

На правах рукописи

Волков

Волков Мария Сабина Александровна

**ИССЛЕДОВАНИЕ КОМБИНАТОРНЫХ СВОЙСТВ И
ОЦЕНКА ВЫЧИСЛИТЕЛЬНОЙ СЛОЖНОСТИ
ЗАДАЧ РЮКЗАЧНОГО ТИПА**

Специальность: 1.2.3. «Теоретическая информатика, кибернетика»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2026

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (МГТУ им. Н. Э. Баумана).

Научный руководитель: **Гордеев Эдуард Николаевич**
доктор физико-математических наук,
профессор кафедры информационной
безопасности Московского государственного
технического университета им. Н.Э. Баумана

Официальные оппоненты: **Романов Дмитрий Сергеевич**
доктор физико-математических наук, доцент,
профессор кафедры математической
кибернетики факультета вычислительной
математики и кибернетики Московского
государственного университета им. М.В.
Ломоносова

Вялый Михаил Николаевич
кандидат физико-математических наук,
старший научный сотрудник отдела №12
Федерального исследовательского центра
«Информатика и управление» Российской
академии наук

Ведущая организация: АО «НПО «Эшелон»

Защита диссертации состоится «__» _____ 2026 г. в __ ч. __ м.
на заседании диссертационного совета 24.1.224.03 при Федеральном
исследовательском центре «Информатика и управление» Российской
академии наук по адресу: 119333, г. Москва, ул. Вавилова, д.42.

С диссертацией можно ознакомиться в библиотеке Федерального
исследовательского центра «Информатика и управление» Российской
академии наук и на сайте www.frccsc.ru.

Автореферат разослан «__» _____ 2026 г.

Ученый секретарь
диссертационного совета 24.1.224.03,
кандидат технических наук



И. А. Рейер

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

Исследование вычислительной сложности фундаментальных комбинаторных задач занимает центральное место в проблематике теоретической информатики. В данном контексте особый интерес представляет класс NP-полных задач, поскольку его изучение непосредственно связано с проблемой соотношения классов P и NP — одной из ключевых открытых проблем современной математики. Канонические NP-полные задачи, к числу которых относится задача о рюкзаке, служат не только объектом исследования сами по себе, но и инструментом для доказательства NP-полноты других задач посредством сведения. Заложенный в работах С. Кука, Л. Левина и Р. Карпа фундамент теории NP-полноты создал строгий базис для классификации вычислительной сложности, а последующие исследования Д. Кнута, Д. Хопкрофта, М. Гэри, Д. Джонсона, А. Бородина, А. Кобхэма, Д. Эдмондса, М. Рабина, А. Мейера, Л. Стокмейера, Л. Адлемана углубили понимание иерархии классов сложности и структуры вычислимых функций.

Исторически мощным катализатором исследований в этой области выступила криптография. Предложенная Р. Мерклом и М. Хеллманом в 1978 году криптосистема, основанная на задаче о рюкзаке, стала одним из первых практических применений NP-трудности для построения асимметричных шифров. Однако последующая атака А. Шамира, использовавшая методы целочисленного программирования, наглядно продемонстрировала, что сама по себе принадлежность задачи к классу NP-трудных не является гарантией стойкости. Это стимулировало формирование двух взаимосвязанных направлений исследований. С одной стороны, работы Б. Шора, Р. Ривеста, Д. Накааче, Ж. Стерна, Т. Окамото, К. Танаки, С. Учиямы, Я. Мураками, Т. Насако, Б. Ван, В. О. Осипяна и В. В. Подколзина были направлены на конструирование модифицированных и усложненных вариантов рюкзачных криптосистем, устойчивых к известным методам решения. Параллельно, исследования Л. Адельмана, А. М. Одлыжко, Д. Лагариаса, Х. Ленстры, М. Костера, А. Жу, С. Палита, С. Синхи, М. Моллы, А. Ханры и Д. М. Мурина концентрировались на разработке новых аналитических методов, в частности, на анализе таких ключевых параметров, как плотность рюкзачного вектора и структурные свойства ассоциированных целочисленных решеток.

Синтез результатов этих двух направлений исследований способствовал формированию более детализированного представления о вычислительной сложности NP-трудных задач, демонстрируя их зависимость от комбинаторных характеристик конкретных экземпляров. Проведенный анализ показал, что практическая трудоемкость решения определяется не только формальной принадлежностью задачи к классу NP-трудных, но и тонкими структурными особенностями ее конкретных экземпляров. Это обусловило актуальность систематической классификации экземпляров задачи на основе их внутренних свойств, включая структуру множества решений, параметры

полноты покрытия диапазона допустимых значений и особенности размещения допустимых решений в пространстве поиска.

Таким образом, комплексное исследование задачи о рюкзаке, направленное на строгий анализ ее комбинаторных свойств, создание формальных критериев для классификации экземпляров и оценку вычислительной сложности соответствующих алгоритмов, является актуальным направлением в современной теоретической информатике. Полученные результаты вносят вклад в развитие теории сложности вычислений, дискретной оптимизации и построение новых алгоритмических парадигм, являясь также методологической основой для прикладных исследований в смежных областях.

Объектом исследования являются задачи рюкзачного типа как представители класса NP-полных задач комбинаторной оптимизации.

Предметом исследования служат комбинаторные свойства задач рюкзачного типа и их влияние на вычислительную сложность решения, включая механизмы перехода от легко решаемых к трудным для решения классам экземпляров при изменении параметров задачи.

Целью работы является исследование взаимосвязи между параметрами и комбинаторной структурой задач рюкзачного типа и их вычислительной сложностью, а также разработка аналитических и алгоритмических методов построения и классификации экземпляров таких задач с заданными свойствами сложности решения.

Для достижения поставленной цели были решены **следующие задачи**:

1. Провести теоретический анализ комбинаторных свойств множества решений задачи о рюкзаке и получить аналитические выражения, связывающие параметры задачи со структурой множества допустимых решений.
2. Исследовать свойства рюкзачных векторов, выделить классы экземпляров с характерными структурными особенностями и построить зависимость вычислительной сложности решения от параметров задачи.
3. Разработать методы построения экземпляров задач с заранее заданными комбинаторными характеристиками, определяющими трудность их решения.
4. Предложить алгоритмы преобразования экземпляров задач рюкзачного типа, позволяющие изменять структуру множества решений и управлять сложностью их решения.
5. Провести алгоритмическую и программную реализацию разработанных методов, подтвердить их эффективность вычислительными экспериментами и продемонстрировать возможности применения в прикладных задачах.

Методы исследования. В исследовании используются аппарат и методы дискретной оптимизации, математической логики, теории информации,

теории алгоритмов и сложности вычислений, объектно-ориентированного и логического программирования.

Научная новизна работы заключается в комплексном исследовании комбинаторных и алгоритмических свойств задач рюкзачного типа, развитии методов их анализа и построении алгоритмов с обоснованной оценкой вычислительной сложности. В ходе исследования получены следующие новые результаты:

1. Найдены и обоснованы новые комбинаторные формулы для анализа множества решений задачи об ограниченном рюкзаке; выведены выражения для среднего числа допустимых решений задач заданной размерности и среднего значения функционала задачи о 0-1 рюкзаке.
2. Введено и исследовано понятие сюръективных линейных форм как специального класса функциональных конструкций; найден алгоритм вычисления всех допустимых решений задач с такими формами и установлены их характеристики, влияющие на сложность вычислений.
3. Разработаны методы построения линейных форм с контролируемым числом решений, позволяющие управлять комбинаторной структурой задачи и конструировать экземпляры с предсказуемым поведением алгоритмов решения.
4. Исследованы линейные формы с разрывами в области допустимых значений и предложен алгоритм их построения; показано, как наличие разрывов влияет на множество допустимых значений и усложняет задачу поиска решения.
5. Разработан алгоритмический метод преобразования легкорешаемых экземпляров задачи о рюкзаке в трудные для решения при сохранении конфигурации множества решений, что открывает возможности применения таких преобразований в задачах защиты информации и анализа сложности дискретных структур.

Обоснованность и достоверность полученных результатов обеспечивается использованием строгого математического аппарата и опорой на фундаментальные исследования в области теории сложности и комбинаторного анализа, представленные в работах отечественных и зарубежных авторов. Предложенные в диссертации теоретические положения подтверждаются строгими доказательствами и согласуются с известными результатами в смежных областях. Практические алгоритмы и аналитические зависимости проверены вычислительными экспериментами, а основные выводы апробированы на профильных научных конференциях и опубликованы в рецензируемых изданиях.

Теоретическая значимость. В диссертации предложен комплекс научных результатов, формирующих основу для анализа комбинаторных и алгоритмических свойств задач рюкзачного типа. Разработанные методы позволяют:

- проводить формальное описание и классификацию рюкзачных задач по их внутренним параметрам;
- оценивать сложность решения экземпляров задачи и их устойчивость к известным методам решения;
- осуществлять выбор параметров, обеспечивающих заданную сложность решения для конкретных алгоритмов.

Полученные результаты вносят вклад в развитие теории сложности и комбинаторного анализа NP-полных задач.

Практическая значимость работы заключается в разработке алгоритмических средств, позволяющих управлять сложностью экземпляров задач рюкзачного типа и исследовать границы их вычислимости.

В частности:

- предложены методы генерации экземпляров задач с заданными комбинаторными характеристиками, что обеспечивает возможность контроля над числом решений и степенью сложности их нахождения;
- разработан алгоритм поиска всех решений для специального класса рюкзачных задач, обладающий линейной сложностью по числу переменных и количеству решений, что существенно повышает эффективность анализа их структуры;
- предложены преобразования, позволяющие сохранять конфигурацию множества решений при переходе от легкорешаемых к труднорешаемым экземплярам, что открывает возможности их использования в задачах синтеза и моделирования вычислительно сложных структур.

Полученные результаты могут быть использованы в задачах дискретной оптимизации, анализа сложности алгоритмов, а также при построении схем для систем кодирования и защиты информации.

Положения, выносимые на защиту:

1. Найдены формулы для анализа множества решений задачи об ограниченном рюкзаке, позволяющие вычислять среднее число допустимых решений по всем экземплярам заданной размерности, и среднее значение целевой функции через характеристики подзадач меньшей размерности.
2. Выделен и исследован класс сюръективных линейных форм, для которых множество достижимых значений представляет собой непрерывный диапазон целых чисел; установлены необходимые и достаточные условия сюръективности, а также параметры, определяющие сложность решения определенными алгоритмами. задач с такими формами в качестве левой части.
3. Разработан алгоритм вычисления всех допустимых решений сюръективных линейных форм, обладающий линейной сложностью по числу переменных и количеству решений.

4. Для линейных форм с разрывами в области значений установлены зависимости числа и расположения разрывов от коэффициентов формы, построен алгоритм формирования линейной формы с заданным числом разрывов.
5. Установлены зависимости между комбинаторными параметрами экземпляров задачи о рюкзаке (плотностью, структурой множества решений, диапазоном коэффициентов) и сложностью их решения определенными алгоритмами.
6. Разработаны методы параметрических преобразований экземпляров задачи о рюкзаке, обеспечивающие переход от легкорешаемых к труднорешаемым случаям при сохранении конфигурации множества решений и показана их применимость в задачах защиты информации.

Апробация результатов. Результаты диссертации докладывались на научных конференциях:

1. XII Международная научно-техническая конференция Безопасные информационные технологии, Москва, 02 ноября 2023 года
2. XIII Международная научно-техническая конференция Безопасные информационные технологии, Москва, 01 ноября 2024 года
3. V Межвузовская конференция аспирантов, соискателей и молодых ученых «Наука, технологии и бизнес», Москва, 18–19 апреля 2023 года

Публикации. Результаты диссертации опубликованы в 9 работах [1-9], из них 5 статей опубликованы в научных журналах, которые включены в перечень рекомендованных ВАК РФ для публикации основных научных результатов диссертаций [1-5], и 3 работы в трудах международных конференций [6-9].

Структура диссертации. Диссертация состоит из введения, трех глав, заключения, списка использованной литературы. Общий объем диссертации составляет 106 страниц, включая 3 таблицы. Список литературы состоит из 77 источников.

Соответствие паспорту специальности. Проведенное исследование соответствует направлениям паспорта специальности 1.2.3 «Теоретическая информатика, кибернетика», в частности: п. 3 «Теория сложности алгоритмов и вычислений» — выполнен систематический анализ вычислительной сложности некоторых классов экземпляров NP-полной задачи о рюкзаке; установлены зависимости сложности решения от комбинаторных параметров задачи и разработаны алгоритмы решения с обоснованными оценками вычислительной сложности; п. 8 «Математическое программирование» — получены аналитические формулы для среднего числа допустимых решений по всем задачам заданной размерности при ограниченных значениях коэффициентов весов и среднего значения функционала задачи о рюкзаке, что позволяет проводить анализ комбинаторных свойств и структуры множества решений дискретных экстремальных задач; п. 25 «Методы высоконадежной обработки информации и обеспечения помехоустойчивости информационных коммуникаций для целей передачи, хранения и защиты информации» —

показано, что предложенные методы позволяют создавать структуры и алгоритмы, основанные на труднорешаемых экземплярах задач о рюкзаке, которые могут использоваться для обеспечения защиты и устойчивости данных при передаче и обработке.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснованы важность и актуальность темы диссертации, сформулированы цели исследования и решаемые задачи, определены научная новизна и теоретическая и практическая значимость работы, раскрыты основные положения, выносимые на защиту.

Первая глава посвящена формированию теоретических предпосылок и постановке задачи исследования, направленного на анализ комбинаторных свойств и вычислительной сложности классов экземпляров задач рюкзачного типа.

В общем виде задача о 0–1 рюкзаке формулируется следующим образом. Пусть задано множество предметов $I = \{1, 2, \dots, n\}$, каждому предмету i соответствует вес $a_i > 0$ и стоимость $c_i > 0$. Требуется найти $\max_{x_i \in \{0,1\}} \sum_{i=1}^n c_i x_i$, при условии $\sum_{i=1}^n a_i x_i \leq b$.

Упорядоченный по возрастанию набор весов $A = (a_1, \dots, a_n)$ называется рюкзачным вектором, а n — его размерностью. В теории сложности вычислений чаще рассматривается форма распознавания существует ли $x_i \in \{0,1\}$, такое, что $\sum_{i=1}^n a_i x_i = b$?

Задача о рюкзаке рассматривается как одна из фундаментальных NP-полных задач, используемая как в теоретических исследованиях сложности вычислений, так и в прикладных задачах обработки информации. В главе акцент сделан на внутренних характеристиках задачи — структуре множества допустимых решений и их комбинаторных характеристиках.

Анализ известных алгоритмов показал, что вычислительная сложность задачи определяется не только ее размерностью n , но и характеристиками конкретного экземпляра. Экземпляры со специальной структурой коэффициентов, такие как сверхрастущие последовательности, относятся к классу легко решаемых. В противоположность этому, равномерно распределенные значения коэффициентов a_i и целевого значения b порождают экземпляры, близкие к худшему случаю.

Определение 1.6. Вектор $A = (a_1, \dots, a_n)$ называется сверхрастущим, если выполняется условие $\forall k = 2, \dots, n \ a_k > \sum_{i=1}^{k-1} a_i$.

Ключевым параметром, характеризующим сложность экземпляра, является плотность рюкзачного вектора.

Определение 1.7. Плотностью вектора $A = (a_1, \dots, a_n)$ называется величина $d(A) = \frac{n}{\log_2(\max_i a_i)}$, характеризующая соотношение между размерностью задачи и битовой длиной ее максимального коэффициента.

В главе рассмотрены классические результаты о применимости методов редукции решеток к задачам рюкзачного типа. Теорема Лагариаса и Одлыжко устанавливает, что при наличии решения и плотности $d(A) < 0,64$ алгоритм LLL с высокой вероятностью восстанавливает решение. Последующие работы Костера и соавторов расширили эту границу до $d(A) < 0,94$. При этом анализ теоретических и экспериментальных результатов свидетельствует о росте вычислительной сложности при приближении плотности к единице и ее превышении.

Отмечено, что вычислительная сложность экземпляров задачи о рюкзаке определяется не только их NP-трудностью, но и величинами коэффициентов, плотностью и структурой множества решений. В связи с этим формулируется основная научная задача диссертационного исследования, заключающаяся в разработке аналитического аппарата для описания комбинаторных свойств задачи, включая число и структуру допустимых решений и параметры, определяющие переход от легко разрешимых к трудным экземплярам.

В качестве основного инструмента выбран аппарат производящих функций и метод коэффициентов, позволяющий получить аналитическое представление множества решений и проводить асимптотический анализ его усредненных характеристик для широкого класса задач рюкзачного типа.

Первый раздел второй главы посвящен исследованию комбинаторных характеристик множества допустимых решений задачи об ограниченном рюкзаке и выведению аналитических формул, связывающих структуру этого множества с параметрами задачи.

Для получения основных результатов были выведены выражения для исходных условий задачи в виде формальных степенных рядов и получены выражения их производящих функций.

Для обеспечения большей общности исследуется задача об ограниченном рюкзаке, $\sum_{j=1}^n c_j x_j \rightarrow \max, \sum_{i=1}^n a_i x_i \leq b$, где $x = (x_1, \dots, x_n)$ – n -мерный вектор с целочисленными компонентами $x_i \in \{0, 1, \dots, m\}$, $c_1, \dots, c_n; a_1, \dots, a_n; b$ – неотрицательные целые числа.

Множество допустимых решений V_b представляет собой множество n -мерных векторов x с компонентами $x_i \in \{0, 1, \dots, m\}$, удовлетворяющих неравенству $\sum_{i=1}^n a_i x_i \leq b$. Его мощность $|V_b|$ — это количество таких решений. Для каждой переменной x_k ($1 \leq k \leq n$) введены $m+1$ сечений, где сечение с номером d ($0 \leq d \leq m$) включает решения, удовлетворяющие $\sum_{i=1, i \neq k}^n a_i x_i \leq b - da_k, x_i \in \{0, 1, \dots, m\}$. Эти множества обозначены как V_b^{dk} .

При помощи метода коэффициентов, получено соотношение, выражающее среднее значение функционала задачи через количество решений подзадач меньшей размерности. Будем считать, что все точки множества V_b равновероятны. Тогда значения $f(x_1, \dots, x_n) = \sum_{j=1}^n c_j x_j$ — это случайная величина $\xi = \xi(a_1, \dots, a_n, c_1, \dots, c_n, b)$, обозначим $\mathbb{E}\xi = \frac{1}{|V_b|} \sum_{x \in V_b} f(x)$.

Теорема 2.1. Справедливо соотношение

$$\mathbb{E}\xi = \frac{1}{|V_b|} \sum_{k=1}^n c_k (|V_b^{1k}| + 2|V_b^{2k}| + \dots + m|V_b^{mk}|) \quad (1)$$

Полученная формула позволяет оценивать эффективность алгоритмов решения задачи о рюкзаке через сравнение со средним значением целевой функции. Также формула может быть использована для построения нижних оценок в декомпозиционных методах, таких как метод ветвей и границ.

Выражение $|V_b|$ также можно представить через сумму $|V_b^{dk}|$:

$$|V_b| = |V_b^{0k}| + |V_b^{1k}| + \dots + |V_b^{mk}|$$

Эта формула позволяет сократить количество рассчитываемых значений в формуле (1). Применение формулы позволяет проводить оценки мощности множества допустимых решений задачи о рюкзаке, что имеет фундаментальное значение для анализа поведения алгоритмов и исследования структурных свойств пространства решений задачи.

Для анализа вычислительной сложности задачи о рюкзаке существенное значение имеют оценки среднего числа допустимых решений для всех экземпляров заданной размерности. Такие оценки устанавливают связь между параметрами входных данных и структурой пространства решений, позволяя прогнозировать поведение алгоритмов и классифицировать экземпляры по степени сложности.

Обозначим $|\bar{V}_p|$ — среднее число решений набора задач об ограниченном рюкзаке при фиксированном b с коэффициентами весов $a_i \leq p, i = 1, \dots, n$, где p — заранее заданное значение. Значение этой величины выражается по формуле

$$|\bar{V}_p| = \frac{1}{(p+1)^n} \sum_{0 \leq a_i \leq p, i=1, \dots, n} |V_b(a_1, \dots, a_n)|$$

Пусть число b и размерность задачи n фиксированы, при этом $p = b$, т.е. компоненты вектора весов (a_1, \dots, a_n) принимают значения от 0 до b .

Теорема 2.2. При $x \in \{0,1\}^n$ справедлива формула:

$$|\bar{V}_b| = \frac{1}{(b+1)^n} \sum_{k=0}^n C_n^k C_b^{n-k} (b+2)^k \quad (2)$$

Здесь и далее биномиальный коэффициент C_m^r считается равным нулю при $r > m$ или $r \notin \mathbb{Z}$.

Данная формула может быть полезна для выбора оптимального подхода к решению задач, определения вероятности их успешного решения, а также при оценке сложности решения задач методами перебора. Эта формула использована в главе 3 для сравнительного анализа сложности решения различных классов экземпляров задачи о рюкзаке.

Также был рассмотрен вопрос о среднем значении мощности множества допустимых решений в более общем случае. Была найдена производящая функция, которая выражает число решений каждой задачи размерности n с компонентами вектора весов (a_1, \dots, a_n) , принимающими значения в диапазоне от 0 до p .

Для случая $x \in \{0,1,2\}^n$ получена аналитическая, аналогичная случаю $x \in \{0,1\}^n$ в теореме 2.2.

Теорема 2.4. При $x \in \{0,1,2\}^n$ справедлива формула:

$$|\bar{V}_b| = \sum_{k=0}^n C_n^k (b+1)^{k-n} \sum_{t=0}^{n-k} C_{n-k}^t 2^t \left(\frac{C_{n-k+t+b-1}^{n-k} [n-k+t+b \pmod{2} \equiv \equiv 1]}{2} + \frac{C_{n-k+t+b}^{n-k} [n-k+t+b \pmod{2} \equiv 0]}{2} \right),$$

где $[P]$ – скобка Айверсона, равная 1, если условие P выполняется, и 0 в противном случае.

Полученная формула может быть полезна для выбора оптимального подхода к решению для многозначных постановок задачи о рюкзаке и определения вероятности их успешного решения.

Во втором разделе второй главы рассматриваются линейные формы специального вида, названные сюръективными. Под линейной формой понимается отображение $L: \mathbb{Z}^n \rightarrow \mathbb{Z}$, $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$, заданное над кольцом целых чисел. Рассматривается частный случай задачи $\sum_{i=1}^n a_i x_i = b$, при котором частичные суммы компонентов рюкзачного вектора принимают все целочисленные значения от 0 до $\sum_{i=1}^n a_i$. Без ограничения общности далее предполагается, что $a_1 \leq a_2 \leq \dots \leq a_n$.

Везде далее будем рассматривать значения линейной формы $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ на множестве $\{0,1\}^n$. При отсутствии неопределенности будем также использовать обозначение $L(x)$. Множество значений формы на данном множестве обозначим через $L^*(x) = \{L(x) \mid x \in \{0,1\}^n\}$. Уравнение $L(x) = b$ при $x \in \{0,1\}^n$ разрешимо тогда и только тогда, когда $b \in L^*(x)$.

Определение 2.1. Линейная форма $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ называется сюръективной на множестве $\{0,1\}^n$, если $\forall b \in \{0, 1, 2, \dots, \sum_{i=1}^n a_i\} \ b \in L^*(x)$. Далее будем обозначать $L^*(x) = [0, \sum_{i=1}^n a_i]$, понимая под этим совпадение множества значений формы с отрезком целых чисел.

Важность сюръективных линейных форм состоит в том, что проверка разрешимости системы уравнений вида $L(x) = b$, рассматриваемых на множестве $x \in \{0,1\}^n$, является тривиальной: достаточно для каждого уравнения проверить выполнение условия $0 \leq b \leq \sum_{i=1}^n a_i$.

Теорема 2.5. Форма $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ является сюръективной тогда и только тогда, когда

$$a_1 = 1, \text{ и } a_k \leq \sum_{i=1}^{k-1} a_i + 1 \ \forall k = 2, \dots, n \quad (3)$$

Аналогичные условия можно получить и для общего случая.

Следствие 2.2. Пусть $a_1 \leq a_2 \leq \dots \leq a_n$ и $x \in \{0, 1, \dots, p\}^n$. Форма $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ сюръективна на $\{0, 1, \dots, p\}^n$ тогда и только тогда, когда $a_1 = 1, a_k \leq p \sum_{i=1}^{k-1} a_i + 1 \forall k = 2, \dots, n$

Данная теорема и ее следствие демонстрируют конструктивный характер класса сюръективных форм. Процедура их построения обладает сравнительной простотой, что открывает возможности для систематического конструирования задач о рюкзаке с заданными вычислительными характеристиками.

Условия (3) будем также называть условиями сюръективности линейной формы. Из них получены важные свойства сюръективных форм.

Утверждение 2.2. Форма $L(x_1, \dots, x_n) = \sum_{i=1}^n 2^{i-1} x_i$ является сюръективной, причем ее коэффициенты достигают верхней границы в (3).

Соответствующий форме $L(x_1, \dots, x_n)$ рюкзачный вектор $A = (a_1, \dots, a_n)$ также назовем сюръективным.

Следствие 2.3. Плотность сюръективного рюкзачного вектора составляет $d(A) \geq \frac{n}{n-1} > 1$.

Важным структурным свойством рюкзачных векторов является инъективность, которая гарантирует, что каждое достижимое значение суммы соответствует единственному решению, что существенно влияет на сложность алгоритмов решения и проверки решений.

Определение 2.2. Рюкзачный вектор $A = (a_1, \dots, a_n)$ называется инъективным, если $\forall A', A'' \subset A, A' \neq A''$ выполняется $\sum_{a' \in A'} a' \neq \sum_{a'' \in A''} a''$.

Соответствующую линейную форму на рассматриваемом подмножестве по аналогии назовем инъективной.

Утверждение 2.3. За исключением случая $L(x_1, \dots, x_n) = \sum_{i=1}^n 2^{i-1} x_i$, ни одна сюръективная линейная форма не является инъективной.

Таким образом, сюръективные рюкзачные векторы являются плотными ($d(A) > 1$), что делает их устойчивыми к алгоритмам редукции базиса решеток, эффективным лишь при малой плотности. Однако для задач с такими формами, как правило, существует несколько допустимых решений.

Утверждение 2.4. Обозначим общее число сюръективных линейных форм от n переменных через $L_c(n)$. Оно может быть посчитано по формуле: $L_c(n) = \sum_{a_2=1}^2 \sum_{a_3=a_2}^{a_2+2} \dots \sum_{a_n=a_{n-1}}^{a_2+\dots+a_{n-1}+2} 1$

Поскольку данное выражение определяет вложенное суммирование, в котором пределы каждого суммирования зависят от предыдущих индексов, оно не может быть преобразовано в выражение замкнутой формы. Тем не менее, для него были найдены верхняя и нижняя оценки.

Следствие 2.4. $n! < L_c(n) < 2^{n \cdot \frac{n-1}{2}}$

Нижняя оценка показывает, что сюръективные линейные формы образуют обширный класс комбинаторных объектов, мощность которого

растет экспоненциально с увеличением размерности задачи. Это позволяет рассматривать их как самостоятельное значимое множество в пространстве всех экземпляров задачи о рюкзаке.

Наиболее интересным свойством сюръективных форм является то, что каждое решение для задачи, содержащей такую форму, может быть найдено за линейное от количества переменных и числа решений время.

Теорема 2.6. Если форма $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ сюръективна, $b \in L^*(x_1, \dots, x_n)$, то все решения задачи $\sum_{i=1}^n a_i x_i = b$ можно найти за время $O(ns)$, где s — число этих решений.

В доказательстве теоремы приведена рекурсивная процедура нахождения всех допустимых решений сюръективной формы.

Для анализа устойчивости структурных свойств сюръективных форм к алгебраическим преобразованиям представляет интерес исследование их поведения при модульных преобразованиях.

Определение 2.3. Пусть дан вектор целых чисел $A = (a_1, a_2, \dots, a_n)$, целое число $M > \sum_{i=1}^n a_i$, и натуральное $t < M$, такое что $\gcd(t, M) = 1$. Тогда говорят, что вектор $B = (b_1, b_2, \dots, b_n)$, где $b_i \equiv (ta_i \bmod M)$, $i = 1, \dots, n$ получен из A сильным модульным умножением относительно пары (M, t) .

Показано, что сильное модульное умножение компонентов сюръективного вектора нарушает упорядоченность и сюръективность соответствующей ему формы, сохраняя при этом комбинаторную структуру множества решений. Установлено, что при усреднении по модульному множителю коэффициенты формы обладают однородными статистическими характеристиками и по главному члену ведут себя как случайная подвыборка из множества $\{1, \dots, M - 1\}$.

Утверждение 2.5. Пусть $A(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ — сюръективна, $a_1 < a_2 < \dots < a_n$ и коэффициенты формы $B(x_1, \dots, x_n)$, полученной из формы $A(x_1, \dots, x_n)$ посредством сильного модульного умножения $b_i = ua_i \pmod{M}$, $(u, M) = 1$, рассматриваются как случайная выборка n элементов из множества $\{1, 2, \dots, M - 1\}$.

Обозначим упорядоченные компоненты формы $B(x_1, \dots, x_n)$ через $b^{(1)} < b^{(2)} < \dots < b^{(n)}$. Тогда математическое ожидание компоненты $b^{(k)}$ равно:

$$\mathbb{E}[b^{(k)}] = \frac{Mk}{n+1}, \quad k = 1, \dots, n.$$

Из данного утверждения следует, что модульное умножение приводит к значительному увеличению наименьших компонентов формы, что делает невозможным покрытие всего диапазона от 1 до $\sum_{i=1}^n b_i$ без пропусков. В общем случае поиск решения для такой формы становится экспоненциально сложной задачей. Из формулы (2) можно получить среднюю оценку числа решений для таких форм, которая для больших M оказывается существенно меньше 1, что подчеркивает сложность решения этой задачи.

В третьем разделе второй главы работы рассматриваются формы с разрывами в области значений, что связано с неоднородностью распределения

решений и влияет на сложность нахождения решений. Построение таких форм открывает возможности для формирования классов задач с контролируемой вычислительной сложностью и анализа границ между легко- и труднорешаемыми случаями.

Определение 2.4. Разрывом области значений линейной формы $L(x)$ назовем интервал $(d, e) = \{z \in \mathbb{Z}, d < z < e\}$, такой, что $d, e \in L^*(x)$, $\forall z \in (d, e) z \notin L^*(x)$.

Определение 2.5. Обозначим через $\nu(L(x_1, \dots, x_n))$ количество разрывов формы $L(x_1, \dots, x_n)$, а через $\mu(L(x_1, \dots, x_n)) = \nu(L(x_1, \dots, x_n)) + 1$ количество ее отрезков сюръективности.

В данной главе показано, что добавление новой компоненты к линейной форме сохраняет сюръективность на краевых участках области значений, тогда как в центральной части структура определяется объединением областей значений исходной и дополненной форм. В леммах доказано, что количество отрезков сюръективности при этом возрастает не более чем вдвое. Показано, что чем раньше в последовательности коэффициентов нарушается условие сюръективности, тем больше разрывов образуется в области значений, что существенно увеличивает сложность решения задачи.

Также была сформулирована и доказана

Теорема 2.7. Для любого $h \in [0, 2^n - 1]$ можно построить линейную форму $L(x_1, \dots, x_n)$ с h разрывами.

Предложенный в доказательстве теоремы алгоритм построения линейной формы с заданным числом разрывов обладает полиномиальной сложностью и может быть использован для генерации экземпляров с контролируемыми характеристиками.

Для устранения разрывов и восстановления сюръективности предложен метод дополнения, основанный на итеративном добавлении компонент вида $a'_i = \sum_{i=1}^{t-1} a_i + 1$ в позиции первого нарушения условия сюръективности. Полученная форма допускает применение эффективных алгоритмов решения с фильтрацией решений, содержащих добавленные компоненты. Таким образом, формы с разрывами могут быть сведены к сюръективным, что расширяет класс практически решаемых задач и углубляет теоретический анализ сложности задачи о рюкзаке.

Первый раздел третьей главы посвящен исследованию влияния размерности, плотности и структуры коэффициентов на вычислительную сложность задачи о рюкзаке. Рассматриваются экземпляры, которые, с одной стороны, устойчивы к алгоритмам редукции базиса решеток, а с другой — сохраняют полиномиальную разрешимость при использовании методов, разработанных в главе 2.

В данной главе ключевым инструментом анализа выступают модульные преобразования. Как показано ранее, задачи о рюкзаке с сюръективными векторами допускают эффективное решение, однако при сильном модульном умножении сюръективность нарушается, но множество допустимых решений

сохраняется. В результате экземпляры, исходно принадлежащий к легкорешаемому классу, становится труднорешаемым для стандартных алгоритмов, включая методы редукции решеток. Это позволяет использовать модульные преобразования для контролируемого изменения сложности и исследования границы между полиномиально и экспоненциально разрешимыми случаями.

Для реализации такого подхода в разделе обосновывается выбор размерности n , плотности d , модуля M и среднего числа решений \tilde{N} , при которых преобразованный экземпляр обладает требуемыми характеристиками.

Выбор длины вектора n . Изменение длины вектора коэффициентов n влияет на размерность пространства решений, структуру экземпляров и эффективность используемых алгоритмов.

В качестве ориентирующего примера рассматриваются сверхрастущие рюкзачные векторы, используемые в классической схеме Меркла–Хеллмана. В данной схеме закрытый ключ задается сверхрастущим рюкзачным вектором, для которого задача о рюкзаке допускает полиномиальное решение, а открытый ключ формируется посредством сильного модульного умножения коэффициентов по фиксированному модулю. В схеме Меркла–Хеллмана использовались векторы длиной $n = 100$, однако современные исследования¹ показывают, что увеличение размерности до $n = 150–200$ существенно снижает эффективность решеточных методов.

В данной работе для анализа экземпляров с высокой плотностью выбрана размерность $n = 200$. Элементы вектора после модульного преобразования лежат в диапазоне от 2^{170} до 2^{200} , что обеспечивает плотность, близкую к $d \approx 1$ — области теоретически наибольшей сложности². Рост n увеличивает диапазон коэффициентов, снижая плотность при фиксированном числе решений и напрямую влияя на вычислительную сложность.

Установлено, что:

- При $d < 0,94$ задача эффективно решается алгоритмом LLL;
- При $d > 1$ нарушается однозначность отображения решений;
- Наибольшая сложность достигается при $d \approx 1 + \frac{\log(n/2)}{n}$.

Для $n = 200$ это соответствует $d \approx 1,033$, что принято для дальнейших расчетов

Трудоемкость базовой реализации LLL-алгоритма при $d \approx 1$ оценивается как: $T(n) = O(n^6 \log^3 a_n) \approx O(n^9)$

Для $n \geq 200$ вычислительные затраты становятся непрактичными, что повышает устойчивость задачи к редукции базиса.

¹ Liu J., Bi J. Xu S. An Improved Attack on the Basic Merkle–Hellman Knapsack Cryptosystem.

² Koskinen, A. (2003). Non-Injective Knapsack Public-Key Cryptosystems. Proceedings of the 3rd Central European Conference on Cryptography (CECC).

Утверждение 3.1. Сложность решения задачи с сюръективной формой при плотности d в среднем случае ограничена снизу: $\Omega(n \cdot 2^{(n(1-1/d)-1)})$

Для сохранения плотности $d \approx 1,033$ при $n > 200$ требуется значительное увеличение среднего числа решений, что резко повышает вычислительные затраты. Таким образом, выбор $n = 200$ представляет собой компромисс между устойчивостью к редукции решеток и практической разрешимостью.

Определение плотности рюкзачного вектора. Согласно исследованиям³, наибольшая сложность задач о рюкзаке достигается при плотности вектора около $1 + \frac{\log_2(n/2)}{n}$. Также подтверждается, что для экземпляров с $n \approx 100$ и $d \approx 1$ эффективные алгоритмы решения неизвестны, особенно при весе решения, близком к $n/2$. Это свидетельствует о повышенной сложности решения экземпляров с плотностью в интервале от 1 до указанного значения.

Для выбранной размерности $n = 200$ расчет по формуле дает: $d \approx 1 + \frac{\log_2(200/2)}{200} \approx 1,033$. Такое значение плотности обеспечивает существенное усложнение решения задачи и повышает устойчивость к алгоритмам редукции базиса решеток.

В приведенной схеме генерации сюръективной последовательности начальные элементы задаются из малого диапазона $a_1 = 1, a_2 \in \{1,2\}$, с рекуррентной границей $a_{k-1} + 1 \leq a_k \leq \sum_{i=1}^{k-1} a_i + 1$. Однако такие предсказуемые начальные фрагменты обладают двумя недостатками:

1. Ограниченное разнообразие комбинаций, что снижает ценность для анализа сложных случаев;
2. Облегчение криптоанализа в прикладных схемах за счет предоставления атакующему опорной информации;

Для устранения этого эффекта в процедуре генерации используется короткий детерминированный начальный фрагмент длины $p = 5$, после чего линейная форма строится по заданному алгоритму, начиная с p -го элемента.

Плотность при этом оценивается по эффективной длине $n' = n - p$: $d(A) \approx \frac{n-p}{\log_2(\max A)}$, что корректно отражает вклад именно случайной части последовательности в формирование вычислительной сложности.

Подбор значения модуля M . Сильное модульное умножение нарушает сюръективность исходной формы, делая часть значений правой части недостижимой, но сохраняет общий характер распределения решений. Это делает его ценным инструментом для анализа перехода от простых к трудным для решения экземплярам.

³ Schnorr, C. P., & Euchner, M. (1994). Lattice Basis Reduction: Improved Practical Algorithms and Solving Subset Sum Problems. *Mathematical Programming*, 66, 181–199.

Плотность преобразованного вектора оценивается как: $d(A) \approx \frac{n'}{\log_2 M}$, где $n' = n - p$. Значение M определяет баланс между плотностью и сложностью решения:

- Слишком большое значение M снижает плотность, упрощая задачу для методов редукции решеток;
- Слишком малое значение M нарушает распределение решений, создавая вырожденные экземпляры;

Для достижения целевой плотности $d^* \geq 1,033$ при $n = 200$ и $p = 5$ модуль должен удовлетворять: $M \leq 2^{n'/d} = 2^{195/1,033} \approx 2^{189}$. При этом для сохранения распределения решений необходимо: $M \geq \sum_{i=1}^n a_i$

В текущей реализации было принято решение выбирать $M \approx \sum_{i=1}^n a_i + k$, где $k \in [0; a_{n-1}]$, поскольку для больших n такая надбавка на порядок меньше значения $\sum_{i=1}^n a_i$ и не оказывает существенного влияния на значение плотности. Данный подход обеспечивает сохранение распределения решений при высокой плотности преобразованного экземпляра.

Выбор среднего числа решений. Основной недостаток алгоритма решения, приведенного в главе 2, заключается в том, что число допустимых решений уравнения $\sum_{i=1}^n a_i x_i = b$ с сюръективной формой может значительно превышать величину n . Например, для формы $L(x_1, \dots, x_{100}) = \sum_{i=1}^{100} i x_i$ число решений задачи $L(x_1, \dots, x_{100}) = 2525$ приблизительно равно $1,73e+27$, что во много раз превышает значение n .

Однако, при построении сюръективной формы, можно ограничить среднее число ее решений на рассматриваемом множестве. Поскольку форма $L(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i$ сюръективна, ее среднее число решений по всем b в интервале $[0, \sum_{i=1}^n a_i]$ выражается как $N(L) = \frac{2^n}{\sum_{i=1}^n a_i + 1}$. Тогда, чтобы среднее число решений не превышало некоторого \tilde{N} и распределение числа решений по интервалу $[0; \sum_{i=1}^n a_i]$ было близко к равномерному, достаточно при ее построении выбирать каждый следующий a_k так, чтобы выполнялось $a_k \geq \frac{2^{k-1}}{\tilde{N}}$. Таким образом, можно строить сюръективные рюкзачные векторы, для которых сложность алгоритма решения не превышает выбранного значения $O(\tilde{N}n)$. Плотность рюкзака, построенного по описанной процедуре, будет $d(A) \leq \frac{n}{(n-1) - \log_2(\tilde{N})}$. Это позволяет управлять балансом между плотностью рюкзака и числом решений, создавая экземпляры задачи с заданной сложностью.

Для обеспечения желаемой плотности $d^* \approx 1,033$ необходимо учитывать следующие соотношения: $2^{189} \geq M \geq \sum_{i=1}^n a_i \geq 2a_n \geq \frac{2^n}{\tilde{N}}$. Из этого выражения можно вывести, что $\tilde{N} \geq \frac{2^{200}}{2^{189}} = 2048$.

Контрольные суммы. Множественность решений в задачах о рюкзаке требует введения избыточности для однозначной идентификации корректного

решения в вычислительных экспериментах. Поскольку совпадения решений чаще происходят по старшим коэффициентам, контрольная последовательность размещается в начале решения, обеспечивая чувствительность к отклонениям в младших разрядах.

Вероятность случайного совпадения контрольных последовательностей оценивается как 2^{-l} , где l — длина последовательности в битах. В реализации используется усеченное хэш-преобразование SHA-256 с $l = 16$. С учетом отброшенных ранее элементов ($p = 5$) общая вероятность коллизии составляет порядка $2^{-21} = 1/2097152$.

При параметрах $n = 200$, $l = 16$, $p = 5$ избыточная часть данных составляет $\sim 10,5\%$ от общего объема. Это обеспечивает баланс между компактностью представления и надежностью идентификации при множественности допустимых решений.

Во втором разделе третьей главы продемонстрировано практическое применение представленных ранее теоретических положений в модифицированной криптосистеме на основе рюкзачных структур с использованием сюръективных форм.

Анализ классической схемы Меркла-Хеллмана показал ее уязвимость к криптоаналитическим атакам, обусловленную структурной предсказуемостью сверхрастущих последовательностей. В качестве альтернативы предложено использовать сюръективные линейные формы, которые сохраняют эффективность обратного преобразования при обеспечении более высокой криптостойкости.

Разработаны модифицированные алгоритмы генерации ключей, шифрования и расшифрования. Основные особенности предложенного подхода:

1. Замена сверхрастущих последовательностей на сюръективные формы в качестве закрытого ключа;
2. Сохранение механизма сильного модульного умножения для преобразования в открытый ключ;
3. Применение рассчитанных ранее параметров ($n = 200$, $p = 5$, $d \approx 1,033$, $\tilde{N} = 2048$) для обеспечения баланса между стойкостью и эффективностью;
4. Использование предложенного в главе 2 алгоритма для поиска решений и верификация по контрольной сумме.

Представлены формальные описания алгоритмов:

- Генерации ключей с построением сюръективной последовательности и модульным преобразованием;
- Шифрования с вычислением взвешенной суммы и добавлением контрольной информации;
- Расшифрования с обратным преобразованием и поиском решений с верификацией по контрольной сумме.

В третьем разделе третьей главы представлены результаты экспериментальной проверки предложенных конструкций.

1. Проверка корректности работы алгоритмов

При параметрах $n = 200$, $l = 16$, $p = 5$, $\tilde{N} = 2000$, $d \approx 1,031$, $\tilde{N} = 2048$ выполнено 100 000 тестовых циклов прямого и обратного модульного преобразования решений. Во всех случаях достигнуто полное соответствие векторов решений, что подтверждает корректность реализации алгоритмов.

2. Сравнительный анализ производительности

Проведено сравнение разработанных алгоритмов расшифрования с эталонной реализацией RSA (2048 бит):

- Генерация ключей: 0,05 мс (против 340 мс у RSA)
- Шифрование: 0,6 мс (сопоставимо с RSA)
- Расшифрование: 26 мс (превышает показатель RSA в 15 раз)

Отмечается, что текущая реализация на Python не оптимизирована, что оставляет потенциал для улучшения производительности.

3. Оценка стойкости к LLL-атакам

Экспериментально исследована устойчивость к решениям на основе редукции базиса решеток. Результаты демонстрируют экспоненциальное снижение эффективности атак с ростом длины блока:

Таблица 1. Зависимость доли успешных решений от длины блока

Длина блока (бит)	Доля успешных атак (%)
8	0,47
12	0,33
16	0,18
32	0,05
64	0,008
128	0,0003

Под длиной блока понимается количество последовательных значений решения, корректно восстановленных после применения LLL-алгоритма. Для плотностей $d > 1$ алгоритм LLL часто находит допустимые, но неполные решения, воспроизводящие лишь часть исходной последовательности. Введенная метрика показывает долю экспериментов, где восстановленный вектор совпадает с эталоном как минимум в N первых битах.

Результаты демонстрируют быстрое уменьшение вероятности успешного восстановления с ростом требуемой длины совпадения, что при высокой размерности становится вычислительно недостижимым.

В Заключении перечислены основные результаты, полученные в диссертации:

1. Получены аналитические формулы для среднего числа допустимых решений и среднего значения целевого функционала, связывающие параметры задачи со структурой множества решений.
2. Выделен и изучен класс сюръективных линейных форм; установлены критерии сюръективности и параметры, влияющие на сложность решения, а также предложен эффективный алгоритм вычисления всех решений.
3. Исследованы линейные формы с разрывами в области значений и разработан метод синтеза форм с контролируемым числом разрывов, позволяющий моделировать переход от простых к трудным для решения экземплярам.
4. Разработаны методы подбора параметров, позволяющие конструировать экземпляры, устойчивые к решеточным методам, но эффективно решаемые предложенными алгоритмами полного перебора.
5. Предложены преобразования рюкзачных экземпляров, позволяющие влиять на сложность решения при сохранении структуры множества решений, что может быть применено в прикладных задачах.
6. Создан программный комплекс, реализующий предложенные алгоритмы, проведены вычислительные эксперименты, подтвердившие корректность и эффективность разработанных методов, а также их применимость в задачах кодирования и защиты информации.

Публикации автора по теме диссертации

1. Волков М. С. А. Комбинаторные свойства задачи об ограниченном рюкзаке // Прикладная дискретная математика. 2024. № 63. С. 117–130. (WOS, Scopus, ВАК).
2. Волков М. С. А., Гордеев Э. Н., Леонтьев В. К. О среднем числе допустимых решений в задаче о рюкзаке // Прикладная дискретная математика. 2025. № 68. С. 103–113. (WOS, Scopus, ВАК).
3. Волков М. С. А., Гордеев Э. Н. Применение неинъективных векторов в ранцевых криптосистемах // Безопасность информационных технологий, 2025. Т. 32. №. 1. С. 122–131. (ВАК).
4. Волков М. С. А. Анализ и реализация криптосистемы на основе неинъективных ранцев // Безопасность информационных технологий, 2025, Т. 32, №. 2. С. 100–111. (ВАК).
5. Леонтьев В. К., Гордеев Э. Н., Волков М. С. А. Классическая непрерывность и ее дискретный вариант. Прикладная физика и математика. 2022. № 1. С. 31–37. (ВАК).
6. Волков М. С. А., Гордеев Э. Н. О непрерывности линейных форм // Безопасные информационные технологии: Материалы XII Международной научно-технической конференции, посвященной 25-летию кафедры ИУ8, Москва, 02 ноября 2023 года. Москва: Издательство МГТУ им. Баумана, 2024. С. 16–20.

7. Волков М. С. А., Гордеев Э. Н., Леонтьев В. К. О свойствах решений обобщенной задачи о рюкзаке // Безопасные информационные технологии: Материалы XII Международной научно-технической конференции, посвященной 25-летию кафедры ИУ8, Москва, 02 ноября 2023 года. Москва: Издательство МГТУ им. Баумана, 2024.
8. Волков М. С. А., Гордеев Э. Н. Исследование и применение сюръективных рюкзаков в криптографии // Безопасные информационные технологии: Материалы XIII Международной научно-технической конференции, Москва, 01 ноября 2024 года. Москва: Издательство МГТУ им. Баумана, 2024. С. 54–56.
9. Volkov, M. S. A. Application of the Method of Coefficients for the Analysis of Combinatorial Properties of the Knapsack Problem / M. S. A. Volkov // Science, Engineering and Business: Collection of materials V Interacademic Conference for Graduate Students and Young Researchers, Moscow, 18–19 апреля 2023 года. Moscow: МГТУ им. Н.Э. Баумана, 2023. P. 303–308.