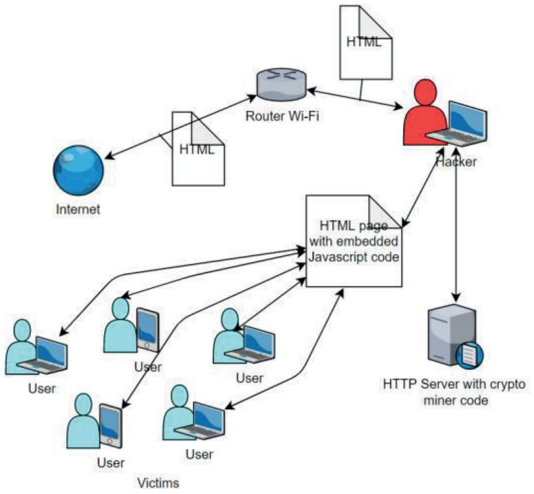
HTML
Router Wi-Fi
HTML
Internet
Hacker
HTML page with embedded Javascript code
User
User
User
HTTP Server with crypto miner code
User
User
Victims

Figure 2.1 - The process of hacking Wi-Fi with BTFAT

Table 3 - Wi-Fi Hacking process with BTFAT

```
Algorithm-3: Wi-Fi hacking process with BTFAT
1. Initialization: {$T_N$: captured network traffic; $U_{ip}$: user's IP address; $H_a$: hacker's MAC address, N: wireless network, $E_m$: embedded in the network miner; $A_{sp}$: ARP spoofing; $P_{ht}$: HTML pages, $S$: iterate through all pages, $JS_c$: JavaScript code; $M_s$: getting started miner; $H_c$: hacker's computer}
2. Input: { $T_N$, $U_{ip}$, $H_a$, $N$}
3. Output: {$E_m$}
4. Set $U_{ip}, H_a, N$
5. Do process $A_{sp} \rightarrow N$
6. Link $U_{ip}$ to $H_a$
7. For $P_{ht}$=0 $to$ $P_{ht} = S$
8. If $P_{ht} \in T_N$
9. Set $JS_c \rightarrow T_N = M_s$
10. End if
11. Deploy $P_{ht}$ $on$ $H_c$
12. Do $JS_c \rightarrow P_{ht}$
```

Algorithm-3 explains the process of hacking Wi-Fi with BTFAT. In step 1, the variables are initialized for the process of hacking Wi-Fi. Steps 2-3 give the input and output processes, respectively. In Step 4, the network for hacking, the user's IP address and the hacker's MAC address are determined. Step 5 is an ARP spoofing attack on the network. Step 6 is linking the hacker's MAC address to the user's IP address to direct traffic from the user's IP address to the hacker. Step 7 is iteration through each page directed to the user. In Step 8 we check that the HTML page belongs to the captured traffic. In Steps 9-10 we embed the JavaScript code into the captured traffic, thus triggering the miner. In Step 11, we deploy the HTTP server on the hacker's computer to serve the miner. In Step 12, JavaScript code is embedded in HTML pages for mining.

A large amount of data contained in the captured packets is analyzed in order to get the password.

Condition: m packets are used to analyze n amounts of data. Let's make the parameters of i-th packets as $b_i(i = 1,2,3, \dots, m)$. This parameter is set to analyze packets by parameters when searching for an encrypted password.

The analysis is performed for each packet from the 1st to the i-th, and a single volume of the j-th amount of data is used. Let the use of the i-th order packet to analyze a single volume of the j-th amount of data be given as $a_{ij}(i = 1,2,3, \dots, m; j=1,2,3, \dots, n)$.

Packet analysis takes a certain time the amount of which depends on the speed of analysis. Therefore, let the speed from analyzing the unit volume of the j-th amount of data be set as $c_j(j = 1,2,3, \dots, n)$.

This mathematical model explains a data analysis plan that provides the maximum speed of analysis under the specified restrictions on data packets.

The volume of analysis of the j-th amount of data is set as a vector of variables equal to $X = (X1, X2, X3, \dots, Xn)$, where $x_j(j = 1,2,3, \dots, n)$ is the volume of analysis of the j-th type of data.

When analyzing packets, a restriction is imposed on their number. It follows from the variable vector that the restriction on using packages for analyzing all data is set by the expression:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \le b_i \qquad (27)$$

Based on the speed of analysis of the j-th data quantity $c_j x_j$, the objective function is calculated using the equation (28):

$$Z(X) = c_1x_1 + c_2x_2 + \dots + c_nx_n \qquad (28)$$

A mathematical model for analyzing the m number of packets that contain n amount of data is defined as a system of the following expressions:

$$M = \begin{cases} Z(X) = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow max, \\ a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \le b_2, \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \le b_m, \\ x_j \ge 0, j = 1,2,3, \dots, n. \end{cases} \qquad (29)$$

Where M: the process of analyzing packages; m: the number of the packets; n: the amount of data.

Let the set of attacks made by the BTFAT be given by the expression (30):

$$K \in K_1 \times K_2 \times K_3 \ldots \times K_b \qquad (30)$$

Where $K_b\left(\overline{1, b+1}\right)$: the set of values of the i-th parameter of a particular attack that determines the type of attack. Each attack $\vec{k} \in K$ is a vector $(k_1, k_2, \ldots, k_{b+1})$, where $\overrightarrow{k_b} \in K_b$.

Rainbow tables are defined as an expression (31):

$$\vec{y} \in Y, Y \in Y_1 \times Y_2 \times Y_3 \ldots \times Y_j \qquad (31)$$

Where $Y_j(j = \overline{1, n})$: set of values of the j-th parameter of the rainbow table.

The network for hacking is indicated by the expression (32):

$$\vec{g} \in G, G \in G_1 \times G_2 \times G_3 \ldots \times G_f \qquad (32)$$

Where $G_f(f = \overline{1, m})$: the set of values of the f-th parameter of the wireless network.

The success of hacking the network using the BTFAT is related to the attack used to break into the wireless network and the formation of rainbow tables in the process of decrypting the password. Thus, the function that sets the level of successful hacking of the network by an attack $\vec{k} \in K$ c the application of rainbow tables $\vec{y} \in Y$ to hack the wireless network $\vec{g} \in G$ is denoted by the expression (33):

$$\delta: K \times Y \times G \to [0; 1] \qquad (33)$$

Where δ: the function that sets the level of successful hacking; K: the attack; Y: the rainbow tables; G: the wireless network.

The function that determines the degree of success from applying an attack to a wireless network is calculated by the expression (34):

$$\beta: G \times K \to [0; 1] \qquad (34)$$

Where β: the function that determines the degree of success from applying an attack to a wireless network; G: the wireless network; K: the attack.

The probability of a successful application of a hacker attack with rainbow tables is calculated:

$$\gamma: Y \times K \to [0; 1] \qquad (35)$$

Where γ: the probability of a successful application of a hacker attack with rainbow tables; K: the attack; Y: the rainbow tables.

Thus, based on the expressions (33), (34), (35), the function δ is expressed as:

$$\delta(\bar{k}, \bar{y}, \bar{g}) = \beta(\bar{g}, \bar{k}) * \gamma(\bar{y}, \bar{k}) \qquad (36)$$

Where $\delta(\bar{k}, \bar{y}, \bar{g})$: the function that sets the level of successful hacking; $\beta(\bar{g}, \bar{k})$: the function that determines the degree of success from applying an attack to a wireless network; $\gamma(\bar{y}, \bar{k})$: the probability of a successful application of a hacker attack with rainbow tables.

Define the function $\beta(\bar{g}, \bar{k})$. To do this, consider a family of functions:

$$\beta_{uh}: G_g \times K_h \to R_+ \qquad (37)$$

Where R_+: the set of non-negative real numbers; β_{uh}: a function that sets the level of mutual influence of the wireless network parameter G_g and the attack parameter k_h on the network:

$$\beta_{uh}(g, k) = 0, \qquad (38)$$

if an attack with the value of the parameter $k \in K_h$ is not applicable to a wireless network with the c $\in G_g$ parameter value.

$$0 < \beta_{uh}(g, k) < 1, \qquad (39)$$

if the value of the wireless network parameter c $\in G_g$ reduces the probability of a successful attack with the value of the parameter $k \in K_h$.

$$\beta_{uh}(g, k) = 1, \qquad (40)$$

if the value of the wireless network parameter c $\in G_g$ does not affect the applicability of the attack with the parameter $k \in K_h$.

$$\beta_{uh}(g, k) > 1, \qquad (41)$$

if the value of the wireless network parameter c $\in G_g$ indicates that an attack with the parameter $k \in K_h$ is applicable for hacking.

Denote by $\overline{\beta_{uh}}: G_g \times K_h \to [0;1]$ the function:

$$\overline{\beta_{uh}}(g,k) = \frac{\beta_{uh}(g,k)}{\sum_{\varepsilon \in C_g} \beta_{uh}(\varepsilon,k)} \qquad (42)$$

Then, based on the expression (18), the success rate of applying the attack $\vec{k} \in K$ to the wireless network $\vec{g} \in G$ is calculated:

$$\beta(\vec{g},\vec{k}) = \min_{h=\overline{1,b+1}} \prod_{g=1,s} \overline{\beta_{uh}}(g_u,k_h) \qquad (43)$$

Where the attack and wireless network are set by the parameters $(k_1,k_2,\dots,k_{b+1})$ and $(g_1,g_2,\dots,g_f)$, respectively.

The function $\gamma(\bar{y},\bar{k})$ is expressed similarly to the function $\beta(\vec{g},\vec{k})$:

$$\gamma(\bar{y},\bar{k}) = \min_{h=\overline{1,b+1}} \prod_{t=1,s} \overline{\gamma_{th}}(y_t,k_h) \qquad (44)$$

Where the attack and rainbow table are set by the parameters $(k_1,k_2,\dots,k_{b+1})$ and $(y_1,y_2,\dots,y_j)$, respectively.

Thus, the function that sets the level of successful hacking of the network by an attack $\vec{k} \in K$ c the application of rainbow tables $\vec{y} \in Y$ to hack the wireless network $\vec{g} \in G$ takes the form:

$$\delta(\bar{k},\bar{y},\bar{g}) = \min_{h=\overline{1,b+1}} \prod_{g=1,s} \overline{\beta_{uh}}(g_u,k_h) * \min_{h=\overline{1,b+1}} \prod_{t=1,s} \overline{\gamma_{th}}(y_t,k_h) \ (21)$$

The reliability of Wi-Fi hacking is characterized by the probability of password decryption, which is determined by the equation (45):

$$P_c(t) = \frac{N_0 - \sum n_i}{N_0} \qquad (45)$$

Where $P_c(t)$: reliability of Wi-Fi hacking; N_0: the number of initially captured packets; $\sum n_i$: the number of denied de-authorization requests.

The probability of decrypting the password from the received handshake is equal to the product of the probabilities of successful processing of elements of the Wi-Fi hacking process (packet capture, requests for client deauthorization, half-baked handshake):

$$P_c = P_1 \times P_2 \times P_3 \dots \times P_n \qquad (46)$$

Where P_c: the probability of decrypting the password; P_n: the probabilities of successful processing of elements of the Wi-Fi hacking process.

Theorem-3: The time of cracking the Wi-Fi (T_c) depends on the complexity of the password, which is selected from the space of possible passwords ($P = L^c$).

Proof: A password is selected from the space of possible passwords. The size of the space P is determined by the expression (47):

$$P = L^c \qquad (47)$$

Where P: the size of the possible password space; L: the length of characters in the password; C: the number of characters in the password.

Thus, the time is calculated by the expression (48):

$$T_c = \frac{P}{10^9 \times 3600} \qquad (48)$$

Where T_c: the time of cracking Wi-Fi; P: the size of the possible password space.

Hypothesis-3: Hacking a network using the BTFAT tool takes less memory, less processing power, and less time as compared to other tools designed to hack a network.

Proof: To break into the network, a hacker needs to get a handshake containing an encrypted password and to decrypt the password. The W function converts the encrypted password $e(P)$ into a new password $W\big(e(P)\big)$. The encrypted password in the handshake is written in binary notation, and the password is written as numbers in the notation Q, where Q: the number of possible characters for passwords. The C function then converts the data from the binary number system to the Q number system. For each encrypted password $e(P)$, the function calculates a new password $W\big(e(P)\big)$. The BTFAT for hacking a wireless network has the ability to use rainbow tables, which speed up the process of decrypting the password while spending less computer resources. Using the W function, it is possible to precompute data tables (rainbow tables).

To generate a data point in the rainbow table, a possible password P_0 is assigned, an encrypted password $e(P_0)$ is calculated, then a possible password $W(e(P_0))$ is calculated, which becomes P_1. This process continues until the encrypted password starts with twenty 0 ($e(P_n)$). Such an encrypted password occurs 1 time in about 10^6 encrypted passwords. The pair $[P_0, e(P_n)]$ that contains an encrypted password starting with twenty 0 is stored in the table.

The set of such pairs is calculated. Each pair contains a sequence of possible passwords $P_0, P_1, \dots, P_n$ and encrypted passwords. However, there may be spaces, meaning some passwords may not be present in all calculations. For a good database without spaces, the memory required to store the calculated pairs is small. Presumably, in the captured packets, passwords have a certain type: 12 characters, taken from 26 letters of the alphabet. The encrypted password d_0 in the captured packet data set is used to identify the associated password. To do this, first calculate $e\big(W(d_0)\big)$ to get the new encrypted password d_1, then calculate $e\big(W(d_1)\big)$ to get d_2, and so on until the encrypted password starting with twenty 0 (d_m) is displayed. The table is then checked to see which source password, P_0, the encrypted fm password is associated with. Based on P_0, the password and encrypted passwords $e_1, e_2, \dots$ are calculated until the original encrypted password d_0, denoted d_k, is generated. The password that the hacker is looking for is the one that gave rise to d_k, i.e. $(W(d_k - 1)$, which is one step earlier in the chain of calculations.

The required computation time is what it takes to find the d_m in the table plus the time it takes to compute the sequence of encrypted passwords from the corresponding password $(e_1, e_2, \dots, e_k)$ which is about a million times less than the time it takes to compute the table itself. Thus, performing a preliminary calculation and storing the results allows a hacker to get any password with a known encrypted password in a reasonable amount of time. This process takes a few seconds.

The process of decrypting the password by the rainbow tables of the novel BTFAT is presented in expression (26):

$$R_t = \begin{cases} G_0 \xrightarrow{e} e(G_0) \xrightarrow{W} G_1 \xrightarrow{e} e(G_1) \xrightarrow{W} G_2 \xrightarrow{e} \dots \xrightarrow{e} e(G_n) \\ F_0 \xrightarrow{e} e(F_0) \xrightarrow{W} G_1 \xrightarrow{e} e(F_1) \xrightarrow{W} F_2 \xrightarrow{e} \dots \xrightarrow{e} e(F_n) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \\ J_0 \xrightarrow{e} e(J_0) \xrightarrow{W} G_1 \xrightarrow{e} e(J_1) \xrightarrow{W} J_2 \xrightarrow{e} \dots \xrightarrow{e} e(J_n) \end{cases} \qquad (49)$$

Where R_t: the process of decrypting the password by the rainbow tables of the BTFAT; G_0, F_0, J_0: possible passwords; $e(G_0), e(F_0), e(J_0)$: the encrypted passwords; W: a function to convert the encrypted password.

Corollary-3: A hacker can hack any wireless network by getting a handshake. Starting with the first stolen encrypted password (B_s), the hacker applies the functions W and e repeatedly, calculating a series of encrypted passwords and final passwords, until he reaches the encrypted password with twenty 0 in front of it. The hacker then searches for this last encrypted password in the table (encrypted password E) and identifies the corresponding password (password E)**.**

Table 4 - Rainbow table

Password Q	Encrypted password Q
Password W	Encrypted password W
Password E	Encrypted password E
Password R	Encrypted password R

The hacker then applies the W and e functions again, starting with the identified password, continuing until one of the received encrypted passwords in the chain matches the stolen encrypted password:

$$Password\ E \rightarrow Encrypted\ password\ 1 \rightarrow Password\ 2 \rightarrow Encrypted\ password\ 2$$
$$\rightarrow Password\ 3 \dots \rightarrow \cdots Password\ 33 \rightarrow Encrypted\ password\ 34 \begin{bmatrix} a\ match\ to\ encrypted \\ password\ B_s \end{bmatrix}$$

Where $Password\ E, Password\ 1{,}2{,}3 \dots$: intermediate and final password required for hacking Wi-Fi; $Encrypted\ password\ 1, 2, \dots$: encrypted password located in the handshake.

An encrypted password that matches (*Encrypted password 34*) will mean that the previous password (*Password 33*) from which it was obtained is associated with the stolen encrypted password. To set the first and last columns of the rainbow table, you need to perform a lot of calculations. They store only the data in these two columns, and by recalculating the chain, hackers can identify any password by its encrypted password located in the handshake.

Experimental results

This section contains the proposed BTFAT. To demonstrate the advantages of choosing this tool for calculating the values of such characteristics as reliability, efficiency, and time of user de-authorization during hacking of a wireless network, these data were also calculated for three other tools (Reaver, Wifite, Wireshark).

Network hacking requires the following components, which are described in Table 5.

Table 5 - Components for hacking Wi-Fi

Components	Version/The name of the system
Personal computer	x64
Operation system	Linux Kali 5.9.0
Wireless access point	D-linkDIR-615
Resolution	1920x1080 px
Processor	Intel(R) Core (TM) i7-8750H
Maker	Acer
RAM	2048 MB
Video memory	16 MB
HARD Disk	39,9 GB (/dev/sda1)
CPU MHz	2208.002
Cash size	9216 KB

Based on the results, the following metrics are measured.

- Effectiveness of packet capture
- Client de-authorization time
- Reliability
- Processing performance of sent requests.

A. Effectiveness of Packet-capture

The effectiveness of using a particular method when hacking Wi-Fi is calculated by the equation:

$$E = \frac{P}{T} \times 100\% \qquad (50)$$

Where *E:* the effectiveness of packet capture; *P:* the number of packets captured; *T:* the time taken for a packet capture.

The data for calculating the packet capture effectiveness is given in Table 3 and Table 4. When calculating the effectiveness, the number of captured packets is considered. Table 3 and Table 4 show the number of packets captured by various tools within 50 seconds. The largest number of packets in a time equal to 50 seconds was captured by the BTFAT (48.5), the smallest number of packets – by Wifite (41). Figure 4 shows the effectiveness of packet capture using the BTFAT, Reaver, Wifite, and Wireshark tools. Figure 4 shows that BTFAT (97%) has the highest effectiveness. Figure 4 also shows that the effectiveness of the BTFAT increases over time.

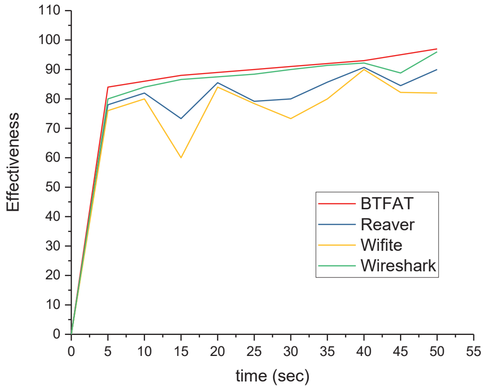

Figure 4 - The effectiveness of packet capture

B. Client de-authorization time

The client de-authorization time depends on the number of requests made by the hacker and the responses received from the client, as well as the speed of sending requests. The de-authorization time is calculated using the equation:

$$t = \frac{N_r \times N_a}{V_c} \qquad (51)$$

Where *t:* the client de-authorization time; N_r: the number of requests; N_α: the responses received from the client; V_c: the speed of sending requests.

Data for calculating the de-authorization time are given in Table 5 and Table 6. Figure 5 shows the client de-authorization time for each tool. If the speed of sending requests is the same for all tools, then calculating the de-authorization time by the equation (51), it is noticeable that the de-authorization time increases with the passage of time and the requests sending. Figure 5 shows that the BTFAT sent 50 requests and the de-authorization time took 116.6 seconds. Thus, the BTFAT can complete the authorization process in a shorter time compared to other tools, which contributes to faster handshake establishment for password decryption.

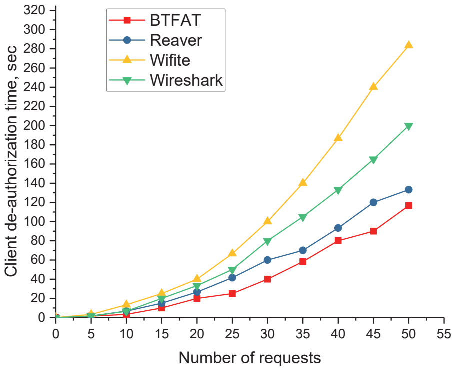

Figure 5 - Client de-authorization time

Figure 6 shows that for BTFAT, even with an increased number of requests, the pre-authorization time is minimal compared to other tools.

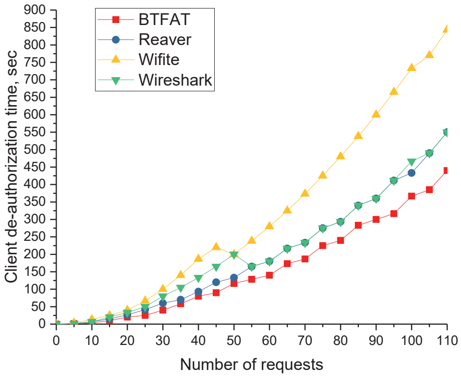

Figure 6 - Client de-authorization time

Figure 7 shows the relationship between the number of requests sent to the user and the number of responses received from the user. Figure 7 shows that the smallest number of responses received was accepted by the BTFAT (6). This means that the BTFAT requires less resources and time to intercept the handshake, as fewer requests are processed.

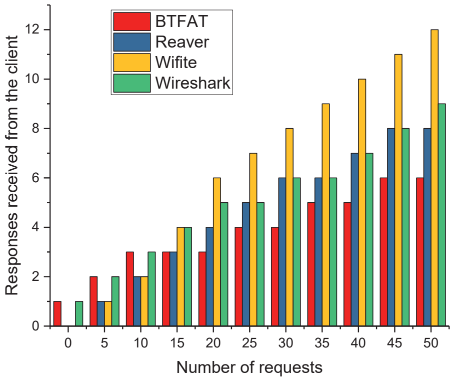

Figure 7 - Elements of the de-authorization process

Figure 8 shows the relationship between user requests and responses. The duration of a network hacker attack depends on the number of responses received as a result of requests sent by the hacker. The fewer responses received from the client and the user are de-authorized, the faster the user processes requests and the wireless network is hacked.

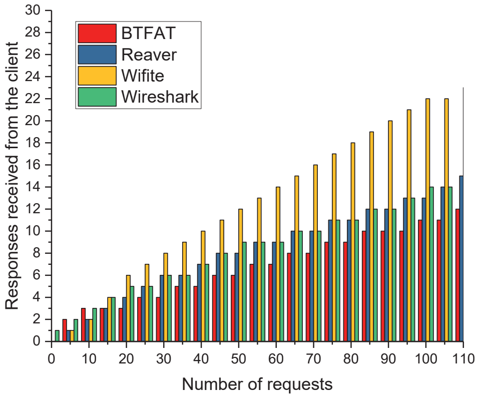

Figure 8 - Elements of the de-authorization process

During the de-authorization process, the time of this process depends on such elements as the number of requests and responses, and the speed of sending requests. Figure 9 shows the correlation between speed and time, as well as between the time and number of client responses. Figure 9 shows that the correlation values in the upper graph are less scattered, which means a higher correlation. In the lower graph, the values are more scattered, which means a high correlation. Table 9, showing the correlation coefficient of each element of the de-authorization process, demonstrates a 92% correlation between the time spent on client de-authorization and the number of responses received as a result of requests. This means that the de-authorization time is highly dependent on the number of responses received. The correlation between the speed of sent packets and the de-authorization time is 53%, and an average noticeable relationship is formed. This means that the de-authorization time is weakly dependent on the speed of sending packets.

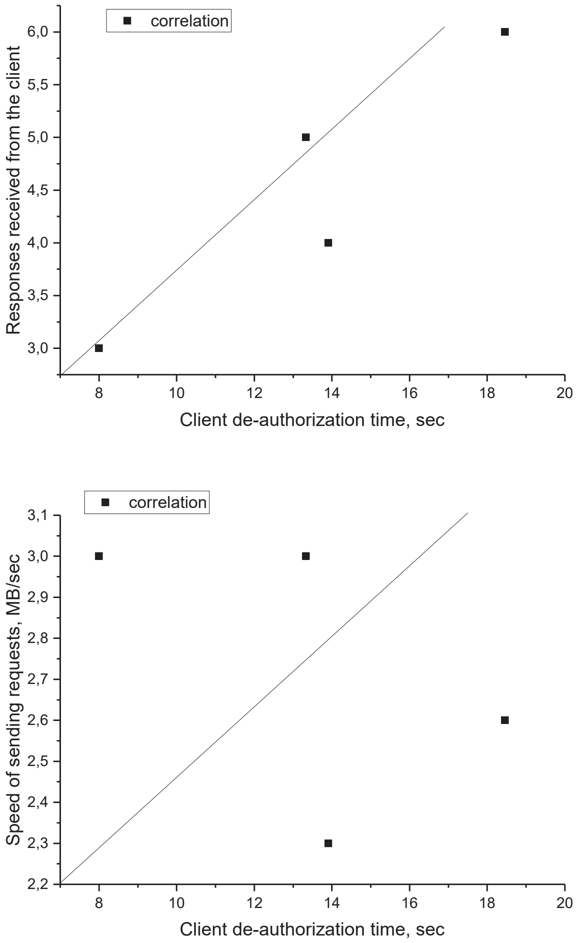

Figure 9 - Correlation dependence

C. Reliability

The reliability of Wi-Fi hacking is characterized by the probability of password decryption, which is determined by the following equation (52):

$$P_c(t) = \frac{N_0 - \sum n_i}{N_0} \qquad (52)$$

Where $P_c(t)$: reliability of Wi-Fi hacking; N_0: the number of initially captured packets; $\sum n_i$: the total number of requests.

Figure 10 presents the percentage of reliability of the network hacking process for each tool. Data for calculating the reliability of using each tool are shown in Table 10 and Table 11. Figure 10 shows that the BTFAT has the highest reliability (86%), and the Wifite tool has the lowest reliability (66%). Also, Figure 8 reveals that over time, the reliability of packet capture using the BTFAT remains higher than with other tools.

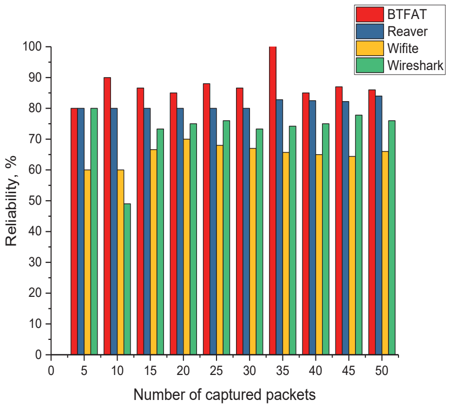

Figure 10 - Reliability of the network hacking process

D. Processing performance of sent requests

The processing performance of the requests sent to the user affects the time of the de-authorization process, as well as the process of handshake interception. Therefore, this parameter affects the total time of Wi-Fi hacking. The higher the performance, the faster a hacker can crack the Wi-Fi. The processing performance of the requests sent to the user is calculated using the equation:

$$P_p = \frac{\beta \times C}{\omega} \times 100\% \qquad (53)$$

Where P_p: the processing performance of sent requests to the user; β: the number of processed responses; C: the channel capacity; ω: the request processing time.

Figure 11 shows that the processing performance of requests sent by the BTFAT is stable compared to other tools and is equal to 85%. Also, the BTFAT has the highest performance, which contributes to the fastest Wi-Fi hacking. The data for calculating performance is described in Tables 12 and 13.

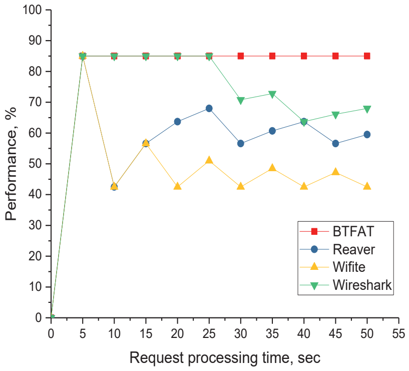

Figure 11 - The processing performance of the requests sent to the user

Discussion of results

The proposed BTFAT consists of three stages. The first stage is packet capture, the second is user de-authorization, and the last is Wi-Fi Blockchain-Featured Hacking Process. The advantages of using BTFAT is the use the features of Blockchain technology that capture the packets effectively, reduction of the user de-authorization, and the reliability. The packet capture effectiveness is 97%, which is higher than that of the other tools. The user de-authorization time with the BTFAT is more effective as compared to other tools. This time is minimal when compared with other state-of-the-art tools. Thus, it proves that the BTFAT tool takes less time to perform de-authorization of the user, so the minimum amount of time is needed to intercept a handshake. The reliability of hacking a wireless network with BTFAT is 86%, which is the highest indicator. Table 3 shows the comparative analysis of the proposed BTFAT tool and other contending tools.

Another advantage of this tool is that BTFAT works with any wireless network adapters whose driver supports the monitoring mode. Also, the advantage of this tool is its extensive functionality. In addition to cracking WEP/WPA/WPA2 keys, BTFAT can decrypt intercepted traffic with a known key, analyze traffic, create a virtual tunneling interface, create encrypted packets for injection, provide techniques for attacking the client, remove WEP masking from PCAP files, store and manage lists of ESSIDs and passwords, calculate paired master keys, and open access to the wireless network card from other computers. However, this method of hacking Wi-Fi has disadvantages. The main disadvantages are the slow speed of password search and the lack of tables with pre-calculated hashes for password selection.

Table 3 - Comparative analysis of the proposed BTFAT, Reaver, Wifite and Wireshark tools

Name of tools	Effectiveness of packet capture	Client de-authorization time		Responses received from the client	Responses received from the client	Reliability	Processing performance of sent requests
		55 Request	110 Request	55 Requests	110 Requests	Maximum captured 50 packets	Request processing 50 seconds
BTFAT	97%,	116.6	192	06	12	86%	85%
Reaver	87.4%	133.3	551.2	08	16	84%	59.5%
Wifite	74.3%	283.3	796	12	24	66%	47.2%
Wireshark	94.3%	200	552	9	18	76%	66.1%

Conclusion

This paper introduces a Blockchain-featured BTFAT for controlling the hacking of the wireless network. It also provides a detailed description of the wireless network hacking process. The Wi-Fi hacking process occurs in three phases. In the beginning, packets are captured by monitoring and saved to a file, then the user is de-authorized, and the handshake is recorded in a previously saved file. The last phase consists of Blockchain technology features, which are used for controlling the hacking process of the wireless network and decrypting the password. The advantage of the proposed BTFAT is that the BTFAT can hack networks that use Blockchain technology, given that networks with such technology have very high security. Another advantage is the speed of using this method, its efficiency, and reliability.

The expressions have been used to calculate the values of efficiency, reliability, and user de-authorization time, and a comparative analysis of several tools for hacking Wi-Fi was performed. The reliability of using the BTFAT is 86%, efficiency - 97%, the request processing performance time is 85%. Furthermore, the time of detecting the Wi-Fi-hacking is minimal compared to other existing state-of-the-art tools. These results show that the proposed BTFAT is the best choice for Wi-Fi-hacking prevention. In the future, we will model the pen-testing process with BTFAT for evaluating the wireless network security metrics.

REFERENCES

[1] Cisar, P., and S. Maravic Cisar. "Ethical hacking of wireless networks in kali Linux environment." Annals of the Faculty of Engineering Hunedoara 16.3 (2018): 181-186.

[2] Astudillo, Karina. Wireless Hacking 101. Babelcube Inc., 2017.

[3] Karagiannis, Konstantinos. "Hacking Blockchain." (2017).

[4] Venkatesh, V. G., et al. "System architecture for blockchain based transparency of supply chain social sustainability." Robotics and Computer-Integrated Manufacturing 63 (2020): 101896.

[5] Werbach, Kevin. The blockchain and the new architecture of trust. Mit Press, 2018.

[6] Sinha, Sanjib, Sanjib Sinha, and Karkal. Beginning Ethical Hacking with Kali Linux. Apress, 2018.

[7] Ansari, Juned Ahmed. Web penetration testing with Kali Linux. Packt Publishing Ltd, 2015.

[8] Guo, Rui. "Survey on WiFi infrastructure attacks." International Journal of Wireless and Mobile Computing 16.2 (2019): 97-101.

[9] Pimple, Nishant, et al. "Wireless Security—An Approach Towards Secured Wi-Fi Connectivity." 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2020.

[10] Noshad, Zainib, Nadeem Javaid, and Muhammad Imran. Analyzing and securing data using data science and blockchain in smart networks. Diss. MS thesis, COMSATS University Islamabad (CUI), Islamabad 44000, Pakistan, 2019.

[11] Swedan, AbedAlqader, et al. "Detection and prevention of malicious cryptocurrency mining on internet-connected devices." Proceedings of the 2nd International Conference on Future Networks and Distributed Systems. 2018.

[12] Kabanov, P. A., and Mikhail Sergeevich Sukhodoev. "Overview of hacking tools and protection of modern ICT devices." 14th International Forum on Strategic Technology (IFOST-2019), October 14-17, 2019, Tomsk, Russia:[proceedings].—Tomsk, 2019.. 2019.

[13] Goyal, Piyush, and Anurag Goyal. "Comparative study of two most popular packet sniffing tools-Tcpdump and Wireshark." 2017 9th International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, 2017.

[14] Astudillo, Karina. Wireless Hacking 101. Babelcube Inc., 2017.

[15] Vance, William. Linux for Hackers: A Comprehensive Beginners Guide to the World of Hacking using Linux. joiningthedotstv, 2020.

[16] Таганов, П. А. "Исследование алгоритма атаки на беспроводную сеть Wi-Fi." Организатор конференции. 2018.

[17] Parasram, Shiva VN, et al. Kali Linux 2018: Assuring Security by Penetration Testing: Unleash the full potential of Kali Linux 2018, now with updated tools. Packt Publishing Ltd, 2018.

[18] Baloch, Rafay. Ethical hacking and penetration testing guide. CRC Press, 2017.

[19] Carranza, Aparicio, et al. "Automated Wireless Network Penetration Testing Using Wifite and Reaver." Global Partnerships for Development and Engineering Education: Proceedings of the 15th LACCEI International Multi-Conference for Engineering, Education and Technology, July 19-21, 2017, Boca Raton, FL, United States. Latin American and Caribbean Consortium of Engineering Institutions, 2017.

[20] Carranza, Aparicio, et al. "Automated Wireless Network Penetration Testing Using Wifite and Reaver." Global Partnerships for Development and Engineering Education: Proceedings of the 15th LACCEI International Multi-Conference for Engineering, Education and Technology, July 19-21, 2017, Boca Raton, FL, United States. Latin American and Caribbean Consortium of Engineering Institutions, 2017.

[21] Martin, Alexander, Basiru Mohammed, and Rajkumar Ramadhin. "WEP VS WPA2 Encryptions." (2019).

[22] Alassouli, Hidaia Mahmood. Hacking of Computer Networks. Dr. Hidaia Mahmood Alassouli, 2020.

[23] Al Neyadi, Eiman, et al. "Discovering Public Wi-Fi Vulnerabilities Using Raspberry pi and Kali Linux." 2020 12th Annual Undergraduate Research Conference on Applied Computing (URC). IEEE, 2020.

[24] Pimple, Nishant, Tejashree Salunke, Utkarsha Pawar, and Janhavi Sangoi. "Wireless Security—An Approach Towards Secured Wi-Fi Connectivity." In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 872-876. IEEE, 2020.

[25] Pandikumar, T., and Mohammed Ali Yesuf. "Wi-Fi Security and Test Bed Implementation for WEP and WPA Cracking." International Journal of Engineering Science 13571 (2017).

[26] Sharma, Himanshu. Kali Linux-An Ethical Hacker's Cookbook: Practical recipes that combine strategies, attacks, and tools for advanced penetration testing. Packt Publishing Ltd, 2019.

[27] Bullock, Jessey, and Jeff T. Parker. Wireshark for Security Professionals: Using Wireshark and the Metasploit Framework. John Wiley & Sons, 2017.

[28] Li, Lei, Zhigang Li, Hossain Shahriar, Rebecca Rutherfoord, Svetana Peltsverger, and Dawn Tatum. "Ethical Hacking: Network Security and Penetration Testing." (2018).

[29] Sinha, Sanjib. "Hashes and Passwords." Beginning Ethical Hacking with Kali Linux. Apress, Berkeley, CA, 2018. 323-345.

[30] Santo Orcero, David. Kali Linux. Grupo Editorial RA-MA, 2018.

[31] Sinha, Sanjib. "MAC Address." Beginning Ethical Hacking with Python. Apress, Berkeley, CA, 2017. 191-194.

Разак А., Әділ А.Ж., Аманжолова С.Т.

Блокчейн технологиясына негізделген Wi-Fi хакерін анықтаудың жаңа құралы

Аңдатпа. Wi-Fi бизнес, білім беру, өнеркәсіп және т.б. көптеген салаларда маңызды рөл атқарады, екінші жағынан, Wi-Fi осалдықтары пайдаланушылардың мәліметтерінің құпиялылығына зиян келтіреді, егер осалдықтар дұрыс өңделмесе. Кейбір хакерлер бұзу процесіне әкелетін Wi-Fi осалдығын пайдалану үшін Linux құралын пайдаланады. Бұл мақалада Wi-Fi желісінің қауіпсіздігін жақсарту үшін Blockchain Technology-Featured Novel Air-Cracking tool (BTFAT) ұсынылған. Құрал құнды функциялардан тұрады (мысалы, бақылау, сканерлеу, бұзу және тестілеу). Бұл функциялар желінің осалдықтарын анықтауға көмектеседі, BTFAT C тілінде бағдарламаланған. Эксперимент нәтижелеріне сүйене отырып, BTFAT басқа қолданыстағы әдістермен салыстырғанда жоғары өнімділікті қамтамасыз етеді.

Кілт сөздер: Wi-Fi, осалдық, BTFAT, құпиялылық, сенімділік, тестілеу, Blockchain технологиясы.

Разак А., Әділ А.Ж., Аманжолова С.Т.

Новый инструмент для обнаружения взлома Wi-Fi на основе технологии блокчейн

Аннотация. Wi-Fi играет важную роль во многих областях, таких как бизнес, образование, промышленность и т. д. С другой стороны, уязвимости Wi-Fi наносят ущерб конфиденциальности данных пользователей, если уязвимости не обрабатываются должным образом. Некоторые хакеры используют инструмент Linux для использования уязвимости Wi-Fi, которая приводит к процессу взлома. В этой статье представлен новый инструмент Blockchain Technology-Featured Novel Air-Cracking tool (BTFAT) для улучшения безопасности сети Wi-Fi. Инструмент состоит из ценных функций (например, мониторинг, сканирование, взлом и тестирование). Эти функции помогают обнаружить уязвимости сети, запрограммирован на языке С. Основываясь на результатах эксперимента, BTFAT обеспечивает более высокую производительность по сравнению с другими существующими методами.

Ключевые слова: Wi-Fi, уязвимость, BTFAT, конфиденциальность, надежность, тестирование, технология блокчейн.

Авторлар туралы мәлімет:

Абдул Разак, «Киберқауіпсіздік» кафедрасының профессоры, Халықаралық ақпараттық технологиялар университеті.

Әділ Алтынай Жанарбекқызы, «Компьютерлік инженерия» кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті.

Аманжолова Сауле Токсановна, «Киберқауіпсіздік» кафедрасының меңгерушісі, Халықаралық ақпараттық технологиялар университеті.

Валиев Бахытжан Бауржанович, «Компьютерлік инженерия» кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті.

Сведения об авторах:

Абдул Разак, профессор кафедры «Кибербезопасность», Международный университет информационных технологий.

Әділ Алтынай Жанарбекқызы, магистрант кафедры «Компьютерная инженерия», Международный университет информационных технологий.

Аманжолова Сауле Токсановна, заведующая кафедрой «Кибербезопасность», Международный университет информационных технологий.

Валиев Бахытжан Бауржанович, магистрант кафедры «Компьютерная инженерия», Международный университет информационных технологий.

About the authors:

Razaque A., Professor, Department of Cybersecurity, International Information Technology University.

Adil A.Zh., Master student, Department of Computer Engineering, International Information Technology University.

Amanzholova S. T., Head of the Department of Cybersecurity, International Information Technology University.

Valiyev B.B., Master student, Department of Computer Engineering, International Information Technology University.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 38–41
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.004
УДК 336; 336.61

Auken V.M.
The Standing Committee on Public Accounts, Nur-Sultan, Kazakhstan.
vauken@hotmail.com

INTERACTION ANALYSIS OF GOVERNMENT REVENUE AND AUDIT

Abstract. The article examines the relationship between unforeseen government revenue and the actual audit conduct. It has been found that overprofits can worsen the public organizations' work, while government audits increase their effectiveness. The findings were obtained through macroeconomic management modeling that was structurally evaluated using econometric formulas and quasi government data. Based on the estimated model, the correlation between over income and the number of audits on the public administration's effectiveness has been determined.

Keywords: government auditing, government efficiency, management model, super profits, Bernoulli formula, optimal strategy, game theory, determinism, correlation, exogenity, endogenity

Introduction

The existing annual reports on the performance of regional audit commission budgets were highly useful for examining the role of government auditing in the work of local governments in Kazakhstan. At the same time, "the subject of the content of the analysis was the posting of the conclusion to the Annual Report on the performance of regional budgets for 2017 or 2018 on the online resource of the audit commission." The study found that "only four audit commissions have placed conclusions to the annual performance reports of regional budgets for 2018, despite the fact that the annual reports of all local budgets have been approved" [1].

Some regional mayors have been subjected to repeated checks. Consequently, it is an excellent research opportunity to compare the levels of audits of local government bodies, which vary in the nature of the work carried out, depending on the regional specifics. And in the future, it is an opportunity to assess the impact of past audits on the subsequent work of local administrations, which face the same expected probability of conducting a government auditing.

Purpose of the research

To build a model of the influence of the government auditing, its unscheduled inspection of the work of the local administration. Then, to examine how the audit and the likelihood of a new one will force the state administration to change its beliefs and own actions in relation to the subsequent financial transactions and the work in general.

Materials and research method

In the model, the factor of the unpredictability of government auditing is important, which subsequently has a significant impact on the number of financial violations. Pre-alert audit in a short period of time: in a few days – has a positive effect on the effectiveness of the audit. Otherwise, given a longer waiting time, the public administrations are more carefully prepared for audit in an attempt to cover up violations. This conclusion is not new to the model in question, as well as to others.

Background and data

Kazakhstan is a centralized unitary country, in which regional bodies are headed by officials appointed from the Center, with the provision of certain transfers from the national budget. The mayors of the regions make decisions together with local legislatures, oil districts, on the expenditure of budget funds. Taking into account the probabilistic central oversight by the Standing Committee on Public Accounts, in the established practice, there are certain conflicts on the development of funds annually. In fact, "local government in Kazakhstan is inefficient and under-demanded by the state" [2].

General violations can occur in unfinished community service, paid but unfinished, and the use of forged documents and fictitious organizations that exist only nominally. Financial irregularities, however, are due to scams in the procurement of goods and services, diversion of funds and excessive billing for goods and services.

As a rule, such distribution is obtained with large transfers from the national budget with the existing judicial procedures of insolvency, as well as minimal media activity on the ground [3].

Studies by Eric Avis and Frederico Finana of the University of California, Berkeley, and Claudio Feraz of the Pontifical Catholic University in Rio de Janeiro found that the level of corruption is about 8% lower among local government agencies audited compared to those not exposed. It is also shown that auditing can lead to significant legal costs. It has been consistently proven that the disciplinary effects of trials can account for the 72% reduction in corruption on the ground. In accordance with this, model estimates in this case document an increase in the probability of legal action by 20% after the audit [4].

In the conditions of Kazakhstan, "the basis for the implementation of control are exclusively the annual and quarterly plans of the Standing Committee for Public Accounts to monitor the implementation of the national budget" [5]. Control on behalf of the President of the Republic of Kazakhstan, the Administration of the President of the Republic of Kazakhstan, and the requests of the deputies of the Parliament of the Republic of Kazakhstan is carried out based on the relevant amendments and additions to the quarterly plan of the Accounts Committee. Based on preliminary examination of the monitoring facility, a monitoring plan is drawn up and an external government auditing program is being developed. The Standing Committee may re-examine the activities of the monitoring facility if the previous audit of the standards of state financial control does not comply [5].

In accordance with most models of public administration, the audit plays a positive role in its suddenness and the timing of the notification of the holding. Thus, local governments will only be able to calculate the probability of the next public audit, and their alert within two days will affect the effectiveness of work. Faced with this uncertainty, it seems likely that local governments will seek to take into account the risks of such a sudden audit in their work, extracting information from both their own experience and the neighboring regional administrations.

Interpreting the main conclusions of the E. Avis, F. Finan and K. Feraz model, it is intuitive that reducing the level of corruption due to the conduct of government auditing will lead to increased responsibility of the administration, decisions and their efficiency. In turn, such a check, especially at the time of appointment or reassignment, at the same time, has a positive effect on the quality of personnel, the so-called selection effect, and on improving the efficiency of the administration as a whole [6].

Finally, the existing research on the relationship between appointed responsibility and political action reveals that incumbent officials are responding to incentives for reassignment. Brazil's Bolsa Escola program, a local-oriented conditional remittance program, has performed far better in those administrations where leadership has been encouraged to reelect.

In Kazakhstan, "the fact that the modern development" of the country "has a clear regional context," it is still necessary to conduct a "state regional policy that will aim to smooth the differences between regions" [7].

Based on the data obtained by the quasi-public sector (further quasi-sector) we will build our own model, starting with the calculation of the parameters of the linear equation of multiple regression:

$$y = \alpha + b_1x_1 + b_2x_2 + \varepsilon \qquad (1)$$

where y – is a profitable part, x – is a consumable part consisting of x_1, capital turnover and x_2, used capital, $\forall\varepsilon > 0$, a small amount with a normal distribution of mathematical expectation of $\mu = 0$ and an average deviation σ^2_ε.

In this case, the parameters of the equation (1) will be found from the solution of the system of equations:

$$\begin{cases} \bar{y} = a + b_1\overline{x_1} + b_2\overline{x_2} \\ \overline{yx_1} = a\overline{x_1} + b_1\overline{x_1^2} + b_2\overline{x_1x_2} \\ \overline{yx_2} = a\overline{x_2} + b_1\overline{x_1x_2} + b_2\overline{x_2^2} \end{cases} \qquad (2)$$

Model

In order to understand how the government auditing affects the work, we will use the existing model of E. Avis, F. Finan and K. Feraz and the model of political responsibility on the basis of the career ladder of B. Holmstrom, T. Persson and G. Tabellini [8]. To simplify, we will develop a continuous model at the initial stage, taking into account the impact of the uncertainty factor of unscheduled audits. Local government administrations will be characterized by personality traits throughout their tenure:

$$X_i(\text{gender, education, position}), \qquad (3)$$

where $i = \overline{1, n}$ – the number of local administration workers dependent on three variables.

Then *F's* functional ability will be expressed through the personality qualities of X_i and cognitive ε_i:

$$F_i(X) = X_i + \varepsilon_i, \tag{4}$$

where $a \in R$ and $a>0;$ $\forall \varepsilon i>0$, small value with normal distribution of mathematical expectation $\mu = 0$ and mid-square deviation σ_ε^2.

$$R_i^t(p,X) = F_i^t(X) + E_i^t(p,X), \tag{5}$$

where E_i – efforts to work, depending on personality and probabilities, p, conducting time checks, t, by Bernoulli: $p_i^t = 1$, if the government auditing has taken place in time t and $p_i^t = 0$, if not.

The expression (5) can also be rewritten as a

$$R_i^t(p,X) = \alpha X_i + \varepsilon_i + E_i^t(p,X) \tag{6}$$

If financial irregularities are detected in the course of audit conduct with the probability, p_i^t, for the time, t, the local administration will be punished, expressed function of costs:

$$C_i^t(p,X) = c_i^t(p_i^t, E_i^t(p,X) \tag{7}$$

At the same time, the result of the audit will affect the overall efforts of the administration and efficiency in general, as well as the expected administrative management decisions.

In general, the local administration solves the task of optimizing the utility function, U, maximizing their performance by improving efficiency and minimizing costs, being more responsible for decisions:

$$max_x U(p,X) = R_i^t(p,X) - \ C_i^t(p,X) \tag{8}$$

It is important to note that local state bodies are politically “juggling” between the central authorities, represented by the controlling structures, and the appointees, who, in turn, are guided by the rating of the effectiveness of management.

Results and discussion

Preliminary results and calculations of the model show that the local administration, in which the government auditing was conducted, commits 7.9% less financial violations than those in which it was not.

However, the lack of information of government agencies and the high probability of repeated unplanned inspections significantly increases their effectiveness and the functioning of the administration in general. Of course, this estimated effect only covers short-term equilibrium effects. If there are side effects, estimates and their true impact in the long term are likely to be underestimated. All of this will create a need for further external resources, such as the media and certain judicial procedures, to consolidate the short-term results of the government auditing.

Despite the importance of legal liability for financial irregularities, understanding how to conduct government audits more effectively remains limited for the time being, especially when such violations are endemic. In this sense, more research is urgent here to better understand the nature of public auditing in today’s environment, as well as the emerging conflicts caused by audits when financial irregularities are detected.

Conclusion

This article shows that government audits, unscheduled inspections can be the main policy in the fight against financial violations in the country. In the case of the local government authorities, where such audits were carried out, there subsequently occurred about 8% less violations.

REFERENCES

1. Айтенова Ш.А., Спанов М.У. Контент-анализ официальных интернет ресурсов ревизионных комиссий // Государственный аудит. – 2019. - №2 (43). – С. 33.

2. Положение о Счетном комитете по контролю за исполнением республиканского бюджета. Указ Президента Республики Казахстан от 5 августа 2002 года №917.

3. From Kazakhstan General Prosecutor's office website

4. Eric Avis, Claudio Ferraz, Frederico Finan. 2017. "Do Government Audits Reduce Corruption? Estimating the Impact of Exposing Corrupt Politicians".

5. From Kazakhstan Standing Committee on Public Accounts website

6. Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." The Quarterly Journal of Economics 123 (2):703–745.

7. Жамкеева 7М.К., Тузубекова М.К., Жумагулова А.К. Проблемы государственного регулирования развития регионов // Государственный аудит. – 2019. - №2 (43). - С. 36

8. Persson, Torsten and Guido Tabellini. 2000. Political Economics. Cambridge: MIT Press.

9. Auken, V.M. 2017. "University Performance Measurement in Quality Assurance." Modern mathematics and its applications: Papers of the International Scientific-Practical Conference, 18-20 May 2017, Ufa.– Part I. / Chief Editor S.A. Mustafina. – Sterlitamak: Sterlitamak Branch of the Bashkir State University. – 416 p. – ISBN 978-5-86111-585-8.

Аукен В.М.

Мемлекеттік кірістер және аудиттің өзара әсерлері

Андатпа. Мақалада күтпеген мемлекеттік кірістер мен нақты аудит жүргізудің арақатынасы қарастырылады. Артық пайда қоғамдық ұйымдардың жұмысын нашарлатады, ал мемлекеттік аудит олардың тиімділігін арттырады. Нәтижелер квазимемлекеттік деректер пайдаланылған эконометрикалық формулалар арқылы құрылымдық бағаланған макроэкономикалық басқаруды модельдеу арқылы алынды. Болжалды үлгінің негізінде мемлекеттік басқару тиімділігіне тексерулер саны мен артық кіріс арасындағы корреляция анықталды.

Түйін сөздер: Мемлекеттік аудит, мемлекеттік тиімділік, басқару моделі, супер пайда, Бернулли формуласы, оңтайлы стратегия, ойын теориясы, детерминизм, корреляция, экзогендік, эндогендік.

Аукен В.М.

Анализ взаимодействия государственных доходов и аудита

Аннотация. В статье исследуется взаимосвязь непредвиденных государственных доходов и фактического проведения аудита. Выяснилось, что сверхприбыль может ухудшить работу общественных организаций, а государственный аудит повысит их эффективность. Результаты были получены с помощью моделирования макроэкономического управления, которое было структурно оценено с использованием эконометрических формул с использованием квазигосударственных данных. На основе оценочной модели была определена корреляция между избыточным доходом и количеством проверок эффективности государственного управления.

Ключевые слова: государственный аудит, эффективность государственного управления, модель управления, сверхприбыли, формула Бернулли, оптимальная стратегия, теория игр, детерминизм, корреляция, экзогенность, эндогенность.

Авторлар туралы мәлімет:

Аукен Вилмур Мұратұлы, экономика ғылымдарының докторы, PhD. Мемлекеттік есеп жөніндегі тұрақты комиссия, ғылыми хатшы. Нұр-Сұлтан, Қазақстан.

Сведения об авторе:

Аукен Вильмур Муратович, доктор экономических наук, PhD. Постоянный комитет по государственным счетам, Ученый секретарь. Нур-Султан, Казахстан.

About the author:

Auken Vilmur Muratovich, Doctor of Economic Sciences, PhD; Secretary for Science, Standing Committee on Public Accounts, Nur-Sultan, Kazakhstan.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 42–48
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.005
UDC 378.1

Berdykulova G.M.
International Information Technology University, Almaty, Kazakhstan
g.berdykulova@edu.iitu.kz

METHODOLOGY OF TEACHING THE ECONOMIC DISCIPLINES IN DIGITAL ERA

Abstract. Underestimation of the scientific achievements of the post-industrial theory and disruptive innovations affect the quality of educational curricula and syllabi of the disciplines taught to economic majors, thereby ignoring the requirements of a new civilization and a new paradigm generated by the digital age. The article explains how new economic knowledge should be introduced in the content of such subjects as Economic Theory and Economics and Industrial Engineering. One of the ways to harmonize science and educational practice is updating the methodology of teaching economic disciplines. While teaching the Economic Theory and Economics and Industrial Engineering we introduced original examples of post-industrial society and breakthrough innovations in the process of digitalization in Kazakhstan, the theory of learning and communication management, behavioral science, research methodology, rhetoric model, the principle of specificity and scientific knowledge. To overcome the negative impact of disruptive innovations on the process of digitalizing education, the author suggests the method of synopsis writing based on tabulation, paraphrasing and forming questions for each lecture topic followed by online group and other communication activities during practical lessons. Implementation of the improved methodology of teaching economic disciplines in the study of other subjects at IITU and at an educational center in Almaty has demonstrated its effectiveness.

Keywords: teaching methodology, economic discipline, digitalization, principle of specificity, principle of scientific knowledge, written rhetoric, oral rhetoric

Introduction

The fourth technological revolution has brought about transformations in social and economic life associated with digitalization. There have taken place radical changes in almost all spheres of society, including the system of education. The university life is paramount to socio-economic development of the state, because it is higher educational institutions that form the intellectual, cultural, creative potential of society and are responsible for the preparation of the future specialists of a high level. The development of society is inextricably linked with the progress of education. Educational activities of the higher schools should be targeted on fostering a generation that is able to bring the best values of human society into daily life.

The article published in the Goggle Scholar base has identified four trends connected with the introduction of digital technologies and tools into the educational process: the formation of a blended learning model; transition to online learning; creation of a virtual (digital) educational environment; changing the approach to the management of educational organizations [1]. Higher education can lead to many benefits, including a prosperous career and financial security. In the 21st century, education plays an even more significant role in other aspects of personal life. Attaining a higher education level can increase opportunities and improve the overall quality of life [2]. Earlier, higher education was meant to provide the labor market with qualified personnel and many Americans viewed higher education as a path to a “good job.” Employers’ requirements to university graduates include teamwork, written and oral communication, ethical decision-making, critical thinking skills and the ability to apply knowledge in real-world settings. With the passage of time and under the influence of rapid changes in this century, the learning outcomes also include those skills that, in the opinion of many higher educational institutions, will prepare graduates not only for the job but also for an active life and interested citizenship [3].

A need to find out how changes in the methodology of economic science should be reflected in the content of economic disciplines, specifically, Economic Theory and Economics and Industrial Engineering became a driver of writing this article. The State Program for the Development of Education and Science of the Republic of Kazakhstan for 2020-2025 prioritizes raising the status of the teaching profession, modernizing teacher education, introducing a vertical system of administration and financing of education [4]. The tasks of the previous program implemented before 2020 seem to be quite relevant. In particular, modernization of the system of technical and vocational education; achieving a high level of quality in higher education that meets the needs of the labor market, the tasks of industrial and innovative development of the country and the individual, compliance with the best world practices in the field of education [5].

A review of the relevant literature in the public domain shows that the focus is on the problems of teaching methods and techniques. Whereas teaching methods represent a tool in cognition and assimilation of knowledge, the educational methodology is the doctrine in the organization of purposeful training of students. Consequently, the research methodology of Economics is fundamental for the renewal of the content of education, training and teaching.

Purpose of the research

Teaching in a new technological reality of the 21st century should meet the requirements of a new civilization structure and a new paradigm of the world community development. A rapidly changing reality requires the compliance of the teaching methodology of economic disciplines with the updated research methodology of economic theory and the innovative pedagogical and educational content of academic courses. After modifying the existing methodology of teaching economic disciplines, the quality of education will increase based on the targeted impact of updated principles, methods, technologies and a new didactic system on the cognitive process of training a competitive specialist and a creative personality. Thus, the purpose of this article is to conceptualize ways of modifying the methodology of teaching economic disciplines based on the updated economic research and new pedagogical and educational technologies in the digital era.

Materials and research methods

To achieve the goal of the study, several objectives were set. Among them, *literature review* of the manuscripts posted in the Kazakhstan Citation Base (KazBC), the Russian Science Citation Index (RSCI), Scopus bibliographic and abstract database, as well as the results of scholarly research in the open space. Acquaintance with the sources has shown that both domestic and foreign authors focus on the study of individual components of the methodology.

The directions of research cover the topical problems of teaching economic disciplines, modern forms and methods of organizing the educational process, including distance learning, the most pressing methodological issues of teaching related to the students' cognitive ability, the development of their practical skills, systemic thinking, deepening and expanding their competences. Teachers share their experience and offer their vision of the existing problems and ways of solving them. The table below presents a classification of literature on the topics related to different content components of methodology.

Table 1 - Classification of literature by teaching methodology components

Author	Source	Content
	Methodology and techniques of teaching	
L.I. Podderegina.	Methodology of teaching economic disciplines in a technical university.	Principle of methodology. System approach. Dependence on the important aspects of learning process.
L. I. Podderegin, E. M.Gainutdinov.	Methodological foundations of teaching economic disciplines (including marketing and management) in the higher education system.	Lack of domestic methodological developments.
Editorial team: T.I. Trubitsyna E.V. Ogurtsova	Techniques of teaching Economics: experience and problems.	Distance learning. Current methodological issues of teaching. Cognitive activity of students. Development of practical skills. Systemic thinking.
	Methods	
D.M.Senkebaeva.	Applying new methods of teaching Economics to college students.	The focus on new teaching methods and new technologies.
V.G.Budashevsky, K.V. Krinichansky O. N. Pastukhova.	The development of disciplines with a flexible subject area in the digital age: logical and heuristic methods and model.	Human and technological progress. Flexible subject. Logical methods. Heuristic methods. Model. Digital age.
	Principle of learning	
T Green.	Flipped classrooms: An agenda for innovative marketing of education in the digital era.	Flipped teaching, learning, and assessment ideas for marketing educators.

Oparaocha Gospel Onyema, Pokidko Daniil.	Educating the 21st century learners: are educators using appropriate learning models for honing skills in the mobile age?	A new digital environment potentiating a global diaspora of highly interactive entrepreneurial and intrapreneurial commerce. The UNESCO ICT-CFT Model.
Educational Policy		
Mohamed Ally.	Competency profile of the digital and online teacher in future education.	The forces shaping education in the future and the competencies required for the digital teacher to function effectively.
S.L. Hoe.	Digitalization in practice: the fifth discipline advantage.	Systems thinking, personal mastery, mental models, shared vision and team learning in the context of the current digitalization megatrend.
Technologies		
Patricia Altass. Sean Wiebe.	Re-imagining education policy and practice in the digital era.	Technology automation and digital Taylorism. Technology, changing communication, collaboration and knowledge creation.
Pedagogical innovation		
Orit Avidov-Ungar,Alona Forkosh-Baruch.	Professional identity of teacher educators in the digital era in the light of pedagogical innovation.	Vis-à-vis technology-integrated teaching. Institutional support. Professional identity construction of innovative teacher educators.

Compiled by the author based on sources [5], [6, [7], [8], [9], [10], [11, [12], [13], [14], [15], [16].

Proceeding from the *principles of specificity*, the original examples of teaching the disciplines "Economic Theory" and "Economics and Industrial Engineering" were used to reveal the research problem. The teaching uses the fundamental provisions of the theories of post-industrial society (PIS) and disruptive innovations (DI), constituting the basis of new scientific knowledge. Thus, the updated knowledge of teaching methods, the implementation of t*he principle of scientific knowledge,* which has enriched the teaching methodology, have significantly affected the educational process.

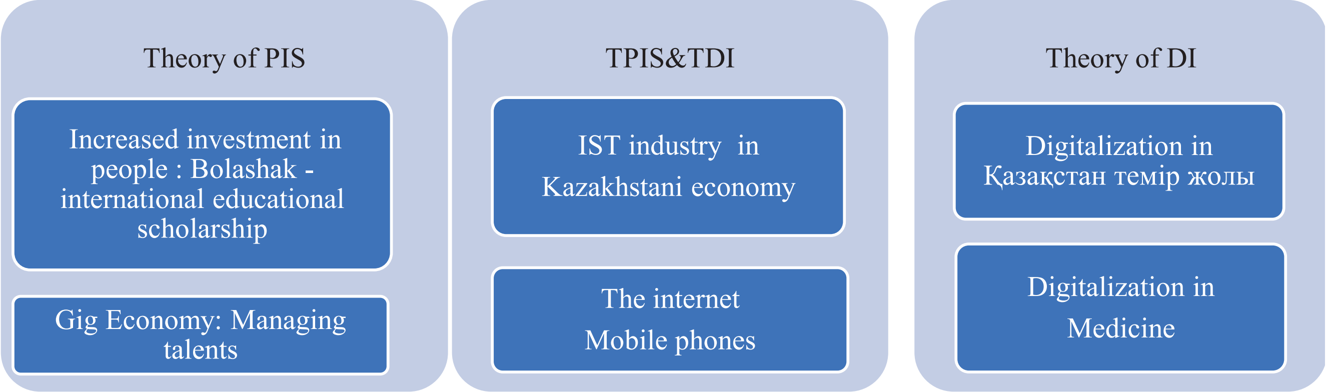
Theory of PIS
Increased investment in people : Bolashak - international educational scholarship
Gig Economy: Managing talents
TPIS&TDI
IST industry in Kazakhstani economy
The internet
Mobile phones
Theory of DI
Digitalization in Қазақстан темір жолы
Digitalization in Medicine

Figure 1 – Examples of issues of post-industrial society and disruptive innovations

In a rapidly changing world, students ought to receive advanced knowledge, therefore, the standard curricula of disciplines should be periodically revised, and scientific achievements in the field of Economics must be implemented. Hence, the *comparison of research as a model of rhetoric* becomes one of the bases of teaching methodology.

The new knowledge embraces various types of communicative interaction in oral and written speech (including the traditional rhetoric: bene dicendi scientia "the science of good speech"), natural and artificial languages. Such a methodological requirement has become especially relevant in the context of a pandemic and the transition to online education. The standard educational policy of teaching the courses "Economic Theory" and "Economics and Industrial Engineering" requiring that students must be disciplined, educated and polite in accordance with university policy and social requirements of the community was enriched during the quarantine with the development of appropriate online communication skills and competencies.

Power Point Presentation of the first introductory lecture contains a slide on online communication and management. Assignment for students' self-study is based on the recommended MOOC courses "What is communication and effective online communication: why is it important in management"? Mandatory online learning outcomes include reporting on progress in the acquisition of online communication skills [17].

Communication management and online communication

Communications management is the systematic planning, implementing, monitoring, and revision of all the channels of communication within an organization, and between organizations; it includes the organization and dissemination of new communication directives connected with an organization, network, or communications technology.

Online communication is how people communicate, connect, and transact to send, retrieve, or receive information of any kind via the internet using digital media. All the communication that is carried out via the internet is known as online communication. Because of our increasing presence online, this type of communication is becoming as important as offline communication.

https://www.youtube.com/watch?v=S7CN9Trw43w&

https://www.youtube.com/watch?v=nIQhHEWpdW

Figure 2- Communication management and online communication in teaching Economics and Industrial Engineering

Written rhetoric in digital times is extremely important due to issues of disruptive innovations. The original methodology consists in compiling a Synopsis of the lecture based on paraphrasing and questionnaires. The purpose of this activity is to make students fully understand the content of each lecture. The objectives include the requirements of keeping a personal copybook, in which the student should reflect a summary of PPT of each lecture in a tabular form, based on reading each slide, its paraphrasing, and setting appropriate questions to the text of the slide [17].

Table 2 - Sample of "My Synopsis" of the lecture with paraphrasing and questions

Number of the slide	Text of the slide	Paraphrasing	Question	Note
1.	Economics and Industrial Engineering.	The title of the introductory lecture of the discipline.	What is the name of discipline? Who delivers the lecture?	Two
2.	Ancient scholars on economy.	Explanation of the nature of economy by the ancient scholars: Hesiod, Xenophon, Aristotle and al-Farabi; their main book.	Who is the first to mention the term "economy"? Name the peculiarities of the economic concepts of the ancient scholars. Trace the links between Aristotle and al-Farabi economic heritage. Who talks on natural needs of people, values, and rules of economy?	Four.
3-15	Corresponding text.	Corresponding paraphrasing.	Corresponding questions.	Corresponding meaning.

The constant updating of academic curricula is a warrantee of the long-term practice of teaching economic disciplines, including Economic Theory and Economics and Industrial Engineering. The development of teaching methodology reflects changes in the methodology of economic theory. Therefore, the digitalization of all aspects of life, including the economic one, requires the inclusion of a section and the essence of changes caused by

ICT and digital technologies. The original updated course content includes a discussion of various emerging trends in economic theory with real-life examples and case studies. This becomes possible after examining such issues as ICT and economy, Information Kazakhstan, Digital Kazakhstan and Industrial and Innovation Policy in Kazakhstan, Post industrial Society, Digital Economy, Digital Entity and Digital University.

Results

Research on methodology of teaching economic disciplines in the digital era has shown the conceptualized ways of modifying the methodology of teaching economic disciplines based on the updated research into economic theory and new pedagogical and educational technologies in the digital era. The research findings may be categorized into the following sections.

First, the analyzed experience demonstrates examination of different aspects of methodology and techniques of teaching, principle of learning, educational policy, methods and technologies. Among them the methodology of teaching economic, marketing and management disciplines in technical universities, the cognitive activity of students, practical skills and systemic thinking, using the mobile age models to hone skills, the benefits of ICT in education for teacher’s professional development, flipped learning and teaching as an educational principle, digitalization in practice and digital Taylorism. The area of technologies is of interest in re-imagining education policy, practice, and technology-integrated teaching in the digital era.

Second, from the viewpoint of effective pedagogy, enrichment of academic programs curricula and syllabi with relevant and new knowledge, use of the principles of specificity and scientific knowledge, the traditional oral and written rhetoric as "the science of good speech" contribute to the development of teaching methodology in general and, in particular, of teaching economic disciplines. Based on these provisions, the author's contribution to the methodology of teaching economic disciplines has been tested for several years. Generalization of this experience makes it possible to systematize the research results.

Table 3 - Systematization of research results

Problem	Theory fundamentals	Issue	Methodology	
Underestimation of the new knowledge.	Post-industrial society.	Knowledge- strategic resource.	Principle of specificity.	Principle of scientific knowledge.
		Intellectual property and intellectual capital. Digitalization and freelancing.	Increased investment in people: international educational scholarship Bola1shak. Gig Economy: Managing talents.	Ongoing updating of academic curricula and syllabi.
	Disruptive innovations.	NA	Digitalization in “Қазақстан темір жолы”. Digitalization in Medicine.	Ongoing updating of academic curricula and syllabi.
Negative influence of disruptive innovation.	Learning theory.	Reluctance to write, poor handwriting, no written speech skills.	Written rhetoric.	Oral rhetoric.
		Inability to concentrate, work with text, analyze, be attentive, and diligent.	Regular execution of tasks in writing in the discipline notebook. Make notes in the discipline notebook during each lecture. Work on the lecture texts in the form of notes based on paraphrasing and formulating questions for self-examination.	NA
	Behavioral science. Theory of communication management.	Insufficient level of general culture, education and communication.	NA	Development of online communication and skills by working in groups through a defined channel of communication, discussion, group presentation, peer-review.

The following learning outcomes testify to the effectiveness of the proposed teaching methodology:

- The positive students' feedback.

- Application of the original methodology "My synopsis" to the study of other disciplines taught by other teachers.

- Implementation of the author's teaching methodology in the educational center of Almaty.

- Application of new knowledge in reports to the employer and in the situations where a company needs to overcome crisis.

- Creating a friendly atmosphere during joint online learning, building social networks, mastering time management and online communication skills, as well as familiarization with and application of cultural and ethical values in everyday life.

Conclusion

According to the results of the study, it was found that the existing educational issues in teaching Economic Theory and Economics and Industrial Engineering disciplines were associated with the need to update the curriculum. Therefore, the sections on the theory of post-industrial society and the theory of disruptive innovations have been included in the author's teaching methodology for a number of years. For these purposes, there were used the principle of specificity and the principle of scientific knowledge, written and oral rhetoric as components of the methodology of teaching economic disciplines. It became possible to deliver and receive the necessary new knowledge in the field of Economics, as well as use this methodology in the study of other subjects and in teaching at other educational institutions.

REFERENCES

1. Minina V.N. Tsifrovizatsiya Vyschego obrazovaniya i eyey sotsialnye resultaty. // Vestnik Sant-Peterburskogo universiteta. Sosiologia. 2020. T. 13. Vyp. 1. S. 84–101.

2. The Importance of Higher Education in the 21st Century. Vista College. November 26, 2019. https://www.vistacollege.edu/blog/resources/higher-education-in-the-21st-century/

3. The Integration of the Humanities and Arts with Sciences, Engineering, and Medicine in Higher Education/ Higher Education and the Demands of the Twenty-First Century. Washington (DC): National Academies Press (US); 2018 May 7.

4. Gosudarstvennaya proramma razvitia obrazovania i nauki v Respyblike Kazahstan na 2020-2025 gody. https://primeminister.kz/ru/gosprogrammy/gosudarstvennaya-programma-razvitiya-obrazovaniya-i-nauki-respubliki-kazahstan-na-2020-2025-gody-9114129

5. Gosudarstvennaya proramma razvitia obrazovania i nauki v Respyblike Kazahstan na 2011-2020 gody. https://nao.kz/blogs/view/2/105

6. Podderegina L.I. Metodika prepodavania ekonomicheskih discipline v technicheskom vuze. - 2014 г.

7. Podderegina L.I., Gainutdinov E. M. Metodicheskie osnovy prepodavania ekonomicheskih discipline v systeme vyschego obrazovania. https://rep.bntu.by/handle/data/35129

8. Trubitsina T., Ogurtsova E.V. Metodika prepodavania ekonomiki: opyt i problem. Sb. Metodich. statei. Vyp. 1 / Pod red. dozenta E.V. Ogurtsovoi. - Saratov: Izdatelski tsentr «Наука», 2010. – 72 s.

9. Senkebaeva D.M. Primenenie novyh metodov prepodavania ekonomikictudentam. Respublikanski informatsionno-metodicheski sentr "Obrazovanie" https://agartu.com/

10. Budashevsky V. G., Krinichansky K.V., Pastukhova O.N. The Development of Disciplines with a Flexible Subject Area in the Digital Age: Logical and Heuristic Methods and Model. From the book Human and Technological Progress Towards the Socio-Economic Paradigm of the Future. Berlin, Boston: De Gruyter Oldenbourg, 2020, pp. 207-216.

11. Green. Flipped classrooms: An agenda for innovative marketing education in the digital era. - Marketing Education Review, 2015 - Taylor & Francis.

12. Oparaocha G. O., Pokidko D., Educating the 21st century learners: are educators using appropriate learning models for honing skills in the mobile age? Journal of Entrepreneurship Education. Volume 20, Issue 2, 2017.

13. Ally M. Competency Profile of the Digital and Online Teacher in Future Education. International Review of Research in Open and Distributed Learning. Volume 20, Number 2, April 2019.

14. Hoe S.L., Digitalization in practice: the fifth discipline advantage. The Learning Organization, 2019. Volume 27 Issue 1.

15. Altass P. Sean, Wiebe S. Re-imagining Education Policy and Practice in the Digital Era. Journal of the Canadian Association for Curriculum Studies Vol. 15 No. 2 (2017). The University of Prince Edward Island.

16. Avidov-Ungar O., Forkosh-Baruch A. Professional identity of teacher educators in the digital era in light of demands of pedagogical innovation. Teaching and Teacher Education. Volume 73, July 2018, Pages 183-191.

17. Berdykulova G.M. Economics and Industrial Engineering. https://dl.iitu.edu.kz/course/view.php?id=13581

Бердіқұлова Ғ.М.

Цифрлық дәуірде экономиканы оқыту әдістемесі

Андатпа. Постиндустриалды қоғам теорияларының ғылыми жетістіктерін және жаңашыл инновацияларды бағаламау білім беру бағдарламалары мен экономикалық мамандықтар пәндерінің бағдарламаларының сапасына әсер етеді, осылайша цифрлық дәуір тудырған жаңа өркениет пен жаңа парадигманың талаптарын елемейді. Мақалада экономикалық теория мен экономика және өнеркәсіптік инженерия сияқты пәндердің мазмұнына жаңа экономикалық білімдерді қалай енгізу керектігі түсіндірілген. Ғылым мен оқу практикасын үйлестіру әдістерінің бірі - экономикалық пәндерді оқытудың жаңартылған әдістемесі. Экономикалық теорияны оқытуға постиндустриалды қоғам мен Қазақстандағы цифрландыру үдерісіндегі жаңа инновациялар, оқыту мен қарым-қатынасты басқару теориясы, мінез-құлық ғылымы, зерттеу әдістемесі, риторикалық модель, спецификалық принципі мен ғылыми білімдер енгізілді. және экономика. және өнеркәсіптік құрылыс. Білім беруді цифрландыру үдерісіне серпінді инновациялардың теріс әсерін жою үшін автор лекцияның әр тақырыбы бойынша кесте қоюға, перифрадациялауға және сұрақ қалыптастыруға негізделген авторлық дәріс жазбаларын жазу әдісін қолданды. Практикалық сабақтар кезінде жазбаша және ауызша риторикаға негізделген онлайн -топтық іс -шаралар мен онлайн -байланыс жүзеге асырылды. Зерттеудің тиімді нәтижесі экономикалық пәндерді оқытудың жетілдірілген әдістемесін басқа пәндерді оқуда және Алматыдағы оқу орталығында қолдану деп санауға болады.

Түйін сөздер: оқыту әдістемесі, экономикалық тәртіп, цифрландыру, ерекшелік принципі, ғылыми таным принципі, жазбаша риторика, ауызша шешендік өнер.

Бердыкулова Г.М.

Методология преподавания экономических дисциплин в цифровую эру

Аннотация. Недооценка научных достижений теорий постиндустриального общества и прорывных инноваций сказывается на качестве образовательных программ и программ дисциплин экономических специальностей, игнорируя тем самым требования новой цивилизации и новой парадигмы, порожденной цифровой эпохой. В статье объясняется, как новые экономические знания должны быть внесены в содержание таких предметов, как экономическая теория и экономика и промышленная инженерия. Одним из способов гармонизации науки и образовательной практики является обновленная методика преподавания экономических дисциплин. Оригинальные примеры постиндустриального общества и прорывных инноваций в процессе цифровизации в Казахстане, теория обучения и коммуникационного менеджмента, поведенческая наука, методология исследования, риторическая модель, принцип специфичности и научного знания были внедрены в преподавание экономической теории и экономики. и промышленное строительство. Для преодоления негативного влияния прорывных инноваций на процесс цифровизации образования была использована авторская методика написания конспекта лекций на основе табулирования, перефразирования и формирования вопросов по каждой теме лекции. Реализована онлайн-групповая деятельность и онлайн-общение на основе письменной и устной риторики во время практических занятий. Эффективным результатом исследования можно считать применение усовершенствованной методологии преподавания экономических дисциплин при изучении других предметов и в образовательном центре в г. Алматы.

Ключевые слова: методология преподавания, экономическая дисциплина, цифровизация, принцип специфичности, принцип научного познания, письменная риторика, устная риторика.

Автор туралы мәліметтер:

Бердіқұлова Ғалия Мертаевна, экономика ғылымдарының кандидаты, «Экономика және бизнес» кафедрасының профессорының м.а Халықаралық ақпараттық технологиялар университеті, ORCID: **0000-0002-3000-1675.**

Сведения об авторе:

Бердыкулова Галия Мертаевна, кандидат экономических наук, и.о.профессора кафедры «Экономика и бизнес», Международный университет информационных технологий, ORCID: **0000-0002-3000-1675.**

About the author:

Galiya M. Berdykulova, Candidate of Economic Sciences, Professor, Department of Economics and Business, International Information Technology University, ORCID: **0000-0002-3000-1675.**

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 49–54
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.006
УДК 004.02

Элле В.Ж.[*], Мелисова Л.Т., Куандыков А.А., Куатбаева А.А., Аманбайқызы З.

Международный университет информационных технологий, Алматы, Казахстан
*abu.kuandykov@mail.ru,
**ali.venera15@gmail.com

СВОЙСТВА РЕАЛЬНЫХ БИЗНЕС-ПРОЦЕССОВ С ТОЧКИ ЗРЕНИЯ ПРОЕКТИРОВАНИЯ

Аннотация. В статье дается свое определение бизнес-процесса, которое адекватно соответствует для решения преследуемой исходной задачи. Описаны этапы проектирования начальной стадии процесса проектирования номинального процесса. Развитие работы сводится к рассмотрению концепций и методов управления бизнес-процессом при всех производственных ситуациях, связанных с возникновением инцидентов и проблемности. В данной работе такие термины как «процесс» бизнеса, «процесс для бизнеса», «процесс бизнеса», «бизнес-процесс» считается, что несет одинаковый смысл, одинаковое понятие. Поэтому все эти термины заменены одним термином «бизнес-процесс».

Ключевые слова: Процесс, бизнес-процессы, проектирование процессов, компоненты операций и процессов, метамодель процессов, ресурсов процесса, средства выполнения операций, инфраструктура, институциональное обеспечение

Введение

Бизнес-процесс является совокупностью объектов, которые выполняются для определенной цели производства по заданной последовательности, исходя из каких-то условий и/или заданного расписания. Реальный бизнес-процесс может быть представлен в разных моделях, в частности в следующих видах [1-7]:

- однопоточный не интегрированный бизнес-процесс;
- многопоточный интегрированный бизнес-процесс;
- однопоточный интегрированный бизнес-процесс.

Процесс проектирования всех трех видов моделей бизнес-процессов разбиты на три стадии: начальный, детальный, завершающий. Тогда в качестве операций бизнес-процесса выступают действия или задачи, выполняемые над каждым объектом.

Методикой проектирования является соединение желания лица, принимающего решение (ЛПР) и пути ее достижения с имеющимися ресурсами предметной области и системы управления (причем методы использования ресурсов могут быть разными, в том числе обучаемыми или не обучаемые) [8-16].

Особенности проектирования реальных бизнес-процессов

Процесс автоматизации процессов управления бизнес-процессом предполагает:

• во-первых, проектирование самого объекта управления, т.е. в нашем случае в качестве объекта управления выступает процесс для бизнеса или бизнес-процесс;

• во-вторых, системы управления, куда входит несколько системы, например, система прогнозирования поведения объекта управления, т.е. бизнес-процесса, система организации процессов управления, субъекта управления и т.д.

В связи с тем, что в данной работе планируется разработка системы управления бизнес-процессов, выполнение (или функционирование) сопровождаемые инцидентами, то следует проектировать два вида классов проектов (модели) одного и того же бизнес-процесса [17-24]:

Класс номинального варианта бизнес-процесса и тех же классов бизнес-процесса (которые являются сателлитами номинальных бизнес-процессов), но подвергаемых влиянию инцидента.

Номинальный бизнес-процесс необходим для установления или определения всех функциональных операций бизнес-процесса, которые должны выполниться в нормальных условиях. Функциональная операция имеет следующие компоненты: метамодель операции, предмет труда, средства труда и инфраструктуры, институциональное обеспечение.

Модель бизнес-процесса, подверженного влиянию инцидента может быть одна или набор в зависимости от применения стратегии к процессу проектирования [25-26]. Например, если инцидентом является неопределенность, то один вид модели, а если инцидентом является риски, то другой вид модели, в случае когда инцидентом является нарушения, то третий вид модели.

Все эти модели предполагают, что бизнес-процесс состоит из одного потока, т.е. бизнес-процесс сводится к однопоточной модели. Но реальный бизнес-процесс может быть представлен в моделях разного количество потоков, в частности в следующих видах:

- однопоточный не интегрированный процесс или бизнес-процесс;
- многопоточный интегрированный процесс или бизнес-процесс;
- однопоточный интегрированный процесс или бизнес-процесс.

Для всех трех видов модели процессов могут быть три стадии проектирования: начальный, детальный, завершающий. Так как в данной работе большое внимание уделено вопросам начальной стадии проектирования бизнес-процессов из «сырого» и/или неоформленных фрагментов процессов предприятия, т.е. в несистематизированном виде, для начала процесса проектирования следует отдельно выделить: метамодель процессов, предмет труда и средство труда.

Ресурсы процесса состоят из средств труда и других видов ресурсов, например, финансы, исполнители, инфраструктура, институциональное обеспечение процесса и т.д. Поэтому в работе средства труда включены в состав ресурсов процесса. Далее рассматриваются вопросы выбора оптимальных средств труда, а именно средства транспортировки, предмет и продукции труда для операций и/или процесса в целом.

Таким образом, в работе более подробно рассматриваются особенности проектирования бизнес-процессов. Структуру процесса или бизнес-процесса схематично в виде модели можно представить графически как на рисунке 1. В данной модели бизнес-процесс является номинальной детерминированной, т.е. без инцидентов и сводится к одному потоку операций:

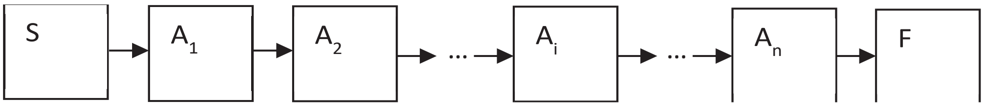

Рисунок 1 - Графическая модель детерминированного процесса

где S – начало процесса;

A_i – функциональные операции;

F – конец процесса.

В данном случае функциональные операции (A_i) включают в себя: метамодель, объект, предмет труда, инфраструктура, институциональное обеспечение и т. д.

Представление бизнес-процесса показывается только последовательностью выполнения операций в составе процессов. Следует выделить компоненты процессов отдельно. Отсюда вытекает процесс проектирования в общем виде.

Итак, начальная стадия процесса проектирования процессов/бизнес-процессов состоит из следующих этапов.

Первый этап – определение миссии процесса или бизнес-процесса. Необходимо определить миссию производства, т.е. назначение (конечная цель) процесса, которую определяет ЛПР.

Миссию для процесса ЛПР выбирает следующим образом:

1) исходя из возможности имеющихся ресурсов, которыми процесс может распоряжаться;

2) исходя из возможности тех ресурсов, поступление которых прогнозируется и которыми планируется наполнить инфраструктуру процесса или бизнес-процесса.

Второй этап – планирование выполнения миссии путем назначения цели для всех видов циклов процесса или бизнес-процесса, которые возможны.

Выполнение миссии надо спланировать исходя из возможностей имеющихся объектов и ресурсов, которыми инфраструктура процесса владеет и исходя из выполнения требования институционального обеспечения.

Поэтому надо оценить, операции процесса при выполнении каких действий над какими объектами можно достичь выполнения или приблизится к выполнению миссии процесса.

Третий этап – оформление операций процесса. Данный этап соответствует второй или детальной стадии процесса проектирования процесса/бизнес-процесса, где ведется оформление операций процессов или бизнес-процессов.

Уровень оформленности операций может быть следующей:

- на уровне простых одноактных действии;
- на уровне процедуры;
- на уровне производственные задачи;
- на уровне управленческие задачи.

Каждая операция или бизнес-операция позволяет приблизиться к выполнению миссии или ее подцели цикла процесса.

Автоматизация может проводиться на различном уровне проекта процесса, в зависимости от условий решения задачи автоматизации она может иметь предпочтения. Модели представления бизнес-процессов для автоматизации:

Детерминированный – номинальный бизнес-процесс

- Одномерный детерминированный процесс
- Многомерный детерминированный процесс
- Интегрированный детерминированный процесс

Инцидентный

- Одномерный инцидентный процесс
- Многомерный инцидентный процесс
- Интегрированный инцидентный процесс

В свою очередь, учет инцидентности процессов может быть следующим:

- Идеальная;
- Измерение возмущения;
- Неопределенное состояние;
- Состояние рисковые (патологические):
- риск группы 1 – несущественный;
- риск группы 2 – существенный;
- риск группы 3 – очень существенный;
- риск группы 4 – аварийный;
- риск группы 5 – катастрофический или ЧС.
- Нарушение нормального функционирования или выполнения процесса и его операций.

Система управления может быть построена на основе любого варианта представления. В связи с этим возникает проблема установки различных вариантов представления и их особенностей.

В зависимости от миссии и сложности окружающей среды для компании необходимо тот или иной уровень автоматизации процессов. Поэтому рассмотрены варианты представления операций и процессов, существующих для автоматизации.

Заключение

В данной работе раскрыты свойства реальных бизнес-процессов в точки зрения проектирования, а именно рассмотрены этапы начальной стадий процесса проектирования номинального процесса или бизнес-процесса без учета инцидентов и проблемности, которые могут возникнуть в ходе выполнения операций. А также развитие работы сводится к описанию процесса автоматизации процессов управления.

Таким образом начальная стадия проектирования состоит из трех этапов, которые определяют миссию процесса, планирование выполнения миссии за счет назначении цели и оформление операции процесса.

СПИСОК ЛИТЕРАТУРЫ

1. Платформа цифровой трансформации бизнес процессов национальной экономики [Текст]: отчет о НИР (промежуточ.): АО МУИТ; рук. Ускенбаева Р.К.– А., 2018. – 70 с. – Исполн.: Куандыков А.А. и др. – №BR05236517. – Инв. № 0218РК01240.

2. Платформа цифровой трансформации бизнес процессов национальной экономики [Текст]: отчет о НИР (промежуточ.): АО МУИТ; рук. Ускенбаева Р.К.– А., 2019. – 113 с. – Исполн.: Куандыков А.А. и др. – №BR05236517. – Инв. № 0219РК00837.

3. Платформа цифровой трансформации бизнес процессов национальной экономики [Текст]: отчет о НИР (закл.): АО МУИТ; рук. Ускенбаева Р.К.– А., 2020. – 113 с. – Исполн.: Куандыков А.А. и др. – №BR05236517. – Инв. № 0220РК00894.

4. Uskenbayeva R., Kuandykov A., Kalpeyeva Z., Kassymova A. Formation of Order Packages for Planning

of the Orders Implementation Process in E-Commerce. Proceedings of 19th International Conference on Control, Automation and Systems (ICCAS 2019) Oct. 15~18, 2019; ICC Jeju, Jeju, Korea, pp. 29-33.

5. Uskenbayeva R., Abu K., Sabina R., Aigerim B. Research of the Relationship Between Business Processes in Production and Logistics Based on Local Models. In: Le Thi H., Le H., Pham Dinh T. (eds) Optimization of Complex Systems: Theory, Models, Algorithms and Applications. WCGO 2019. Advances in Intelligent Systems and Computing, vol 991. Springer, Cham, pp. 861-870.

6. Uskenbayeva R.K., Kuandykov A.A., Rakhmetulayeva S.B., Bolshibayeva A.K. Basics of creating platforms for automation of business processes of logistics // Proceedings of 18th International Conference on Control, Automation and Systems (ICCAS 2018), Oct. 17~20, 2018; YongPyong Resort, PyeongChang, GangWon, Korea

7. Uskenbayeva R.K., Rakhmetulayeva S.B., Bolshibayeva A.K., Managing Business Process Based on the Tonality of the Output Information, (2020) Advances in Intelligent Systems and Computing, 991, pp. 882-890

8. Uskenbayeva R., Moldagulova A., Mukazhanov N.K. Creation of Data Classification System for Local Administration. In: Le Thi H., Le H., Pham Dinh T. (eds) Optimization of Complex Systems: Theory, Models, Algorithms and Applications. WCGO 2019. Advances in Intelligent Systems and Computing, vol 991. Springer, Cham, pp. 761-768.

9. Kurmangaliyeva B.K.,Uskenbayeva R.K., Cho Y.I., Bektemyssova G.B., Mukazhanov N.K. Kozhamzharova D.K., Multidimensional indexing structure development for the optimal formation of aggregated indicators in OLAP hypercube//Proceedings of the 14th International Conference on Control, Automation and Systems (ICCAS 2014). - Gwangju, Korea. October 22-25, 2014. – P. 1466-1470.

10. Kurmangaliyeva B.K., Uskenbayeva R.K., Kuandykov A, Cho Young Im., Bektemyssova G.U. A Conceptual Approach to Construction of Large Information Systems International//Proceedings of the 14th International Conference on Control, Automation and Systems (ICCAS 2014). - Gwangju, Korea. October 22-25, 2014. – P. 1317-1320.

11. Uskenbayeva R.K., Kurmangaliyeva B.K., Yedilkhan D. Situational management for process implementation of working operations of the business process, Society of Instrument and Control Engeneers of Japan (SICE), 34th Chinese Control Conference and SICE Annual Conference. Hangzhou, China, 2015. – C.292-297.

12. Kurmangaliyeva B.K., Uskenbayeva R, Yedilkhan D, Kassymova A. Principles for Achieving the Optimal Performance of the Input Tasks Flow of a Business Process and Optimal Performance of the Business Process.// SICE Annual Conference 2015, Hangzhou, China, Luly 28-30, 2015, P.909-914.

13. Kurmangaliyeva B.K., Uskenbayeva R. Mobile business process construction principles on the mobile robot system platform.//Proceedings of the 15th International Conference on Control, Automation and Systems (ICCAD 2015). –BEXCO, Busan, Korea, Oct.13-16, 2015, P.648-650.

14. Kassymova A., Uskenbaeva R. Questions for storage and processing of unstructured data // Proceedings of the 12th International Scientific Conference Information Technologies and Management 2014. – Riga; Latvia, 2014. - P.113-114.

15. Uskenbayeva R., Chinibayev Y., Kassymova A., Temirbolatova T., Mukhanov K. Technology of integration of diverse databases on the example of medical records // Proceedings of the 14th International Conference on Control, Automation and Systems (ICCAS 2014). - Gyeonggi-do; Korea, 2014. – P. 282-285.

16. Kassymova A., Uskenbayeva R., Kurmangaliyeva B., Yedilkhan D. Principles for achieving the optimal performance of the input tasks flow of a business process and optimal performance of the business process // Proceedings of the 34th SICE Annual Conference. – Hangzhou; China, 2015. - P. 909-914.

17. Kassymova A., Young Im Cho, Uskenbayeva R., Kuandykov A. Methods of representation data for the integration // Proceedings of the ISIS 2015 The 16th International Symposium on Advanced Intelligent Systems. – Mokpo; South Korea, 2015. - P.337-339.

18. Kassymova A., Uskenbayeva R., Kalpeyeva Zh. Organization of computational processes in distributed cloud environments // Proceedings of the 34th SICE Annual Conference. – Hangzhou; China, 2015, july 28-30. - P. 909-915.

19. Kassymova A., Uskenbayeva R., Young Im Cho, Uskenbayeva Z., Bektemyssova G., Temirbolatova T. Recursive decomposition as a method for integrating heterogeneous data sources // Proceedings of the 15th International Conference on Control, Automation and Systems (ICCAS 2015). – Busan; South Korea, 2015, october 13-16. – P.2076-2079.

20. Kuandykov A., Rakhmetulayeva S., Senbay D., Saparkhojayev N., Bektemisova G. Package service flow applications in GRID-system using genetic algorithm.Computer Science and Network Technology (ICCSNT) «2nd International Conference on China».–Changchun, 2012. – P.25–28.

21. Kuandykov A., Rakhmetulayeva S., Saparkhojayev N. Genetic algorithm of the batch processing of the request stream in grid system.// International Journal of Mathematics and Physics. Quarterly Journal of al-Farabi Kazakh National University. –Almaty, 2012. – Volume 3. – №2. – С.107–111.

22. Kuandykov A.A., Rakhmetulayeva S.B., Baiburin Y.M., Nugumanova A.B/ Usage of singular value decomposition matrix for search latent semantic structures in natural language texts//The 34th Chinese Control Conference and SICE Annual Conference 2015 (CCC&SICE2015), Hangzhou, China, July 28 to 30, 2015.

23. Uskenbayeva R.K., Cho Y.I., Bektemyssova G.B., Mukazhanov N.K., Kozhamzharova D.K.. Kurmangaliyeva B.K. Multidimensional indexing structure development for the optimal formation of aggregated indicators in OLAP hypercube. // Proceedings of the 14th International Conference on Control, Automation and Systems (ICCAS 2014). - Gwangju, Korea. October 22-25, 2014.

24. Kurmangaliyeva B.K., Uskenbayeva R.K., Kuandykov A.A., Methods for Functionality Improvement of Business Process Support System in Government Bodies//Proceedings of the International Scientific-Practical Conference «Smart Government: Science and Technology» Astana, Kazakhstan, October 7-8, 2014. – P.170-178.

25. Kassymova A., Uskenbayeva R., Kurmangaliyeva B., Yedilkhan D. Principles for achieving the optimal performance of the input tasks flow of a business process and optimal performance of the business process // Proceedings of the 34th SICE Annual Conference. – Hangzhou; China, 2015. - P. 909-914.

26. Kuandykov A.A., Rakhmetulayeva S.B., Baiburin Y.M., Nugumanova A.B/ Usage of singular value decomposition matrix for search latent semantic structures in natural language texts//The 34th Chinese Control Conference and SICE Annual Conference 2015 (CCC&SICE2015), Hangzhou, China, July 28 to 30, 2015.

Элле В.Ж.[1], Мелисова Л.Т.[2],Куандыков А.А.[3], Куатбаева А.А.[4], Аманбайқызы З.[5]

Жобалау тұрғысынан нақты бизнес-процестердің қасиеттері

Андатпа: Бұл мақалада бизнес-процестің өзіндік анықтамасы берілген, ол ұсынылған бастапқы мәселені шешуге сәйкес келеді.

Номиналды процесті жобалау процесінің бастапқы кезеңін жобалау кезеңдері сипатталған. Жұмыстың дамуы инциденттер мен проблемалардың пайда болуымен байланысты барлық өндірістік жағдайларда бизнес-процесті басқарудың тұжырымдамалары мен әдістерін қарастырудан басталады.

Бұл жұмыста «бизнес процесі», «бизнеске арналған процесс», «бизнестің процесі», «бизнес-процесс» сияқты терминдер бір мағыналы, бір ұғымды білдіреді. Сондықтан бұл терминдердің барлығы бір ғана «бизнес-процесс» терминімен ауыстырылды.

Түйін сөздер: Процесс, бизнес-процестер, процесті жобалау, операциялар мен процестердің құрамдас бөліктері, процестердің метамоделі, процесс ресурстары, операцияларды орындау құралдары, инфрақұрылым, институционалдық қамтамасыз ету.

Elle V.[1], Melissova L.[2], KuandykovA.A.[3], Kuatbayeva A.A.[4], Amanbaikyzy Z.[5]

Properties of real business processes from a design point of view

Abstract: The article gives its own definition of the business process, which assists in the solution of the initial task set and describes the design of the initial stage of the nominal design process. The research work focuses on consideration of the concepts and methods of business process management in all production situations associated with the occurrence of incidents and problems.

In this work, such term as "process" of business, "process for business", "business process" are considered to have the same meaning, the same concept. Therefore, all of these terms have been replaced by a single term "business process".

Key words: process, business process, process design, components of operations and processes, metamodel of processes, process resources, means of performing operations, infrastructure, institutional support.

Авторлар туралы мәлімет:

Элле Венера Жанатқызы, «Ақпараттық жүйелер» кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0002-3863-942X.

Мелисова Лиана Тулешевна, «Ақпараттық жүйелер» кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0003-0732-6313.

Куандыков Абу Абдикадырович, т.ғ.д., «Ақпараттық жүйелер» кафедрасының профессоры, Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0002-0055-2513.

Куатбаева Акмарал Алихановна, PhD, «Ақпараттық жүйелер» кафедрасының ассистент-профессоры, Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0002-2143-3994.

Аманбайқызы Зульфия, «Ақпараттық жүйелер» кафедрасының докторанты, Халықаралық ақпараттық технологиялар университеті, ORCID: 0000-0001-6688-9920.

Сведения об авторах:

Элле Венера Жанаткызы, магистрант кафедры «Информационные системы», Международный университет информационных технологий, ORCID: 0000-0002-3863-942X.

Мелисова Лиана Тулешевна, магистрант кафедры «Информационные системы», Международный университет информационных технологий, ORCID: 0000-0003-0732-6313.

Куандыков Абу Абдикадырович, д.т.н. профессор кафедры «Информационные системы», Международный университет информационных технологий, ORCID: 0000-0002-0055-2513.

Куатбаева Акмарал Алихановна, PhD, ассистент-профессор кафедры «Информационные системы», Международный университет информационных технологий, ORCID: 0000-0002-2143-3994.

Аманбайкызы Зульфия, докторант кафедры «Информационные системы», Международный университет информационных технологий, ORCID: 0000-0001-6688-9920.

About the authors:

Venera Zh. Elle, Master’s student, Department of Information Systems, International Information Technology University, ORCID: 0000-0002-3863-942X.

Liana T. Melissova, Master's student, Department of Information Systems, International Information Technology University, ORCID: 0000-0003-0732-6313.

Abu A. Kuandykov, D.F.S., Professor, Department of Information Systems, International Information Technology University, ORCID: 0000-0002-0055-2513.

Akmaral A. Kuatbayeva, PhD, Associate Professor, Department of Information Systems, International Information Technology University, ORCID: 0000-0002-2143-3994.

Zulfiya Amanbaikyzy, Doctoral student, Department of Information Systems, International Information Technology University, ORCID: 0000-0001-6688-9920.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp. 55–61
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.007
УДК 530.1, 681.3.06

Koshimbay A.B., Moldagulova A.N.

International Information Technology University, Almaty, Kazakhstan

RESEARCH METHOD OF ANALYZING AND PROCESSING SOCIAL NETWORK DATA IN ORDER TO DETERMINE THE TONALITY

Abstract. A wide spread of social online services and the advancement of Big Data technologies poses a challenge to utilize data from social media in numerous circles. Nowadays, the «social listening» and substance examination advances pick up ubiquity in Data Science. The sentiment analysis of the text is one of the especially important tasks in the field of natural language processing. It is used in different spheres. This article discusses the main methods of identifying emotions in text data and analyzes the current achievements in the field of computer analysis of emotions in text data. At the moment, there are many unresolved problems in the field of automatic text analysis to determine the emotional coloring of the vocabulary in social media texts.

Keywords: sentiment-analysis, tonality of text, text vectorization, machine learning, support vector machine, tone dictionary

Introduction

The majority of people nowadays can't imagine their lives without social media. According to the research results, social media platforms such as Facebook, Instagram, Twitter, and VKontakte have become an indispensable part of everyone's life during the last decade. A significant number of businesses are establishing themselves on social media platforms. Consumer relationships can be maintained, requests and feedback can be responded to rapidly, and marketing campaigns can be run through social media. As a consequence, social media have a significant impact on the socialization of modern individuals. The rise of social media has signified not just the transfer to new data sources, but also the possibility for each user to be there. Now, if a user comes across something interesting, he or she can put it in his or her profile [1].

Quite a substantial shift in the way information is distributed offers enormous possibilities for individuals and society as a whole. In general, the online world transports a massive amount of data, including user-to-user communication. It should be emphasized that major media firms use public evaluation based on the number of likes, postings, and comments to explore public opinion. In social networks, there is a massive amount of data. For instance, Facebook, which had 2.23 billion active monthly users at the end of summer 2018, posts over half a million comments per minute and over 100,000 photos per day. Every day, Twitter's audience sends out more than half a billion tweets, amounting to roughly 200 billion messages every year. This is a massive amount of unstructured data that has never been seen before. Likes and dislikes ratings may potentially become public tools for researching user opinions in networks [1]. Hence, there is a need for a tool that can automatically extract relevant information from publications, distinguish reviews from advertisements, and determine, among other things, the users' attitudes toward the topic of interest. By processing such data, mass media companies can improve the quality of their content, identify their target audience, and assess the attitude to materials in a comprehensive perspective [2].

Social media posts come in a wide range of content formats. Video (YouTube, images (Instagram, Pinterest, and text (other social media platforms are the most common types of material (Twitter, Facebook. This report will concentrate primarily on textual information and the methods for processing it, specifically, based on the semantic analysis of texts [2].

The purpose of the article is to give an overall overview of various methods for monitoring and evaluating social media opinion in order to perform an in-depth analysis of unstructured data and extract negativity and threat from text arrays. The ability to recognize emotionally colored vocabulary and analyze the user's appraisal of the product can be determined based on the tone of textual messages [2].

It was previously impossible to find an automated solution to this issue. Presently, computer linguistics' functionalities enable the extraction of information from texts using computer technology and accurate models. Assessment of the emotional coloring of a text is one of the objectives of this research (text tone analysis, content analysis, sentiment analysis. As a result of the research, there has been defined a variety of open text sources that present data about people's perspectives on a variety of topics. There is a need for greater research in the field of text tone analysis in order to obtain more complicated statistical data [3].

Sentiment analysis of texts is defined as a set of approaches for automatically selecting an emotionally colored language in writings and assessing the writers' emotional reactions (opinions) to the items mentioned in the text. Sentiment analysis is utilized in marketing research, audience loyalty tracking for various themes and brands, and other applications [3].

The study of emotions in textual data seems to be applicable to a variety of fields, including evaluation of the properties of products and services based on online customer feedback, analysis of unfavorable emotions in messages, prediction of stock market trends, the political environment based on news streams. In order to analyze such a large amount of data, many methods for automatically determining the tone of the text have been introduced in recent years, which will be reviewed in this article.

Materials and research methods

Text tonality determination is a difficult task whose outcome is highly dependent on the context, study region, and amount of text data. There are a few fundamental techniques for determining a text's tone. The three most widely acknowledged ways to determine the tone of a sentence can be divided into several categories. (Figure 1)

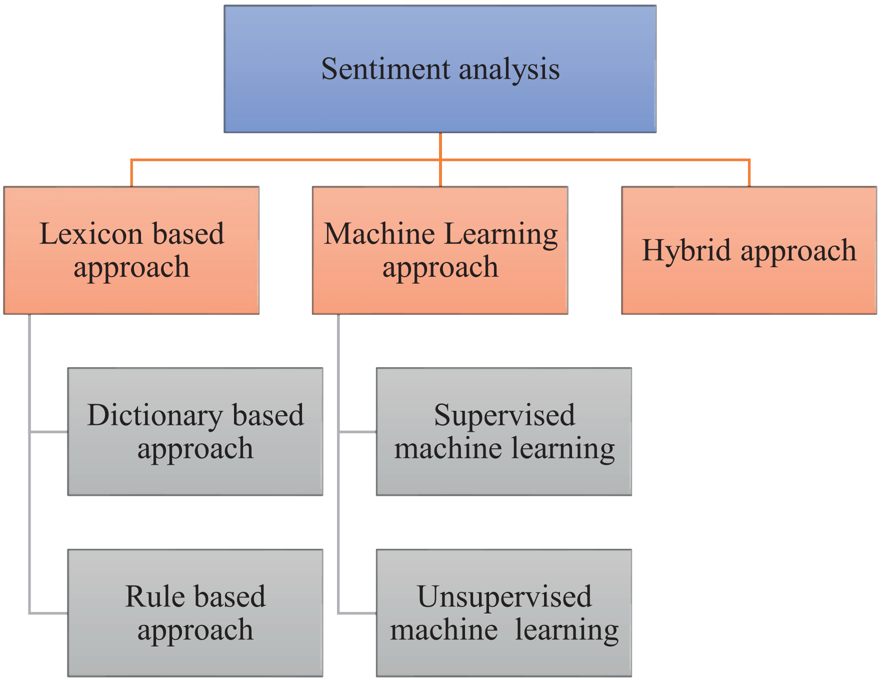
Sentiment analysis
Lexicon based approach
Machine Learning approach
Hybrid approach
Dictionary based approach
Rule based approach
Supervised machine learning
Unsupervised machine learning

Figure 1– Classification of approaches to sentiment analysis

1. Methods based on rules and dictionaries

The conversational style of speech is common in text messages published on social media. The inclusion of slang, changed words, typos, and other elements that make it difficult to interpret the text are the main challenges in analyzing users' emotions. The text quality is strongly influenced by the social network in question. Also, it should be borne in mind that subjective and objective texts should be differentiated. A subjective text, as opposed to the objective one, contains information on an event without the author's personal perspective. It is subjective assessments that are of interest, as they allow of the extraction of a certain author's viewpoint. The majority of objective articles are informational in nature, hence they are removed from the bulk of the material under study [4].

Rule-based approaches typically specify a set of rules in a particular scripting language that denotes subjectivity, polarity of opinion. This approach entails the creation of numerous rules, based on "if-then" rules. For instance, if the particle "not" appears before a positive adjective, the construction is considered to be negative. This strategy also includes the use of tone dictionaries, in which the words examined belong to a specific category (positive, negative, neutral, and so on). The rules accept a variety of inputs, including standard NLP (Neurolinguistic Programming) approaches like stemming, tokenization, voice tagging, and syntactic analysis. Two sets of polarized terms, such as negative words like «awful», «worse», «ugly», and positive words like "good," "best," and "beautiful," are a fundamental illustration of these techniques [4].

The text is given. The number of favorable words in the text is counted. There is also a count of bad terms.

When the number of positive words in a text exceeds the number of negative ones, the positive emotion is the polar opposite of the negative feeling. Otherwise, the result is neutral. This approach is extremely basic because it ignores the word combinations in the sequence. The more advanced processing analysis is feasible; however, such systems are extremely difficult to construct, because they require additional support for new terminology and language, despite the fact that new rules are more difficult to enforce. Furthermore, adding new rules may result in unfavorable outcomes because such systems necessitate a large investment in manual configuration and adherence to them [4].

The second linguistic approach is using tone dictionaries. The main concept of this method is to create tonal glossaries, which are collections of words and emotional states with a numerical positive or negative tonal score. When applying this method, linguists conduct the so-called rule-based tone analysis. The outcome of this study is a collection of rules (also known as a lexicon or sentiment lexicon analysis) that classify words as positive or negative, along with the corresponding intensity measure [5].

Furthermore, attribution of this or that text to a specific tone class is calculated on the basis of all collected weights of the text terms. The arithmetic mean of the weights is most commonly employed for this, with the sum of the weights or artificial neural networks being utilized in rare circumstances [5]. The text is regarded as positive if the number of positive words exceeds the number of negative ones, and vice versa. Clearly, this method is inefficient. It necessitates the creation of a big word dictionary, which must be updated on a regular basis. Furthermore, the fact that the number of positive words outnumbers the number of negative words is not a trustworthy standard by which the work might be evaluated properly [5].

2. Machine learning approach.

Supervised machine learning algorithms

A supervised machine learning approach necessitates a training set of texts that are marked up inside an emotional space and used to build a statistical or probabilistic classifier. A set of training samples is required to solve the sentiment analysis problem. A set of pre-labeled review texts is provided for supervised learning. A pair of feature vectors, which is a representation of a single text, and the tonality of the text constitute an individual instance of this set. The marked set of texts (the training sample) is evaluated, and a statistical pattern is developed for use in classifying new input vectors in supervised machine learning algorithms. There are several algorithm classifiers based on the supervised machine learning approach, such as linear classifier, SVM, decision tree classifier, Naive Bayesian classifier, etc. The Naive Bayesian classifier and the support vector machine approach are the most widely used methods in the field of tone analysis [6].

A Naive Bayesian classifier (NB) is a probabilistic classifier that relies on the Bayes theorem and assumes class independence. The Naive Bayes approach ensures that no relationship exists between the system's various parameters. It is particularly useful for huge data sets and, despite its simplicity, outperforms other more complicated algorithms in terms of consistency.

Table 1 represents the advantages and disadvantages of using Naïve Bayes classifier: [6]

Table 1 - Table of advantages and disadvantages of the Naive Bayes classifier

Benefits	Drawbacks
The Naive Bayesian classification method is simple and quick to implement.	One of the most significant drawbacks of the Naïve Bayesian classification is strong independence of features, as in actual life it is quite difficult to have a set of features that are totally independent of one another.
It will converge more quickly than such discriminative models as regression models.	Another issue with Naive Bayesian classification is that it has a "zero-frequency" feature, which implies that if a classifier has a category but is not observed in the training data set, then the Naive Bayesian model would give it as a zero probability and will be unusable for prediction.
It requires less training data.	
It's extremely scalable because the number of predictors and data points scales linearly. It has the ability to produce probabilistic predictions and can work with both continuous and discrete data.	

Support vector machines (SVMs) are linear classifiers. Regression predicts a continuous value, while linear classification predicts a label or a group. The method's basic idea is to create a hyperplane that isolates sample items as well as possible. SVM classifies input data by identifying a hyperplane that separates classes in n-dimensional spaces. One of the benefits of SVM is its versatility, as it can solve problems using a variety of kernel functions. SVM classifiers also have a high level of accuracy and can handle enormous data sets. Because SVM classifiers only use a fraction of training points, they require extremely little memory. However, they have long learning times, so in practice, they are not suitable for large datasets. Another disadvantage is that SVM classifiers do not work well with overlapping classes. [6]

Unsupervised machine learning

In contrast to the methods described above, the unsupervised learning method determines relations and patterns between objects without labeled data. Such methods include Gaussian mixture and k-nearest neighbor models.

K-means is an algorithm that finds k training examples within the shortest distance to the provided sample. The class of the object of interest will be the most common among k objects. To implement it, the algorithm needs a training sample of marked reviews. It was necessary to calculate the distance between the vector of this review and vectors from the training sample in order to establish the class of review from the test sample and determine the minimum distance between k items in the training sample (k is given by the expert or chosen according to efficiency estimates). The input vector's class is the one where more than half of the nearby k vectors are members. As an advantage, the k-mean algorithm has high accuracy, insensitivity to outliers, and no assumptions about data entry. Yet, it has drawbacks such as high temporal complexity and high spatial complexity. Once the sample is unbalanced, one class's sample size is quite huge while the sample size of the other classes is quite tiny. If a sample is entered, then the class with the largest sample size is the one with the most K adjacent values, which causes issues in classification [7].

3. **Hybrid approach**

The techniques that combine some of the methods outlined above are known as hybrid methods. A hybrid method employed in learning models by A.C. Koenig and E. Brill included the method based on tone dictionaries and the method of reference vectors. Using this strategy, the authors attained a learning accuracy of 72 % [8].

Results and discussion

According to the reviewed articles and works, scientists generally merge approaches to achieve the best outcomes. V.G. Vasiliev, S. Davydov, and M.V. Khudyakova in their works use a linguistic strategy complemented with machine learning approaches to adjust individual classification rules through training [9]. Numerous studies have demonstrated that integrating linguistics to overcome the tonality problem produces beneficial consequences. Rule-based algorithms provide more accurate results than machine learning approaches that apply the statistics and probability theory because their operation is intimately tied to the meaning of words. However, as previously indicated, the linguistic method has some significant limitations. While comparing different algorithms for tonality analysis, it must be highlighted that rule-based methods produce more accurate outcomes than machine learning approaches. It can be asserted that each approach is unique, and the combination of different approaches results in an increased accuracy of the training model. Table 2 shows a comparison of methodologies based on the primary criteria for selecting a tonality analysis method.

Table 2- Comparison of tonality analysis methods

	Accuracy	Automation	Training data	Easy to apply	Applicability in commercial systems
Rule based approach	High accuracy	automated	No data required	-	+
Dictionary based approach	Not unique	Automated within the same subject area	Data required	+	-
Supervised machine learning	Moderate accuracy	Automated	Data required	+/-	+
Unsupervised machine learning	Low accuracy	Automated	No data required	+	+

Tang and colleagues (Tang et al.) claim that: [10]

«To differentiate and evaluate comments on Internet reviews, most of the available methods rely on the natural language processing algorithms. However, while comments on Internet evaluations are less formal than those on news stories or magazine articles, they still demand poor accuracy. Many of the sentences in the books contain grammatical faults as well as unknown components not found in dictionaries.»

Using the sentence syntax method, the researcher was able to reach the best level of accuracy. All the aforementioned algorithms seem to be unable to function with word order, since saving the words necessitates the creation of a text array in the form of a matrix, with each row representing a vectorized word. However, only convolutional neural networks are capable of using it. Convolutional neural networks, on the other hand, are only implemented when the maximum classification accuracy is required. This approach is rarely applied in reality, but it improves the model's accuracy by 2% on average when compared to other classifiers [11].

M.V. Chernyshevich in his work analyzes the main existing types of classifying the tone of opinions and offers his opinion scale, which operates with both absolute and comparative evaluations. He conducted a study of the emotional tone of readers' comments. In his research, 38 newspaper articles were selected and a library of readers' comments was compiled. The study discovered that negative remarks heavily dominated in the Russian-language comments (59.3%), whereas neutral comments made up the largest group in the English-language comments (46.1 %). Both Russian and English-language comments had the lowest percentage of good remarks, although Russian-language comments had twice as many of them as the English-language comments (17.7% and 8.4 %, respectively). The researchers note that Russian-speaking users frequently employ negative evaluative language, as well as a lot of irony, sarcasm, and insults directed towards the state and other commentators, whereas comments on English-language articles have a neutral tone [12].

In constructing an application for assessing the tone of messages from social networks, Bobyakova D.A. used several approaches: supervised machine learning, specifically the Naïve Bayesian classifier, and the dictionary-based approach. In the beginning, the frequency dictionary included 100 terms with the highest frequency of words in Russian. Furthermore, the majority of the terms in the compiled dictionary turned out to be pronouns, connectives, and prepositions, resulting in a vocabulary of only 29 items. The tests were performed on a Twitter database of 100,000 texts. To pre-process the texts, the author removed stop words, links, hashtags, and words with the highest frequency of occurrence. Bigrams and unigrams were used to show the documents. This method demonstrated 86.6 % training accuracy and 89.1 % recall [13].

In general, the automatic text tonality analysis is a sufficiently impartial and effective method that can be used successfully in both sociological and linguistic research.

Conclusion

This article investigates the textual features of messages in public networks in the context of developing ways to assess the emotional coloring of opinions and discusses strategies to analyze text messages.

Natural language analysis tasks are gaining popularity as the volume of unstructured textual data grows. Modern techniques can be used to handle problems of this type, given the availability of open machine learning libraries. The number of conferences in the field of sentiment analysis is growing year by year, and so does the number of publications in Russian and other foreign languages on text analysis. According to the Google Academy, about 700 works on the sentimental analysis of Russian-language texts were published in 2018, as opposed to 8,500 works on the English-language texts. It should be noted that, based on the above material, we can determine that the researchers obtained the accuracy of the tonal analysis of Russian-language texts of about 80%, and in English-language texts, the accuracy reaches 96% [14].

The existence of numerous works on the topic of sentiment analysis indicates that this topic is relevant today and is in demand in many areas, such as the economic market, politics, marketing, etc. The approach proposed in the work can be used for marketing, sociological and political research. It also allows monitoring the loyalty of the audience to a particular topic or brand, which gives the management an opportunity to make timely decisions. However, as the analytical study demonstrates, sentimental analysis algorithms for Russian-language texts are less developed than those for foreign-language texts.

The investigation of various approaches of tonality analysis by different experts assisted in the selection of the appropriate method of analysis for developing a system for analyzing Russian-language comments. At the moment, a training sample for training a multi-class classifier of Russian-language Internet texts is being developed and the maximum weighted average f1- score, which is the average of accuracy values, is approaching 50%. The hybrid methods consist of supervised machine learning and a bag-of-words rule-based approaches. This model uses a hybrid approach to analyze the emotional coloring of the text tone. More work needs to be done to search and develop the optimal method for training a model for Russian language text messages. Thus, it can be concluded that that there is a high demand in the modern world for automatic sentiment analysis of texts, as well as an increase of its application possibilities.

REFERENCES

1. Kolmogorova A.V. (2019) The use of texts of the "Internet revelation" genre in the context of solving the problems of sentiment analysis. [Bulletin of NGU. Series Linguistics and Intercultural Communication] 3.71-82.

2. Moussa, M., Mohamed, E. & Haggag, M. (2018) A survey on opinion summarization techniques for social media. Future Computing and Informatics Journal. 3 (1). 82–109 PP.

3. Boudad, N. et al. (2018) Sentiment analysis in Arabic: A review of the literature. Ain Shams Engineering Journal. 9 (4). 2479–2490.

4. Taboada, Maite & Brooke, Julian & Tofiloski, Milan & Voll, Kimberly & Stede, Manfred. (2011). Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics. 37. 267-307.

5. Kaur, Fatehjeet & Bhatia, Rekha. (2016). Sentiment Analyzing by Dictionary based Approach. International Journal of Computer Applications. 152. 32-34.

6. Kolmogorova A.V., Kalinin A.A., Malikova A.V. (2018) Linguistic principles and methods of computational linguistics for solving problems of sentiment analysis of Russian texts. [Series: Actual problems of philology and pedagogical linguistics] 1, 139–148

7. Yousefpour A., Ibrahim R., Hamed H.A. (2017) Ordinal-based and frequency-based integration of feature selection methods for sentiment analysis. [Expert Systems with Applications] Vol. 75. 80–93 PP.

8. Khudyakova M.V., Davydov S., Vasilyev V.G. (2017) Classification of feedback of users using fragment rules.

9. Poecze F., Ebster C., Strauss C. (2018) Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts. [Proceedings of the 9th International Conference on Ambient Systems, Networks and Technologies (ANT)]. Vol. 130,660–666 PP.

10. Pang B., Lee L. (2018) Opinion mining and sentiment analysis. [Foundations and Trends in Information Retrieval]. Vol. 2, 1-135P.

11. Stieglitz S., Mirbabaie M., Ross B., Neuberger C. (2018) Social media analytics – Challenges in topic discovery, data collection, and data preparation. [International Journal of Information Management]. Vol. 39,156–168 PP.

12. Chernyshevich M.V. (2018) Sentiment classification for the task of automatic text sentiment analysis. Vestnik Uchenye zapiski UO VGU im. P.M. Masherova 28, 136–140

13. D.A. Bobyakova (2017) Development of an application for analyzing the sentiment of texts from social networks. [Final qualification work of ITMO University, St. Petersburg, Russia] 2017 – 74 P.

14. K. Archipenko, I. Kozlov., Y. Trofimovich, K. Skornyakov., A. Gomzin., D. Turdakov (2017) Comparison of neural network architectures for the analysis of Russian tweets. [Proceedings of the International Conference on Computational Linguistics and Intelligent Technologies.] URL: http://www.dialog-21.ru/media/3380/arkhipenkoetal.pdf

Көшімбай А.Б., Молдагулова А.Н.

Тоналдылықты анықтау мақсатында әлеуметтік желілердің деректерін талдау және өңдеу әдісін зерттеу

Аңдатпа. Әлеуметтік онлайн-сервистердің кең таралуы және Үлкен Деректер технологияларының дамуы әлеуметтік желілердегі деректерді әртүрлі салаларда қолдануға деген қызығушылықтарды артыруда. Қазіргі таңда контент анализ және "әлеуметтік желілерді мониторлау"технологиялары танымалдылыққа ие болуда. Атап айтқанда, мәтіннің тоналдылығын талдау табиғи тілді өңдеу саласындағы маңызды міндеттердің бірі болып табылады. Ол әртүрлі салаларда қолданылады. Ұсынылған мақалада мәтіндік деректердегі эмоцияларды сәйкестендірудің негізгі әдістері қарастырылады. Мәтіндік деректердегі эмоцияларды компьютерлік талдау саласындағы қазіргі жетістіктер талданды. Қазіргі уақытта әлеуметтік медиа мәтіндерінің лексикасының эмоционалды түсін анықтау үшін автоматты талдау саласында көптеген шешілмеген мәселелер бар.

Түйін сөздер: әлеуметтік желілерді мониторлау, комментарийлерге анализ жасау, әлеуметтік көңіл-күй, пайдаланушының қабылдауын бағалау, машиналық әдіспен оқыту

Кошимбай А.Б. Молдагулова А.Н

Исследование метода анализа и обработки данных социальных сетей с целью определения тональности

Аннотация. Быстрое распространение общественных онлайн-сервисов и эволюция технологий «Больших данных» инициировали внимание к применению сведений из общественных сеток во всевозможных секторах экономики. На сегодняшний момент, известность технологии завоевывают, технологии как «прогноз социальных сетей» (social listening) и контент анализа. В особенности анализ тональности текста является одной из важных задач в области обработки естественного языка. Необходимо подчеркнуть, что данная технология применяется в разных областях. В предоставленной статье рассматриваются основные методы идентификации эмоций в текстовых данных. Исследованы и проанализированы существующие достижения в области компьютерного анализа эмоций в текстах. В результате исследования, на данный момент существует множество нерешенных проблем в области автоматического анализа для определения эмоциональной окраски текстов в социальных сетях.

Ключевые слова: анализ тональности текста, сентимент-анализ, классификатор, метод опорных векторов, тональный словарь

Авторлар туралы ақпарат

Көшімбай Айасем Бекжанқызы – Халықаралық ақпараттық технологиялар университетінің «Ақпараттық жүйелер» кафедрасының магистранты.

Молдагулова Айман Николаевна – PhD, Халықаралық ақпараттық технологиялар университетінің «Ақпараттық жүйелер» кафедрасының қауымдастырылған профессоры.

Сведения об авторах

Кошимбай Айасем Бекжанкызы – магистрант кафедры “Информационные системы” Международного университета информационных технологий.

Молдагулова Айман Николаевна – PhD, ассоц. профессор кафедры «Информационные системы» Международного университета информационных технологий.

About the authors

Aiasem B. Koshimbay – Master's student, Department of Information Systems, International University of Information Technology.

Aiman N. Moldagulova – PhD, Assoc. Professor, Department of Information Systems, International University of Information Technology.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08(2021). Pp. 62–69
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.008

УДК 004.5

Базарбеков И.М.[*], Шарипов Б.Ж.
Международный университет информационных технологий, Алматы, Казахстан

РАЗРАБОТКА БИЗНЕС-ПРОЦЕССА ДЛЯ ПОЛУЧЕНИЯ ОНЛАЙН УСЛУГ В ОРГАНИЗАЦИИ ОБРАЗОВАНИЯ

Аннотация. В данной статье рассмотрены и разработаны процессы получения онлайн услуг в организациях образования на примере университета АО «МУИТ». Под онлайн услугами подразумеваются такие услуги, как: получение справки с места обучения, получение транскрипта, подача заявлений онлайн, ликвидация задолженности. Были разработаны карта процессов и модели нотации BPMN «AS IS» и «TO BE».

Ключевые слова: “Smart campus”, Business Process Model and Notation, “AS IS”, “TO BE”, онлайн услуги, система, электронная цифровая подпись

Современные университеты имеют разнообразную и большую инфраструктуру. Университет занимается не только образовательным процессом, но и административно-организационными составляющими. Если для образовательного процесса требуется всего одна система дистанционного обучения (moodle), то функционал для административно-организационных процессов может быть весьма разнообразен. В условиях карантина без подобных систем не обойтись. Ключевым звеном университетской инфраструктуры является кампус университета как коммуникативная среда взаимодействия студентов, докторантов, преподавателей и научных работников, что является неотъемлемой составляющей учебного процесса. В соответствии с доминирующими тенденциями, определяющими развитие современной системы образования и связанными с внедрением новых информационных технологий и формированием единого научно-образовательного пространства, коммуникативная среда кампуса должна базироваться на применении современных ИТ-решений. В такой постановке университетский электронный кампус («Smart campus») становится важным инфраструктурным элементом с полным циклом автоматизации важнейших задач деятельности университета, предоставлением персонализированного информационного пространства и соответствующих информационных услуг.

В настоящее время такие услуги, как получение справки с места обучения в АО «МУИТ», получение транскрипта, ликвидация задолженности, подача заявлений в деканат, не автоматизированы в полной мере. Например, процесс получения справки с места обучения происходит следующим образом: обучающийся отправляет на почту заявку с указанием ФИО, группы и курса обучения. После этого, в течение трех рабочих дней, студент забирает документ из университета либо получает отсканированную копию на почту. Однако сотрудники университета сканируют или вносят изменения в справки вручную. Также больших временных затрат требует подписание справки руководством университета. Остальные процессы, такие, как подача заявлений, ликвидация академических задолженностей, происходят через почту и обрабатываются вручную, на регистрацию каждого письма и отправки ответа уходит большое количество времени. Система «Smart campus» позволит автоматизировать данные процессы и будет генерировать справки, регистрировать входящие номера для заявлений моментально, сразу после подачи заявки студентами, что значительно сократит нагрузку на административный персонал и улучшит условия для обучающихся.

Моделирование процессов

А. Карта процессов

Карта процессов «Получение онлайн услуг» - инструмент планирования и управления процессом, который используется для повышения эффективности выполняемых действий в течении всего процесса. Карта процессов отражает следующие виды процессов:

- Основные процессы.
- Процессы управления.
- Вспомогательные процессы [1].

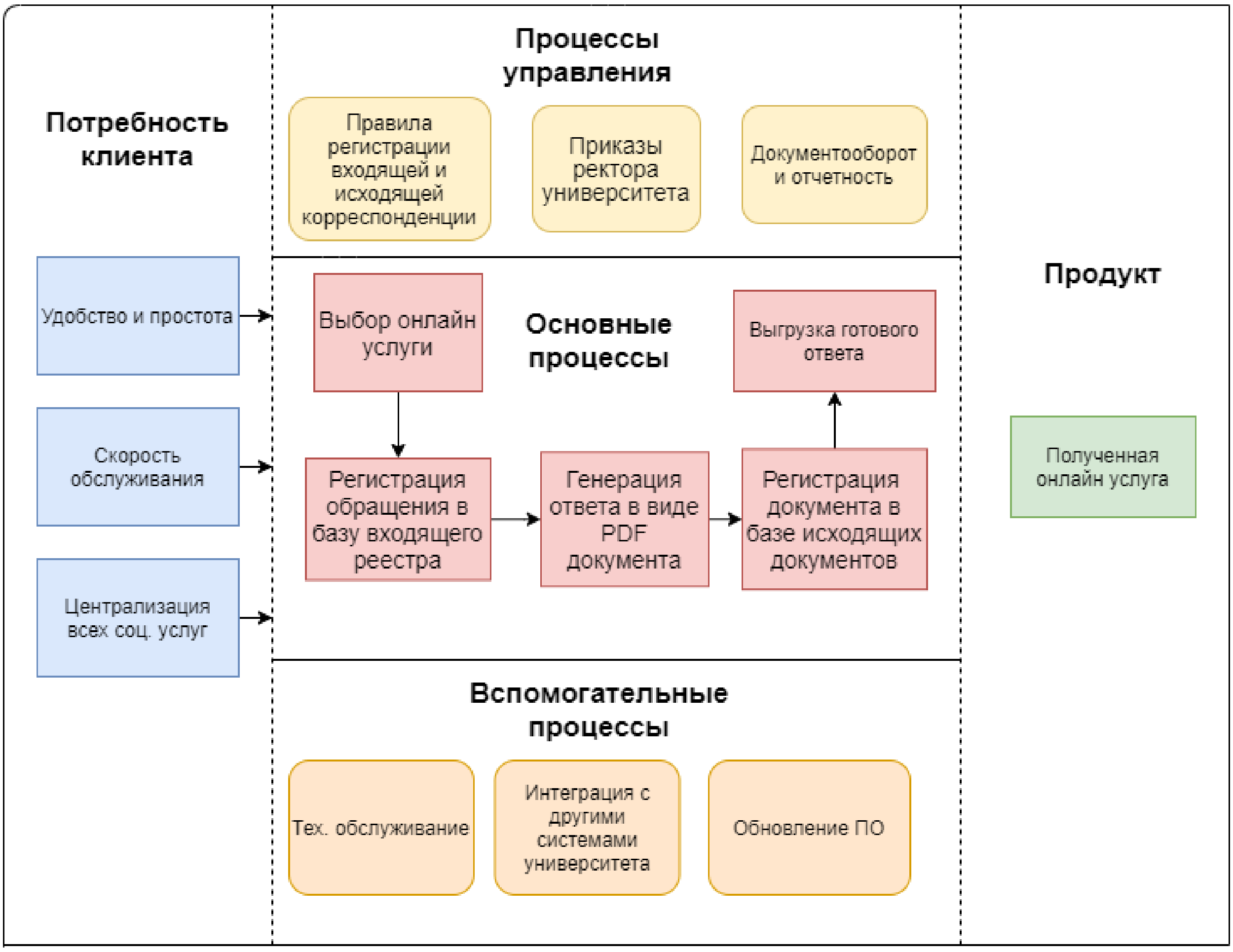
Процессы управления
Потребность клиента
Правила регистрации входящей и исходящей корреспонденции
Приказы ректора университета
Документооборот и отчетность
Продукт
Удобство и простота
Выбор онлайн услуги
Основные процессы
Выгрузка готового ответа
Скорость обслуживания
Регистрация обращения в базу входящего реестра
Генерация ответа в виде PDF документа
Регистрация документа в базе исходящих документов
Полученная онлайн услуга
Централизация всех соц. услуг
Вспомогательные процессы
Тех. обслуживание
Интеграция с другими системами университета
Обновление ПО

Рисунок 1 - Карта процесса «Получение онлайн услуг в системе “Smart-campus”пан»

На рисунке 1 приведена карта процесса «Получение справки с места обучения». На карте процесса показаны следующие процессы:

Основные процессы:

- Выбор онлайн услуги;
- Регистрация обращения в базе входящего реестра;
- Генерация ответа в виде PDF документа;
- Регистрация номера выданного документа в базе исходящих документов;
- Выгрузка готового ответа.

Процессы управления:

- Правила регистрации входящей и исходящей корреспонденции;
- Документооборот и отчетность;
- Приказы ректора университета.

Вспомогательные процессы:

- Тех. обслуживание;
- Обновление ПО;
- Интеграция с другими системами университета.

B. BPMN модель бизнес-процессов

BPMN модель включает в себя набор концепций, необходимых для моделирования процессов. BPMN модель является связующим звеном между техническими разработчиками и бизнес-пользователями процесса. Модель также отражает все существующие интеграции и связи различных систем [2]. Ниже, на рисунке 2, приведена BPMN модель процесса получения онлайн услуг в системе «Smart campus».

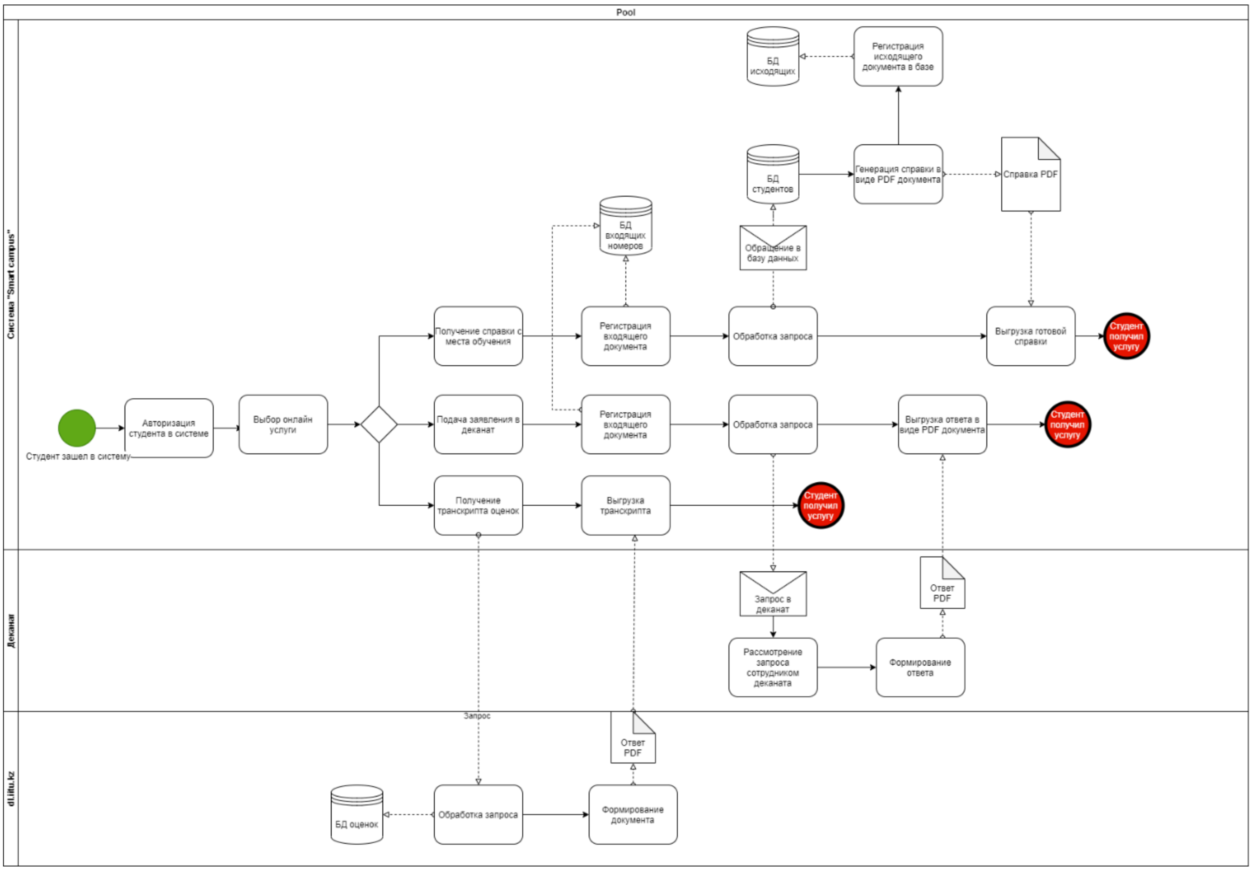

Рисунок 2 - BPMN модель процесса «AS IS»

Регламент бизнес-процесса

Цель процесса — автоматизировать и упростить порядок получения справок, транскриптов, отчетов по оплате, подачи заявлений и других услуг посредством их получения в режиме онлайн, без ожидания своей очереди на прием к ответственным за это людям. Помимо облегчения процесса для обучающихся автоматизация получения данных услуг экономит ресурсы персонала университета, значительно сокращая их задачи. Владельцем процесса является университет в лице директора департамента технического сопровождения и IT-поддержки.

Владелец процесса несет ответственность за:

- бесперебойную работу системы в любое время суток;
- быстрое реагирование на запросы технических неполадок и их решение;
- своевременное обновление данных и инструментов системы;
- интеграцию системы с другими департаментами и внутренними системами;
- выполнение процесса и его результат.

Обычно определение подхода к управлению процессами включает следующие этапы:

- Документирование процесса для понимания того, как работа проходит через процесс;

- Присвоение права собственности на процесс с целью установления управленческой подотчетности;

- Управление процессом для оптимизации некоторых показателей эффективности процесса;

- Улучшение процесса для повышения качества продукции или показателей эффективности процесса;

- Управляя процессами, можно лучше интегрировать перспективы и приоритеты с ресурсами;

- Многие новые управленческие инициативы требуют управления процессами, и их невозможно реализовать;

- Управление процессами открывает двери для творческих и новаторских подходов к улучшению организационной деятельности;

- Управление процессами позволяет эффективно внедрять современные системы и стандартное программное обеспечение [3].

Можно сделать вывод, что для поддержки инноваций в данном процессе необходим систематический подход к проектированию и анализу. Ниже, в таблицах 1 и 2, показаны описание процесса получения онлайн услуг и характеристика данного процесса.

Таблица 1 –Описание процесса в системе “Smart-campus”

Название процесса	**Тип процесса**	**Цель / назначение процесса**	**Владелец процесса**
Получение онлайн услуг в системе “Smart campus”.	Основной	Упростить порядок получения справок, транскриптов, отчетов по оплате, подачи заявлений и других услуг посредством их получения онлайн, не ожидая очереди на прием к ответственным за это людям.	ВУЗ, администрация ВУЗа, департамент технического сопровождения и IT-поддержки.

Таблица 2- Характеристика процесса в системе “Smart-campus”

Основное событие начала	**Основной вход**	**Основной поставщик**	**Основное событие окончания**	**Основной продукт**	**Основной клиент**
Регистрация в системе “Smart campus” / выбор онлайн услуги	1) Запрос на получение онлайн услуги (получение справки, транскрипта, информация об оплате); 2) Данные клиента(студента)	Система “Smart-campus”	Клиент(студент) получил онлайн-услугу (готовый документ)	Справка об обучении, транскрипт, отчет по оплате	Обучающийся в организации образования (ВУЗа)

Модель бизнес-процессов «TO BE». Модель бизнес-процессов «TO BE», то есть «как должно быть» отражает оптимизированные и доработанные аспекты модели «как есть»

[4]. Перед построением данной модели «как должно быть» необходимо исследовать плюсы и минусы модели «как есть», проанализировать возможные варианты улучшения эффективности процессов и выбрать подходящие методы решения данных проблем. Ниже, в таблице 3, приведены основные проблемы и их решения.

Таблица 3 – Проблемы модели «как есть»

Проблемы	Решения
Сбой системы	1. Принимаются меры по устранению департаментом технического сопровождения и IT поддержки. 2. В случае затяжных технических работ необходимо прописать в инструкции, что заявка подается альтернативными способами (e-mail, whats app и т.д.), затем заявка обрабатывается вручную сотрудником деканата.
Проверка подлинности документа	Для того чтобы справка имела юридическую силу, необходимо внедрить ЭЦП, а для проверки подлинности в углу документа генерировать QR код, который откроет с устройства электронный документ по адресу официального сайта университета [5].
Аутентификация пользователя	Добавление номера телефона и смс-подтверждение, а также внедрение ЭЦП.

После исследования всех элементов модели «как есть» BPMN модель бизнес-процессов «Получение онлайн услуг в системе “Smart campus”» была оптимизирована. Модель после оптимизации «TO BE» («как должно быть») показана ниже, на рисунке 3.

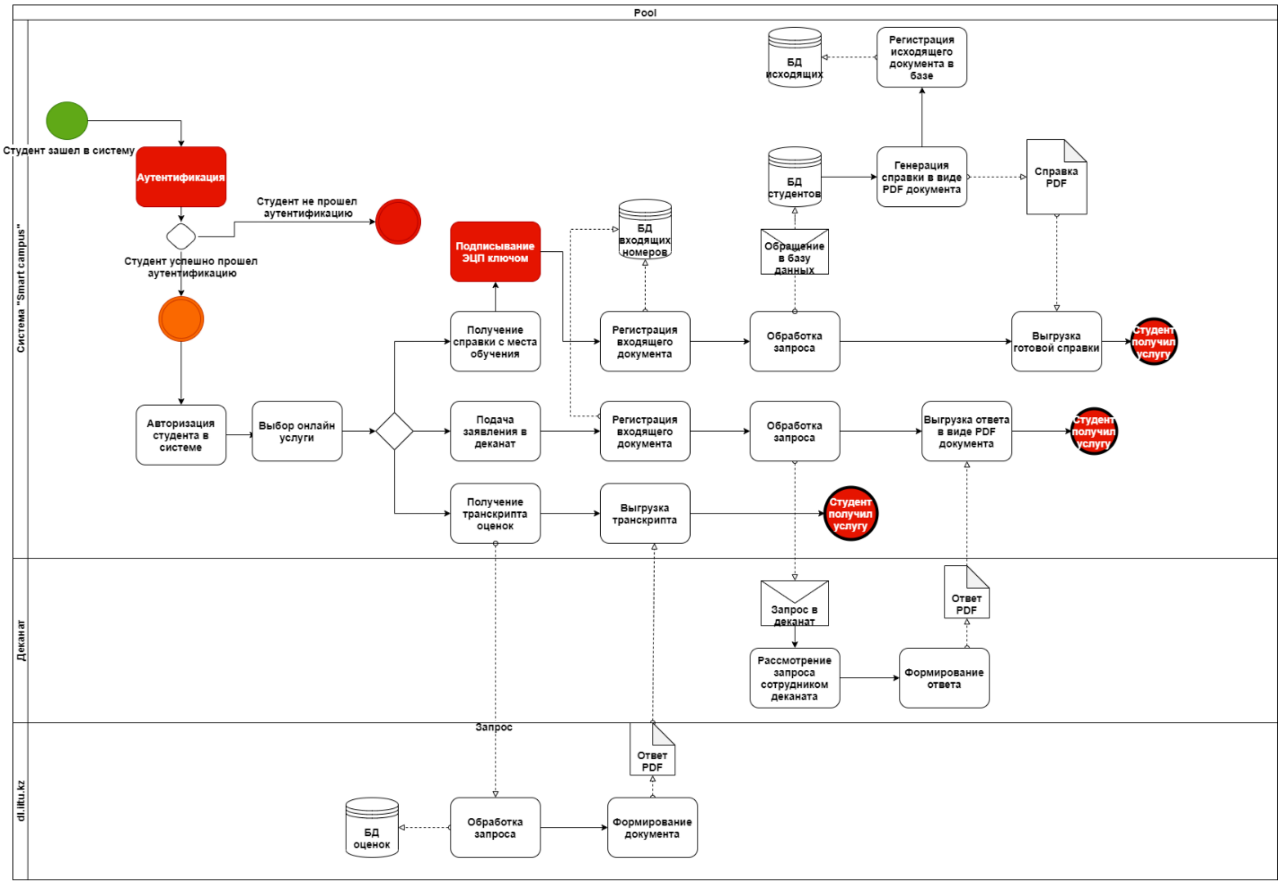

Рисунок 3 - BPMN модель процесса «TO BE»

В заключении отметим, что в данной статье была разработана модель бизнес-процессов получения онлайн услуг в система «Smart campus», также разработанная модель была оптимизирована после исследования и анализа. Были описаны основные бизнес-процессы предоставления онлайн услуг в университете, выявлены проблемы существующих моделей и предложены пути их решения.

СПИСОК ЛИТЕРАТУРЫ

1. Веб-приложение учета и регистрации абитуриентов АО «МУИТ»: система «CAMPUS IITU», [Электронный ресурс] URL: https://campus.iitu.kz. (дата обращения: 10.04.2021)
2. BPMN 2.0 ИЗ ЧЕГО СОСТОИТ МОДЕЛЬ БИЗНЕС ПРОЦЕССА, [Электронный ресурс] URL: https://rzbpm.ru/knowledge/bpmn-2-0-iz-chego-sostoit-model-biznes-processa.html (дата обращения: 11.04.2021)
3. Регламент бизнес-процесса. Глоссарий, [Электронный ресурс] URL: https://www.businessstudio.ru/articles/article/glossariy_reglament_protsessa (дата обращения: 15.04.2021)
4. Портал СМК ВКТУ, [Электронный ресурс] URL: https://www.ektu.kz/abouttheuniversity/qms/processmap.aspx?lang=ru (дата обращения: 15.04.2021)
5. Портал электронного правительства - EGOV,[Электронный ресурс] URL: https://www.egov.kz (дата обращения: 15.04.2021)

REFERENCES

1. *Veb-prilozhenie ucheta i registracii abiturientov AO "MUIT": sistema "CAMPUS IITU"*, [The web application of accounting and registration of enrollees of JSC "IITU"], [Electronic resource] URL: https://campus.iitu.kz (accessed: 10.04.2021)
2. *BPMN 2.0 IZ CHEGO SOSTOIT MODEL BIZNES PROTSESSA*, [BPMN 2.0 WHAT IS A BUSINESS PROCESS MODEL] [Electronic resource] URL: https://rzbpm.ru/knowledge/bpmn-2-0-iz-chego-sostoit-model-biznes-processa.html. (accessed: 11.04.2021)
3. *Reglament biznes-protsessa. Glossariy*, [Business process regulations. Glossary], [Electronic resource] URL: https://www.businessstudio.ru/articles/article/glossariy_reglament_protsessa (accessed: 15.04.2021)
4. Portal SMK VKTU, [SMK VKTU portal], [Electronic resource] URLhttps://www.ektu.kz/abouttheuniversity/qms/processmap.aspx?lang=ru (accessed: 15.04.2021)
5. Portal elektronnogo pravitelstva - EGOV, [E-government portal- EGOV], [Electronic resource] URL: https://www.egov.kz (accessed: 15.04.2021)

Базарбеков И.М., Шарипов Б.Ж.

Білім беру ұйымында онлайн қызмет көрсету үшін бизнес-процессін дамыту

Аңдатпа. Бұл мақала «ХАТУ» АҚ Университетінің мысалында білім беру ұйымдарында онлайн-қызметтерді пайдалану процестерін дамыту туралы. Онлайн-қызметтер дегеніміз: оқу орнынан сертификат алу, транскрипт алу, өтінімдерді онлайн режимінде беру, қарызды жою. Технологиялық карта және BPMN «AS IS» және «TO BE белгілеу модельдері» әзірленді.

Түйінді сөздер: «Smart-кампус», Business Process Model and Notation, "AS IS", "TO BE", онлайн қызмет көрсету, жүйе, электрондық цифрлық қолтаңба.

Bazarbekov I.M., Sharipov B.Zh.

Development of a business process for obtaining online services in the organization of education

Abstract. This article presents a case study of IITU JSC illustrating expansion in the use of online services at educational organizations. Online services mean such services as: obtaining a certificate from the place of study, receiving a transcript, submitting applications online, eliminating debt. To streamline the process of rendering such services online we developed a process map and some BPMN notation models such as "AS IS" and "TO BE".

Keywords: Smart-campus, Business Process Model and Notation, “AS IS”, “TO BE”, online services, system, electronic digital signature.

Авторлар туралы мәлімет:

Базарбеков Икрам Медеуұлы «Ақпараттық жүйелер» кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті.

Шарипов Бахыт Жапарович, т.ғ.к., п.ғ.д, «Ақпараттық жүйелер» кафедрасының профессоры, білім беру инновациясы және SMART оқыту орталығының директоры, ХИНА академигі, РЖҒА шетелдік академигі, Халықаралық ақпараттық технологиялар университеті.

Сведения об авторах:

Базарбеков Икрам Медеуұлы, магистрант кафедры «Информационные системы», Международный университет информационных технологий.

Шарипов Бахыт Жапарович, к.т.н., д.п.н., профессор кафедры «Информационные системы», академик МАИН, иностранный академик РАЕН, директор центра образовательных инноваций и SMART обучения Международного университета информационных технологий.

About the authors:

Ikram M. Bazarbekov, master student, Department of Information Systems, International Information Technology University.

Bakhyt Zh. Sharipov, Dr of P.Sci, Cand. of Techn.Sci. , academician of IAIN, foreign academician of RANS, Director of the Center for Educational Innovation and SMART Training, Professor, Department of Information Systems, International Information Technology University.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2.Is . 4. Number 08(2021). 69–74
Journal homepage: https://journal.iitu.edu.kz

https://doi.org/10.54309/IJICT.2021.08.4.009

УДК 53.06

Жунусов Д.О.[*], Алиаскаров С.Ж.

Международный университет информационных технологий, Алматы, Казахстан

МЕТОД КЛАССИФИКАЦИИ ТЕКСТОВ НА ОСНОВЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Аннотация. В этой работе были описаны способы расчета и анализа метода классификации текстов, а также были выявлены основные параметры применения алгоритмов машинного обучения для обработки естественных языков. Для этого были проанализированы и показаны некоторые реализации алгоритмов машинного обучения. Для разработки системы по анализу и категоризации текстов были применены алгоритмы по обработке текстовых данных, а также результаты исследования системы обработки естественного языка.

Ключевые слова: обработка, естественные языки, классификация, анализ, категоризация, распознавание, понимание

Введение

В цифровом мире важность обработки информации возрастает с каждым днем. Быстрое развитие технологий и разработка систем по мониторингу социальных сетей определили одно из ключевых направлений деятельности индустрии ИТ — создание систем по обработке естественного языка, которые могут обрабатывать любые виды текстовой информации: как новостные публикации и статьи, так и их комментарии. На сегодняшний день это проявляется в виде таргетированных предложений, составления портрета пользователя и прогнозирования действий на основе понимания текстов. Одной из самых востребованных технологий являются алгоритмы, основанные на NLP (natural language processing), поскольку данная сфера еще мало изучена. Обработка естественного языка — область, которая возникла в результате синхронизации таких наук, как лингвистика и математика [1]. Но одной классификации для точного понимания текста недостаточно. Необходимо категоризировать текст для дальнейшей обработки. Классификация текстов — это процесс синтаксического анализа текстов в различных сферах для дальнейшей работы с данными [2]. Далее следует визуализация данных и их предобработка для принятия решений в информационных системах. Однако для классификации текстов необходимы точные алгоритмы машинного обучения и большие корпусы размеченных данных, которые создают риски потери семантического значения категорий. Поэтому для обработки текстов, в частности для классификации текстов по категориям, требуется более эффективный метод анализа. Использование комплексных, более сложных методов анализа текстов может уменьшить вероятность ошибок, но повлечет увеличение требований к ресурсам. Предполагается, что есть необходимость в повышении точности и использовании более продвинутых алгоритмов для классификации.

Методология

В данной статье будут рассматриваться методы классификации текстов, поэтому объектом исследования является анализ текста. Не существует стандартизированных шаблонов для анализа текстов. Анализ текста является примером области междисциплинарного исследования, соответственно, понятие анализа трансформируется в зависимости от области и от того, каков общий контекст статьи, связанной с данным понятием.

Классификация — довольно распространенная задача в машинном обучении. Она исследуется в таких сферах, как: распознавание и обработка изображений, сегментация объектов, нахождение дистанций и параметров, медицинские задачи и проблемы науки. Специалисты машинного обучения, работающие над сложными алгоритмами нейронных сетей, стремятся обеспечить их необходимое функционирование в определенных, заранее неизвестных условиях, то есть добиваются от разрабатываемых систем и алгоритмов необходимых форм поведения. Лингвисты изучают информацию, определяющую составное функционирование текста, а также единицу каждого слова в нужном контексте. Ключевыми функциями когнитивной лингвистики являются особенности усвоения и обработки информации [3]. Условно говоря, обработка текстовой информации — это нахождение последовательностей в текстах, поступивших из различных информационных источников. Это определение обработки информации в данной статье взято за основу, то есть текст рассматривается как единица вычисления, а классификация текстов подразумевает анализ того, как они себя показывают в разных условиях. Вслед за определением понятий нужно также выявить измеримые и высчитываемые параметры. В данном случае это контекст, в котором единицы участвуют для дальнейшей предобработки.

В рамках данного исследования метод классификации выступает как основа для дальнейших взаимодействий в системе, дополняя систему в качестве полноценного цикла понимания текста. Поэтому первичным действием, определяющим категорию текста, является сегментация. Каждая единица текста имеет определенный вес и будет рассматриваться в рамках всего контекста, в котором будут выстраиваться связи.

Материал для изучения был получен методом выгрузки корпуса больших данных по заметкам в Википедии и данных, находящихся в открытом доступе. В общей сложности было получено 600,000 оригинальных текстов. Для успешной реализации, поставленной задачи корпус данных был расширен методом исследования социальных сетей для выгрузки наибольшего числа оригинальных словосочетаний и предложений. Анализ сплошного содержания заметок был проведен методом выявления главных тем, методом прибавления первичного корпуса тегов исследования. Теги исследования задавались для поиска по социальным сетям и по новостным порталам. Роль человека в сборе была минимальной. В корпус больших данных методом оценки уникальности добавлялись предложения и создавали кластеры оригинальных услуг. Основная масса данных, собранных по тегам, была на русском и казахском языках. Все наборы данных не форматировались и не обрабатывались в процессе сбора, и в корпус могли попадать стоп-слова и малозначимые тексты. В общей сложности были задействованы программы для сбора данных из двенадцати социальных сетей и восьмисот новостных порталов, что позволило повысить качественный и количественный состав корпуса для дальнейшей обработки.

Для сбора данных из открытых источников использовалась техника «поиск-ключ», которая работает на основе языка python. Эта техника включает в себя инструменты для выгрузки html варианта страницы и поиска по тегам. Теги, соответствующие теме, были получены через поисковые системы методом поиска внутри источников.

Для сбора данных из новостных порталов был применен поиск по темам и дальше по тегам. Портал выгружался отдельным скриптом через определенные периоды времени и сохранялся в корпус для последующей обработки. Поиск слов проводился по новостным публикациям и по комментариям под ними путем установления фильтрации по времени (от новых к старым).

В общей сложности было собрано 21,231 текстов по социальным сетям и новостным порталам. Например, проводился поиск по тегам на политические и общественные темы, статистика обрисовывала настроение населения в определенный отрезок времени.

Классификации по типу были разбиты на несколько частей по функциям и базисным тегам. При делении на категории учитывались смежные элементы. Впоследствии выявлялось процентное соответствие близости текста к конкретному классу. Лингвистические особенности всех текстов различались и были проанализированы.

Для автоматической классификации текстов деление собранного материала производилось посредством инструментов, которые выдавали более близкую, к слову, категорию для прибавления.

Из библиотеки keras были применены модели и layers. Keras — открытая нейросетевая библиотека, написанная на языке Python [4]. Целью Keras является оперативная работа с глубоким обучением. Собранные данные были приведены в исходную форму в массиве словосочетаний, а вслед затем и текстов. Для обеспечения очередности в перечне по схожести длины была применена pad_sequences из keras. По умолчанию это делается путем добавления 0 в начале каждой последовательности, пока каждая последовательность не будет иметь такую же длину, как и самая длинная последовательность. Для выявления категорий был применен алгоритм машинного обучения Multinomial Naive Bayes. Тренировка позволяет минимизировать издержки.

В рамках реализации нашего инструмента для систематизации по категориям мы сделали наш классификатор. Для этого были применены техники удаления стоп-слов, замены прописных букв на строчные и приведения текстов в исходную форму. Дальше была задействована функция Tokenizer из библиотеки keras. Tokenizer преобразовывал слова в очередности и переводил их в векторное представление как числовое. Дальше с поддержкой pad_sequences из библиотеки keras preprocessing мы возвращали очередности с интервалами и метками.

Для наибольшей сложности систематизации категорий полученные результаты из нашего классификатора по работе с входными данными мы запускали в функцию predict нашей модели, построенной на основе MultinomialNB (Multinomial Naive Bayes). Multinomial Naive bayes — один из вариантов классификаторов, построенных на алгоритме NB, который используется в основном в задачах обработки полиномиально распределенных данных, например, в классификации текста [5]. Для оптимизации рабочего процесса мы предварительно сохраняли подборки итога тренировки модели в формате h5 и загружали итог через функцию библиотеки keras load_model. В итоге получали результат, с поддержкой индексирования итога выводили категорию. Первые результаты категоризации показывают точность модели около 80% (табл. 1).

Таблица 1. Первый результат категоризации текстов на примере большого корпуса данных.

	precision	recall	f1-score	support
0	0.78	0.83	0.8	2777
1	0.82	0.77	0.8	2763
micro-avg	0.8	0.8	0.8	5540
macro-avg	0.8	0.8	0.8	5540
weighted-avg	0.8	0.8	0.8	5540

Чтобы обрабатывать разные случаи ввода, мы можем собирать данные из различных источников и сфер деятельности, так повысится качество собранного набора данных, что даст положительный эффект в работе модели. В процессе внедрения модели мы увеличивали точность модели с использованием логистической регрессии, и результат точности увеличился до 84%. Метод будет хорошо работать на различных типах данных, включая иронию и сатиру, так как постоянно идет улучшение за счет количества и качества данных.

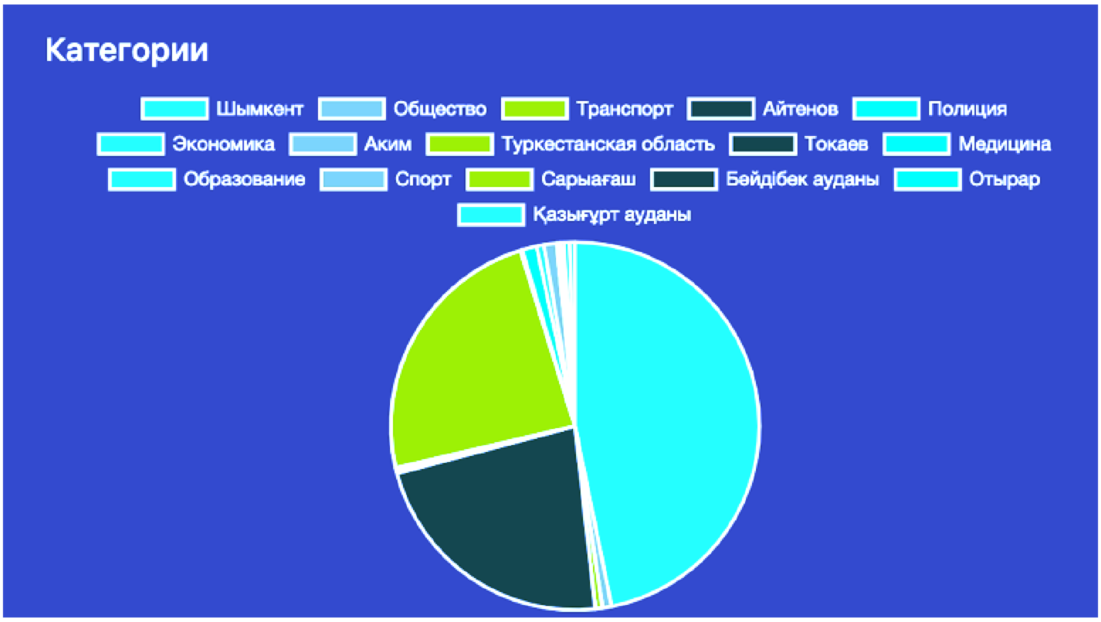

Рисунок 2 - Результат категоризации текстов на примере большого корпуса данных после улучшений модели

Как и все остальные системы, теоретически систему классификации, основанную на анализе текста, можно обманывать, если тексты приходят не в подготовленных шаблонах, но для этого есть процент погрешности. Несмотря на это, на практике система работает стабильно и показывает хорошие результаты по классификации текстов. Тексты, приходящие из информационных систем, а также из различных источников, классифицируются по категориям и дают возможность для последующей обработки и анализа естественного языка. Эта работа основана на гипотезе, что тексты являются неструктурированными и собираются из разных источников, но есть процент погрешности выявления категорий в зависимости от самой области контекста. Решением данной проблемы является увеличение корпуса данных, тренировка моделей и работа по чистке данных и выявлению ключевых параметров.

Заключение

В заключении нужно отметить, что целью классификации текстов по их признакам является обеспечение целостного понимания и обработки естественных языков. Классифицировать текст можно только сравнивая данные, получаемые от единиц текста, которые являются частью контекста. Так как общий контекст играет ключевую роль в классификации текста, необходимо учесть особенности языкового корпуса, токенизацию слов и приведение в изначальную форму. Для того, чтобы не было отклонений в точности результатов, необходимо работать над чистотой корпуса и уделять существенное внимание проработке модели и выстраиванию параметров. Таким образом данная модель будет иметь процент погрешности в расчетах, но будет постоянно самообучаться и давать результаты лучше с каждым разом, так как корпус текста нужно будет увеличивать и уделять внимание чистке.

СПИСОК ЛИТЕРАТУРЫ

1. Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. Автоматическая обработка текстов на естественном языке и анализ данных, НИУ ВШЭ, Москва, 2017, 124-128 с.

2. Рафанов С.М. К проблеме классификации текстов в машинном переводе, КРСУ, Москва, 2013, 36-42 с.
3. Попова З.Д., Когнитивная Лингвистика, Федеральное агентство по образованию Воронежский Государственный Университет, Воронеж, 2007, 7-12 с.
4. Ketkar, Nikhil. (2017). Introduction to Keras. 10.1007/978-1-4842-2766-4_7.
5. Xu, Shuo & Li, Yan & Zheng, Wang. (2017). Bayesian Multinomial Naïve Bayes Classifier to Text Classification. 347-352. 10.1007/978-981-10-5041-1_57.

REFERENCES

1. Bol'shakova E.I., Voroncov K.V., Efremova N.E., Klyshinskij E.S., Lukashevich N.V., Sapin A.S. *Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i analiz dannyh* [Automatic Natural Language Processing and Data Analysis] NIU VSHE, Moscow, 2017, 124-128 c.
2. Rafanov S.M. K probleme klassifikacii tekstov v mashinnom perevode [the problem of text classification in machine translation], KRSU, Moscow, 2013, 36-42 c.
3. Popova Z.D., Kognitivnaya Lingvistika [Cognitive Linguistics], Federal'noe agentstvo po obrazovaniyu Voronezhskij Gosudarstvennyj Universitet, Voronezh, 2007, 7-12 c.
4. Ketkar, Nikhil. (2017). Introduction to Keras. 10.1007/978-1-4842-2766-4_7.
5. Xu, Shuo & Li, Yan & Zheng, Wang. (2017). Bayesian Multinomial Naïve Bayes Classifier to Text Classification. 347-352. 10.1007/978-981-10-5041-1_57.

Жунусов Д.О., Алиаскаров С.Ж.

Машинналық оқыту алгоритмдері негізінде мәтіндер классификациясының әдісі

Аңдатпа. Бұл жұмыста мәтінді жіктеу әдісін есептеу және талдау әдістері сипатталған. Сонымен қатар, табиғи тілді өңдеуге арналған машиналық оқыту алгоритмдерін қолданудың негізгі параметрлері анықталды. Бұл жұмыс үшін машиналық оқыту алгоритмдерінің кейбір енгізілімдері талданды және көрсетілді. Мәтіндерді талдау және санаттарға бөлу жүйесін құру үшін мәтіндік деректерді өңдеу алгоритмдері, сонымен қатар табиғи тілді өңдеу жүйесін зерттеу нәтижелері қолданылды.

Түйінді сөздер: өңдеу, табиғи тілдер, жіктеу, талдау, санаттарға бөлу, тану, түсіну.

Zhunissov D.O., Aliaskarov S.Zh.

Method for text classification based on machine learning algorithms

Abstract. This work describes the methods of calculating and analyzing the text classification and identifies the main parameters of applying machine learning algorithms for natural language processing. The authors analyzed some implementations of machine learning algorithms, applied such algorithms for processing text data to develop a system for the analysis and categorization of texts and demonstrated the results of their study of a natural language processing system.

Keywords: processing, natural languages, classification, analysis, categorization, recognition, understanding.

Авторлар туралы мәлімет:

Алиаскаров Серик Женисханович, «Ақпараттық жүйелер» кафедрасының докоранты, Халықаралық ақпараттық технологиялар университеті.

Жунусов Дархан Ондасынулы, «Ақпараттық жүйелер» кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті.

Сведения об авторах:

Алиаскаров Серик Женисханович, докторант кафедры «Информационные системы», Международный университет информационных технологий.

Жунусов Дархан Ондасынулы, магистрант кафедры «Информационные системы», Международный университет информационных технологий.

About the authors:

Serik Zh. Aliaskarov, doctoral student, Department of Information Systems, International Information Technology University.

Darkhan O. Zhunussov, master student, Department of Information Systems, International Information Technology University.

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 2. Is. 4. Number 08 (2021). Pp.75–79
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/10.54309/IJICT.2021.08.4.010
УДК 615.035.4)

Синчев Б.

Международный университет информационных технологий, Алматы, Казахстан

О ПОЛИНОМИАЛЬНОЙ РАЗРЕШИМОСТИ КЛАССА NP-COMPLETE

Аннотация. В множестве четных и нечетных неотрицательных целых чисел X^n мощности n найти подмножество $X^k \subseteq X^n$ мощности $K=[X^k]$, сумма элементов которого равна сертификату S. Поставленная задача относится к классу NP-complete. Доказаны леммы о полиномиальной разрешимости задачи о сумме подмножеств X^k с мощностью k, удовлетворяющей условиям $n \geq 3, k(n) \geq 3 \ \vee \ n > 3$, $k(n) \leq n < k^2(n) - k(n)$ и найдено фиксированное значение мощности k из выше приведенного интервала. Предлагаемый полиномиальный метод решения задачи о сумме подмножеств обеспечивает решение других проблем класса NP-complete с помощью сводящих функций и справедливость равенства классов $P = NP$ на базе известной теоремы: если некоторая NP-полная задача разрешима за полиномиальное время, то $P = NP$.

Ключевые слова: класс NP-complete, полиномиальная разрешимость, задача о сумме подмножеств

Введение

Любая задача из класса NP может быть решена полным перебором. При этом, даже если вычисление целевой функции от каждого конкретного возможного решения задачи может быть осуществлено за полиномиальное время, в зависимости от количества всех возможных решений полный перебор может потребовать экспоненциального времени работы. В теории алгоритмов известны несколько широко применимых общих концепций. Метод полного перебора является одной из них. Детальный обзор последних работ [1,2] о псевдополиномиальных и экспоненциальных методах показывает, что важная задача о сумме подмножеств в теории сложности алгоритмов относится к основной из трудных проблем класса NP-complete. В работах [3,4] предложены полиномиальные алгоритмы решения задачи о сумме подмножеств. Однако в задаче перебора и задаче о сумме подмножеств очень много нерешенных проблем. К этим проблемам можно отнести:

- существование мощности *k* подмножества $X^k \subseteq X^n$;
- нахождение нижней и верхней границы мощности *k*;
- определение фиксированного значения мощности *k*;
- взаимосвязь между *n* и *k*;
- определение минимального значения мощности подмножества, получаемого при расщеплении исходного множества X^n на подмножества.

Постановка задачи

Задача о сумме подмножеств формулируется в виде:

$$\sum_{i=1}^{n} \alpha_i x_i = S, \alpha_i \in \{0,1\}, x_i \in X^n, \ \ i \in N, \qquad (1)$$

где X^n – множество целых четных и нечетных неотрицательных чисел, мощность n – $[X^k]$, $x_i < +\infty$, *N*-множество натуральных чисел с мощностью $n = |N|, n < +\infty$. Предполагается, что $S - x_i > 0, x_i \in X^n, \ i \in N$.

Параметризованной формой задачи (1) будем называть следующую задачу:

$$\sum_{i=1}^{k} \alpha_i x_i = S, \alpha_i \in \{0,1\}, x_i \in X^n, \alpha_i = 1, i \in K, \alpha_i = 0, i \in N \setminus \ K \cdot \qquad (2)$$

где $X^k \subseteq X^n$, $k = |X^k|$, $k \leq n$, $K \subseteq N$, $k = |K|$, K – подмножество индексов всех выбранных переменных $x_i \in X^n$, N\K-подмножество индексов всех остальных (невыбранных) переменных $x_i \in X^n$ подзадачи (2). Отметим, что задача (1) является частным случаем задачи (2), когда подмножество $K = \emptyset$.

Полиномиальный разрешимость NP-complete

Известно много разных обобщений классической проблемы Варинга[5] для полиномов. Мы будем подразумевать под проблемой Варинга для полиномов следующую задачу: для данного натурального числа n найти минимальное число k=k(n), для которого любой полином $g \in C[x]$ может быть представлен в виде $g = f_1^n + f_2^n + \cdots + f_k^n$, где $f_i^n \in C[x]$. При решении проблемы Варинга достаточно ограничиться случаем $g(x) = x$. Действительно, если $x = f_1^n(x) + f_2^n(x) + \cdots + f_k^n(x)$ и $h(x)$-произвольный полином, то $h(x)$= $f_1^n\big(h(x)\big) + f_2^n\big(h(x)\big) + \cdots + f_k^n\big(h(x)\big)$. Тождество $\left(x + \frac{1}{4}\right)^2 - \left(x - \frac{1}{4}\right)^2 = x$ показывает, что k(2)=2.

Лемма1. Если множество X^n с заданной мощностью n, то существует некоторый сертификат, представимый в виде $\tilde{S} = x_1 + x_2 + \cdots + x_k$ с элементами из подмножества X^k мощности k (X^k⊆X^n), удовлетворяющей неравенствам:

$$n \geq 3,\ k(n) \geq 3\ \vee n > 3,\ k(n) \leq n < k^2(n) - k(n). \qquad (3)$$

Доказательство. Представим эквивалентную задаче(1) полиномиальную постановку задачи о сумме подмножеств: необходимо найти подмножество X^k⊆X^n с сертификатом $S = \sum_{i=1}^{k} x_i$, равным второму коэффициенту a_1 полинома

$$a^k(x) = x^k - Sx^{k-1} + a_2x^{k-2} + \cdots + a_k, \qquad (4)$$

удовлетворяющего следующим соотношениям:

$$a^k(x)b^{n-k}(x) = c^n(x), \qquad (5)$$

где известный полином $c^n(x)$ степени n с заданными коэффициентами,

$$c^n(x)\text{=}x^n - Qx^{n-1} + c_2x^{n-2} + \cdots + c_n,\ Q = \textstyle\sum_{i=1}^{n} x_i. \qquad (6)$$

Коэффициенты полинома(6) находятся основе теоремы Виета, корнями которого являются элементы $x_i \in X^n$. Полином $b^{n-k}(x)$ находится на основе соотношения(5)

$$b^{n-k}(x)\text{=}x^{n-k} - (Q - S)x^{n-k-1} + b_2x^{n-k-2} + \cdots + b_{n-k}. \qquad (7)$$

Коэффициенты полинома(7) определяются на основе деления полинома $c^n(x)$ на полином $a^k(x)$ с применением алгоритма Евклида.

Согласно проблеме Варинга для полиномов и теореме Неймана-Слейтера [19] существует полином $h(x)$= $f_1^n\big(h(x)\big) + f_2^n\big(h(x)\big) + \cdots + f_k^n\big(h(x)\big)$, где f_i^n –полиномы степени n, которые формируются на основе полинома $c^n(x)$ с разными знаками коэффициентов. В силу произвольности этого полинома $h(x)$ мы можем подобрать полином

$$a^k(x)\text{=}\ h^k(x) \qquad (8)$$

степени k, причем второй коэффициент $h_1 = -S$ полинома $h^k(x)$, а в силу теоремы Виета величина $S = x_1 + x_2 + \cdots + x_k$. Здесь важным является то, что этот коэффициент h_1 состоит из k произвольных элементов множества X^n. Последнее показывает существование подмножества X^k мощности k из множества X^n мощности n. Эти мощности k, n удовлетворяют неравенствам(3).

Задача о сумме подмножеств требует решения вопроса о существовании сертификата S в виде суммы с ограниченным (минимальным) числом элементов. Классическая проблема Варинга заключается в том, чтобы для данного натурального числа n найти минимальное число k=k(n), для которого любое натуральное число m может быть представлено в виде $m = m_1^n + m_2^n + \cdots + m_k^n$, где $m_1, m_2, \ldots, m_k$ –целые неотрицательные числа. Эта проблема Варинга была доказана Д. Гильбертом в 1909 году.

Лемма2. Пусть существуют параметр k, число $m = x_1^n + x_2^n + \cdots + x_k^n$ и элементы x_i из множества X^n. Тогда для задачи(1) найдутся сертификат $S < m$ и подмножество X^k мощности k (X^k⊆X^n), сумма элементов которого равна $x_1 + x_2 + \cdots + x_k = S$.

Доказательство. Согласно классической гипотезе Варинга справедливо

$$m/k = (x_1^n + x_2^n + \cdots + x_k^n)/k. \qquad (9)$$

Далее воспользуемся известными неравенствами:

$$\frac{x_1^n+x_2^n+\cdots+x_n^n}{n} \geq x_1x_2 \ldots x_n,\quad \frac{x_1+x_2+\cdots+x_n}{n} \geq (x_1x_2 \ldots x_n)^{\frac{1}{n}}. \qquad (10)$$

Если заменить элементы x_i^n в равенстве(9) на $x_i^{1/n}$ при учете неравенства $(x_1x_2 \ldots x_k)^{1/k} \geq (x_1x_2 \ldots x_k)^{\frac{1}{n}}$, так как $k \leq n$. Тогда с учетом неравенств(10) получим

$$\frac{x_1+x_2+\cdots+x_k}{k} \geq (x_1 x_2 \ldots x_k)^{1/k}. \qquad (11)$$

Из неравенства(11) имеем, что $\frac{S}{k} = \frac{x_1+x_2+\cdots+x_k}{k} \geq (x_1 x_2 \ldots x_k)^{\frac{1}{k}}$. Это показывает, что найдется подмножество X^k, состоящее из k элементов, причем их сумма $x_1 + x_2 + \cdots + x_k = S$.

Полученный результат коррелирует с классической проблемой Варинга в теории чисел. В работах [6] и [3,4] найдены интервалы изменения мощности k подмножества X^k

$$0 \leq k \leq n/2, \; 0 \leq k \leq n/4 \qquad (12)$$

соответственно.

Предложенные леммы позволяют определить значение мощности k подмножества X^k при заданной мощности n множества X^n и сертификате S. Найденные интервалы изменения мощности k $(k(n) \geq 3 \;\vee\; k(n) \leq n < k^2(n) - k(n))$ имеют важное значение при решении задачи о сумме подмножеств (1)-(2).

Действительно, найдены уточненные пределы изменения мощности k подмножества $X^k \subseteq X^n$ по сравнению с ранее полученными интервалами (12).

В частности, из лемм получаем, что при n =1024 мощность k принадлежит интервалу $3 \leq k < 32$, что несравнимо с интервалами $0 \leq k \leq 512$, $0 \leq k \leq 256$, следующих из (12). В этом случае длина входных данных n удовлетворяет неравенству $32 \leq n \leq 998$. Данный интервал позволяет разбить исходное множество на подмножества меньшей размерности на основе метода «разделяй и властвуй». При этом наименьшая размерность подмножества может быть равна 32. Тем самым, мы показали способ разбиения множества X^n на подмножества X^k с определением мощности k.

Предложенные леммы позволили решить поставленные проблемы во введении, причем

$$n \geq 3, \; k(n) \geq 3 \;\vee\; n > 3, \; k(n) \leq n < k^2(n) - k(n). \qquad (13)$$

Однако эти леммы не дают окончательного определения фиксированного значения мощности k подмножества X^k поставленной задачи(1). Поэтому для окончательного определения фиксированного значения мощности k подмножества X^k воспользуемся алгеброй полиномов[7].

При нахождении корней полинома $c^n(x)$ используется классическое преобразование

$$y = x - c_1/n \qquad (14)$$

для перевода полинома $c^n(x)$ в полином $c^n(y)$ с другим аргументом, аналогично может быть использовано преобразование(14) для полиномов $a^k(x)$, $b^{n-k}(x)$ в виде

$$y = x - \frac{a_1}{k}, y = x - \frac{b_1}{n-k} \qquad (15)$$

соответственно.

Из (14), (15) на основе формул (4-8) имеем

$$x^{c}_{arithm} = Q/n, \qquad (16)$$

$$x^{a}_{arithm} = \frac{S}{k}, \quad x^{b}_{arithm} = \frac{Q-S}{n-k}. \qquad (17)$$

Лемма3. Если $S \leq Q - S$, то мощность k подмножества X^k из множества X^n с заданной мощностью n определяется следующим образом:

$$k = [S/x^{c}_{arithm}] \vee k = \left[\frac{S}{x^{c}_{arithm}}\right] + 1 \wedge k \leq n - k \qquad (18)$$

Доказательство. Применение теоремы Чолша[5] и формул(16-17) позволяют найти фиксированное значение мощности k подмножества X^k по формулам(18) на основе соотношения (5) и интервалов(13).

Примечание. 1. Совместное использование отображения $\tau = (S - x)x$ из [3] и преобразования(15) позволяет снизить мощность $k(n)$ на единицу. 2. Если $S > Q - S$, тогда в лемме сертификат S заменяется на величину $Q - S$, при этом должны выполняться условия $n - k < k, Q - S - x_i > 0, x_i \in X^n, i \in N$, и тем самым облегчается решение поставленной задачи(1). 3. Теорема Чолша гласит: пусть корни полиномов f_1 и f_2 лежат в кругах K_1 и K_2, радиусы которых равны r_1 и r_2, а центры находятся в точках c_1 и c_2.

Тогда все корни произвольного полинома $f = f_1 \, f_2$ лежат либо в K_1 либо K_2, либо в круге K радиуса $\frac{n_2 r_1 + n_1 r_2}{n_2 + n_1}$ с центром в точке $c = \frac{n_2 c_1 + n_1 c_2}{n_2 + n_1}$, где $n_1 = degf_1, n_2 = degf_2$. Здесь $c_1 = x^{a}_{arithm}, c_2 = x^{b}_{arithm}, \; c = x^{\mathrm{c}}_{arithm}$.

Заключение

Сделаем вывод, хотя вопрос о равенстве классов P и NP до сих пор не решен, многие ученые склонны считать, что они не равны. Это утверждение справедливо для поставленной Куком знаменитой задачи, в которой время работы проверочного алгоритма всегда меньше времени работы решающего алгоритма для задачи о сумме подмножеств.

Предложены леммы о полиномиальной разрешимости задачи о сумме подмножеств. Особо отметим, что предлагаемый метод решения задачи о сумме подмножеств не разделяет на проверочные и решающие алгоритмы, которые имеются в самой постановке задачи Кука. Существование полиномиальных методов и полнота задачи о сумме подмножеств обеспечивает решение других проблем класса NP-complete и справедливость равенства классов P = NP (известная теорема: если некоторая NP-полная задача разрешима за полиномиальное время, то P = NP). И наконец, предлагаемый подход уменьшает время обработки больших данных (Big Data), связанных с набором признаков *VVV* (*volume, velocity, variety*).

СПИСОК ЛИТЕРАТУРЫ

1. Konstantinos Koiliaris, Chao Xu. A Faster pseudopolynomial time algorithm for subset sum. To appear in SODA ’17, 2017. //arXiv:1610.04712v2[cs.Ds] 8 Jan 2017.-18p.

2. Karl Bringmann. A near-linear pseudopolynomial time algorithm for subset sum. To appear in SODA ’17, 2017. //arXiv:1610.04712v2[cs.Ds] 8 Jan 2017.-18p.

3. B. Sinchev, A.B. Sinchev, J. Akzhanova, A.M. Mukhanova. New methods of information search. I. // News of the National Academy of Sciences of Kazakhstan, Series of Geology and Technical Sciences, Volume 3, Number 435 (2019), pp. 240-246

4. B. Sinchev, A.B. Sinchev, J. Akzhanova, Y. Issekeshev, A.M. Mukhanova. Polynomial time algorithms for solving NP-complete problems . // News of the National Academy of Sciences of Kazakhstan, Series of Geology and Technical Sciences, Volume 3, Number 441 (2020), pp.97-101

5. В.В. Прасолов. Многочлены. _М.:МЦНМО, 2001. -336с.

6. E. Horowitz, S. Sanni. Computing Partitions with Application to the Knapsack Problem //Journal of the ACM(JACM), 1974, T21, pp.277-292

7. А. Г. Курош. Курс высшей алгебры-М.: Наука, 1975.-432с.

REFERENCES

1. Konstantinos Koiliaris, Chao Xu. A Faster pseudopolynomial time algorithm for subset sum. To appear in SODA ’17, 2017. //arXiv:1610.04712v2[cs.Ds] 8 Jan 2017.-18p.

2. Karl Bringmann. A near-linear pseudopolynomial time algorithm for subset sum. To appear in SODA ’17, 2017. //arXiv:1610.04712v2[cs.Ds] 8 Jan 2017.-18p.

3. B. Sinchev, A.B. Sinchev, J. Akzhanova, A.M. Mukhanova. New methods of information search. I. // News of the National Academy of Sciences of Kazakhstan, Series of Geology and Technical Sciences, Volume 3, Number 435 (2019), pp. 240-246

4. B. Sinchev, A.B. Sinchev, J. Akzhanova, Y. Issekeshev, A.M. Mukhanova. Polynomial time algorithms for solving NP-complete problems . // News of the National Academy of Sciences of Kazakhstan, Series of Geology and Technical Sciences, Volume 3, Number 441 (2020), pp.97-101

5. V.V. Prasolov. Polynomials. -M.: MTSNMO, 2001.-336p.

6. E. Horowitz, S. Sanni. Computing Partitions with Application to the Knapsack Problem //Journal of the ACM(JACM), 1974, T21, pp.277-292

7. A.G. Kurosh. A course of higher algebra.-M .: Science, 1975.-432p.

Синчев Б.

NP-complete сыныптың полиномиялық шешімі туралы

Аңдатпа. N кардиналының X^n жұп және тақ теріс емес бүтін сандарының жиынтығында элементтерінің қосындысы S сертификатына тең болатын $K=[X^k]$ нықталғандықтың $X^k \subseteq X^n$ ішкі жиынын табыңыз. Қойылған мәселе NP толық сыныпқа жатады. Леммалар $n \geq 3, k(n) \geq 3 \ \vee \ n > 3, \ k(n) \leq n < k^2(n) - k(n)$ шарттарын қанағаттандыратын k координаталық X^k жиынтықтарының қосындысының есебінің полиномдық шешімділігі бойынша дәлелденді және жоғарыдағы аралықтан k қуатының тұрақты мәнін тапты. Есептерді ішкі жиындардың қосындысына шешудің ұсынылған полиномдық әдісі белгілі теорема негізінде P = NP кластарының теңдігін азайту функцияларын және NP толық кластың басқа есептерін шешуді ұсынады: егер кейбір NP болса -толық есеп көпмүшелік уақытта шешіледі, содан кейін P = NP.

Түйінді сөздер: NP класы толық, полиномдық шешімділік, ішкі жиындардың қосындысына есеп

Sinchev B.

On polynomial decision of class NP-complete

Abstract. In the set of even and odd non-negative integers $X^n$of cardinality n, find a subset $X^k \subseteq X^n$ of cardinality $K=[X^k]$, the sum of whose elements is equal to the certificate S. The problem posed belongs to the NP-complete class. Lemmas are proved on the polynomial solvability of the problem of the sum of subsets $X^k$with cardinality k satisfying the conditions $n \geq 3, k(n) \geq 3 \ \vee \ n > 3, \ k(n) \leq n < k^2(n) - k(n)$ and a fixed value of the power k from the above interval is found. The proposed polynomial method for solving the problem on the sum of subsets provides a solution to other problems of the NP-complete class using reducing functions and the equality of the classes P = NP on the basis of the well-known theorem: if some NP-complete problem is solvable in polynomial time, then P = NP.

Keywords: NP-complete class, polynomial solvability, subset sum problem

Автор туралы ақпарат:

Синчев Бахтгерей, Техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университетінің Ақпараттық Жүйелер кафедрасының профессоры.

Сведения об авторах:

Синчев Бахтгерей, доктор технических наук, профессор кафедры информационных систем Международного университета информационных технологий.

About the author:

Sinchev Bakhtgerey, Doctor of Technical Sciences, Professor, Department of Information Systems, International Information Technology University.

INTERNATIONAL JOURNAL OF INFORMATION AND
COMMUNICATION TECHNOLOGIES

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ
КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

Ответственный за выпуск	Есбергенов Досым Бектенович
Редакторы	Медведев Евгений Юрьевич
Компьютерная верстка и дизайн	Жадыранова Гульнур Даутбековна

Редакция журнала не несет ответственности за
недостоверные сведения в статье и
неточную информацию по цитируемой литературе

Подписано в печать 15.12.2021 г.
Тираж 500 экз. Формат 60х84 1/16. Бумага тип.
Уч.-изд.л. 6.5. Заказ №170

Издание Международный университет информационных технологий
Издательский центр КБТУ, Алматы, ул. Толе би, 59