

АКАДЕМИЯ НАУК РЕСПУБЛИКИ УЗБЕКИСТАН

---

ИНСТИТУТ ХИМИИ И ФИЗИКИ ПОЛИМЕРОВ

*На правах рукописи*

УДК 541:536.63  
+541.64:532.73

**Нургалиев Ильнар Накипович**

**ТОПОЛОГИЧЕСКОЕ (QSPR) МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ  
СВОЙСТВ ПОЛИМЕРНЫХ СИСТЕМ ПОСРЕДСТВОМ  
ОПТИМАЛЬНЫХ ДЕСКРИПТОРОВ**

01.04.19 - Физика полимеров

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата физико-математических наук

Ташкент-2010

Работа выполнена в Институте химии и физики полимеров Академии наук Республики Узбекистан

<b>Научный руководитель</b>	кандидат физико-математических наук Торопов Андрей Андреевич
<b>Научный консультант</b>	доктор химических наук, профессор, академик АН РУз Рашидова Сайёра Шарафовна
<b>Официальные оппоненты:</b>	доктор физико-математических наук Тураева Нигора Назаровна  кандидат физико-математических наук, доцент Сидоренко Олег Евгеньевич
<b>Ведущая организация</b>	Отдел теплофизики Академии наук Республики Узбекистан

Защита состоится " 18 " июня 2010 года в 14<sup>00</sup> часов на заседании специализированного совета ДК 015.24.01 при Институте химии и физики полимеров АН РУз по адресу: 100128, г. Ташкент, ул. А. Кадыри 7б, телефон: (998-71) 241-85-94, факс: (998-71) 244-26-61, e-mail: carbon@uzsci.net

С диссертацией можно ознакомиться в библиотеке Института химии и физики полимеров АН РУз

Автореферат разослан " 15 " мая 2010 г.

Ученый секретарь  
специализированного совета,  
кандидат химических наук

Усманова М.М.

## ОБЩАЯ ХАРАКТЕРИСТИКА ДИССЕРТАЦИИ

**Актуальность работы.** В настоящее время наблюдается широкое и плодотворное внедрение информационных и компьютерных технологий в различные сферы наук, в том числе и в материаловедение. Такие направления применения теории графов, как моделирование структуры, синтеза, физико-химических и биологических свойств соединений, лекарственных средств, наноструктур, наночастиц, а также различных химических процессов и реакторов уже завоевали прочные позиции в современной науке и обеспечили повышение эффективности различного рода исследований. Поэтому роль различных теоретических моделей и вычислительных методов в конструировании и оценке свойств полимеров должна заключаться, в частности, в поиске связи «структура-свойство», реализованной в соответствующих принципах, физико-математических и компьютерных моделях.

Наиболее перспективным в этом отношении является подход, основанный на методологии QSPR (Quantitative Structure-Property Relationships - количественные соотношения «структура – свойство»). Основным направлением данных исследований является: количественный анализ связи «структура-свойство», разработка принципиально новых подходов и алгоритмов в области QSPR, создание новых дескрипторов; создание оригинальных компьютерных программ. В последние десятилетия для описания структуры молекул было предложено множество различных дескрипторов, которые включают индексы, характеризующие топологию, электронную структуру, энергию, форму молекул и т.д., а внедрение персональных компьютеров в практику теоретических и прикладных исследований, вызвало необходимость создания специализированных программных продуктов для решения задач. Однако работ по установлению взаимосвязи «структура-свойство» для высокомолекулярных соединений мало. В связи с этим разработка новых дескрипторов, описывающих одно- и двухкомпонентные полимерные системы является актуальной задачей в данном направлении исследований.

**Степень изученности проблемы.** В настоящее время существует три основных теоретических метода количественной оценки свойств полимеров, которая осуществляется на основе их строения. Первый подход, развитый Ван Кревеленом, основан на идее «групповых вкладов». Данный подход чисто эмпирический, как отмечает сам автор, позволяет с хорошей точностью рассчитывать физические свойства многих линейных полимеров. Второй подход, развитый А.А. Аскадским и Ю.И. Матвеевым, является полуэмпирическим. Согласно этому подходу, уравнения для расчета физических свойств получены на основании представлений физики твердого тела, а калибровка метода осуществляется с помощью физических характеристик полимерных стандартов, свойства которых хорошо изучены. Третий подход, предложенный Дж. Бицерано, основан на так называемых индексах связности, что на практике свелось к поиску различных корреляций

физических свойств с множеством правил, как находить коэффициенты корреляционных зависимостей.

Для нахождения количественных корреляционных зависимостей между структурой и свойствами химических соединений, с использованием третьего подхода, используются различные дескрипторы, в том числе структурные. Исследований же по проблеме моделирования «структура-свойство» на основе топологических дескрипторов применительно к полимерным системам, где прогнозируются, такие физические характеристики, как температура плавления, стеклования, параметр взаимодействия «полимер-растворитель», вязкости полимерных растворов практически нет. Для описания двухкомпонентных полимерных систем нет универсального дескриптора, на основе которого возможно описать структуру полимера и растворителя.

#### **Связь диссертационной работы с тематическими планами НИР.**

Диссертация выполнена в соответствии с тематическими планами научно-исследовательских работ Института химии и физики полимеров ЗФ: «Научные основы формирования наночастиц и наноструктур в полимерных системах», ФПФИ 103-06 “QSPR моделирование параметра Флори-Хаггинса полимерных систем с использованием представления молекулярной структуры в виде "наносегментов"” и ФА-3-Ф-Т-100 “Наноструктуры в полимерах. Пути создания и закономерности их проявления в специальных свойствах материалов”.

**Цель исследования** - разработка топологического QSPR моделирования для прогнозирования физических свойств полимеров на основе анализа информации об их структуре с использованием оптимальных дескрипторов, основанных на статистических закономерностях как теоретической основы для определения свойств и целенаправленного синтеза полимеров.

#### **Задачи исследования:**

-разработка новой методики к моделированию взаимосвязи «структура-свойство» на основе оптимальных дескрипторов, рассчитываемых при помощи оптимизации корреляционных весов локальных инвариантов графов: индексов связности (нулевого, первого, второго и третьего порядков) и кодов ближайшего соседства.

-расчет оптимальных дескрипторов, рассчитываемых при помощи оптимизации корреляционных весов локальных инвариантов графов для прогнозирования физических свойств (температуры стеклования, плавления, параметра Флори-Хаггинса, нижней критической температуры смешения, характеристической вязкости) исследуемых систем, построение различных прогностических моделей в зависимости от используемых дескрипторов.

-применение новой методики расчетов для моделирования параметра Флори-Хаггинса, нижней критической температуры смешения и характеристической вязкости двухкомпонентных систем полимер-растворитель.

-сравнительный анализ моделей физических свойств, полученных с использованием оптимальных дескрипторов, рассчитываемых на основе локальных инвариантов молекулярных графов, с точки зрения их

статистических характеристик, а также влияния того или иного локального инварианта на прогнозирующую способность моделей.

**Объект и предмет исследования.** Объектами исследования являются синтетические полимеры, а также низкомолекулярные соединения. Предметом исследования было установление количественной взаимосвязи «структура-свойство» в выбранном ряду соединений и систем.

**Методы исследования.** Для построения зависимостей «структура-свойство» была опробована оптимизация корреляционных весов локальных инвариантов графов. По сути, данный подход является гибридом аддитивной схемы и регрессионного анализа на основе структурных дескрипторов.

**Основные положения, выносимые на защиту:**

-разработана новая методика к моделированию «структура-свойство» на основе оптимальных дескрипторов, кодирующих структуру повторяющегося звена полимера и рассчитываемых при помощи оптимизации корреляционных весов локальных инвариантов графов, индексов связности (нулевого, первого, второго и третьего порядков) и кодов ближайшего соседства применительно к одно- и двухкомпонентным системам;

-использование оптимальных дескрипторов для математического моделирования температуры стеклования, плавления, параметра Флори-Хаггинса, нижней критической температуры смешения, характеристической вязкости для одно- и двухкомпонентных систем, образованных различного класса синтетических полимеров;

-математические модели температуры стеклования, плавления, параметра Флори-Хаггинса, нижней критической температуры смешения, характеристической вязкости для одно- и двухкомпонентных систем.

**Научная новизна** состоит в разработке новой методики проведения изысканий по проблеме «структура-свойство», позволяющей проводить этот поиск не только для индивидуальных веществ, но и для смесей полимер-растворитель. Впервые использованы оптимальные дескрипторы для установления корреляций «структура-свойство».

Оценено использование данных дескрипторов, основанных на оптимизации корреляционных весов локальных инвариантов графов для прогнозирования физических свойств одно- и двухкомпонентных полимерных систем.

Впервые на основе оптимальных дескрипторов были получены модели температур стеклования полиариленоксидов, температур плавления олигофениленов, коэффициентов распределения октанол-вода органических веществ, а также параметров Флори-Хаггинса, нижних критических температур смешения и характеристической вязкости двухкомпонентных систем полимер-растворитель.

**Научная и практическая значимость результатов исследования** состоит в том, что предложен новый метод расчетов физических характеристик полимеров, представляющий практический интерес для прогнозирования исследуемых свойств без привлечения экспериментальных методов. Полученные в ходе расчетов модели различных характеристик

полимерных систем позволят предсказывать указанные характеристики для широкого ряда одно- и двухкомпонентных систем полимер-растворитель, расчетным путем, что делает необязательным проведение сложных и дорогостоящих, а иногда просто технически трудоемких экспериментов.

Методика исследования, полученные модели «структура–свойство» и алгоритмы прогнозирования, могут использоваться:

- при анализе наборов полимерных структур с целью отбора перспективных из них, для технологической реализации практических задач в области физики и химии высокомолекулярных соединений;
- в научной работе в области материаловедения при исследовании зависимости «структура–свойство» для высокомолекулярных одно- и двухкомпонентных систем;
- в образовательном процессе материаловедения в высших учебных заведениях.

#### **Реализация результатов:**

- разработанный QSPR метод с использованием оптимальных дескрипторов, нашел применение в выполнении научных фундаментальных тем: ЗФ-1 «Научные основы формирования наночастиц и наноструктур в полимерных системах» (2003-2007 гг.); ФПФИ 103-06 «QSPR моделирование параметра Флори-Хаггинса полимерных систем с использованием представления молекулярной структуры в виде "наносегментов"» (2006-2007 гг.); ФА-3 Ф-Т-100 «Наноструктуры в полимерах. Пути создания и закономерности их проявления в специальных свойствах материалов».

**Апробация работы.** Материалы диссертации представлены и обсуждены на международных и республиканских конференциях: Республиканской конференции "Полимеры-2002", Ташкент, 2002 г.; Республиканской конференции "Актуальные проблемы науки о полимерах", Ташкент, 2004 г.; Республиканской конференции "Химия и физика высокомолекулярных соединений", Ташкент, 2002 г.; International Conference "Nanochemistry: New approaches to creation of polymeric systems with specific properties", Tashkent, 2003 г.; Международной конференции "Биологически-активные полимеры: синтез, свойства, применение", Ташкент, 2003 г.; Международной конференции "Наноматериалы в химии и биологии", Киев, 2004 г.; 8-ая Международной Пушинская школе-конференции молодых ученых "Биология-наука XXI века", 2004 г.; Международной конференции «Теоретические аспекты формирования полимерных наноструктур», Ташкент, 2004 г.; Международной конференции «Актуальные проблемы химии и физики полимеров», Ташкент, 2006 г.; Четвертой Всероссийской Каргинской конференции, посвященной 100-летию со дня рождения академика В.А. Каргина «Наука о полимерах 21-му веку», Москва, МГУ, 2007 г.; Республиканская научно-техническая конференция «Получение нанокмозитов, их структура и свойство», Ташкент, 2007 г.

**Опубликованность результатов.** По материалам диссертации опубликовано статей - 7, тезисов докладов -7.

**Структура и объём диссертации.** Диссертационная работа состоит из введения, 3 глав, включая обзор литературы, выводов, библиографического списка, включающего 99 источников. Материалы диссертационной работы изложены на 120 страницах текста и иллюстрированы 28 таблицами и 19 рисунками.

## ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

**Во введении** сформулированы возможности теоретических и компьютерных методов для предсказания свойств веществ на основе их строения. Обоснована актуальность работы, определены цель, задачи и практическая значимость результатов работы.

**В первой главе** (обзоре литературы) проанализированы основные концепции выявления связи между строением веществ и их физико-химическими свойствами, описано применение теории графов для кодирования молекулярных структур, основные понятия теории графов, подходы для установления теоретических зависимостей «структура–свойство/активность», особенности дескрипторного описания молекулярной структуры химических соединений; проанализированы известные зависимости, связывающие строение веществ и их свойств; обсуждены современные программные средства моделирования QSPR, QSAR и системы прогноза свойств соединений.

**Во второй главе** описан метод, и основные этапы компьютерных экспериментов поставленных в работе для получения QSPR моделей физико-химических свойств полимерных систем. Основное внимание уделено основным этапам компьютерного моделирования, а также алгоритма расчетов. Представлены формулы по расчету оптимальных дескрипторов, основанных на оптимизации корреляционных весов локальных инвариантов графов.

**В третьей главе** представлено использование оптимальных дескрипторов, основанных на локальных инвариантах графов для получения QSPR моделей температур стеклования полиариленоксидов, температур плавления олигофениленов, являющихся олигомерами. Получена модель параметра гидрофобности для низкомолекулярных растворителей. Рассмотрено использование дескрипторов, учитывающих информацию о строении, как мономерного звена полимера, так и растворителя, также апробирован дескриптор системы мономерных звеньев. На основе оптимальных дескрипторов получены модели параметра Флори-Хаггинса, нижней критической температуры смешения и характеристической вязкости двухкомпонентных систем полимер-растворитель.

---

*\*Автор выражает глубокую признательность академику Рашидовой С.Ш. и к.ф.-м.н. Торопову А.А. за научные консультации и неоценимую помощь, оказанные при выполнении диссертации.*

## ОБСУЖДЕНИЕ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

**Количественные соотношения «структура-свойство».** Неформализованные задачи физики полимеров, решению которых посвящена данная работа, являются практическими приложениями фундаментальной задачи современной полимерной науки – установления зависимостей «структура-свойство».

Для построения зависимостей «структура-свойство» была опробована оптимизация корреляционных весов локальных инвариантов графов. По сути, данный подход является гибридом аддитивной схемы и регрессионного анализа на основе структурных дескрипторов. Корреляционные веса – это альтернатива аддитивным вкладам молекулярных фрагментов, но под моделью подразумевается не сумма инкрементов различных молекулярных фрагментов и(или) функциональных групп (что характерно для аддитивных схем), а корреляционная связь величины представляющего интерес физико-химического параметра с дескриптором (что характерно для моделирования посредством регрессионного анализа). Для проведения компьютерных экспериментов были осуществлены следующие этапы:

1. Ввод и хранение молекулярных структур. Программа-редактор позволяет конструировать формулы, включающие: мономерные звенья полимеров; одинарные; двойные; тройные связи; ароматические; гетероциклические; алифатические циклы с любым числом атомов в цикле.

2. Ввод и хранение выборки соединений. Далее весь набор структур был поделен на два поднабора: тренировочный и тестовый. Разделение структур хаотично, но так, чтобы все инварианты, которые присутствуют в структурах обучающей выборки, присутствовали в контрольной.

3. Выбор дескриптора как функции локальных инвариантов (ЛИ) графов. В качестве дескрипторов были использованы индексы связности, рассчитываемые на основе оптимизации корреляционных весов локальных инвариантов графов. В данной работе использованы два локальных инварианта, расширенная связность (морганова степень вершины), и так называемые коды ближайшего соседства. Расширенная связность ("обычная" степень) вычисляется по формуле

$${}^0EC_i = \sum_{j'}^n a_{ij'} \quad (1)$$

где  $j'$ -номер вершины молекулярного графа, представляющего мономерное звено полимера, не являющейся образом водорода или электронной пары ( $ee$ );  $a_{ij'}$ -элемент матрицы смежности  $A(G)$ , на основе которой вычисляется локальный инвариант;  $n$ -число вершин в графе.

В настоящей работе использовалась расширенная связность первого и второго порядков, вычисляемая по рекуррентному соотношению

$${}^kEC_i = \sum_{(i,j)}^n {}^{k-1}EC_j \quad (2)$$

где  $k=1,2$ ;  $(i,j)$ -обозначение ребра молекулярного графа.

Второй тип локальных инвариантов, рассмотренных в данной работе - так называемые коды ближайшего соседства (*КБС*). *КБС* для *i*-ой вершины - определяется как:

$$КБС_i = 100 \times (N_T) + 10 \times (N_C) + N_H \quad (3)$$

где *i*-число вершин в графе,  $N_T$ -общее число атомов, соседних с *i*-ой вершиной графа,  $N_C$ -число атомов углерода, соединенных с атомом, представляющим *i*-ую вершину графа, и  $N_H$ -число атомов водорода, соединенных с атомом, представляющим *i*-ую вершину графа.

Для получения корреляции между строением структур и величинами свойства использовали дескрипторы вида

$$DMS({}^k EC) = D_m(ЛИ) + D_s(ЛИ) \quad (4)$$

где  $D_m(ЛИ)$ ,  $D_s(ЛИ)$ -дескрипторы, вычисляемые для структур мономера и растворителя по формуле

$$D_x(ЛИ) = \sum_{i=1}^n CW(a_i) \times CW(ЛИ_i) \quad (5)$$

где  $CW(a_i)$ -корреляционный вес присутствия атома  $a_i$  (С, Н, О, Cl, ее);  $CW(ЛИ_i)$ -корреляционный вес присутствия расширенной связности *k*-го порядка ( $k=1, 2$ ); *n*-число атомов;  $x=m, s$  (*m*-мономер, *s*-растворитель).

Формулы (4) и (5) применяли для оптимизации корреляционных весов структур мономерных звеньев полимеров. Для оптимизации корреляционных весов локальных инвариантов графов, представляющих систему мономерных звеньев, использован следующий дескриптор

$$DMS_{system}({}^k EC) = DMS({}^k EC) \times \frac{1000}{\sum_{k,j} BL(k, j)} \quad (6)$$

где  $BL(k, j)$  - расстояние между *k* и *j* вершинами (сумма ковалентных радиусов).

4. Посредством оптимизационной процедуры Монте Карло расчет так называемых корреляционных весов локальных инвариантов молекулярных графов. Корреляционные веса подбирались таким образом, что бы обеспечить как можно большую величину коэффициента корреляции между исследуемым свойством и специальными дескрипторами, вычисляемыми по этим корреляционным весам. Данная процедура это оптимизация корреляционных весов локальных инвариантов графов (ОКВЛИ).

5. На основе метода наименьших квадратов определение зависимости вида

$$Свойство = C_0 + C_1 \times Дескриптор, \quad (7)$$

где  $C_0$  и  $C_1$ - коэффициенты линейной регрессии.

6. Оценка прогнозирующей способности модели на структурах контрольной выборки.

7. Анализ данных полученных в результате означенных расчетов.

**Моделирование температуры стеклования полиариленоксидов.** Полиариленоксиды – важный класс конденсационных полимеров, на основе которых получен ряд инженерных термопластов с высокими

эксплуатационными характеристиками. Температура стеклования ( $T_g$ ) является одной из характеристик определяющих отличие данного класса полимеров от полимеров близкого к их строению. Очевидно, что температура стеклования является важной характеристикой, на основе которой можно сделать вывод о проявление того или иного свойства полимера. Поэтому нами была поставлена задача получения модели данной характеристики полиариленоксидов.

В качестве локальных инвариантов в настоящей работе использовалась расширенная связность первого и второго порядков. Структуру молекулярных графов рассматриваемых мономерных звеньев полиариленоксидов из литературы, кодировали посредством линейной номенклатуры. В таблице 1 представлено статистическое качество ОКВЛИ моделей, основанных на дескрипторах  $D(^0EC)$ ,  $D(^1EC)$ ,  $D(^2EC)$  и  $D(^3EC)$ . Из таблицы 1 видно, что лучшая модель температур стеклования получена для дескриптора  $D(^2EC)$ , поскольку модели, основанные на  $D(^0EC)$  и  $D(^1EC)$ , имеют статистическое качество хуже, чем модель, основанная на  $D(^2EC)$ , а при использовании дескриптора  $D(^3EC)$ , происходит «переучивание», когда высокое качество модели для обучающей выборки, сопровождается ухудшением качества модели на контрольной выборке. На рисунке 1 представлена зависимость экспериментальных и рассчитанных величин температур стеклования

Среднее значение величины разницы между экспериментальными и рассчитанными значениями для структур контрольной выборки составляет  $9.475 \text{ } ^\circ\text{C}$ . Согласно литературным данным точность экспериментального определения температуры стеклования составляет  $10\text{-}20 \text{ } ^\circ\text{C}$ . Таким образом, точность модели на контрольной выборке сопоставима с точностью экспериментального определения температуры стеклования.

Таблица 1

Статистическое качество ОКВЛИ моделей температур стеклования при различных порядках степеней вершин Моргана ( $n$  - число структур,  $C_0$ ,  $C_1$  - коэффициенты линейной регрессии,  $r$  - коэффициент корреляции,  $S$  - стандартное отклонение,  $F$  - обозначение  $F$  - отношения Фишера)

Типы дескрипторов	$C_1$	$C_0$	Обучающая выборка, $n=50$			Контрольная выборка, $n=18$		
			$r$	$S, \text{ } ^\circ\text{C}$	$F$	$r$	$S, \text{ } ^\circ\text{C}$	$F$
$D_0(^0EC)$	0.38	97.922	0.6791	31.625	41	0.8236	20.97	34
$D_1(^1EC)$	3.166	136.580	0.8365	23.612	112	0.7926	25.764	27
$D_2(^2EC)$	4.634	144.85	0.9469	13.847	417	0.9288	12.757	100
$D_3(^3EC)$	1.877	130.939	0.967	10.982	691	0.7343	23.594	19

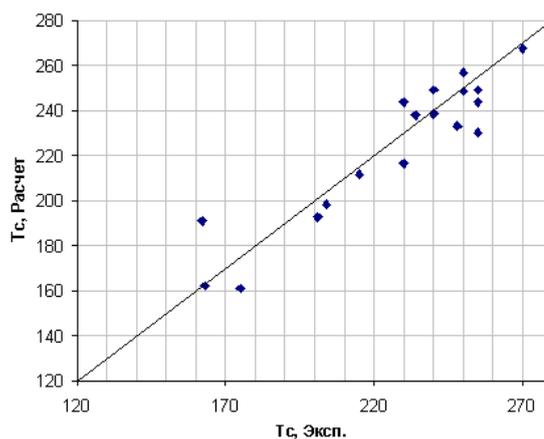
Из расчетов видно, что наиболее применимым локальным инвариантом следует считать расширенную связность второго порядка, так как модель, полученная на ее основе имеет лучшие статистические характеристики относительно других моделей. Таким образом, показано, что оптимизация корреляционных весов локальных инвариантов молекулярных графов, представляющих структуру мономеров полиари-

леноксидов, может быть использована для оценки температур стеклования полиариленоксидов, по которым нет экспериментальных данных по упомянутому параметру. Моделирование температуры стеклования полиариленоксидов производилось на основании информации о химическом строении элементарного звена. Однако физико-химическое свойство материала не всегда определяется только этим параметром. Оно может зависеть также и от других характеристик. Известно, в частности, что на температуру плавления олигомеров оказывает влияние их молекулярная масса. В связи с этим, нами предпринята попытка моделирования температур плавления олигофениленов различной степени полимеризации, близких по строению элементарного звена.

#### **Моделирование температур плавления олигофениленов.**

Температура плавления ( $T_{пл}$ , °C) является физической характеристикой, которая наиболее трудно поддается расчету. Существуют два метода в решении этой задачи, один из них основан на отношении температуры стеклования к температуре плавления. Однако согласно правилу Бимена,  $T_c/T_{пл} \approx 2/3$  и детальный анализ большого количества полимеров разнообразного строения, показал, что это отношение варьируется в широких пределах, хотя для большой группы полимерных систем оно составляет  $\sim 2/3$ . Второй подход основан на рассмотрении повторяющегося звена полимера как набора ангармонических осцилляторов. Однако, несмотря на достаточно хорошую сходимость расчетных и экспериментальных данных следует отметить, что данные методы имеют ограниченное применение.

Компьютерные эксперименты производились на 19 структурах олигофениленов. Данные о химическом строении, степени полимеризации и температурах плавления взяты из литературы. Для расчетов использована расширенная связность нулевого, первого, второго и третьего порядков. В таблице 2 приведены параметры статистического качества моделей, полученных с использованием расширенной связности различного порядка.



**Рис. 1. Экспериментальные и рассчитанные значения  $T_c$  (°C), для структур контрольной выборки**

Таблица 2

Статистическое качество моделей ОКВЛИ температуры плавления ( $^{\circ}\text{C}$ ), основанных на расширенной связности нулевого ( ${}^0\text{EC}$ ), первого ( ${}^1\text{EC}$ ), второго ( ${}^2\text{EC}$ ) и третьего ( ${}^3\text{EC}$ ) порядка ( $r$  — коэффициент корреляции,  $S$  — стандартное отклонение,  $F$  — величина отношения Фишера)

ЛИ	$C_0$	$C_1$	Обучающая выборка			Контрольная выборка		
			$r$	$S, ^{\circ}\text{C}$	$F$	$r$	$S, ^{\circ}\text{C}$	$F$
${}^0\text{EC-1}$	20,179	190,7895	0,8524	63,298	19	0,5526	92,286	4
${}^1\text{EC-1}$	9,3588	192,0217	0,8524	63,309	19	0,5503	88,748	3
${}^2\text{EC-1}$	6,4697	115,8158	0,9651	31,699	95	0,9511	33,832	76
${}^3\text{EC-1}$	1,7649	20,3145	0,9958	9,720	952	0,9884	18,362	297
Обратное направление								
${}^3\text{EC-3}$	1,8962	36,875	0,998	7,662	1752	0,969	26,278	123

Результаты показывают, что с увеличением степени связности наблюдается улучшение статистического качества модели как по обучающей, так и по контрольной выборкам. Наиболее достоверной является модель, основанная на расширенной связности третьего порядка (рис. 2).

Причем высокое статистическое качество сохраняется, если поменять местами обучающую и контрольную выборки (таблица 2, обратное направление). Известно, что величины температур плавления для олигомеров различной молекулярной массы могут быть выражены следующей формулой:

$$1/T_{nl} = 2R/\Delta HP + 1/T_{nl}^* \quad (8)$$

где  $T_{nl}^*$  — температура плавления образца с бесконечно большой молекулярной массой;  $R$  — газовая постоянная;  $\Delta H$  — теплота плавления в расчете на мономерное звено цепи;  $P$  — степень полимеризации.

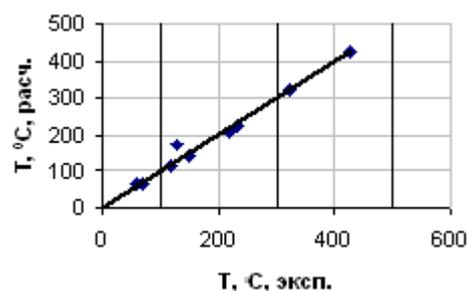


Рис. 2. Экспериментальные и рассчитанные значения  $T_{nl}$  ( $^{\circ}\text{C}$ ), для структур контрольной выборки

Таблица 3

Статистическое качество корреляции экспериментальных и вычисленных по уравнению (8) величин температуры плавления ( $r$  — коэффициент корреляции,  $S$  — стандартное отклонение,  $F$  — величина отношения Фишера)

Олигомеры	$r$	$S$	$F$
<i>n</i> -Олигофенилены	0,9471	45,7	17
<i>m</i> -Олигофенилены	0,9193	23,2	30
<i>o</i> -Олигофенилены	0,8513	61,1	11
Циклические олигофенилены	0,9878	19,1	121

Однако, как видно из сопоставления таблиц 2 и 3, моделирование экспериментальных данных величин температур плавления посредством уравнения (8) имеет статистическое качество существенно хуже, чем модели, полученные посредством ОКВЛИ (таблица 2).

Таким образом, ОКВЛИ может быть использована для предсказания величин температур плавления олигофениленов различной молекулярной массы. Следующим шагом были поставлены компьютерные эксперименты по моделированию коэффициента распределения октанол/вода для ряда органических низкомолекулярных соединений, так как в последствие нам нужно было оценить применимость ОКВЛИ для моделирования свойств двухкомпонентных систем полимер-растворитель. На основании результатов применения ОКВЛИ к низкомолекулярным соединениям можно будет сделать вывод о целесообразности использования бинарного дескриптора, учитывающего информацию, как о мономерном звене полимера, так и о молекулярной структуре низкомолекулярного органического растворителя.

Поэтому данный метод показал плохие результаты в определении параметра гидрофобности. В связи с этим мы применили ОКВЛИ для моделирования данной характеристики, а также для сравнения с другими результатами, приведенными в литературе.

**Моделирование коэффициента распределения октанол-вода посредством корреляционных весов кодов ближайшего соседства.** Гидрофобность (коэффициент распределения октанол-вода) - важный параметр органических соединений с точки зрения химии, химической индустрии, биологии, экологии и других. Использование параметров гидрофобности основано на предположении, что полярный растворитель может служить эталонной системой при исследовании взаимодействия активных молекул с липидной фазой. Данная характеристика наиболее часто используется в исследованиях «структура-активность». Для базы получения QSPR моделей используют мультилинейный регрессионный анализ. Для получения модели использовали литературные данные (Basak S. C., Niemi G. J., Veith G. D. // J. Math. Chem. - 1990. – Vol. 4. - P. 185-205). Для расчетов использовали два локальных инварианта, расширенная связность (морганова степень вершины), и так называемые коды ближайшего соседства.

Результаты трех проведенных компьютерных экспериментов, основанных на ОКВЛИ приведены в таблице 4. Из таблицы 4 видно, что лучшие статистические характеристики полученных моделей - с использованием кодов ближайшего соседства. Статистические характеристики моделей, полученных в литературе, следующие:  $n=137$ ,  $S=0.26$ ,  $r^2=0.97$ . Модели величин параметра гидрофобности, полученные на основе наших расчетов следующие:  $n=138$ ,  $S=0.09$ ,  $r^2=0.996$ . Видно, что модель, полученная в нашей работе, заметно лучше литературных данных. Данное исследование показало, что ОКВЛИ, основанные на КБС, более применимы в установлении корреляций между структурой и коэффициентом распределения октанол-вода.

Результаты ОКВЛИ-моделирования коэффициента распределения октанол-вода, основанные на различных локальных инвариантах графов

Пробы ОКВЛИ	Коэффициенты уравнения (7)		Обучающая выборка $n=69$			Контрольная выборка $n=69$		
	$C_1$	$C_0$	$r$	$S$	$F$	$r$	$S$	$F$
${}^0EC-1$	0.147	-0.545	0.9829	0.273	1906	0.9771	0.301	1410
${}^1EC-1$	0.191	-0.501	0.9893	0.216	3073	0.9811	0.257	1724
${}^2EC-1$	0.152	-0.671	0.9969	0.116	10929	0.9800	0.280	1628
${}^NNC-1$	0.196	-0.023	0.9992	0.059	42061	0.9966	0.113	9847

Большое число экспериментальных данных накоплено для растворов полимеров, поэтому описание и предсказание свойств данных систем является полезным для полимерной науки. К созданию новых дескрипторов побуждает углубление представлений о молекулярной структуре соединений. Как уже упоминалось выше, следующим шагом стала апробация так называемого бинарного дескриптора для установления связи «структура-свойство» для двухкомпонентных систем. Для двухкомпонентных систем полимер-растворитель прогноз физических свойств может быть осуществлен с помощью аддитивной схемы. При этом наличие в полимере или в растворителе сильно полярных групп или водородных связей усложняет характер зависимости «структура-свойство». Другой подход к выявлению связи «структура-свойство» – конструирование корреляционных соотношений, связывающих физико-химические параметры с величинами дескрипторов, вычисляемым по молекулярным графам.

**Моделирование параметра Флори-Хаггинса бинарных смесей полимер-растворитель.** Согласно Гильдебранду, параметр Флори-Хаггинса ( $\chi$ ) может быть рассчитан на основе параметров растворимости смеси полимеров следующим образом:

$$\chi_{FH} = V_{seg} \frac{(\delta_A - \delta_B)^2}{RT} \quad (9)$$

где  $V_{seg}$  - объем одного сегмента полимерной цепи,  $\delta_A$ , и  $\delta_B$  - параметры растворимости компонента  $A$  и  $B$ . Параметры растворимости для малых молекул растворителя могут быть надежно определены экспериментально. Однако есть относительно малое количество данных о полимерах, и известно, что эта величина изменяется значительно. Параметр Флори-Хаггинса играет важную роль в теории растворов полимеров и обычно он рассматривается как параметр взаимодействия между полимером и растворителем. В этой связи количественное определение данной характеристики является одной из задач теории.

С помощью принципа аддитивности, развитые Ван Кревеленом и Федорсом, включающие информацию о групповых вкладах каждой функциональной группы макромолекулы полимера, был проведен расчет  $\chi$  для ряда полимеров. Недавно был развит новый метод, основанный на индексах связности, он позволяет производить прогноз параметров

растворимости полимеров произвольного строения, устраняя одно из главных недостатков в использовании параметра растворимости. Поэтому нашей задачей была апробация «бинарных» дескрипторов для установления связи «структура-свойство» для смесей полимер – растворитель.

Цель данной работы - разработка подхода к моделированию параметра Флори–Хаггинса для двухкомпонентных систем полимер-растворитель, основанного на дескрипторах, вычисляемых для системы мономерных звеньев макромолекулы полимера, вместо дескрипторов мономерного звена, применявшихся ранее. Экспериментальные значения параметра Флори-Хаггинса были взяты из литературы.

Были изучены 60 систем полимер (полипропилен, полиизобутилен и поли(этилэтилен))-растворитель. Из них 30 структур (полимер-растворитель) использовали в качестве тренировочного набора, и 30 в качестве тестового. Для проведения расчетов в качестве локального инварианта была выбрана расширенная связность нулевого порядка. В таблице 5 сопоставлено статистическое качество моделей параметра Флори-Хаггинса, основанных на мономерных звеньях. Из таблицы 5 видно, что модель, основанная на системе мономерных звеньев, заметно лучше.

Таблица 5

Сопоставление статистического качества моделей параметра Флори - Хаггинса, полученных на основе представления структуры полимеров в виде мономерных звеньев и в виде системы мономерных звеньев ( $r$  - коэффициент корреляции,  $S$  - стандартное отклонение,  $F$  - отношение Фишера; число систем полимер-растворитель в каждой выборке равно 30)

Дескриптор	$r$ (обучение/контроль)	$S$ (обучение/контроль)	$F$ (обучение/контроль)
$DMS^{(k)EC}$	0.9818/0.9801	0.168/0.177	748/682
$DMS_{system}^{(k)EC}$	0.9968/0.9993	0.071/0.113	4338/2064

Следующим шагом апробации разработанного дескриптора  $DMS_{system}^{(k)EC}$  было использование более высоких степеней расширенной связности. Для проверки достоверности результатов нами проводилась также замена обучающей выборки контрольной. Статистические характеристики моделей параметра Флори-Хаггинса для бинарных смесей представлены в таблице 6. Проведенные расчеты показали, что при использовании «обычного» дескриптора наилучшие статистические характеристики модели параметра Флори-Хаггинса получены при использовании расширенной связности второго порядка как для структур обучающей, так и контрольной выборок. Расчеты, основанные на дескрипторе  $DMS_{system}^{(k)EC}$ , показали также наилучшие результаты при использовании расширенной связности второго порядка.

Таблица 6

Статистические характеристики модели параметра Флори-Хаггинса двухкомпонентных систем полимер-растворитель, полученные на основе использования дескрипторов  $DMS^{(k)EC}$  и  $DMS_{system}^{(k)EC}$

ЛИ	Обучение, $n=30$			Контроль, $n=30$		
	$r$	$S$	$F$	$R$	$S$	$F$
Дескриптор $DMS^{(k)EC}$						
${}^0EC-1$	0.9819	0.168	753	0.9800	0.182	680
${}^1EC-1$	0.9888	0.142	1232	0.9886	0.144	1206
${}^2EC-1$	0.9985	0.049	9123	0.9971	0.076	4774
Дескриптор $DMS_{system}^{(k)EC}$						
${}^0EC-1$	0.9967	0.071	4250	0.9937	0.099	2202
${}^1EC-1$	0.9968	0.433	4191	0.9935	0.416	2127
${}^2EC-1$	0.9977	0.060	6182	0.9958	0.384	3346
Дескриптор $DMS_{system}^{(k)EC}$ (замена обучающей выборки контрольной)						
${}^0EC-1$	0.9951	0.091	2814	0.9950	0.089	2793
${}^1EC-1$	0.9956	0.084	3167	0.9947	0.097	2612
${}^2EC-1$	0.9984	0.221	8587	0.9946	0.219	2557

Однако для упомянутого дескриптора  $DMS_{system}^{(k)EC}$  статистические характеристики модели с использованием расширенной связности различного порядка различаются незначительно, что характерно для структур обучающей (0.9967-0.9977) и контрольной выборок (0.9937-0.9958). Коэффициенты корреляции при использовании дескриптора  $DMS^{(k)EC}$  имеют заметное различие для обучающей (0.9819-0.9985) и контрольной (0.9800-0.9971) выборок. Замена обучающей выборки на контрольную, с использованием дескриптора  $DMS_{system}^{(k)EC}$ , показало, что статистические характеристики модели также различаются незаметно, хотя наилучшие результаты получены при использовании в качестве ЛИ расширенной связности нулевого порядка.

Таким образом, моделирование, основанное на описании молекулярной структуры полимеров в виде графов, представляющих большой по сравнению с мономерным звеном фрагмент цепи, при использовании расширенной связности нулевого порядка, имеет более высокое качество, как для обучающей, так и для контрольной выборок.

Дескрипторы  $DMS^{(k)EC}$  и  $DMS_{system}^{(k)EC}$  могут быть названы как бимолекулярные дескрипторы. Данные дескрипторы могут быть использованы как инструмент в QSPR исследованиях систем, состоящих из двух и более компонентов. Расчеты показали применимость бинарного дескриптора, учитывающего информацию о структуре мономенного звена и молекулы растворителя, однако многие исследователи разрабатывают или дополняют уже используемые дескрипторы. Однако, на характеристики модели свойства влияет также и количество структур в выборке, поэтому для оценки бинарного дескриптора нами были поставлены компьютерные

эксперименты по оценке модифицированного бинарного дескриптора, а также с увеличенной выборкой, которая состояла из 7 полимеров и ряда растворителей.

**Моделирование параметра Флори-Хаггинса квадратичными корреляционными весами локальных инвариантов графов.** В данной работе были получены QSPR модели параметра Флори-Хаггинса бинарных смесей полимер-растворитель, экспериментальные значения были взяты из литературы. Смесей полимер-растворитель были представлены 7 полимерами (полиэтилен, полиизобутилен, полиметилметакрилат, полиэтиленоксид, полихлоропрен, полиэпихлорогидрин, поливинилацетат) и 15 растворителями.

В таблице 7 представлены коэффициенты корреляции между дескриптором и параметром Флори-Хаггинса ( $\chi_{12}$ ) для обучающей ( $r_{обуч}$ ) и контрольной ( $r_{контр}$ ) выборок. Как видно из таблицы лучший результат был получен с  $DSM(^2EC, 2)$ .

Таблица 7

Коэффициенты корреляции для обучающей и контрольной выборок  $r_{обуч}$  и  $r_{контр}$  между параметром Флори-Хаггинса и значениями  $DSM(^xEC, \mu)$

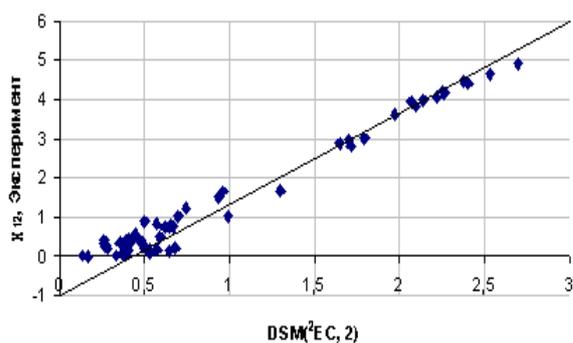
Дескриптор	$r_{обуч}$	$r_{контр}$
$DSM(^0EC, 1)$	0.825	0.783
$DSM(^1EC, 1)$	0.849	0.837
$DSM(^2EC, 1)$	0.863	0.890
$DSM(^3EC, 1)$	0.866	0.864
$DSM(^0EC, 2)$	0.933	0.883
$DSM(^1EC, 2)$	0.940	0.881
$DSM(^2EC, 2)$	0.988	0.985
$DSM(^3EC, 2)$	0.989	0.965

Модель  $\chi_{12}$  рассчитывается как:

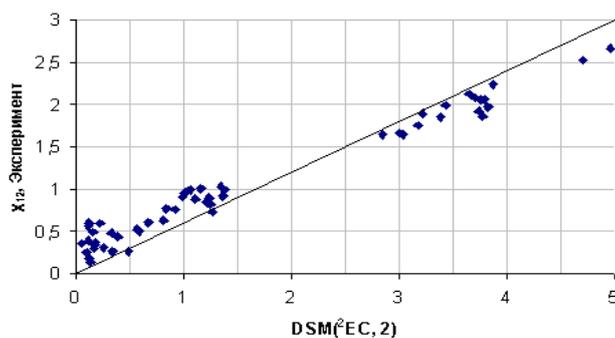
$$\chi_{12} = 0.473 \times DSM(^2EC, 2) + 0.292 \quad (10)$$

$n = 52$ ,  $r = 0.9879$ ,  $S = 0.121$ ,  $F = 2021$  (обучающая выборка)

$n = 53$ ,  $r = 0.9850$ ,  $S = 0.123$ ,  $F = 1666$  (контрольная выборка)



**Рис. 3. Зависимость рассчитанных значений  $DSM^2(EC, 2)$  от экспериментальных значений параметра Флори-Хаггинса для обучающей выборки**



**Рис. 4. Зависимость рассчитанных значений  $DSM^2(EC, 2)$  от экспериментальных значений параметра Флори-Хаггинса для контрольной выборки**

В данной работе проведена оптимизация молекулярного дескриптора  $DSM^2(EC, 2)$  который показал хорошую корреляцию между химической структурой систем полимер-растворитель и параметром Флори-Хаггинса. Хотя данный дескриптор показал надежные результаты в моделировании параметра Флори-Хаггинса, нами были поставлены опыты в моделировании нижней критической температуры смешения (НКТС) ряда систем синтетический полимер-растворитель. Существуют три группы методов для предсказания НКТС. Первая группа использует модели, которые базируются на чисто теоретических расчетах, используя экспериментальные данные систем жидкость – жидкость или пар - жидкость. Эти методы основаны на экспериментальных данных, для установки корреляций свойства с неизвестными параметрами, в результате они ограничиваются прогнозирующей способностью. Другое приближение использует эмпирические уравнения, которые устанавливают корреляцию НКТС с физико-химическими свойствами, такими как плотность, критические свойства и т.д., но неудобно в том, что эти свойства не всегда определяемы. Предложенный нами подход использует линейные модели для предсказания НКТС с использованием оптимальных дескрипторов, значения которых зависят только от структур молекул растворителя и полимера. Однако о применимости данного подхода можно судить по прогностической способности полученной в ходе расчетов модели.

**Моделирование нижней критической температуры смешения систем полимер-растворитель.** Устойчивость системы (фазовое равновесие), как правило, определяется степенью термодинамического сродства компонентов, зависит от их химического состава и строения, внешних условий, в частности, от температуры. Фазовое расслоение в полимерной смеси может наступать и с повышением и с понижением температуры. Возможны случаи, когда с возрастанием температуры взаимная растворимость компонентов сначала увеличивается, а затем снова уменьшается. Например, на диаграммах состояния растворов полимеров наблюдается наряду с верхней критической температурой смешения (ВКТС), представляющей собой максимум на

кривой, еще и нижняя критическая температура смешения (НКТС) – в минимуме.

В области между ВКТС и НКТС компоненты раствора смешиваются полностью. Взаимное расположение критических температур от характера сил (дисперсионные, полярные и т.д.), действующих между полимером и растворителем; растворимость полимера в одном растворителе с увеличением температуры может повышаться, а в другом снижаться. Поэтому, компьютерный прогноз данной характеристики имеет важное значение для определения устойчивости полимерных растворов.

В настоящей работе поставлены компьютерные эксперименты для получения моделей величин НКТС систем полимер-растворитель на основе оптимизации корреляционных весов локальных инвариантов графов. Была использована база данных из 82 систем, состоящих из 8 полимеров и 47 низкомолекулярных растворителей. Экспериментальные данные были взяты из литературы (Melagraki G., Afantitis A., Sarimveis H., Koutentis P. A., Markopoulos J., Igglessi-Markopoulou O. A novel QSPR model for predicting  $\theta$  (lower critical solution temperature) in polymer solutions using molecular descriptors // J. Mol. Model. – 2007. – Vol. 13. – P. 55–64).

Результаты проведенных компьютерных экспериментов представлены в таблицах 8 и 9. По данным таблиц видно, видно, что наилучшие статистические характеристики при использовании дескриптора  $DMS(^kEC)$  для контрольной выборки равно 0,8863, а для дескриптора  $DMS_{system}(^kEC)$  – 0,6544. Использование дескриптора  $DMS(^kEC)$  для построения модели «структура-величина НКТС» показало наиболее лучшую корреляцию предсказанных величин свойств систем контрольной выборки с экспериментальными данными ( $r=0,8863$ ). Однако при сравнении характеристик моделей, с рассчитанными значениями НКТС, взятых из литературы, показали лучшие ( $r=0,9497$ ) результаты по сравнению с данными, полученными в ходе собственных расчетов.

Таблица 8

Статистические характеристики модели НКТС, полученные с использованием дескриптора  $DMS(^kEC)$

ЛИ	Обучающая выборка			Контрольная выборка		
	$r$	$S, ^\circ K$	$F$	$r$	$S, ^\circ K$	$F$
$EC^0$	0.8131	37.205	90	0.7596	37.04	44
$EC^1$	0.9381	22.132	338	0.8465	28.971	81
$EC^2$	0.9742	14.411	858	0.8863	26.03	117
$EC^3$	0.9815	12.248	1206	0.8472	30.904	81

Таблица 9

Статистические характеристики модели НКТС, полученные с использованием дескриптора  $DMS_{system}({}^kEC)$

ЛИ	Обучающая выборка			Контрольная выборка		
	$r$	$S, {}^\circ K$	$F$	$r$	$S, {}^\circ K$	$F$
$EC^0$	0.7876	39.382	75	0.5477	46.317	14
$EC^1$	0.8724	31.235	14	0.4811	51.058	10
$EC^2$	0.9596	17.988	535	0.5885	52.151	17
$EC^3$	0.9652	16.712	627	0.6544	61.391	24

Для получения QSPR моделей в литературном источнике использовали программу ChemSar входящую в программный пакет ChemOffice. В качестве дескрипторов описывающих структуру полимера были отобраны следующие дескрипторы: HOMO energy, ShpC, dipole length (DPLL), radius (Rad). Всеобъемлющее описание молекулярной структуры с помощью какого-либо одного дескриптора практически невозможно. Один дескриптор может учесть преимущественно одну особенность электронной или пространственной структуры, либо один вид межмолекулярных взаимодействий. По этой причине описание молекулярной структуры несколькими дескрипторами и построение моделей «структура-свойство», основанных на нескольких независимых переменных, представляется разумным. Простота алгоритма расчета дескрипторов использованных нами, основанного на ОКВЛИ, не требующего длительного машинного времени и большой мощности, представляется перспективным.

Конформационные характеристики полимерных цепей обычно определяются в разбавленных растворах полимеров. Основные типы измерений проводимых в растворах это – измерение вязкости, светорассеяние, рентгеновское рассеяние под малым углом и осмотическое давление. Согласно Бицерано, конформационные характеристики играют ключевую роль в определении свойств растворов полимеров, как в их синтезе (т.е. полимеризация), так и в обработке (отливание пленок из растворов).

В частности, характеристическая вязкость  $[\eta]$  полимерных растворов - мера объема, связанного данной массой полимера при разбавлении, в термодинамическом равновесии. Далее, для получения моделей с

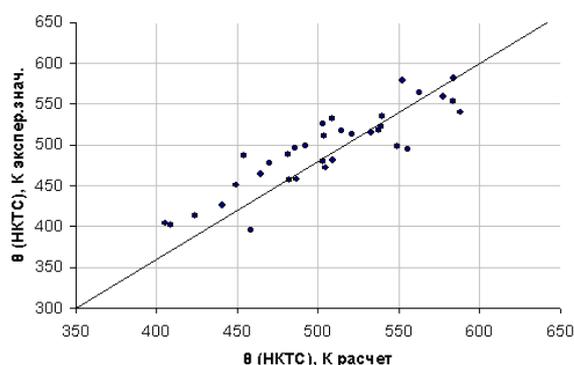


Рис. 5. Зависимость экспериментальных от рассчитанных с использованием дескриптора  $DMS({}^kEC)$  значений НКТС,  ${}^\circ K$

прогнозирующей способностью, применили ОКВЛИ и сравнили два метода в оценке характеристической вязкости.

**Моделирование характеристической вязкости систем полимер-растворитель.** Ван Кревелен оценивал различные уравнения для расчета вязкости растворов и вывел следующее эмпирическое уравнение (11) для оценки характеристической вязкости, когда экспериментальные данные не могут быть получены в  $\theta$  условиях:

$$\eta \approx 0.99328 \left( \frac{K}{M} \right)^2 \frac{\exp(8.5a^{10.3}) M_v^\alpha}{M_{cr}^{(\alpha-0.5)}} \quad (11)$$

В данном уравнении  $M_{cr}$  - критический молекулярный вес,  $M_v$  ( $M_v=2.5 \times 10^5$ ), - средневязкостная молекулярная масса,  $K$  - молярная функция неподвижности,  $M$  - молекулярный вес повторяющегося звена повторения, и параметр  $\alpha$  определяемый из следующего уравнения:

$$\alpha \approx 0.8 - 0.1 \text{abs}(\delta_{\text{растворитель}} - \delta_{\text{полимер}}) \text{ для } \text{abs}(\delta_{\text{растворитель}} - \delta_{\text{полимер}}) \leq 3$$

$$\alpha \approx 0.5 \text{ для } \text{abs}(\delta_{\text{растворитель}} - \delta_{\text{полимер}}) > 3$$

где параметр растворимости ( $\delta$ ) это  $\sqrt{J/c\text{с}}$

Используя уравнение (11) Ван Кревелен рассчитал характеристическую вязкость для 65 систем полимер-растворитель, с коэффициентом корреляции  $r^2=0.324$ . Однако, после удаления из систем 7 растворителей, которые могли образовывать водородные связи с полимерами в растворе, соответствующий расчет дал более высокий коэффициент корреляции ( $r^2=0.483$ ). Данное уравнение довольно хорошо работает, в тех случаях, когда полимер имеет тенденцию к высокой кристалличности в объеме. Как недостаток, эти наблюдения были отмечены также Бицерано.

В данной работе предпринята попытка использования ОКВЛИ для построения моделей «структура-свойство» ряда систем полимер-растворитель. Экспериментальные значения характеристической вязкости были взяты из литературы. Как и для моделирования значений НКТР были апробированы 2 дескриптора (4 и 6). Результаты апробаций и полученных в результате статистических характеристик компьютерных моделей характеристической вязкости показаны в таблицах 10 и 11.

Таблица 10

Статистические характеристики модели характеристической вязкости  $[\eta]$ , полученные с использованием дескриптора (4)

ЛИ	Обучающая выборка				Контрольная выборка			
	$r$	$r^2$	$S, \text{дл/г}$	$F$	$r$	$r^2$	$S, \text{дл/г}$	$F$
$EC^0$	0.6652	0.4424	26.292	21	0.6960	0.4844	19.3295	11
$EC^1$	0.8210	0.6740	20.105	56	0.7085	0.5019	26.0264	12
$EC^2$	0.9001	0.8102	15.34	115	0.5649	0.3191	61.033	6
$EC^3$	0.9589	0.9194	11.515	308	0.7034	0.4947	24.3582	12

Таблица 11

Статистические характеристики модели характеристической вязкости  $[\eta]$ , полученные с использованием дескриптора (б)

ЛИ	Обучающая выборка				Контрольная выборка			
	$r$	$r^2$	$S, \text{дл/г}$	$F$	$r$	$r^2$	$S, \text{дл/г}$	$F$
$EC^0$	0.6793	0.4614	25.841	23	0.7505	0.5632	199.902	15
$EC^1$	0.6890	0.4747	25.5214	24	0.7672	0.5886	30.1303	17
$EC^2$	0.9088	0.8259	14.6919	128	0.5606	0.3142	62.4552	5
$EC^3$	0.9651	0.9314	9.2223	367	0.6675	0.4455	39.2781	10

Как видно по данным таблиц наилучшие статистические характеристики получены с использованием дескриптора (б), однако при получении моделей НКТР наиболее подходящим был дескриптор (4), что подтверждено расчетом. Далее сравнили подход, разработанный и предложенный нами, с результатами расчетов значений характеристической вязкости по уравнению (11). Проведя расчет коэффициента корреляции для структур, входящих в контрольную выборку, для полученной модели коэффициент корреляции составляет – 0,7672, а для тех же систем по данным работы последний составляет – 0,4061, что заметно хуже модели, полученной с использованием дескриптора (б), на основе ОКВЛИ (табл. 11).

Из данной работы, очевидно, что в настоящее время существует несколько методов выявления связи между строением веществ и их физико-химическими параметрами. Однако, учитывая распространенность в природе и в технологических процессах различных многокомпонентных полимерных смесей, информация о количественной связи «структура-свойство/активность» применительно к многокомпонентным полимерным системам может иметь, по крайней мере, не меньшее число приложений использования оптимизации корреляционных весов локальных инвариантов графов, чем к однокомпонентным. Апробированный в данной работе подход, в сущности, есть предложение рассматривать смеси полимер-растворитель как пару молекулярных графов, представляющих структуру смеси, аналогично тому, как принято рассматривать структуры органических соединений в виде одного молекулярного графа. Это дает возможность использовать линейный регрессионный анализ, многократно проверенный и давший многочисленные конструктивные результаты для выборок индивидуальных веществ, как инструмент выявления количественных соотношений «структура-свойство» для упомянутых бинарных систем.

## ЗАКЛЮЧЕНИЕ

1. Разработана новая методика для проведения исследований «структурасвойство» на основе оптимальных дескрипторов, рассчитываемых при помощи индексов связности и кодов ближайшего соседства для однокомпонентных систем.
2. Проведено компьютерное моделирование температуры стеклования для полиариленоксидов по предлагаемой методике. На основе статистических данных анализа результатов показано, что модель описывает температуру стеклования полиариленоксидов удовлетворительно с коэффициентом корреляции  $r=0,9288$ .
3. Проведено компьютерное моделирование температуры плавления олигофениленов по предложенной методике. На основе статистических данных результатов показано, что модель описывает температуру плавления олигофениленов хорошо с коэффициентом корреляции  $r=0,9884$ .
4. Проведено компьютерное моделирование коэффициента распределения октанол-вода с помощью двух локальных инвариантов – расширенной связности и кодов ближайшего соседства, и на основе анализа результатов оптимизации показано, что модель с использованием кодов ближайшего соседства отлично описывает параметр гидрофобности органических соединений с наилучшими статистическими характеристиками ( $r=0,9966$ ).
5. Разработан подход к моделированию свойств для двухкомпонентных систем, основанный на оптимизации бинарных дескрипторов двух видов для системы мономер и растворитель. В результате моделирования по данному подходу параметра Флори-Хаггинса для 60 систем полимер-растворитель показано, что корреляционная зависимость на основе предлагаемого автором бинарного дескриптора с использованием расширенного индекса первого порядка дает наилучшие статистические характеристики с коэффициентом корреляции  $r=0,9958$ .
6. Компьютерный эксперимент по выявлению корреляционной зависимости нижней критической температуры смешения (НКТС) для 82 систем, состоящих из 8 полимеров и 47 низкомолекулярных растворителей, на основе предлагаемого подхода показал, что оптимальные дескрипторы описывают НКТС указанных систем неудовлетворительно с коэффициентов корреляции  $r=0,8863$ , что связано со спецификой взаимодействия в исследованных системах полимер-растворитель.
7. Компьютерный эксперимент по выявлению корреляционной зависимости характеристической вязкости для ряда систем полимер-растворитель на основе предлагаемого подхода с использованием двух видов оптимальных дескрипторов показал, что оптимальные дескрипторы описывают характеристическую вязкость указанных систем с коэффициентов корреляции  $r=0,7672$ , что заметно лучше, чем результаты, полученные другими авторами. Полученные модели различных характеристик полимерных систем могут быть использованы для предсказания свойств

широкого ряда одно- и двухкомпонентных систем полимер-растворитель, расчетным путем, что делает необязательным проведение сложных и дорогостоящих, а иногда просто технически трудоемких экспериментов.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Торопов А.А., Нурғалиев И.Н., Воропаева Н.Л., Рубан И.Н., Рашидова С.Ш. QSPR моделирование свойств двухкомпонентных систем полимер-растворитель // "Полимеры-2002": Тез.докл. Респ. конф. -Ташкент, 2002. -С.43.
2. Toropov A.A., Nurgaliev I.N., Voropaeva N.L., Ruban I.N., Rashidova S.Sh. QSPR modeling of Flory-Huggins parameter based on weighting of paths of length two // "Полимеры-2002": Тез.докл. Респ. конф. -Ташкент, 2002. -С.62.
3. Нурғалиев И.Н., Балахоненко О.И. QSPR моделирование посредством корреляционного взвешивания локальных инвариантов графов // "Химия и физика высокомолекулярных соединений": Тез.докл. Респ. конф. -Ташкент, 2002. -С. 10-11.
4. Toropov A.A., Nurgaliev I.N., Voropaeva N.L., Ruban I.N., Rashidova S.Sh. On some solved and unsolved problems of nanoparticles // "Nanochemistry: New approaches to creation of polymeric systems with specific properties": International Conference: Abstracts. -Tashkent, 2003. -P.21.
5. Toropov A.A., Balakhonenko O.I., Nurgaliev I.N., Voropaeva N.L., Ruban I.N., Rashidova S.Sh. QSPR-modeling of the Flory-Huggins parameter by correlation weighting of local graph invariant // Revue Roumaine De Chimie. - Bucuresti, 2003. - vol. 48. -№ 3. -P. 821-828.
6. Нурғалиев И.Н., Торопов А.А., Рубан И.Н., Рашидова С.Ш. QSPR моделирование физико-химических свойств полимеров на основе "наносегментов" полимера // "Наноматериалы в химии и биологии":Тез. докл. Межд. конфер. -Киев, 2004. -С.73.
7. Ruban I.N., Toropov A.A., Nurgaliev I.N. Nanofragment o polymer molecule as an alternative of the monomer unit in QSPR analysis//«Теоретические аспекты формирования полимерных наноструктур» ":Тез. докл. Межд. конфер. -Ташкент, 2004. -С.25.
8. Торопов А.А., Нурғалиев И.Н., Балахоненко О.И., Воропаева Н.Л., Рубан И.Н., Рашидова С.Ш. QSPR моделирование температур стеклования полиариленоксидов // Журнал Структурной Химии. –Новосибирск, 2004. -Т 45. -№ 4. -С. 741–747.
9. Toropov A.A., Nurgaliev I.N., Balakhonenko O.I., Voropaeva N.L., Ruban I.N., S.Sh. Rashidova QSPR-modeling of octanol-water partition coefficient by means of the nearest neighboring codes correlation weighting // Cellulose Chemistry and Technology. – Bucuresti, 2005. -vol. 39. -№ 1-2. -P. 37-48.
10. Рубан И.Н., Торопов А.А., Нурғалиев И.Н., Воропаева Н.Л., Рашидова С.Ш. QSPR-моделирование параметра Флори-Хаггинса на основе

«наносегментов» полимера // Высокомолекулярные соединения, серия Б. – Москва, 2005. -Т. 47. -№10. -С. 1887-1890.

11. Нургалиев И.Н., Торопов А.А., Кудышкин В.О., Рубан И.Н., Воропаева Н.Л., Рашидова С.Ш. QSPR-моделирование температуры плавления олигофениленов // Журнал Структурной Химии. –Новосибирск, 2006. -Т 47. -№ 4. –С. 368-373.

12. Нургалиев И.Н., Рашидова С.Ш. Моделирование нижней критической температуры смешения ряда двухкомпонентных смесей полимер-растворитель, с использованием дескрипторов, основанных на корреляционных весах // «Получение нанокмозитов, их структура и свойство»: Матер. Респ. науч.-техн. конф. - Ташкент, 2007. -С. 232.

13. Нургалиев И.Н. Моделирование характеристической вязкости систем полимер-растворитель // Доклады Академии наук РУз. – Ташкент, 2009. -№ 3-4. -С. 82-84.

14. Нургалиев И.Н., Рашидова С.Ш. Топологическое моделирование нижней критической температуры смешения систем полимер-растворитель посредством оптимальных дескрипторов // Узбекский физический журнал. - Ташкент, 2009. -Т 11. –№ 23. -С. 224-229.

Физика-математика фанлари номзоди илмий даражасига  
талабгор Нургалиев Ильнар Накиповичнинг  
01.04.19 - Полимерлар физикаси ихтисослиги бўйича  
“Оптималь дескрипторлар воситасида полимер тизимларнинг физик  
хоссаларини топологик (QSPR) моделлаштириш”  
мавзусидаги диссертациясининг

### РЕЗЮМЕСИ

**Таянч сўзлар:** структура – хосса миқдорий боғланиши (QSPR), дескриптор, полимер, эритувчи, молекуляр граф.

**Тадқиқот объектлари:** синтетик полимерлар ҳамда паст молекуляр бирикмалар тадқиқот объектлари ҳисобланади. Тадқиқот предмети этиб танланган бирикмалар қатори ва тизимларида “структура-хосса” ўзаро миқдорий боғлиқлигини ўрганиш белгиланган.

**Ишнинг мақсади:** полимерларнинг хоссаларини аниқлаш ва уларни маълум мақсадларга йўналтирилган тарзда синтез қилиш учун назарий асослар сифатида статистик қонуниятларга таянган оптималь дескрипторлардан фойдаланиб, структуралари ҳақидаги маълумотларни таҳлил қилиш асосида полимерларни физик хоссаларини олдиндан айтиб бериш учун топологик QSPR моделлаштиришни ишлаб чиқишдан иборатдир.

**Тадқиқот методлари:** “структура-хосса” боғланишини тузиш учун локал инвариант графлар корреляцион вазнларини оптималлаштириш синовлари ўтказилган. Моҳияти бўйича ушбу усул структуравий дескрипторларга асосланган аддитив чизма гибриди ва регрессион таҳлилдир.

**Олинган натижалар ва уларнинг янгилиги:** индивидуал моддалар ва полимер – эритувчи аралашмаларнинг “структура-хосса” ўрганиш муаммоси бўйича тадқиқот ўтказиш учун янги усулни ишлаб чиқаришдан иборатдир. Илк маротаба “структура-хосса” корреляцион боғланишини белгилаш учун оптимал дескрипторлар қўлланилган.

Ушбу, локал инвариант графлар корреляцион вазнларини оптималлаштиришга асосланган дескрипторларни бир ва икки компонентли полимер тизимларнинг физик хоссаларини олдиндан айтиб бериш учун қўлланиши баҳоланган.

Илк маротаба оптимал дескрипторлар асосида полиариленоксидларни шишаланиш ҳарорати, олигофенилинлар суюқланиш ҳарорати, октанол-сув органик моддаларни тақсимланиш коэффиценти ҳамда Флори-Хаггинс параметрлари, қуйи критик аралашуш ҳарорати ва икки компонентли полимер-эритувчи тизими учун тавсифли қовушоқлик моделлари ишлаб чиқилган.

**Амалий аҳамияти:** полимер тизимларнинг турли тавсифларини ҳисоблашда олинган моделлар бир ва икки компоненти полимер-эритувчи тизимлар кенг қатори учун таъкидланган тавсифларни ҳисоблашга имкон беради, бу мураккаб, серҳаражат, баъзи ҳолларда техник жиҳатда қийин бўлган экспериментларни ўтказишни талаб этмайди.

**Қўлланиш соҳаси:** тадқиқот усуллари, олинган “структура-хосса” моделлари ва олдиндан айтиб бериш алгоритми қуйидаги:

- юқори молекуляр бирикмалари физикаси ва кимёси соҳаларида амалий вазифаларни технологик амалга ошириш учун истиқболли структураларни танлаш мақсадида полимер структуралар тўпламини таҳлил қилишда;

- бир ва икки компонентли тизимлар учун “структура-хосса” боғланишини тадқиқот этишда материалшунослик соҳасидаги илмий ишларда;

- олий ўқув ютрларда материалшунослик бўйича ўқиш жараёнларида ва бошқа соҳаларда қўлланиши мумкин.

## РЕЗЮМЕ

диссертации Нургалиева Ильнара Накиповича на тему:  
«Топологическое (QSPR) моделирование физических свойств полимерных систем посредством оптимальных дескрипторов» на соискание  
ученой степени кандидата физико-математических наук  
по специальности 01.04.19 - Физика полимеров

**Ключевые слова:** количественная связь «структура – свойство» (QSPR), дескриптор, полимер, растворитель, молекулярный граф.

**Объекты исследования:** объектами исследования являются синтетические полимеры, а также низкомолекулярные соединения. Предметом исследования было установление количественной взаимосвязи «структура-свойство» в выбранном ряду соединений и систем.

**Цель работы:** разработка топологического QSPR моделирования для прогнозирования физических свойств полимеров на основе анализа информации об их структуре с использованием оптимальных дескрипторов, основанных на статистических закономерностях как теоретической основы для определения свойств и целенаправленного синтеза полимеров.

**Методы исследования:** для построения зависимостей «структура–свойство» была опробована оптимизация корреляционных весов локальных инвариантов графов. По сути, данный подход является гибридом аддитивной схемы и регрессионного анализа на основе структурных дескрипторов.

**Полученные результаты и их новизна:** состоит в разработке новой методики проведения изысканий по проблеме «структура-свойство», позволяющая проводить этот поиск не только для индивидуальных веществ, но и для смесей полимер-растворитель. Впервые использованы оптимальные дескрипторы, для установления корреляций «структура-свойство».

Оценено использование данных дескрипторов, основанных на оптимизации корреляционных весов локальных инвариантов графов для прогнозирования физических свойств одно- и двухкомпонентных полимерных систем.

Впервые на основе оптимальных дескрипторов были получены модели температур стеклования полиариленоксидов, температур плавления олигофениленов, коэффициентов распределения октанол-вода органических веществ, а также параметров Флори-Хаггинса, нижних критических температур смешения и характеристической вязкости двухкомпонентных систем полимер-растворитель.

**Практическая значимость:** полученные в ходе расчетов модели различных характеристик полимерных систем позволят предсказывать указанные характеристики для широкого ряда одно- и двухкомпонентных систем полимер-растворитель, расчетным путем, что делает необязательным проведение сложных и дорогостоящих, а иногда просто технически трудоемких экспериментов.

**Область применения:** методика исследования, полученные модели "структура–свойство" и алгоритмы прогнозирования, могут использоваться:

- для анализа наборов полимерных структур с целью отбора перспективных структур для технологической реализации практических задач в области физики и химии высокомолекулярных соединений;
- в научной работе в области материаловедения при исследовании зависимости структура – свойство для высокомолекулярных одно- и двухкомпонентных систем;
- в образовательном процессе материаловедения в высших учебных заведениях.

## RESUME

Thesis of Ilnar Nakipovich Nurgaliev on the scientific degree competition of the doctor of sciences of philosophy in physics and mathematics on speciality 01.04.19 – Polymer physics, subject:  
“Topological (QSPR) modeling physical properties of polymer systems using of optimal descriptors”

**Key words:** quantitative “structure – property” relationships (QSPR), descriptor, polymer, solvent, molecular graph.

**Subjects of the inquiry:** the research objects are synthetic polymers and low-molecular compounds. The subject of the research was determination of the quantitative relationships between structure and property in the selected class of compounds and systems.

**Aim of the inquiry:** development of topologic QSPR modeling for calculation of physical properties of polymers on the basis of information about their structure and with use of optimal descriptors based on statistical regularity as theoretical basis for determination of properties and purposeful synthesis of polymers.

**Method of inquiry:** for determining of the “structure – property” dependencies optimization of correlation weights of local graph invariants has been applied. Per se this approach is a hybrid of the additive scheme and the regression analysis based on structural descriptors.

**The results achieved and their novelty:** for the first time optimal descriptors have been used for the determining of the “structure – property” correlation. Application of local descriptors based on the optimization of correlative weights of local graph invariants for prediction of physical properties of single- and double-component polymeric systems has been estimated. For the first time models of the glass transition temperature of polyarylenoxides, the melting temperature of oligophenilens and octanol-water partition coefficients of organic substances have been estimated as well as the Flory-Huggins parameter, lower critical temperatures of blendings and the intrinsic viscosity of two-componental of “polymer-solvent” systems.

**Practical value:** practical relevancy includes the fact that models of various characteristics of polymeric systems allows to predict by the calculation way above mentioned characteristics for wide range of single- and double-components “polymer-dissolvent” systems obtained which makes unnecessary complicated and expensive and sometimes even technically labor-intensive experiments.

**Sphere of usage:** research method, obtained “structure – property” models and prediction algorithms can be used:

- for analysis of set of polymeric structures in order to select prospective structures for technological implementation of practical tasks in physics and chemistry of high-molecular compounds;

- for researches in the field material science during investigation of the “structure – property” dependencies for high-molecular single- and double-components systems;

- in educational process in material-sciences high education institutes.

***Соискатель:***