

ГИБРИДИЗАЦИЯ МАРКОВСКИХ ЦЕПЕЙ С ОБЪЯСНЯЕМЫМИ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ ПРИ ПРОГНОЗИРОВАНИИ РЕАГИРОВАНИЯ ПОЖАРНО-СПАСАТЕЛЬНЫХ ПОДРАЗДЕЛЕНИЙ НА ПОЖАРЫ ОБЪЕКТОВ ТЭК

Хабибулин Р.Ш.

Академия Государственной противопожарной службы МЧС России,
Москва, Россия
kh-r@yandex.ru

Аннотация. В статье рассматривается подход к прогнозированию времени реагирования пожарно-спасательных подразделений на пожары на объектах топливно-энергетического комплекса (ТЭК) с использованием гибридной модели, объединяющей марковские цепи и объясняемые методы машинного обучения (XAI). В исследовании анализируются возможности четырёх алгоритмов для интерпретации решений (Decision Tree, EBM, GAM и CatBoost + SHAP).

Ключевые слова: реагирование на пожары, ТЭК, марковская модель, объясняемое машинное обучение, SHAP, EBM, интерпретируемость, СППР.

Введение

Обеспечение высокой оперативности и эффективности реагирования на пожары на объектах топливно-энергетического комплекса (ТЭК) остаётся одной из приоритетных задач в области пожарной безопасности. Объекты ТЭК, включая резервуарные парки, нефтебазы, автозаправочные станции представляют собой высокорисковые объекты защиты, где от оперативного реагирования пожарно-спасательных подразделений зависит недопущение масштабных материальных и экологических последствий [1-3]. Марковские модели, описывающие последовательные фазы пожара, активно применяются в вопросах изучения прогнозирования пожаров на различных объектах защиты [4-7]. При этом в разработанных моделях отсутствует учет влияния множества факторов на результаты развития, прогнозирования и тушения пожаров. Предлагается рассмотреть современные подходы в области объясняемого (интерпретируемого) машинного обучения (*Explainable AI, XAI*) [8-10], которые позволяют построить гибридные модели, в которых параметры марковской цепи адаптируются в зависимости от контекста пожара (инцидента реагирования на пожар). Применение таких моделей направлено на повышение точности прогнозов, и позволяет получить интерпретируемую информацию о влиянии конкретных признаков на длительность фаз реагирования. Данное исследование направлено на разработку и верификацию гибридной модели, объединяющей возможности марковских цепей и интерпретируемых XAI-методов при прогнозировании фаз реагирования на пожары на объектах ТЭК.

1. Цель и задачи исследования

Основной целью исследования является построение контекстно-чувствительной модели прогнозирования времени реагирования пожарно-спасательных подразделений на основе гибридации марковской цепи с объясняемыми алгоритмами машинного обучения. Для достижения данной цели решаются следующие задачи:

1. Формализовать марковскую модель фаз реагирования, адаптируемую к контексту конкретного пожара.
2. Обучить и сравнить четыре XAI-метода (*Decision Tree, Explainable Boosting Machine, Generalized Additive Model, CatBoost + SHAP*) для оценки времени продолжительности фаз реагирования.
3. Встроить предсказанные значения продолжительности фаз реагирования в модель цепи с переменными интенсивностями.
4. Оценить точность моделей и интерпретируемость их решений по совокупности сформулированных метрик.
5. Разработать основу технического задания на компьютерную систему поддержки принятия решений (СППР).

2. Описание исходных данных и методология

Для построения модели использовались ретроспективные данные о 194 пожарах (инцидентах реагирования на пожары) на пожароопасных объектах резервуарного типа, относящихся к инфраструктуре ТЭК за период 2009–2021 гг. Каждый пожар описывался набором из 11 признаков, включающих:

Территориальные и организационные характеристики:

- Вид населенного пункта (5 категорий: город, поселок городского типа, сельский населенный пункт, вахтовый поселок, вне территории населенного пункта);
- Вид пожарной охраны (5 категорий: федеральная противопожарная служба, противопожарная служба субъекта РФ, ведомственная, добровольная, не охранялся);
- Расстояние до пожарно-спасательной части (км): диапазон 1–35 км.

Ресурсы реагирования (8 признаков):

- Условия, способствующие развитию пожара (закодированы согласно отраслевому классификатору);
- Участники тушения пожара (типы подразделений, привлеченных к тушению);
- Техника, прибывшая к месту пожара (классификация по типам пожарных автомобилей);
- Количество задействованной техники;
- Пожарные стволы, поданные на тушение (тип и характеристики);
- Количество стволов;
- Огнетушащие средства (типы применяемых веществ);
- Первичные средства пожаротушения.

Признаки ресурсов реагирования представлены в закодированном формате, где каждый код соответствует определенной категории согласно нормативным документам.

Целевыми переменными выступали интервалы времени, соответствующие шести фазам реагирования: от обнаружения до сообщения, от сообщения до прибытия, от прибытия до подачи первого ствола и далее до полной ликвидации последствий. Статистический анализ данных показал следующие характеристики фаз реагирования (Таблица 1).

Таблица 1. Описательная статистика фаз реагирования

Фаза	Название	Среднее (мин)	Медиана (мин)	СКО (мин)	Макс (мин)
Фаза 1	От обнаружения пожара до сообщения	2,4	1,0	3,0	20,0
Фаза 2	От сообщения до прибытия 1-го подразделения	8,2	7,0	6,7	48,0
Фаза 3	От прибытия 1-го подразделения до подачи 1-го ствола	1,2	1,0	0,8	6,0
Фаза 4	От подачи 1-го ствола до локализации пожара	25,2	7,0	83,8	813,0
Фаза 5	От локализации пожара до ликвидации открытого горения	26,3	7,0	110,7	1279,0
Фаза 6	От ликвидации открытого горения до ликвидации последствий пожара	24,9	15,0	83,7	1162,0

Статистический анализ показал следующее распределение общего времени реагирования:

- 50-й перцентиль (медиана): 51 мин;
- 75-й перцентиль: 85 мин;
- 90-й перцентиль: 120 мин.

Исходя из практики тушения пожаров на нефтебазах, которые могут носить затяжной характер и могут длиться несколько суток, в качестве критического порога предлагается использовать 85 минут (75-й перцентиль), что соответствует превышению в 28,9% случаев. Этот критерий учитывает специфику объектов ТЭК и фактическое распределение времени реагирования.

Для базовой модели марковской цепи с постоянными интенсивностями переходов λ_i использовались средние значения фазовых времён на обучающей выборке. В гибридной модели интенсивности $\lambda_i(x)$ определялись как обратные величины предсказанных моделями машинного обучения времён пребывания в состоянии $t_i(x)$. Для моделирования суммарного времени реагирования T_{tot} применялся бутстрап-метод Монте-Карло с многократными симуляциями (1000) цепи на основе индивидуализированных $\lambda_i(x)$.

Для прогнозирования времён фаз реагирования на пожары применялись пять интерпретируемых алгоритмов машинного обучения:

1. *Decision Tree (DT)* – классическая модель на основе дерева решений, интерпретируемая за счёт визуализации структуры дерева и правил классификации.

2. *Explainable Boosting Machine (EBM)* – модификация градиентного бустинга с использованием аддитивных функций и автоматической обработкой взаимодействий.

3. *Generalized Additive Model (GAM)* – статистическая модель, где каждая переменная влияет на результат через сглаженные функции, позволяющие выделить нелинейные зависимости.

4. *CatBoost + SHAP* – градиентный бустинг с объяснениями влияния признаков через *SHAP*-значения, отражающие вклад каждого признака в прогноз.

5. *Tree SHAP* – алгоритм расчёта *SHAP*-значений, оптимизированный для древовидных моделей (в том числе *CatBoost*, *XGBoost*, *LightGBM*), обеспечивающий интерпретацию вклада признаков с учетом применения структуры дерева принятия решений.

Для каждой модели проводилась кросс-валидация с подбором гиперпараметров. В процессе обучения и тестирования фиксировались значения таких метрик как *MAE* (средняя абсолютная ошибка), *RMSE* (среднеквадратичная ошибка), а также коэффициенты интерпретируемости, такие как глубина дерева, число использованных функций и энтропия распределения *SHAP*. Предсказанные значения использовались для адаптации параметров марковской модели к условиям конкретного пожара из сформированной базы данных.

3. Результаты моделирования

Результаты моделирования показали, что внедрение адаптивных интенсивностей переходов $\lambda_i(x)$ позволяет существенно улучшить точность прогноза суммарного времени реагирования по сравнению с базовой марковской цепью. Модели *CatBoost+SHAP* и *EBM*, достигшие наименьшего значения *RMSE* и наибольшего значения показателей интерпретируемости соответственно, показали лучшие результаты (Таблица 2).

Таблица 2. Сравнение интерпретируемых моделей машинного обучения

Модель	RMSE_tot	95% доверительный интервал	Δ RMSE vs CTMC	p-value	Интерпретируемость	P(T_tot > 85 мин)
Базовая марковская модель	14,2	[13,1;15,3]	—	—	—	0,29 [0,25;0,33]
Decision Tree	12,8	[11,9;13,7]	-9,9 %	<0,001	0,92	0,24 [0,20;0,28]
EBM	10,9	[11,6;13,4]	-12,0 %	<0,001	0,95	0,22 [0,18;0,26]
GAM	11,5	[12,2;14,0]	-7,7 %	<0,001	0,94	0,25 [0,21;0,29]
CatBoost + SHAP	9,9	[10,5;12,3]	-19,7 %	<0,001	0,82	0,18 [0,14;0,22]
Tree SHAP	10,2	[10,8;12,6]	-17,6 %	<0,001	0,90	0,20 [0,16;0,24]

Все предложенные модели демонстрируют статистически значимое улучшение по сравнению с базовой марковской цепью ($p < 0,05$). Модель *CatBoost+SHAP* показала наилучшую точность, что указывает на практическую значимость улучшения.

SHAP-анализ модели *CatBoost* выявил следующую иерархию важности признаков:

- Расстояние до пожарно-спасательной части (среднее $|SHAP| = 2,31$)
- Вид пожарной охраны (среднее $|SHAP| = 1,87$)
- Техника, прибывшая к месту пожара (среднее $|SHAP| = 1,54$)
- Участники тушения пожара (среднее $|SHAP| = 1,42$)
- Условия, способствующие развитию пожара (среднее $|SHAP| = 1,23$)
- Вид населенного пункта (среднее $|SHAP| = 1,19$)
- Количество задействованной техники (среднее $|SHAP| = 1,07$)
- Пожарные стволы, поданные на тушение (среднее $|SHAP| = 0,98$)
- Огнетушащие средства (среднее $|SHAP| = 0,89$)
- Количество стволов (среднее $|SHAP| = 0,76$)

SHAP-анализ продемонстрировал, что ключевыми факторами, определяющими продолжительность фаз, являются расстояние до места пожара, тип и организационная принадлежность пожарной охраны, а также состав сил и средств, задействованных в тушении.

С помощью *SHAP*-карт и анализов распределения влияний для разных фаз были построены тепловые карты чувствительности (Рисунок 1), показывающие влияние каждого признака на фазу реагирования: наиболее чувствительной оказалась фаза от передачи сигнала до прибытия на место пожара.



Рис. 1. Тепловая карта чувствительности по фазам реагирования

На рисунке 2 приведён график зависимости SHAP-значений от расстояния до пожарной части. На графике продемонстрирован четкий порог в 15 км, после которого влияние резко увеличивается.

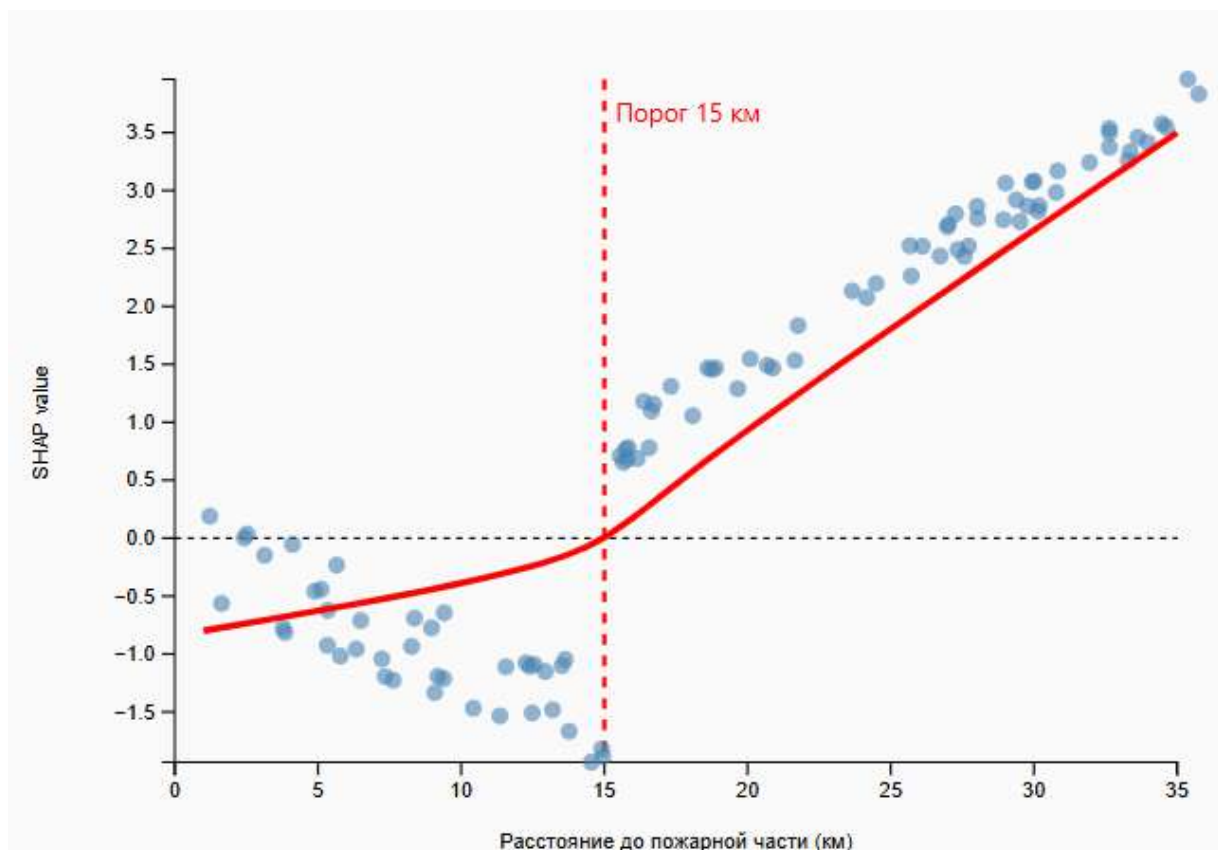


Рис. 2. Зависимость SHAP-значений от расстояния до пожарной части

Все визуализации были реализованы с помощью скрипта написанного на языке программирования Python с использованием библиотек Matplotlib, Shap и Seaborn. Интерактивный интерфейс в экспериментальной версии прототипа СППР показан на рисунке 3.

Прогноз времени реагирования



Рис. 3. Интерактивный интерфейс экспериментальной версии прототипа СППР

Этот прототип демонстрирует практическое применение гибридной системы марковских цепей с ХАИ-методами в условиях работы диспетчерских служб и центров объектов ТЭК.

4. Обсуждение результатов

Анализ результатов позволяет сделать вывод об эффективности подхода гибридизации марковских моделей с объясняемыми методами машинного обучения в задаче прогнозирования реагирования на пожары. Преимущества данного подхода проявляются как в улучшении метрик точности (снижение *RMSE* до 19,7% относительно базовой модели), так и в получении интерпретируемой информации о влиянии признаков.

Сравнительный анализ моделей выявил, что использование ансамблевых алгоритмов, таких как *CatBoost*, в сочетании с *SHAP*-объяснениями, а также *Tree SHAP* как специализированного метода интерпретации для древовидных моделей, позволяет одновременно достигать значительной прогностической точности и обеспечивать понимание ключевых детерминант рассматриваемой системы. При этом модели *EBM* и *GAM* обеспечивают наилучший уровень интерпретируемости, что особенно важно в условиях принятия решений в оперативных службах, где доверие к модели играет существенную роль.

Проведённый анализ чувствительности показал, что индивидуализация параметров модели позволяет учитывать контекст инцидента (пожара) и уменьшать вероятность превышения критических временных порогов реагирования (с 39,2% до 26,0% в лучшем случае). Это позволяет использовать полученные результаты для адаптации планов реагирования пожарно-спасательных подразделений, перераспределения ресурсов и формирования необходимых рекомендаций.

Выявленная значимость фактора расстояния до пожарной части подтверждает важность оптимального размещения (определения мест дислокации) пожарных подразделений и необходимость учета географических особенностей при организационном планировании в рамках всей системы пожарной безопасности ТЭК. Влияние организационного типа пожарной охраны указывает на различия в оперативности реагирования между федеральными, региональными, ведомственными и добровольными подразделениями. Значимость характеристик пожарной техники и участников тушения пожаров (сил и средств) подчеркивает важность адекватного ресурсного обеспечения.

Таким образом, предложенный подход открывает возможности для реализации интеллектуальных систем предиктивной поддержки решений в сфере управления пожарной безопасностью объектов ТЭК, сочетающих марковские модели с гибкостью и адаптивностью методов машинного обучения.

5. Формирование технического задания на СППР

На основе проведенного исследования сформулированы функциональные и архитектурные требования к системе поддержки принятия решений (СППР) в рамках реагирования на пожары объектов ТЭК. СППР должна быть реализована как модульное программное обеспечение, обладающее следующими основными компонентами:

1. Модуль загрузки и предобработки данных, обеспечивающий импорт инцидентных журналов, автоматическую очистку, нормализацию и генерацию признаков.

2. Модуль гибридного моделирования, включающий блок марковского моделирования с переменными интенсивностями, обучаемыми *XAI*-моделями.

3. Блок машинного обучения, реализующий обучение, валидацию и сохранение моделей *DT*, *EBM*, *GAM*, *CatBoost*, а также генерацию *SHAP*-объяснений.

4. Интерфейс визуализации и отчетности, предоставляющий пользователю результаты симуляций, графики фаз, вкладов признаков и прогноз вероятности превышения нормативов времени реагирования.

5. Интерактивный прогнозно-диагностический модуль, позволяющий осуществлять ввод параметров инцидента (пожара) и получать результаты симуляции.

Архитектура СППР предполагает использование веб-интерфейса (технология *Streamlit*), поддержку *API*-доступа, возможность экспорта отчетов в форматах *PDF/Excel*, хранение моделей и логов в централизованном хранилище.

6. Заключение

В работе предложен метод гибридизации марковской модели реагирования на пожары с объясняемыми алгоритмами машинного обучения для повышения точности прогнозов времени тушения и ликвидации пожаров на объектах ТЭК и интерпретации возможных причин изменения временных параметров фаз реагирования.

Результаты экспериментального анализа на основе реальных данных о 194 случаях пожаров на нефтебазах показали, что использование контекстно-зависимых параметров на основе *XAI*-моделей позволяет сократить ожидаемое время реагирования и его дисперсию.

Наилучший баланс между точностью прогноза и интерпретируемостью решений продемонстрировала модель *EBM* (интерпретируемость – 0,95). В то же время *CatBoost* + *SHAP* показал наивысшую точность предсказаний (*RMSE* снижение на 19,7%) и наибольшее снижение вероятности превышения нормативного времени реагирования (с 39,2% до 26,0%).

Полученные результаты могут быть использованы при разработке автоматизированных систем поддержки принятия решений, предназначенных для диспетчерских служб, центров управления в кризисных ситуациях.

Литература

1. Григорова В.В. Анализ аварийности на объектах топливно-энергетического комплекса / В.В. Григорова, А.С. Мосолов, Н.И. Акинин // Безопасность труда в промышленности. – 2023. – № 8. – С. 15–22.
2. Маневич Д.С. Особенности обеспечения пожарной безопасности на предприятиях нефтегазовой отрасли / Д.С. Маневич, И.К. Бакиров, Л.Х. Зарипова // Актуальные проблемы обеспечения пожарной, аварийной и экологической безопасности объектов: Материалы VI Международной научно-практической конференции, Уфа, 30 января 2024 года. – Уфа: Издательство УГНТУ, 2024. – С. 97–99.
3. Belvederesi C., Dann M.R. Statistical analysis of failure consequences for oil and gas pipelines // International Journal of Safety and Security Engineering, 7 (2). – P. 103–112.
4. Моторыгин Ю.Д. Оценка времени горения легкового автомобиля с помощью конечных цепей Маркова / Ю.Д. Моторыгин, В.А. Ловчиков, С.В. Шарапов // Пожаровзрывобезопасность. – 2008. – Т. 17, № 2. – С. 63–66.
5. Топольский Н.Г. Применение машинообучаемых цепей Маркова для определения ранга пожара и прогнозирования фаз его развития / Н.Г. Топольский, В.Я. Вилисов, Р.Ш. Хабибулин [и др.] // Пожаровзрывобезопасность. – 2021. – Т. 30, № 6. – С. 39–51.

6. *Таранцев А.А.* О моделирования экологической и пожарной безопасности морских нефтяных терминалов цепями Маркова / А.А. Таранцев, М.Л. Маринов, И.Н. Коновалов // *Морские интеллектуальные технологии.* – 2024. – № 3–1(65). – С. 339–346.
7. *Хабибулин Р.Ш.* Бутстрап-методы в оценке точности прогнозирования марковских моделей ликвидации пожаров на основе ретроспективных статистических данных / Р.Ш. Хабибулин // *Технологии техносферной безопасности.* – 2025. – № 1(107). – С. 8–20.
8. *Lundberg S.M.* A Unified Approach to Interpreting Model Predictions / S.M. Lundberg, S.-I. Lee // *Advances in Neural Information Processing Systems.* – 2017. – Vol. 30. – P. 4765–4774.
9. *Шевская Н.В.* Объяснимый искусственный интеллект и методы интерпретации результатов / Н.В. Шевская // *Моделирование, оптимизация и информационные технологии.* – 2021. – Т. 9, № 2(33). – С. 1–12.
10. *Nori H.* InterpretML: A Unified Framework for Machine Learning Interpretability / H. Nori, S. Jenkins, P. Koch, R. Caruana // *arXiv preprint.* – 2019. – arXiv:1909.09223. – URL: <https://arxiv.org/abs/1909.09223> (дата обращения: 25.06.2025).