УДК 004.896:678

КОНЦЕПЦИЯ АГЕНТА-ХИМИКА В МУЛЬТИАГЕНТНЫХ СИСТЕМАХ ПРОЕКТИРОВАНИЯ ПОЛИМЕРНЫХ МАТЕРИАЛОВ: ЗНАНИЕВАЯ ВАЛИДАЦИЯ В УСЛОВИЯХ ДЕФИЦИТА ДАННЫХ

Хисамиев И. Л., Богданов К. А.

Казанский национальный исследовательский технический университет имени А. Н. Туполева – КАИ

г. Казань, Российская Федерация

Применение методов машинного обучения (МО) для обратного проектирования (reverse engineering) полимерных материалов сталкивается с системным ограничением: статистические модели, обученные на выборках с недостаточным количеством данных, имеют закрытые механизмы принятия решений, которые скрыты от пользователя и часто генерируют математически корректные, но физически нереализуемые рецептуры [1, 2]. В условиях дефицита экспериментальных данных классические алгоритмы оптимизации способны предлагать составы с нарушенной термодинамической совместимостью компонентов или концентрациями, выходящими за пределы фазовых диаграмм.

Современные генеративные языковые модели, применяемые в области химии , сталкиваются с проблемой генерации ошибочных соединений, хоть и корректных с математической точки зрения. Проблему предлагается решить путем использования каркаса мультиагентной системы, в котором центральное место отводится агенту-химику - модулю валидации информации, выступающему в роли физического ограничителя (guardrail) для генеративных моделей.

В отличие от статистического ядра, которое использует для формирования результата - вероятностные распределения, в свою очередь, агент-химик формирует результат на основе базы знаний, которая формализирует предметную область. Архитектура базы знаний Агента-химика включает три уровня:

- Термодинамический и кинетический уровень: правила совместимости эластомеров с вулканизующими агентами, стерические ограничения для макромолекул, пределы растворимости наполнителей в полимерной матрице.

- Технологический уровень: допустимые интервалы концентраций компонентов в зависимости от типа оборудования и метода смешения, ограничения по вязкости и времени переработки.

- Нормативный уровень: жесткие рамки, заданные отраслевыми стандартами и ГОСТами (например, предельно допустимые массовые доли серы или ускорителей для конкретных групп каучуков).

Подход использует механизм двусторонней связи агента-химика с математическим агентом-оптимизатором. При генерации вероятной рецептуры статистической моделью, агент-химик не просто отбрасывает физически невозможные варианты, но и формирует градиент коррекции. Если предложенный состав нарушает правило совместимости, агент-химик возвращает в контур оптимизации не факт ошибки, а уточненное математическое ограничение (штрафную функцию), сужающее пространство поиска в физически обоснованную область. Это позволяет избежать «холостых» итераций и значительно ускоряет сходимость алгоритма даже на малых данных [3].

Дополнительно в систему интегрируется LLM-агент (большая языковая модель), который использует структурированные знания агента-химика для взаимодействия с инженером-технологом [4]. LLM-агент транслирует неформализованные требования (например, «необходима высокая стойкость к озонному старению при сохранении эластичности») в формальные целевые функции, а затем интерпретирует вердикт агента-химика, объясняя технологу, почему та или иная математически оптимальная рецептура была отклонена с точки зрения химии полимеров.

Агент-химик компенсирует недостаток обучающих данных при интеграции экспертных знаний в контур системы. Переход от нацеленного использования большого обучающего набора данных к методу, который основывается на использовании собственной экспертной базы знаний устраняет

проблему неизвестности в работе модели искусственного интеллекта и может служить базой для разработки интеллектуальных систем, помогающих принимать решения при изучении и разработки в области химии полимеров .

Список использованной литературы

1. Schmidt J. Recent advances and applications of machine learning in solid-state materials science / J. Schmidt, M. R. G. Marques, S. Botti, M. A. L. Marques // npj Computational Materials. – 2019. – Vol. 5. – Art. 83. – URL: https://doi.org/10.1038/s41524-019-0221-0 (дата обращения: 29.09.2026).

2. Wu Z. MoleculeNet: a benchmark for molecular machine learning / Z. Wu, B. Ramsundar, E. N. Feinberg [et al.] // Chemical Science. – 2018. – Vol. 9. – P. 513–530. – URL: https://doi.org/10.1039/C7SC02664A (дата обращения: 29.09.2026).

3. Karpatne A. Theory-guided data science: a new paradigm for scientific discovery from data / A. Karpatne, G. Atluri, J. H. Faghmous [et al.] // IEEE Transactions on Knowledge and Data Engineering. – 2017. – Vol. 29, № 10. – P. 2318–2331. – DOI: 10.1109/TKDE.2017.2720168.

4. Bran A. M. Augmenting large language models with chemistry tools / A. M. Bran, S. Cox, O. Schilter [et al.] // Nature Machine Intelligence. – 2024. – Vol. 6, № 5. – P. 525–535. – DOI: 10.1038/s42256-024-00832-8.