

УДК 004.8+004.942

DOI:10.25729/ESI.2025.39.3.015

Применение алгоритмов поиска ассоциативных правил для определения технологически значимых сочетаний компонентов в рецептурах резиновых смесей

Рыбанов Александр Александрович, Каблов Виктор Федорович

Волжский политехнический институт (филиал) Волгоградского государственного технического университета, Россия, г. Волжский, rybanoff@yandex.ru

Аннотация. В статье представлены результаты применения алгоритма поиска ассоциативных правил Apriori для анализа рецептур резиновых смесей с целью выявления технологически значимых сочетаний компонентов. Предложена методология, основанная на интеллектуальном анализе данных, которая позволяет автоматически обнаруживать устойчивые комбинации ингредиентов, сокращать пространство поиска рецептур и формализовать экспертные знания. В исследовании проанализирована база данных из 5065 промышленных рецептур с использованием метрик поддержки, достоверности и лифта. Выявлены ключевые ассоциации между ингредиентами резиновых смесей. Результаты работы демонстрируют потенциал методов ассоциативного анализа для оптимизации рецептур, сокращения времени разработки и цифровизации проектирования резиновых смесей.

Ключевые слова: ассоциативные правила, резиновые смеси, ингредиенты, машинное обучение, интеллектуальный анализ данных, база данных, метрики значимости, достоверность, поддержка, лифт, цифровизация материаловедения, поиск закономерностей

Цитирование: Рыбанов А.А. Применение алгоритмов поиска ассоциативных правил для определения технологически значимых сочетаний компонентов в рецептурах резиновых смесей / А.А. Рыбанов, В.Ф. Каблов // Информационные и математические технологии в науке и управлении, 2025. – № 3 (39). – С. 177-188. – DOI:10.25729/ESI.2025.39.3.015.

Введение. Современное проектирование рецептур резиновых смесей представляет собой сложную многокритериальную задачу, требующую учета множества взаимосвязанных факторов: выбора полимерной матрицы, наполнителей, пластификаторов, вулканизирующих агентов и их оптимальных концентраций. Традиционные подходы, основанные на эмпирическом подборе компонентов, часто оказываются ресурсоемкими и неэффективными, особенно в условиях роста требований к эксплуатационным характеристикам материалов (прочность, износостойкость, термостабильность) и необходимости сокращения сроков разработки [1, 2].

В этом контексте методы интеллектуального анализа данных, например, такие, как поиск ассоциативных правил (association rule mining, ARM), открывают новые возможности для систематизации знаний о взаимосвязях между компонентами смесей и их свойствами [3, 4, 5]. Алгоритмы Apriori, FP-Growth [6, 7, 8] и ECLAT [9, 10, 11], изначально разработанные для анализа потребительских корзин, демонстрируют высокую эффективность в выявлении скрытых паттернов в многокомпонентных системах. Их применение позволяет:

- автоматически обнаруживать устойчивые комбинации ингредиентов, коррелирующие с целевыми характеристиками;
- сокращать пространство поиска за счет исключения заведомо неэффективных составов;
- формализовать экспертные знания для цифровизации процессов разработки.

Несмотря на успешное использование ARM в фармацевтике и материаловедении, потенциал этих методов в области проектирования резиновых смесей остается недостаточно изученным. Современные алгоритмы ассоциативных правил демонстрируют значительные ограничения при обработке химико-технологических данных, особенно в контексте рецептур резиновых смесей, где ключевую роль играют не только частотные, но и количественные зависимости. В работах [12, 13] представлены алгоритмы ассоциативных правил, которые не

учитывают специфику химических данных и интерпретацию правил с учетом технологических ограничений. Алгоритм ProbBF, предложенный в работе [14], заменяет традиционное вычисление поддержки на вероятностные прогнозы, что является актуальным для фиксированных пропорций компонентов, однако это может привести к потере значимых комбинаций. Потенциалом для оптимизации рецептур обладают эволюционные методы [15], но они требуют адаптации к химической терминологии. В существующих работах преобладают узконаправленные исследования, фокусирующиеся на отдельных классах материалов, тогда как комплексный анализ кросс-ингредиентных взаимодействий в многокомпонентных системах требует дальнейшего развития.

Цель данной работы – исследовать методологию применения алгоритмов поиска ассоциативных правил для оптимизации рецептур резиновых смесей через выявление устойчивых комбинаций и синергетических эффектов.

1. Методы исследований. В выполненном исследовании осуществляется анализ технологических процессов на основе методов интеллектуального анализа данных (Data Mining) с применением алгоритма Apriori для выявления ассоциативных правил во взаимосвязях ингредиентов, входящих в состав резиновых смесей.

Методологическая основа исследования базируется на выявлении частых наборов ингредиентов и последующей генерации ассоциативных правил, позволяющих установить значимые взаимосвязи между ингредиентами резиновых смесей. Введем следующие обозначения:

- $I = \{i_1, i_2, \dots, i_n\}$ – конечное множество из n уникальных ингредиентов, используемых при производстве резиновых смесей (натуральные/синтетические каучуки, вулканизирующие агенты, наполнители, пластификаторы, антиоксиданты и др.);
- $D = \{R_1, R_2, \dots, R_m\}$ – конечное множество рецептур (база данных), где каждая рецептура $R_k \in D$ представляет собой подмножество ингредиентов $R_k \subseteq I$, $|R_k| \geq 1$.

Для алгоритмической обработки каждая рецептура резиновой смеси $R_k \in D$ кодируется бинарным вектором длины n для всех $j = 1, \dots, n$:

$$R_k \rightarrow r \in \{0, 1\}^n,$$

где $r[j] = 1$, если ингредиент $i_j \in R$; $r[j] = 0$, если ингредиент $i_j \notin R$.

Ассоциативным правилом называется импликация вида:

$$X \Rightarrow Y,$$

Удовлетворяющая условиям: $X \subset I$ – антецедент (условие); $Y \subset I$ – консеквент (следствие); $X \neq \emptyset$, $Y \neq \emptyset$, $X \cap Y = \emptyset$.

Семантика ассоциативного правила выражает вероятностную зависимость: наличие в рецептуре всех ингредиентов антецедента X предполагает наличие ингредиентов консеквента Y с вероятностью, равной достоверности правила $confidence(X \Rightarrow Y)$.

Рассмотрим метрики значимости ассоциативных правил для анализа рецептур резиновых смесей.

Для заданного набора ингредиентов $X \subseteq I$ поддержка $support(X)$ – это доля рецептур R_k в базе данных D , которые содержат все ингредиенты из X :

$$support(X) = \frac{|\{R_k | R_k \in D, X \subseteq R_k\}|}{|D|}.$$

Набор ингредиентов $X \subseteq I$ называется частым при минимальном уровне поддержки $support_{min}$, если $support(X) \geq support_{min}$.

Поддержка ассоциативного правила $X \Rightarrow Y$ показывает частоту совместного использования комбинации ингредиентов в рецептурах резиновых смесей:

$$\text{support}(X \Rightarrow Y) = \text{support}(X \cup Y) = \frac{|\{R_k | R_k \in D, X \cup Y \subseteq R_k\}|}{|D|}.$$

Достоверность ассоциативного правила $X \Rightarrow Y$ определяет вероятность включения ингредиентов Y в рецептуру резиновой смеси при наличии в ней ингредиентов X и показывает технологическую зависимость:

$$\text{confidence}(X \Rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X)}, \text{ где } \text{support}(X) > 0.$$

Лифт ассоциативного правила $X \Rightarrow Y$ количественно характеризует степень ассоциативной связи между антецедентом X и консеквентом Y . Данная метрика определяется, как отношение наблюдаемой частоты совместного появления ингредиентов к ожидаемой частоте при условии их статистической независимости:

$$\text{lift}(X \Rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X) \cdot \text{support}(Y)} = \frac{\text{confidence}(X \Rightarrow Y)}{\text{support}(Y)}.$$

Применение ассоциативных правил с анализом указанных метрик представляет собой эффективный инструмент для управления рецептурами резиновых смесей на основе данных.

Высокие значения поддержки $\text{support}(X \Rightarrow Y)$ указывают на типовые комбинации компонентов (например, совместное использование оксида цинка и серы в 66,27 % рецептов).

Анализ метрики $\text{support}(X \Rightarrow Y)$ позволяет выявлять статистически значимые комбинации ингредиентов, где высокие значения поддержки указывают на технологически обоснованные и широко применяемые компонентные сочетания, характерные для отрасли, что позволяет:

- идентифицировать стандартные/типовые комбинации ингредиентов (например, базовые системы вулканизации);
- отсеивать редкие и нетипичные сочетания, потенциально ведущие к нестабильности свойств;
- формировать исходные рецептурные шаблоны при разработке новых составов;
- проводить сравнительный анализ технологических традиций разных производителей.

Достоверность $\text{confidence}(X \Rightarrow Y)$ показывает технологическую зависимость: $\text{confidence} = 0.9185$ для правила (Сера, Стеарин технический) $\Rightarrow$ (Оксид цинка)) означает, что в 91,85% рецептов с серой и стеарином присутствует оксид цинка. Данная метрика количественно определяет надежность технологических зависимостей между ингредиентами, позволяя выделять устойчивые причинно-следственные связи в составе рецептов – высокие значения достоверности ($\text{confidence} > 0.8$) свидетельствуют о стабильных комбинациях ингредиентов, которые следует учитывать при модификации составов, тогда как низкие значения ($\text{confidence} < 0.3$) указывают на случайные или слабые взаимосвязи, не имеющие практической значимости для технологического процесса.

Лифт $\text{lift}(X \Rightarrow Y)$ представляет собой количественную меру технологической синергии между ингредиентами состава. Ее значения интерпретируются следующим образом:

- $\text{lift} > 1$ (положительная корреляция): указывает на синергетическое взаимодействие ингредиентов;
- $\text{lift} \approx 1$ (нейтральная связь): ингредиенты используются независимо друг от друга;
- $\text{lift} < 1$ (отрицательная корреляция): свидетельствует о технологической несовместимости ингредиентов.

Анализ метрики $\text{lift}(X \Rightarrow Y)$ позволяет оптимизировать составы резиновых смесей путем сохранения проверенных синергетических комбинаций ($\text{lift} > 1.5$) и исключения технологически несовместимых пар компонентов ($\text{lift} < 0.5$).

На рис. 1 представлена UML-диаграмма последовательности, иллюстрирующая процесс извлечения ассоциативных правил. Диаграмма отражает поэтапное взаимодействие между компонентами системы, включая:

- извлечение данных из базы рецептов резиновых смесей,
- трансформацию в бинарный формат,
- применение алгоритма Apriori для идентификации частых наборов,
- генерацию ассоциативных правил с заданными метриками поддержки (support) и достоверности (confidence).

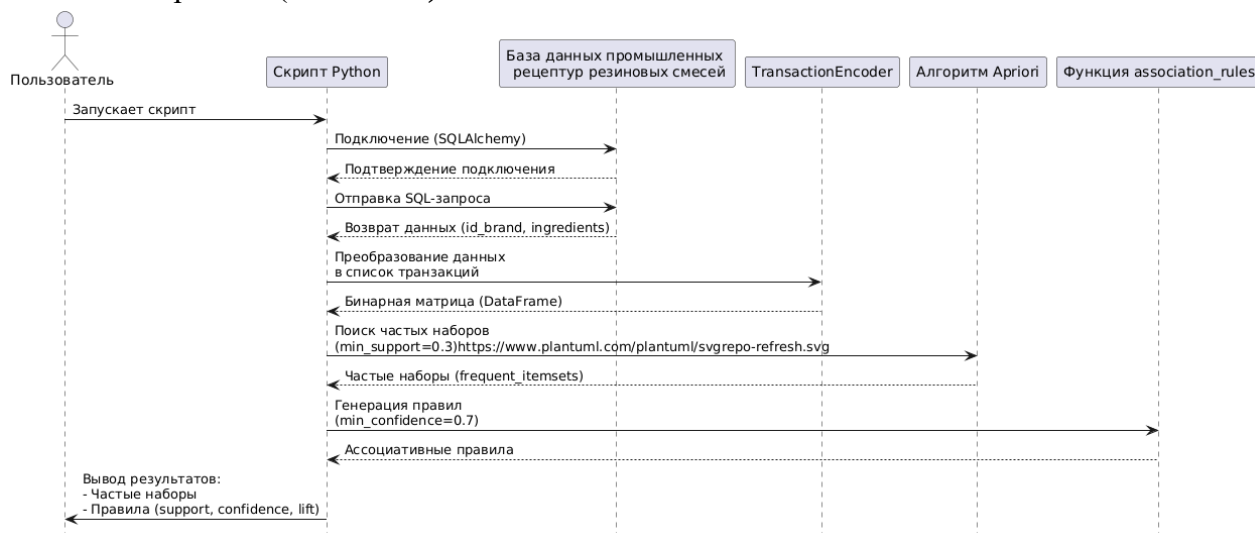


Рис. 1. UML-диаграмма последовательности для процесса поиска ассоциативных правил

Визуализация позволяет проследить логическую последовательность операций, начиная с инициализации процесса пользователем и заканчивая выводом статистически значимых паттернов, что соответствует методологии поиска ассоциативных правил.

В качестве исходных данных в исследовании применялась база данных промышленных рецептов резиновых смесей, охватывающая период с 1989 года по настоящее время. База включает информацию о составе, ингредиентах и свойствах резиновых смесей (с учетом технических синонимов [16, 17], содержащая в общей сложности 5065 уникальных рецептов (рис. 2).

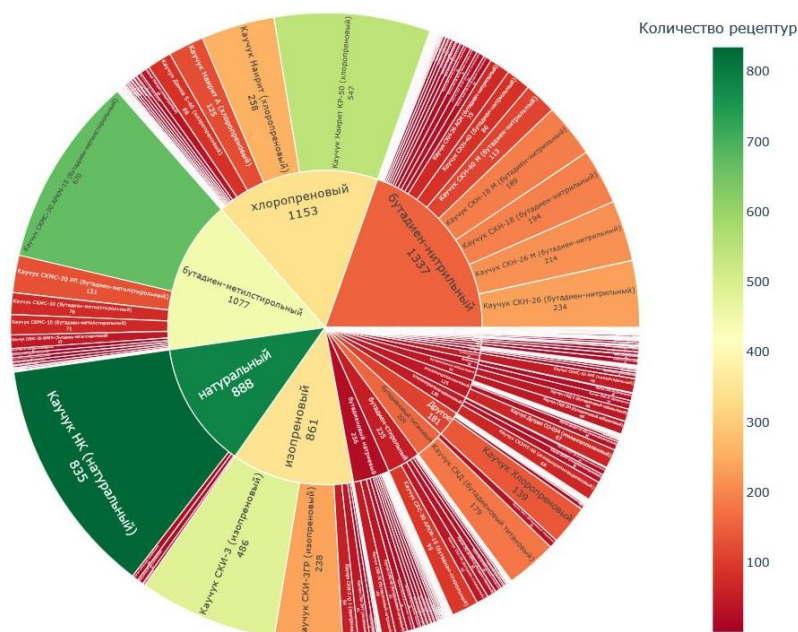


Рис. 2. Иерархическое распределение видов каучука в базе данных

Анализ иерархического распределения видов каучука в рецептурах резиновых смесей БД, представленный на рис. 2, показывает, что хлоропреновый каучук лидирует с 1153 рецептурами, за ним следует натуральный каучук с 888 рецептурами, что свидетельствует об их универсальности и широком применении в различных отраслях. В то же время такие виды, как бутиловый и стирол-бутадиеновый каучук, имеют значительно меньше рецептов, что может указывать на узкую область применения.

Фрагмент логической схемы базы данных, содержащей данные о составе рецептов резиновых смесей, на основе которых была сформирована выборка исходных данных для поиска ассоциативных правил, представлен на рис. 3.

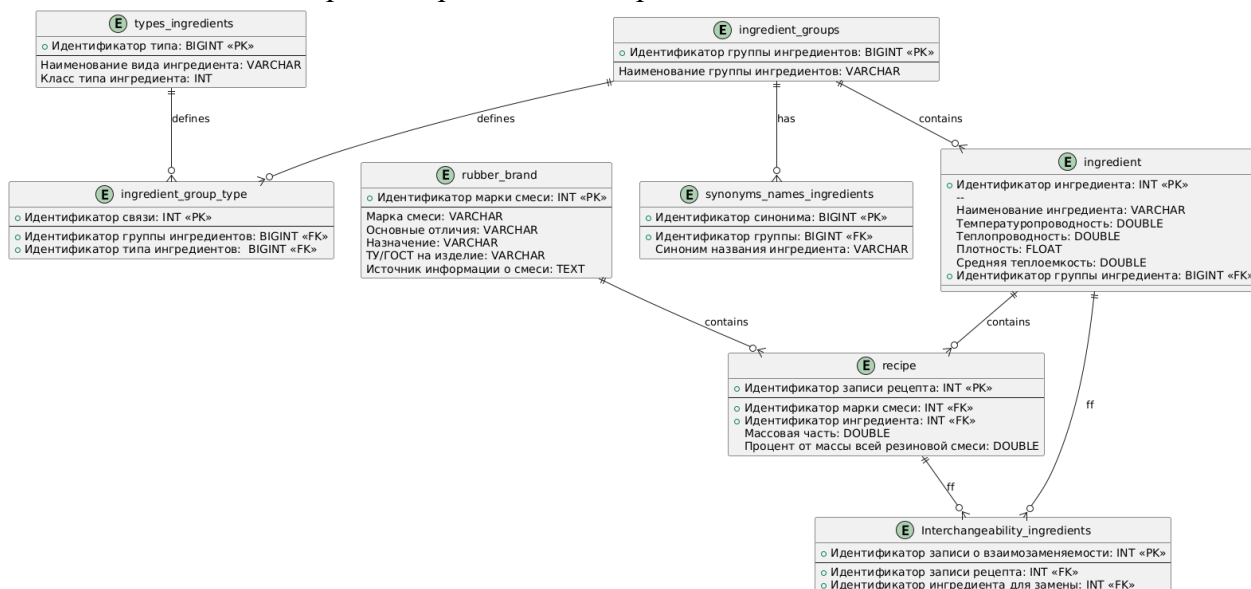


Рис. 3. Фрагмент логической схемы БД с исходными данными по составу рецептов резиновых смесей

Особенностью базы данных является наличие справочников технических синонимов ингредиентов и свойств резиновых смесей, которые упрощают поиск информации и позволяют применять методы машинного обучения для анализа данных

2. Результаты исследований. В рамках проведенного исследования осуществлен анализ 5065 рецептов резиновых смесей с применением методов ассоциативного rule mining для выявления устойчивых взаимосвязей между ингредиентами. Использование алгоритма Apriori позволило идентифицировать частые наборы компонентов и установить статистически значимые ассоциативные правила. Результаты анализа количественно оценены с использованием метрик поддержки (support), достоверности (confidence) и лифта (lift), характеризующих силу и значимость выявленных ассоциаций.

В таблице 1 представлены наиболее распространенные комбинации ингредиентов, ранжированные по уровню поддержки. Анализ данных показал, что наибольшую частоту встречаемости среди одиночных компонентов демонстрируют *оксид цинка* ($support = 0.851$) и *сера* ($support = 0.758$), что свидетельствует об их ключевой роли в составе резиновых смесей.

Также выявлены устойчивые парные и тройные комбинации, такие, как:

- 1) *оксид цинка* + *сера* ($support = 0.663$),
- 2) *оксид цинка* + *технический углерод П-803* ($support = 0.369$),
- 3) *оксид цинка* + *сера* + *стеарин технический* ($support = 0.319$).

Таблица 1. Частые наборы ингредиентов с уровнем поддержки

№	Наборы ингредиентов (itemsets)	Поддержка (support)
1	Нафтам-2	0.372088
2	(Тиурам Д)	0.305771
3	(Технический углерод П-803)	0.415950
4	(Стеарин технический)	0.437566
5	(Мел)	0.338720
6	(Сера)	0.757817
7	(Оксид цинка)	0.851417
8	(Оксид цинка, Нафтам-2)	0.350052
9	(Сера, Технический углерод П-803)	0.324869
10	(Оксид цинка, Технический углерод П-803)	0.368730
11	(Сера, Стеарин технический)	0.347534
12	(Оксид цинка, Стеарин технический)	0.393704
13	(Мел, Оксид цинка)	0.306401
14	(Оксид цинка, Сера)	0.662749
15	(Оксид цинка, Сера, Стеарин технический)	0.319203

Наличие данных сочетаний подтверждает существование стандартизированных подходов к составлению рецептур, где определенные ингредиенты систематически комбинируются для достижения заданных технологических свойств.

В таблице 2 приведены наиболее значимые ассоциативные правила, отобранные на основе комбинации метрик *confidence* и *lift*. Правила с *lift* > 1 указывают на положительную корреляцию между компонентами, при этом значения *lift* > 1.05 свидетельствуют о статистически значимых ассоциациях.

Таблица 2. Ассоциативные правила с показателями достоверности и лифта

№	Причина (antecedents)	Следствие (consequents)	Поддержка (support)	Достоверность (confidence)	Лифт (lift)
1	(Нафтам-2)	(Оксид цинка)	0.350052	0.940778	1.104957
2	(Технический углерод П-803)	(Сера)	0.324869	0.781029	1.030630
3	(Технический углерод П-803)	(Оксид цинка)	0.368730	0.886478	1.041180
4	(Стеарин технический)	(Сера)	0.347534	0.794245	1.048069
5	(Стеарин технический)	(Оксид цинка)	0.393704	0.899760	1.056780
6	(Мел)	(Оксид цинка)	0.306401	0.904585	1.062447
7	(Оксид цинка)	(Сера)	0.662749	0.778408	1.027170
8	(Сера)	(Оксид цинка)	0.662749	0.874550	1.027170
9	(Оксид цинка, Стеарин технический)	(Сера)	0.319203	0.810768	1.069872
10	(Сера, Стеарин технический)	(Оксид цинка)	0.319203	0.918478	1.078765
11	(Стеарин технический)	(Оксид цинка, Сера)	0.319203	0.729496	1.100713

Наибольшую прогностическую силу демонстрируют следующие правила:

1) *Стеарин технический* $\Rightarrow$ *Оксид цинка* (*confidence* = 0.9, *lift* = 1.057) – наличие стеарина в составе с высокой вероятностью предполагает использование оксида цинка.

2) Мел $\Rightarrow$ Оксид цинка (*confidence* = 0.905, *lift* = 1.062) – сочетание мела с оксидом цинка встречается чаще, чем можно было бы ожидать при случайном распределении.

3) Оксид цинка + Стеарин технический $\Rightarrow$ Сера (*confidence* = 0.811, *lift* = 1.07) – данная тройная ассоциация указывает на устойчивую взаимосвязь между указанными компонентами.

На рис. 4 представлена визуализация найденных ассоциативных правил в виде графа.

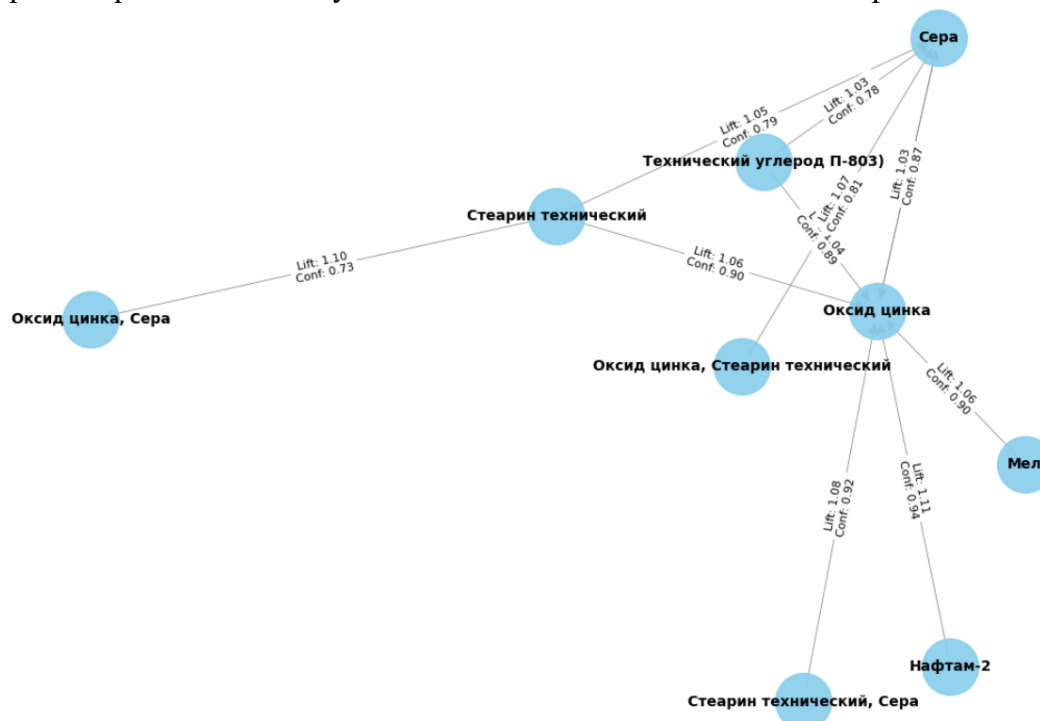


Рис. 4. Граф ассоциативных правил

Графовая структура ассоциативных правил представляет собой ориентированную сеть, узлы которой соответствуют компонентам рецептур резиновых смесей, а дуги отражают статистически значимые взаимосвязи между ними. Весовые коэффициенты связей (*lift* и *confidence*) количественно характеризуют силу и достоверность выявленных ассоциаций между ингредиентами в составе композиций. Данная модель позволяет визуализировать и анализировать закономерности совместного использования компонентов при проектировании резиновых смесей.

3. Обсуждение результатов исследования. Проведенный анализ методами ассоциативного rule mining выявил статистически значимые взаимосвязи между компонентами резиновых смесей, что подтверждает гипотезу о существовании устойчивых комбинаций ингредиентов в промышленных рецептурах. Полученные результаты имеют важное значение как для фундаментального понимания принципов составления резиновых смесей, так и для практического применения в технологии производства резинотехнических изделий.

Интерпретация ключевых закономерностей:

- высокие значения поддержки для оксида цинка (*support* = 0.851) и серы (*support* = 0.758) согласуются с их общепризнанной функцией в процессах вулканизации; однако обнаруженная сильная ассоциация (рис. 5):

оксид цинка $\Rightarrow$ сера (*lift* = 1.027)

и обратная зависимость

сера $\Rightarrow$ оксид цинка (*lift* = 1.027)

указывают на их взаимодополняемость в рецептурах, что может быть связано с необходимостью балансировки активности вулканизирующих систем;

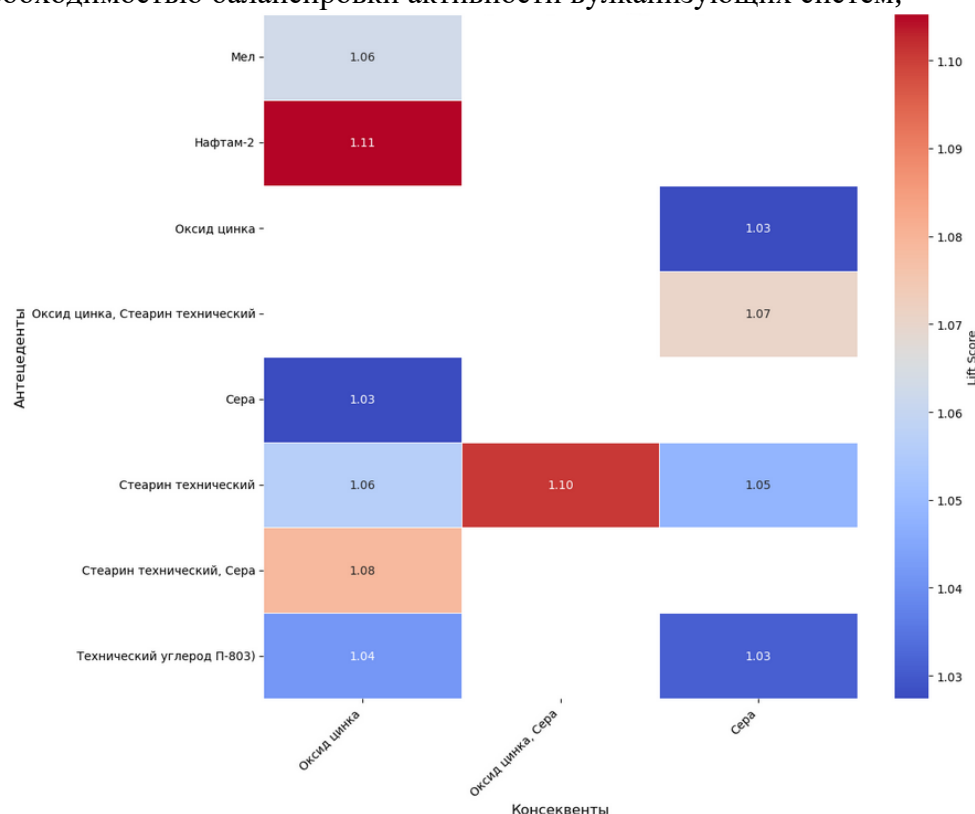


Рис. 5. Тепловая карта lift для наборов данных

- установлено, что стеарин технический ($support = 0.438$) часто сочетается как с оксидом цинка ($confidence = 0.9$, $lift = 1.057$), так и с серой ($confidence = 0.794$, $lift = 1.048$); это подтверждает его роль не только как пластификатора, но и как компонента, влияющего на кинетику вулканизации;
- ингредиент мел демонстрирует неожиданно высокую ассоциацию с оксидом цинка ($confidence = 0.905$, $lift = 1.062$), что может свидетельствовать о его участии в модификации свойств резиновых смесей за пределами традиционной функции наполнителя;
- выявленная взаимосвязь:

оксид цинка + стеарин технический $\Rightarrow$ сера ($lift = 1.07$)

указывает на существование сложных синергетических эффектов, при которых комбинация активатора (ZnO) и пластификатора (стеарин) статистически значимо предопределяет наличие вулканизирующего агента (серы); это может быть связано с оптимизацией рецептур для достижения определенных реологических или механических характеристик.

Можно выделить следующие оригинальные аспекты исследования:

1) обнаружено (рис. 6), что нафтам-2, несмотря на относительно низкую частоту встречаемости ($support = 0.372$), демонстрирует высокую предсказательную силу в отношении оксида цинка ($confidence = 0.941$, $lift = 1.105$); это позволяет предположить его специализированное применение в рецептурах с повышенными требованиями к активации вулканизации;

2) технический углерод П-803 ($support = 0.416$) чаще сочетается с оксидом цинка ($confidence = 0.886$), чем с серой ($confidence = 0.781$), что может отражать его влияние

не только на усиление, но и на диспергирование активных ингредиентов в резиновой матрице (рис. 6).

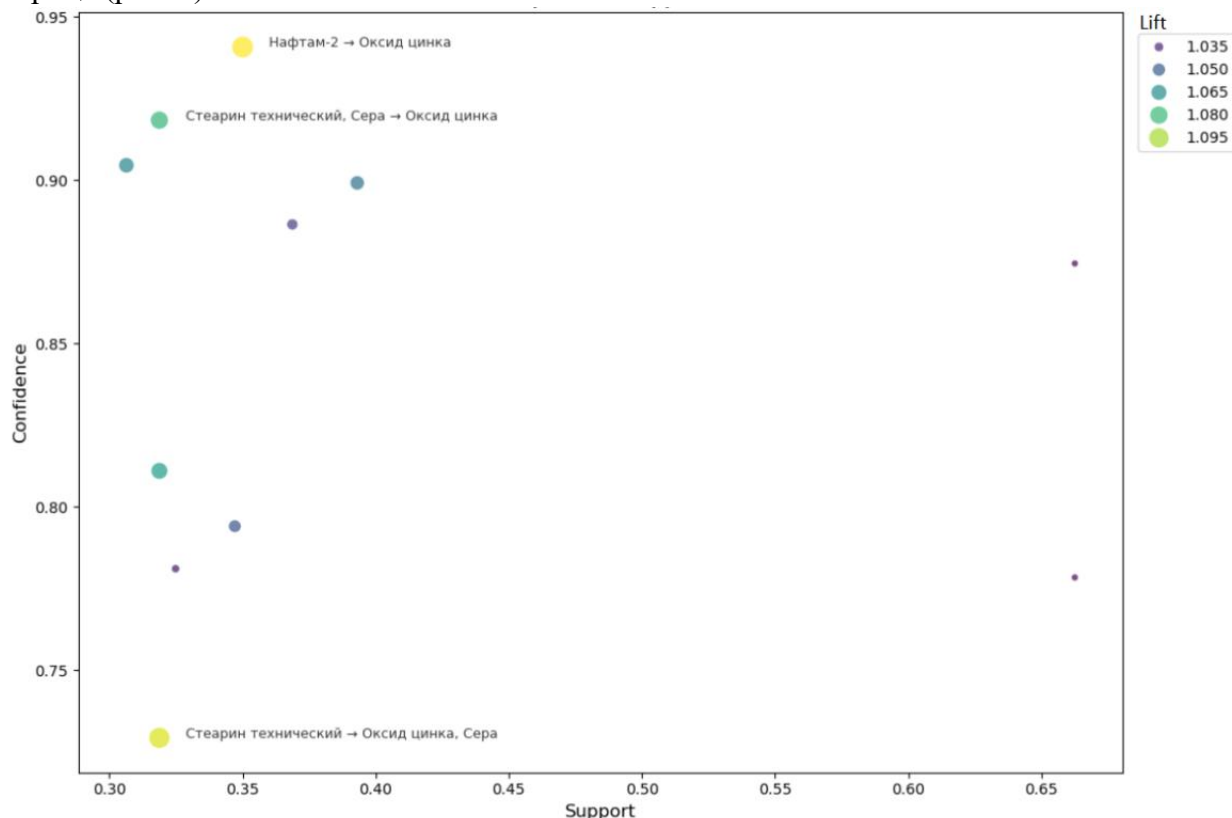


Рис. 6. Визуализация ассоциативных правил

Следует отметить следующие методологические ограничения: во-первых, выявленные закономерности носят специфический характер и привязаны к конкретной выборке рецептурных данных, что потенциально ограничивает их обобщаемость для различных производственных практик; во-вторых, анализ был сосредоточен исключительно на компонентном составе смесей без рассмотрения их эксплуатационно-технических характеристик, что создает определенные сложности в интерпретации взаимосвязей между составом и функциональными свойствами конечных резиновых изделий; в-третьих, хотя показатели $lift > 1$ демонстрируют статистическую значимость обнаруженных ассоциативных правил, их технологическая значимость, особенно для случаев с $lift < 1.1$, требует дополнительного экспериментального подтверждения для валидации практической применимости полученных результатов.

Заключение. Проведенное исследование демонстрирует эффективность методов ассоциативного анализа для выявления скрытых закономерностей в составе резиновых смесей. Полученные результаты имеют значительный потенциал для практического применения в резинотехнической промышленности, особенно в контексте цифровой трансформации производства. Дальнейшие исследования должны быть направлены на экспериментальную верификацию выявленных ассоциаций и разработку на их основе интеллектуальных систем поддержки принятия решений при проектировании новых рецептур. Таким образом, работа вносит вклад в цифровизацию проектирования резиновых смесей, предлагая инструмент для оптимизации состава на основе данных.

Список источников

1. Нигматуллин, В. Р. Использование методов машинного обучения и искусственного интеллекта в химической технологии. Часть I / В. Р. Нигматуллин, Н. А. Руднев // Нефтегазовое дело, 2019. – № 4. – С. 243-268.

2. Каблов В.Ф. Эволюция информационных систем управления рецептурами резиновых смесей: от баз данных к интеллектуальным хранилищам с интеграцией модулей искусственного интеллекта / В.Ф. Каблов, А.А. Рыбанов, М.А. Маслова // Каучук и Резина-2025: Традиции и новации: материалы XIII Всерос. конф. – Москва, 2025. – С. 23-24.
3. Chinnasamy M., Rathanasamy R., Pal S.K. et al. Data Mining in Material Science. Application of Artificial Intelligence in New Materials Discovery, 2023, pp. 24-46, DOI: 10.21741/9781644902530-2.
4. Yu G., Chen J., Zhu L. Data Mining Techniques for Materials Informatics: Datasets Preparing and Applications. Proc. of the Int. Symposium on Knowledge Acquisition and Modeling (KAM 2009), 2009, pp. 189-192, DOI: 10.1109/KAM.2009.98.
5. Rajan K., Rajagopalan A., Suh C. Data mining and multivariate analysis in materials science. Molten salts: from fundamentals to applications, Dordrecht, Springer, 2002, vol. 52, pp. 89-102, DOI: 10.1007/978-94-010-0458-9_8.
6. Пальмов С. В. Алгоритм поиска ассоциативных правил FP-growth / С.В. Пальмов, Е.Н. Французова // Национальная Ассоциация Ученых, 2016. – № 10-1 (26). – С. 27-28.
7. Зуенко А.А. Обзор методов поиска частых паттернов для интеллектуального анализа данных / А.А. Зуенко, О.В. Фридман // Труды Кольского научного центра РАН. Серия: Технические науки, 2024. – Т. 15. – № 3. – С. 82-96.
8. Графеева Н.Г. Реализация алгоритмов машинного обучения средствами реляционных СУБД / Н.Г. Графеева, А.А. Назаров // Компьютерные инструменты в образовании, 2024. – № 2. – С. 58-71.
9. Бурец А.С. Ассоциативные правила и алгоритмы их поиска / А.С. Бурец // Научные вести, 2019. – № 12 (17). – С. 86-96.
10. Suvalka B., Khandelwal S., Patel C. Revised ECLAT algorithm for frequent itemset mining. Proc. of the Int. Conf. on Advanced computing and communication technologies, 2016, pp. 211–220, DOI: 10.1007/978-81-322-2752-6_21.
11. Бова В.В. Оценка эффективности метода поиска ассоциативных правил для задач обработки больших данных / В.В. Бова, Э.В. Кулиев, С.Н. Щеглов // Известия ЮФУ. Технические науки, 2020. – № 2 (212). – С. 66-78.
12. Цымблер М.Л. Параллельный поиск частых наборов на многоядерных ускорителях Intel MIC / М.Л. Цымблер // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика, 2019. – Т. 8. – № 1. – С. 54–70. – DOI: 10.14529/cmse190104.
13. Sadeequllah M., Rauf A., Rehman S.U., Alnazzawi N. Probabilistic support prediction: fast frequent itemset mining in dense data. IEEE Access, 2024, vol. 12, pp. 39330-39350, DOI: 10.1109/ACCESS.2024.3376477.
14. Telikani A., Gandomi A.H., Shahbahrami A.A. A survey of evolutionary computation for association rule mining. Information Sciences, 2020, vol. 524, pp. 318–352.
15. База данных рецептов, ингредиентов и свойств резиновых смесей с учетом технических синонимов: свидетельство о гос. регистрации базы данных № 2025621312 Рос. Федерация: № 2025620796: заявл. 12.03.2025; опублик. 24.03.2025 / В.Ф. Каблов, А.А. Рыбанов, Н.А. Кейбал, и др.; правообладатель ВПИ (филиал) ВолГТУ.
16. Каблов В.Ф. Автоматизированный банк данных нового поколения рецептов и свойств резин – “Раббер С” / В.Ф. Каблов, А.А. Рыбанов, Н.А. Кейбал // Каучук и резина, 2024. – Т. 83. – № 3. – С. 168-173.

Рыбанов Александр Александрович. Кандидат технических наук, доцент, кафедра «Информатика и технология программирования», Волжский политехнический институт (филиал) Волгоградский государственный технический университет, заведующий кафедрой. AuthorID: 234968, SPIN: 9256-6050, ORCID: 0000-0002-8638-9998, rybanoff@yandex.ru. 404111, Россия, Волгоградская область, г. Волжский, пр-кт Ленина, 72.

Каблов Виктор Федорович. Доктор технических наук, профессор, кафедра «Химическая технология полимеров и промышленная экология», Волжский политехнический институт (филиал) Волгоградский государственный технический университет, профессор. AuthorID: 115782, SPIN: 1093-9003, ORCID: 0000-0002-2970-6109, vkablov5@gmail.com. 404111, Россия, Волгоградская область, г. Волжский, пр-кт Ленина, 72.

UDC 004.8+004.942

DOI:10.25729/ESI.2025.39.3.015

Application of association rule mining algorithms to identify technologically significant component combinations in rubber compound formulations

Alexander A. Rybanov, Victor F. Kablov

Volzhsky polytechnic institute (Branch) of Volgograd state technical university,
Russia, Volzhsky, rybanoff@yandex.ru

Abstract. The article presents the results of applying the Apriori association rule mining algorithm to analyze rubber compound formulations in order to identify technologically significant component combinations. A methodology based on data mining is proposed, which enables the automatic detection of stable ingredient combinations, reduces the search space for formulations, and formalizes expert knowledge. The study analyzed a database of 5065 industrial formulations using support, confidence, and lift metrics. Key associations between rubber compound ingredients were identified. The results demonstrate the potential of association analysis methods for optimizing formulations, reducing development time, and digitalizing rubber compound design.

Keywords: association rules, rubber compounds, ingredients, machine learning, data mining, database, significance metrics, confidence, support, lift, digitalization of materials science, pattern discovery

References

1. Nigmatullin V.R., Rudnev N.A. Ispol'zovaniye metodov mashinnogo obucheniya i iskusstvennogo intellekta v khimicheskoy tekhnologii. Chast' I [The use of machine learning and artificial intelligence methods in chemical technology. Part I]. Neftegazovoye delo [Oil and Gas Business], 2019, no. 4, pp. 243-268.
2. Kablov V.F., Rybanov A.A., Maslova M.A. Evolyutsiya informatsionnykh sistem upravleniya retsepturami rezinovykh smesey: ot baz dannykh k intellektual'nym khranilishcham s integratsiyey moduley iskusstvennogo intellekta [Evolution of information systems for managing rubber compound recipes: from databases to intelligent warehouses with integrated AI modules]. Kauchuk i Rezina-2025: Traditsii i novatsii: materialy XIII Vseros. konf. [Rubber and Rubber-2025: Traditions and Innovations: Proceedings of the XIII All-Russian Conference]. Moscow, 2025, pp. 23-24.
3. Chinnasamy M., Rathanasamy R., Pal S.K. et al. Data Mining in Material Science. Application of Artificial Intelligence in New Materials Discovery, 2023, pp. 24-46, DOI: 10.21741/9781644902530-2.
4. Yu G., Chen J., Zhu L. Data Mining Techniques for Materials Informatics: Datasets Preparing and Applications. Proc. of the Int. Symposium on Knowledge Acquisition and Modeling (KAM 2009), 2009, pp. 189-192, DOI: 10.1109/KAM.2009.98.
5. Rajan K., Rajagopalan A., Suh C. Data mining and multivariate analysis in materials science. Molten salts: from fundamentals to applications, Dordrecht, Springer, 2002, vol. 52, pp. 89-102, DOI: 10.1007/978-94-010-0458-9_8.
6. Palmov S.V., Frantsozova E.N. Algoritmy poiska assotsiativnykh pravil FP-growth [FP-growth association rule mining algorithm]. Natsional'naya Assotsiatsiya Uchenykh [National Association of Scientists], 2016, no. 10-1 (26), pp. 27-28.
7. Zuenko A.A., Fridman O.V. Obzor metodov poiska chastykh patternov dlya intellektual'nogo analiza dannykh [A review of frequent pattern mining methods for data mining]. Trudy Kol'skogo nauchnogo tsentra RAN. Seriya: Tekhnicheskiye nauki [Transactions of the Kola Science Centre of the Russian Academy of Sciences. Series: Technical Sciences], 2024, vol. 15, no. 3, pp. 82-96.
8. Grafeeva N.G., Nazarov A.A. Realizatsiya algoritmov mashinnogo obucheniya sredstvami relyatsionnykh SUBD [Implementation of machine learning algorithms using relational DBMS]. Komp'yuternyye instrumenty v obrazovanii [Computer tools in education], 2024, no. 2, pp. 58-71.
9. Burets A.S. Assotsiativnyye pravila i algoritmy ikh poiska [Association rules and algorithms for their mining]. Nauchnyye vesti [Scientific News], 2019, no. 12 (17), pp. 86-96.
10. Suvalka B., Khandelwal S., Patel C. Revised ECLAT algorithm for frequent itemset mining. Proc. of the Int. Conf. on Advanced computing and communication technologies, 2016, pp. 211-220, DOI: 10.1007/978-81-322-2752-6_21
11. Bova V.V., Kuliev E.V., Shcheglov S.N. Otsenka effektivnosti metoda poiska assotsiativnykh pravil dlya zadach obrabotki bol'shikh dannykh [Efficiency evaluation of an association rule mining method for big data processing tasks]. Izvestiya YuFU. Tekhnicheskiye nauki [Izvestiya SFedU. Engineering Sciences], 2020, no. 2 (212), pp. 66-78.

12. Tsymbler M.L. Parallel'nyy poisk chastykh naborov na mnogoyadernykh uskoritelyakh Intel MIC [Parallel frequent itemset mining on Intel MIC many-core accelerators]. Vestnik YuUrGU. Seriya: Vychislitel'naya matematika i informatika [Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering], 2019, vol. 8, no. 1, pp. 54-70, DOI: 10.14529/cmse190104.
13. Sadeequllah M., Rauf A., Rehman S.U., Alnazzawi N. Probabilistic support prediction: fast frequent itemset mining in dense data. IEEE Access, 2024, vol. 12, pp. 39330-39350, DOI: 10.1109/ACCESS.2024.3376477.
14. Telikani A., Gandomi A.H., Shahbahrami A.A. A survey of evolutionary computation for association rule mining. Information Sciences, 2020, vol. 524, pp. 318-352.
15. Baza dannykh retseptur, ingredientov i svoystv rezinovykh smesey s uchetom tekhnicheskikh sinonimov [Database of recipes, ingredients and properties of rubber compounds, accounting for technical synonyms]: svidetel'stvo o gos. registratsii bazy dannykh no. 2025621312 Ros. Federatsiya: no. 2025620796: zayavl. 12.03.2025: opubl. 24.03.2025. V.F. Kablov, A.A. Rybanov, N.A. Keybal, et al.; pravoobladatel' VPI (filial) VolgGTU [Certificate of state registration of the database No. 2025621312 Russian Federation: No. 2025620796: appl. 12.03.2025: publ. 24.03.2025 / V.F. Kablov, A.A. Rybanov, N.A. Keybal, M.A. Maslova, D.A. Kryukova; right holder VPI (branch) of VolgSTU].
16. Kablov V.F., Rybanov A.A., Keybal N.A. Avtomatizirovannyi bank dannykh novogo pokoleniya retseptur i svoystv rezin – “Rabber S” [A new-generation automated databank for rubber recipes and properties - "Rubber S"]. Kauchuk i rezina [Rubber and Rubber], 2024, vol. 83, no. 3, pp. 168-173.

Alexander Aleksandrovich Rybanov. PhD in Engineering, Associate Professor, Head of the Department of "Informatics and Programming Technology", Volzhsky Polytechnic Institute (branch) of Volgograd State Technical University. AuthorID: 234968, SPIN: 9256-6050, ORCID: 0000-0002-8638-9998, rybanoff@yandex.ru. 404111, Russia, Volgograd Region, Volzhsky, 72 Lenin Ave.

Victor Fedorovich Kablov. Doctor in Engineering, Professor, Department of "Chemical Technology of Polymers and Industrial Ecology", Volzhsky Polytechnic Institute (branch) of Volgograd State Technical University. AuthorID: 115782, SPIN: 1093-9003, ORCID: 0000-0002-2970-6109, vkablov5@gmail.com. 404111, Russia, Volgograd Region, Volzhsky, 72 Lenin Ave.

Статья поступила в редакцию 21.05.2025; одобрена после рецензирования 27.05.2025; принята к публикации 02.08.2025.

The article was submitted 05/21/2025; approved after reviewing 05/27/2025; accepted for publication 08/02/2025.