

Применение мультимодального подхода для выявления подобия в многомерных наборах данных с примером использования

Ольга В. Перл, Иван А. Перл

Аннотация—Всё чаще можно встретить исследования из различных областей, посвящённые работе с разноформатными данными, задачами, перспективами исследования или взглядами на объект исследования. Часто эти исследования используют термин *мультимодальный*, который в то же время сильно изменяется от области к области. Так, исследователи из разных предметных областей сталкиваются с одной и той же задачей: выбор подхода и обработка мультимодальных данных. Проблема этой задачи заключается в узкой области применимости предлагаемых решений. Тогда само понятие мультимодальности нуждается в формализации. Одновременно с этим возникает необходимость в выделении структур данных для работы с этим понятием. Важным также является описание подходов для анализа данных в выделенных структурах данных. В данной статье предлагается определение понятия мультимодальных данных, выделяются 4 структуры для работы с ними, а также предлагается метод для выделения наиболее схожих мультимодальных объектов. Помимо этого, приводимые структуры поясняются примерами. Метод для нахождения схожих мультимодальных объектов дополняется описаниями модификаций для того, чтобы его можно было применить для всех 4 структур мультимодальных данных. В статье также приводится демонстрация применения метода на общем с описанием структур данных примере -- исследование городов по характеристикам количества населения, климате и количестве университетов. Пример исследования носит исключительно иллюстративный характер, однако, может быть использован для дальнейших исследований после проверки соответствующими специалистами. В статье приводятся способы конфигурирования метода и рекомендации по работе с ними. Рассчитываемая мера схожести объектов (сила связывания) представляет собой способ задания многомерной метрики над сложной структурой данных. В конце статьи приводятся направления дальнейших исследований, которые уже проводятся авторами в настоящее время.

Ключевые слова— Выявление Связей и Зависимостей, Информационная Модель, Мультимодальные Данные, Структура Мультимодальных Данных

I. ВВЕДЕНИЕ

Проектирование системы, занимающейся

Статья получена 8 октября 2021.

О. В. Перл – Факультет Программной Инженерии и Компьютерной Техники, Университет ИТМО, Санкт-Петербург, Россия (ovkalyonova@itmo.ru).

И. А. Перл – Факультет Программной Инженерии и Компьютерной Техники, Университет ИТМО, Санкт-Петербург, Россия (ivan.perl@itmo.ru).

представлением или обработкой данных разных форматов из гетерогенных источников и, описывающей существенно различные объекты, может оказаться сложной задачей. Одним из наиболее важных аспектов станет выбор правильного метода организации и идентификации данных, а также определение отношений сущностей. Для широкого спектра применений для решения указанных задач могут быть использованы методы мультимодальной обработки данных.

A. Проблема

Важной проблемой является формулировка понятия мультимодальных данных. Несмотря на широкую распространённость термина мультимодальности в различных науках, часто под ним понимаются разные понятия и явления. Так, в логике принято говорить о модальностях речи и о модальностях, как отражении вероятности высказывания [1], [2]; в транспортной логистике говорят о мультимодальности, как о вариативности и комбинировании видов транспорта при доставке грузов [3]; в биологии и медицине говорят о мультимодальности как о множественном представлении изучаемого или используемого объекта (мультимодальный МРТ [4], мультимодальная терапия [5] и пр.); в педагогике говорят как о разнообразии игрушек для развития и обучения ребёнка [6], [7] и так далее.

В информатике понимание мультимодальности так же сильно различаются. Так, говоря о мультимодальной биометрии имеется в виду работа с различными способами идентификации человека: отпечатки пальцев, сетчатка глаза, голосовая идентификация и пр. [8]. Свое применение термин мультимодальности нашёл и в построении виртуальной и дополненной реальности [7]. Другим широко распространённым значением мультимодальных данных является случаи, когда имеют в виду разнообразие форматов данных, в первую очередь это изображения, видео и аудио форматы [9], [10].

Во всех этих науках есть собственные исследования, основанные на понятии мультимодальности и мультимодальных данных. Однако, в связи со специфичностью предметных областей, исследования из одной области крайне трудно использовать для применения в другой.

В свою очередь информатика, как наука, основным фокусом которой в первую очередь является

информация, должна бы решить эту проблему, предоставив аппарат, позволяющий при этом переносить исследования из одной области в другую для ускорения общего прогресса.

Первоочередной задачей тогда становится разработка унифицированного определения понятия мультимодальных данных и выделение структур данных для работы с ними.

Наиболее близким понятием, описывающим вариативность и смешанность данных, подвергаемых анализу и обработке, является гетерогенность. Поэтому понятие мультимодальных данных будет справедливо строить с учётом гетерогенных данных, как частного случая мультимодальных данных.

В. Цель

Целью данной работы является определение метода и понятий, необходимых для анализа гетерогенных наборов данных с нетривиальными связями между их сущностями. Для достижения поставленной цели необходимо рассмотреть следующие вопросы:

1. Сформулировать определение понятия мультимодальных объектов на основе реальных гетерогенных объектов. Это означает, что нам необходимо разработать механизм, объясняющий, как представить гетерогенные объекты из широкого спектра задач в виде мультимодальных объектов, пригодным для дальнейшей обработки и анализа.
2. Определить подход к представлению отношений между мультимодальными объектами, извлеченными из реальных исследуемых объектов.
3. Определить подход к обработке мультимодальных данных для поиска взаимосвязей и корреляций между различными мультимодальными объектами.

II. СТРУКТУРЫ МУЛЬТИМОДАЛЬНЫХ ДАННЫХ

А. Базовая структура мультимодальных данных

Под мультимодальным объектом в данной статье мы будем понимать объект или процесс, который описывается различными характеристиками, представленными в соответствующих модальностях. Под модальностью мы будем понимать один из нескольких атрибутов, которые описывают конкретную характеристику объекта. Совокупность характеристик одного мультимодального объекта будем называть сущностными представлениями или образами. Не обязательно, чтобы каждый мультимодальный объект имел сущностное представление (образ) во всех модальностях.

По своей природе гетерогенные объекты описываются рядом признаков, которые не совместимы друг с другом, например, цветом, весом, ценой, и эти характеристики могут рассматриваться как модальности конкретного объекта. Объект, имеющий несколько описывающих его модальностей, можно назвать мультимодальным.

Проиллюстрируем сформулированное определение на

примере. Выберем в качестве объекта исследования города. У всех городов выберем характеристики, которые их описывают: население, климат и университеты. Тогда положим эти характеристики в качестве модальностей. Сущностным представлением (essensial presentation [11]) назовём значение каждого города в этой модальности. Так, для Саратова сущностным представлением в модальности “Население” является 838,042 человек (по данным на 2020 год), а для Санкт-Петербурга - 5,388,759.

Однако, на практике не всегда можно выделить действительно существующий объект, как город, поэтому согласно принятой терминологии в анализе и представлении данных [12], мультимодальные объекты также называют абстрактными объектами (АО, abstract object) [11].

Определим теперь необходимые виды связей в наборе мультимодальных данных. Продолжая пример с городами, легко заметить, что под связью можно условиться понимать многое: наличие прямого сообщения, являются ли города побратимами и многое другое. Задаваемые извне связи, о которых нам известно перед началом анализа мультимодальных данных, назовём прямыми. Вычисленные аналитически связи, отражающие собой меру подобия (силу связывания w), назовём косвенными связями. Обозначим количество городов, или мультимодальных объектов, как n , количество модальностей m (в нашем случае $m = 3$: население, климат и университеты). Обозначим так же множество прямых связей как r_{dl} , а множество косвенных - r_{il} , тогда общее число связей между мультимодальными данными: $r_v = r_{il} + r_{dl}$. Подробнее о связях в мультимодальных данных описано в [11], как и о методе представления мультимодальных данных для получения косвенных связей, краткое изложение которого приводится в разделе III.

Простая структура мультимодальных объектов показана на рис. 1 а). На верхнем уровне находятся абстрактные объекты, сущностные представления которых находятся в соответствующих модальностях. Сущностные представления внутри каждой модальности выстроены иерархически. Сущностные представления, относящиеся к одному абстрактному объекту, имеют ссылки между собой и с самим абстрактным объектом.

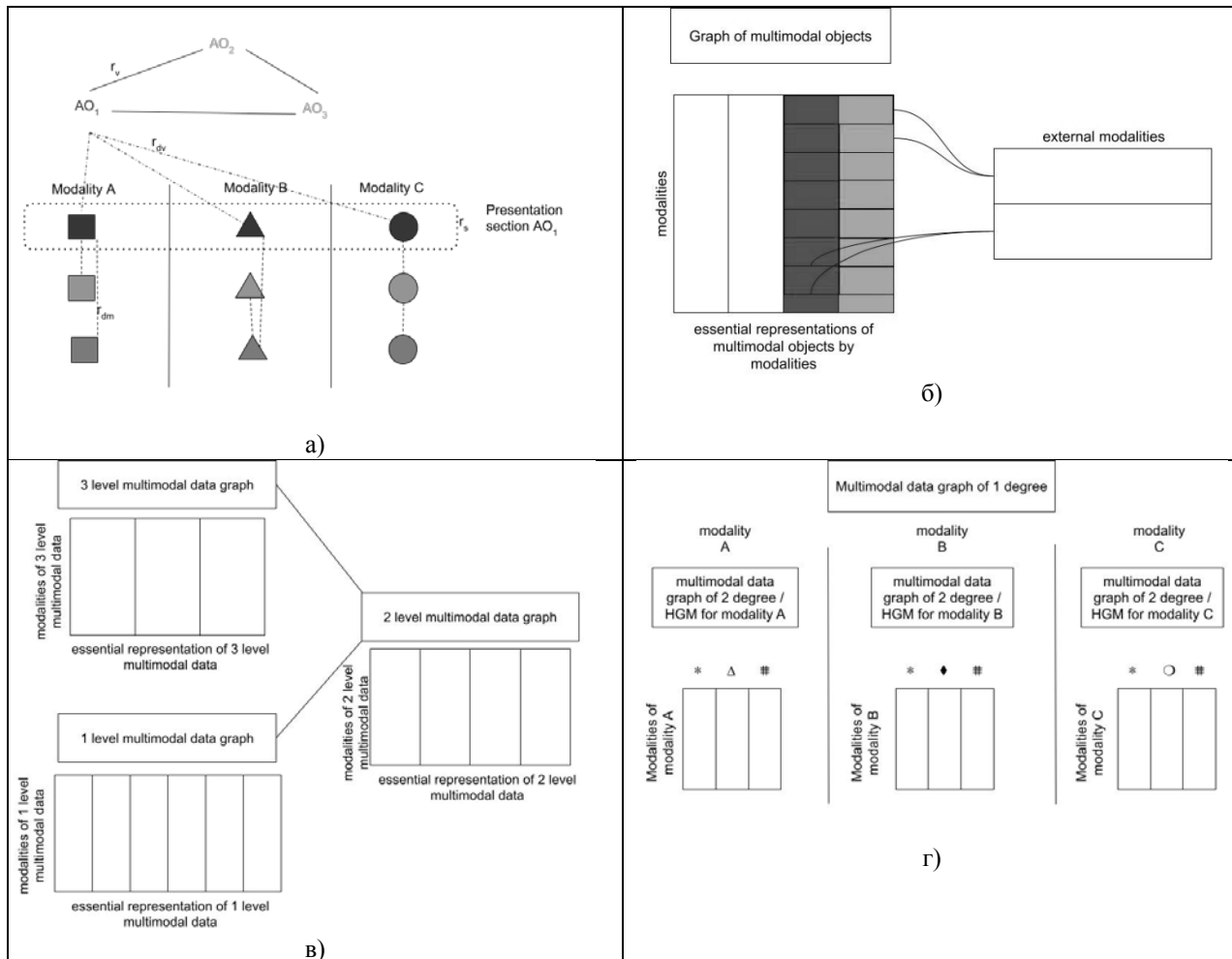
Когда обычных множеств мультимодальных объектов недостаточно для описания комбинаций разнородных объектов, можно использовать более сложные структуры мультимодальных данных. В этой статье мы опишем следующее:

- Мультимодальные данные со связями во внешних модальностях.
- Мультимодальные данные высших порядков.
- Рекурсивные мультимодальные данные.

В. Мультимодальные данные со связями во внешних модальностях

Приступая к работе с мультимодальными данными, мы часто сталкиваемся с вопросом о том, какие виды

задействовать внешнюю модальность. Она даст нам временную шкалу и список временных шагов модели, относительно которых станет возможным выровнять значения в результатах выполнения.



Пример структуры мультимодальных данных со связями во внешних модальностях

Рисунок 1. Структуры мультимодальных данных: а) простая структура б) структура мультимодальных данных со связями во внешних модальностях в) мультимодальные данные высших порядков г) рекурсивные мультимодальные данные.

данных можно использовать в качестве модальности? Можно ли определить модальность, сущностное представление которой будет задано в виде набора значений реального объекта?

В общем случае не принято представлять модальность объекта с набором атрибутов реального объекта, но если это так, то подход рекурсивных мультимодальных данных описывает ее.

Иногда мы можем столкнуться с ситуациями, когда значения мультимодальных данных в одной и той же модальности связаны друг с другом через внешние модальности, которые необходимы для правильного отображения элементов данных друг на друга.

Например, если при исследовании демографических процессов в городах нам захочется задать значения не одним числом, а списком значений в разные годы, то внешней модальностью вероятнее всего будет время в годах.

Такая структура может быть использована для хранения временных рядов, как, например, результаты моделирования, т. к. они не могут быть сопоставлены друг с другом напрямую. Для этого нам нужно

С. Мультимодальные данные высших порядков

Следующим шагом является работа не только с самими объектами, как с мультимодальными объектами, но и со сложными описаниями связей между этими объектами. В некоторых случаях такие связи так же могут быть представлены в виде мультимодальных объектов.

Возвращаясь, к примеру с рассмотрением связей между городами, вспомним, что прямые связи сами по себе бывают очень разнообразными. Если для установления прямых связей в структуре мультимодальных данных была выбрана транспортная сеть, то можно для таких связей выделить собственные модальности: наличие автомобильной дороги, железнодорожное сообщение, авиасообщение, а также морское и речное.

В качестве альтернативного примера можно привести свойства связей между сущностями в теории построения онтологических моделей [13].

Для случая, когда у связей появляются собственные модальности можно предложить два следующих

подхода:

1. Первый подход заключается в интерпретации объектов как связей, а связей как объектов. Тогда мы будем рассматривать связи между изначальными объектами как мультимодальные объекты, а сами изначальные объекты как простые связи без заданных модальностей. Иными словами, объектом исследования становятся связи, в те объекты, которые они связывают. Если исследование больше сфокусировано на изучении связей между городами [13], то в качестве мультимодальных объектов можно выбрать сами эти связи, а города - в качестве связей (без присущих им модальностей населения, климата и университетов) между такими объектами.
2. Второй подход заключается в том, чтобы рассматривать независимо изначальные мультимодальные объекты с их модальностями и связи между ними как мультимодальные объекты, но более высокого порядка с собственным набором модальностей. Соответственно получим 2 порядка мультимодальных данных: на первом уровне мы рассматриваем города, а на втором — связи между городами.

Основным недостатком первого подхода является то, что он не позволяет исследовать сами объекты, поскольку ключевой фокус переносится на связи, а сами объекты упрощаются. Такой подход применим только в тех случаях, когда аналитика интересуют только связи между объектами, а не сами объекты.

Второй подход характеризуется повышением сложности системы. Но в то же время мы можем обрабатывать как связанные мультимодальные объекты, так и связи между ними, которые также представлены как мультимодальные объекты более высокого уровня.

Косвенные связи между городами, отражающие силу связывания или величину их подобия, можно принять в некоторых случаях как одну из модальностей следующего порядка.

При вынесении связей на новый слой мультимодальных данных, мы будем говорить, что мы получили $p + 1$ уровень мультимодальных данных (создали более высокий порядок мультимодальных данных). В то же время связь между мультимодальными объектами различных порядков также может быть вынесена на следующий уровень $p + 2$ порядка мультимодальной структуры данных. Пример структуры показан на рис. 1 в).

D. Рекурсивные мультимодальные данные

Особой структурой мультимодальных данных являются рекурсивные мультимодальные данные. Эта структура является прямым отражением концепции сложных систем. Рекурсивные мультимодальные структуры предполагают, что сущностные представления мультимодальных объектов в определенных модальностях также могут быть мультимодальными

объектами.

Например, при климатических исследованиях краткой характеристики в соответствующей модальности городов может оказаться недостаточно. Тогда сама климатическая характеристика городов может быть описана с помощью мультимодальных объектов. У таких мультимодальных объектов можно выделить собственные модальности, например: температура, влажность и пр.

Важное отличие структуры рекурсивных мультимодальных данных от мультимодальных данных со связями во внешних модальностях заключается в том, что сущностные представления вложенной структуры рекурсивных данных не могут быть определены как перечисления одной модальности, а представляют собой самостоятельные значения, никак не связанные с другими объектами или между собой внутри списка.

Однородность (гомогенность) мультимодальных объектов, которые являются сущностными представлениями различных модальностей или мультимодальных объектов, не гарантируется.

Пример структуры рекурсивных мультимодальных данных показан на рис. 1 г).

III. МЕТОД ПРЕДСТАВЛЕНИЯ МУЛЬТИМОДАЛЬНЫХ ДАННЫХ ДЛЯ ПОИСКА КОСВЕННЫХ СВЯЗЕЙ И ЕГО МОДИФИКАЦИИ

A. Базовый метод представления мультимодальных данных

Данный метод состоит из 4 основных шагов.

1) Получение набора входных данных

На этом этапе у нас есть множество $X^n = \{x_1, x_2, \dots, x_n\}$ входных данных, разделенных по модальностям. Максимальный размер модальности определяется количеством n мультимодальных объектов (для случая, когда все сущностные представления в модальности уникальны).

2) Классификация данных внутри модальностей

На следующем этапе полученные наборы данных передаются классификаторам в соответствии с модальностями. Заметим, что методы классификации зависят от модальностей и не являются принципиальными с точки зрения всего метода. Однако выходной формат *строго определен и должен представлять собой иерархический граф значений в данной модальности.*

Классификация необходима для определения степени взаимосвязи между сущностными представлениями различных мультимодальных объектов, что и выполняется на следующем этапе.

3) Расчет высот общих предков

Результатом шага 2 для каждой модальности является иерархический граф модальности, определяющий взаимосвязь различных сущностных представлений мультимодальных объектов. Обозначим через $h(a, b)$ высоту общего предка для мультимодальных объектов a и b . Это значение будет рассчитываться в соответствии с максимальной ($\max \min(a, b)$) стратегией как расстояние от каждого из выбранных узлов (a и b) до

узла предка.

Помимо этого, могут быть введены дополнительные параметры h_{max} и h_{term} . Параметр h_{max} определяет максимальное сохраняемое значение высоты общего предка, а h_{term} – максимальную высоту общего предка, до которого выполняется поиск. Если достигнуто число h_{term} , а общий предок так и не найден, значит полагаем, что общего предка не существует и расстояние между ними бесконечно. На практике такое расстояние можно обозначать как -1 .

4) Расчет силы связывания (подобия) мультимодальных объектов

На данном этапе выполняется поиск косвенных связей между мультимодальными объектами на основе значений сущностных представлений и знаний о расстояниях сущностных представлений в различных модальностях. По сути данный этап может рассматриваться как задание многомерной метрики между объектами.

На этом этапе вычисляется сила связывания между всеми парами мультимодальных объектов на основе высот общих предков для всех модальностей. В простейшем случае это значение вычисляется следующим образом:

$$w^h = \sum_{i=1}^m \begin{cases} 0, & \frac{h}{h_i(a,b)} \geq 1 \\ 1, & \frac{h}{h_i(a,b)} < 1 \end{cases}$$

где w^h — это сила связывания, при том, что высота общего предка в каждой модальности для исследуемых объектов не будет выше заданного значения h . Помимо этого, данная формула может быть дополнительно настроена при помощи взвешенных коэффициентов для каждой модальности или в виде функций от высоты. Подробнее об этом изложено в разделе IV.

Этот метод может работать в двух основных режимах: полное и инкрементное обучение. Первый режим предполагает, что система будет обучена один раз и не изменится. Второй режим, наоборот, позволяет работать с периодически добавляемыми данными без полного переобучения системы. В этом случае количество проверок увеличивается в зависимости от количества мультимодальных объектов, как показано на рис. 2 для

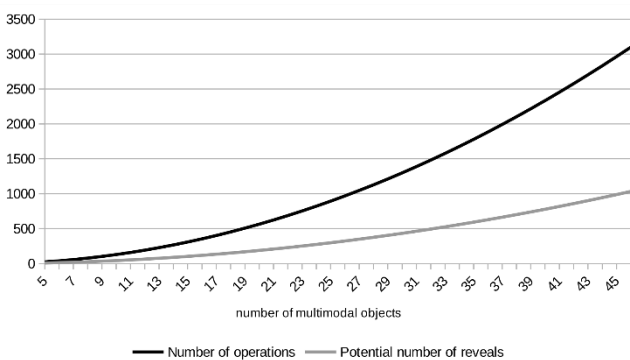


Рисунок 2. Количество операций и количество потенциальных связей при увеличении числа мультимодальных объектов в методе с полным пересчетом при добавлении нового мультимодального объекта.

полного обучения и на рис. 3 для инкрементного

обучения. Инкрементный режим схож с добавлением столбцов в таблицу базы данных без пересоздания всей базы данных.

В. Модификация метода для представления мультимодальных данных со связями во внешних модальностях

Необходима некоторая модификация метода для того, чтобы метод мультимодального представления данных мог обрабатывать наборы данных, содержащие связи между сущностными представлениями мультимодальных объектов через внешние модальности.

Прежде всего, следует понимать, что такая структура предполагает, что в данной модальности сущностное представление каждого мультимодального объекта представляет собой ряд значений, соотносящихся один к одному со значениями внешней модальности. Эта структура показана на рис. 1 б). На рисунке ряд значений сущностных представлений обозначен горизонтальными прямоугольниками. Число внешних модальностей на показанном рисунке $m_f = 2$, и число внутренних модальностей, имеющих связи во внешних модальностях так же равно 2.

Использование внешних модальностей следует учитывать при построении иерархического графа такой внутренней модальности системы, число которых равно m_f на втором этапе метода. Например, для построения иерархического графа модальности могут быть использованы вспомогательные вычисления: проверка монотонности числовых значений, числовой диапазон, аппроксимирующая функция и т. д.

С. Модификация метода для представления мультимодальных данных высших порядков

Для этой структуры существует два основных способа использования метода представления:

1. сверху вниз: от верхнего порядка (слоя) к нижнему;
2. снизу вверх: с нижнего, базового порядка (слоя) к верхнему.

Заметим, что в случае использования метода снизу вверх сила связывания низкоуровневых объектов w может быть использована в качестве модальности для более высоких порядков. Однако при применении метода нисходящего мультимодального представления

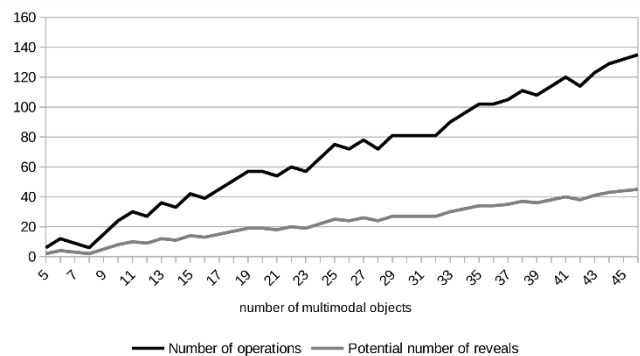


Рисунок 3. Количество операций и количество потенциальных связей при увеличении числа мультимодальных объектов в методе с инкрементным расчётом при добавлении нового мультимодального объекта.

данных шаги 2 и 3 могут выполняться параллельно для всех порядков.

Базовая структура мультимодальных объектов высших порядков показана на рис. 1 в).

D. Модификация метода для представления рекурсивных мультимодальных данных

В структуре рекурсивных мультимодальных данных поиск связей между объектами различных модальностей (A, B, C) может осуществляться только для сущностей первого уровня вложенности и только для общих модальностей (*, #).

Метод может применяться рекурсивно внутрь (от общего к частному) или наружу (от частного к общему).

Базовая структура рекурсивных мультимодальных объектов показана на рис. 1 г).

IV. ДЕМОНСТРАЦИЯ ПРИМЕНЕНИЯ МЕТОДА

Отличительной особенностью гетерогенных данных является фрагментированность источников таких данных и/или фрагментированность компонентов объектов, полученных из этих данных, в том числе неполных данных. Причин такой неоднородности может быть достаточно много, однако проблема в целом сводится к неоднородности методов анализа таких объектов.

Одним из примеров гетерогенных данных могут быть данные, полученные из различных источников, в том числе Интернет-ресурсов.

Для демонстрации работы метода был подготовлен набор данных из описанного примера: для 20 крупнейших городов России собраны данные о количестве населения [15], типу климата [15] и количеству университетов [16]. Также будем предполагать, что информации о прямых связях между городами не задано. В результате работы метода предполагается получить косвенные связи между городами с рассчитанными силами связывания.

Обратим внимание, что проводимое далее применение метода носит исключительно демонстрационный характер и не имеет целью настоящее исследование городов. Все полученные в ходе демонстрации результаты должны быть проверены соответствующими специалистами перед использованием. Про реальное применение метода представления мультимодальных данных для системной динамики подробнее описано в статье [11].

Опишем подробнее процесс применения метода. Для демонстрации был выбран простой набор, отражающий базовую структуру мультимодальных данных. О модификации метода для работы с более сложными структурами подробнее изложено в разделе III B-D.

1. Первым этапом является сбор данных. В нашем случае данные были получены из сети Интернет посредством ручного сбора в связи с ограниченностью размера набора данных. В более масштабных случаях для этих целей могут быть использованы, например, специальные программные агенты (web-crawlers). Полученный набор данных

представлен в таблице 1.

2. Следующим этапом является разработка классификаторов для модальностей. В данном примере для модальностей с числовыми значениями (модальности “Население” и “Университеты”) был выбран подход отображения, результатом которого является иерархический граф, вырожденный в цепочку. Так, все числовые значения были отсортированы в порядке убывания, неуникальные значения были отброшены. Тогда расстоянием между сущностными представлениями будет являться разница по модулю между номерами в цепочке. Полученные иерархии для модальности “Население” и “Университеты” представлены в таблице 1.

Для модальности “Климат” был построен простой иерархический граф модальности, показанный на рис. 4. Так как большинство значений сущностных представлений мультимодальных объектов не уникальны, следует выбрать по одному элементу каждого значения. Тогда становится легко построить иерархический граф модальности. В данном примере граф был построен аналитически. Однако, для более сложных наборов данных с большим числом уникальных сущностных представлений для этой задачи могут быть применены алгоритмы кластеризации и классификации.

3. На следующем этапе для каждой пары мультимодальных объектов в каждой модальности было найдено расстояние — высота общего предка.

Например, для модальности “Население” для пары Москва и Санкт-Петербург $h^{\text{население}}(a, b) = 1$, потому что количеству населения Москва находится на 1 месте, а Санкт-Петербург — на 2. Тогда население(Санкт – Петербург, Москва) = $2 - 1 = 1$. Аналогично для $h^{\text{население}}(\text{Саратов, Санкт – Петербург}) = 17 - 2 = 15$.

Схожим образом расчёт высот общих предков производится для модальности “Университеты”. Отличие лишь в том, что разные города могут иметь одинаковое количество университетов и функция отображения присвоит им одинаковый индекс. Тогда получим: $h^{\text{университеты}}(\text{Москва, Санкт – Петербург}) = 2 - 1 = 1$, а для $h^{\text{университеты}}(\text{Омск, Пермь}) = 8 - 8 = 0$.

Важно заметить, что из-за особенностей данных в модальности “Университеты” могут быть получены 0 для высоты общего предка между двумя значениями сущностных представлений, что указывает на их идентичность, а в модальности “Население” –

нет (пока сами данные не будут представлены в другом виде, например, с более высокой размерностью или при анализе малых городов). Это свойство необходимо учитывать при построении функции расчёта w^h .

Опишем стратегию расчета высот общих предков для модальности «Климат». Как было описано, высота общего предка рассчитывается по максимальной ($max\ min$) стратегии. Это означает, что для каждого a и b мы ищем минимальное расстояние до общего предка $h_a(a, b)$ и $h_b(a, b)$ и из двух значений выбираем наибольшее. Например, для $a = \text{«переходный»}$ и $b = \text{«умеренный морской»}$ общим предком является «умеренно – континентальный». Расстояние от $a = \text{«переходный»}$ до общего предка равно 1, а расстояние от $b = \text{«умеренный морской»}$ – равно 2, тогда:

$$h(\text{«переходный»}, \text{«умеренный морской»}) = \max(\min(h_{\text{переходный}}(\text{«переходный»}, \text{«умеренный морской»}), \min(h_{\text{умеренный морской}}(\text{«переходный»}, \text{«умеренный морской»}))) = \max(1, 2) = 2,$$

следовательно расстояние между Санкт-Петербургом и Краснодаром по модальности «Климат» равно 2.

Для случаев, когда значения существенных представлений совпадают, например для Москвы и Воронежа, расстояние между ними в соответствующей модальности равно 0.

Как видно из полученных значений, стратегия построения иерархических графов модальности и выбор классификаторов является крайне важным для получения корректных высот. Например, для модальностей «Население» и «Университеты» также можно было бы построить дерево, сгруппировав их по выбранной максимальной разнице между существенными представлениями.

Приступим теперь к расчету силы связывания (величины подобия) мультимодальных объектов. Для этого выберем значение максимальной высоты общего предка h , единой для всех модальностей величины, указывающей на то, что значения существенных представлений рассматриваемой пары мультимодальных объектов в заданной модальности достаточно близки, чтобы учитываться при расчёте сил связывания. Это представляет собой порог срабатывания части функции, фильтр.

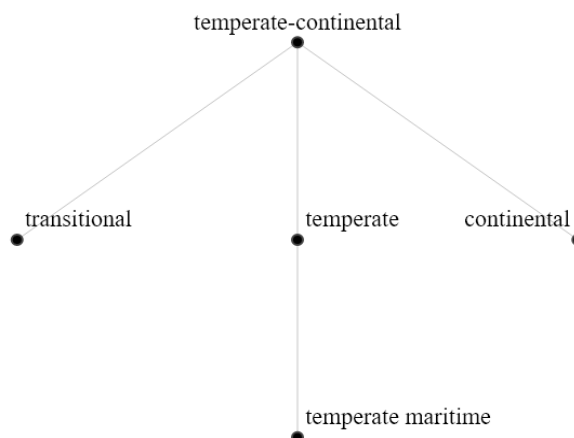
Простейший способ расчета w^h заключается в подсчёте количества модальностей, в которых существенные представления мультимодальных объектов близки при заданном h .

Имеет смысл перебирать различные

значения h в порядке увеличения до тех пор, пока количество пар связанных элементов не будет удовлетворять задаче или желанию аналитика. Обратим внимание на то, что в зависимости от данных в модальностях, некоторые значения h не могут быть получены. Например, в модальности «Население» в рассматриваемом наборе h никогда не будет равным 0, потому что едва ли найдётся 2 города с абсолютно одинаковым значением, только если это не особо малые города или если значение выражено в больших размерностях. Соответственно, при расчёте w^0 для рассматриваемого набора данных, $w_{max} = 2$. Значение w_{max} – максимальное из возможных значений сил связывания. Мультимодальные объекты, для пары которых $w^h(a, b) = w_{max}^h$ обладают наибольшей силой связывания, а, следовательно, наиболее схожи и близки.

При этом для w^1 , $w_{max} = 3$, так как все 3 модальности могут иметь высоту общего предка существенных представлений $h = 1$. Так же не имеет смысла выбирать h для w^h , который больше всех возможных значений высот во всех модальностях, так как он не будет сказываться на получаемом результате. Например, для модальности «Климат» не будут влиять значения $h \geq 3$, т.е. при расчёте w^h не будет изменений в выбираемых данных в модальности «Климат» для функции $w^{h=3}$ и $w^{h=10}$, хотя для остальных модальностей это и будет иметь смысл и так до $w^{h=20}$, когда выбраны абсолютно все значения. Тогда, при дальнейшем увеличении значения h значение функции w^h не изменится.

Для рассматриваемого примера для w^0 было получено 3 пары, для которых $w^0 = w_{max} = 2$: Ростов-на-Дону — Самара, Волгоград — Уфа, Саратов — Тольятти. Аналогично для w^1 было получено 4 связи, для которых $w^1 = w_{max} = 3$, среди них: Казань — Екатеринбург, Казань — Нижний Новгород,



Воронеж — Пермь, Тольятти — Ижевск. Заметим, что пары изменились, так как к функции добавились модальность “Население”. Для $w^2 = w_{max} = 3$ было получено 23 связи.

силы связывания вовсе не всегда говорит о наличии какой бы то ни было закономерности. Корреляции всегда должны подвергаться строгому анализу со стороны аналитика, т. к. в любом, а тем более в большом наборе данных и количестве

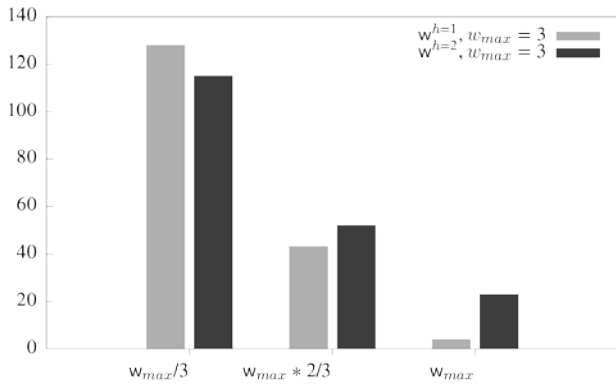
Таблица 1.Использованный набор данных о городах

Город	Население	Климат	Университеты	Иерархия количества населения	Иерархия количества университетов
Москва	12678	умеренно континентальный	266	1	1
Санкт-Петербург	5398	умеренный морской	88	2	2
Новосибирск	1626	континентальный	33	3	5
Екатеринбург	1494	умеренно континентальный	41	4	3
Казань	1257	умеренно континентальный	34	5	4
Нижний Новгород	1252	умеренный	33	6	5
Челябинск	1197	умеренно континентальный	25	7	10
Самара	1157	умеренно континентальный	31	8	6
Омск	1155	континентальный	27	9	8
Ростов-на-Дону	1138	умеренно континентальный	31	10	6
Уфа	1129	умеренно континентальный	23	11	11
Красноярск	1094	континентальный	26	12	9
Воронеж	1058	умеренно континентальный	30	13	7
Пермь	1055	умеренно континентальный	27	14	8
Волгоград	1009	умеренно континентальный	23	15	11
Краснодар	933	переходный от умеренно континентального к сухому субтропическому	30	16	7
Саратов	838	умеренно континентальный	18	17	12
Тюмень	807	континентальный	13	18	14
Тольятти	699	умеренно континентальный	18	19	12
Ижевск	648	умеренно континентальный	16	20	13

Важно обратить внимание не только на пары с максимальными значениями сил связывания, но и на количество пар со всеми остальными значениями. Это может быть использовано для более детального изучения таких связей. График количества связей разного типа при изменении значения h показан на рис. 5.

Напомним, что наличие высокого значения

модальностей (характеризующих признаков), можно найти корреляции, смысл которых будет отсутствовать, а сама корреляция будет обуславливаться случайностью.



V. КОНФИГУРИРОВАНИЕ ВЕСОВОЙ ФУНКЦИИ

Рисунок 5. Количество связей в долях максимально достижимого w_{max} для функций w^h с $h = 1$ и $h = 2$

Важной задачей для улучшения результатов метода является настройка и калибровка основных параметров метода. В первую очередь такими параметрами могут стать h и w .

Вопрос настройки классификаторов выходит за рамки данной статьи, поскольку процесс классификации в большей степени зависит от типов и значений обрабатываемых данных. Вопрос о выборе значений максимальной высоты общего предка для каждой из модальностей также не будет рассматриваться отдельно в данной статье, так как он в большей степени зависит от результирующей классификации модальности и является специфичным для каждой отдельной модальности.

Большое значение для результата работы метода имеет параметр w , так как от него зависит величина связей и количество связей с такими значениями, что

крайне важно для анализа полученных связей.

Стратегия расчета силы связывания w может быть задана следующим образом:

1. Как число модальностей, в которых значения существенных представлений рассматриваемых объектов близки (минимально значение высот общего предка). Соответственно, все модальные коэффициенты равны 1.
2. Как взвешенная сумма, с коэффициентами, определёнными аналитиком. Каждая модальность имеет свой собственный коэффициент, числовое значение веса, сумма которого определяет максимальное значение w . Тогда для расчета конкретной связи необходимо суммировать только те коэффициенты (в отличии от первого случая, когда все коэффициенты равны 1), при которых достигается требуемая близость общего предка h .
3. С помощью заданных функций в зависимости от высоты общего предка: $w = a_1(h_1) + a_2(h_2) + \dots + a_m(h_m)$.

Функции w может быть рассчитана следующим образом:

1. $w_{max} = 3 = 1 + 1 + 1$
2. Положим $a_{население} = 0.9; a_{климат} = 0.5; a_{университеты} = 1$.
Тогда $w_{max} = 0.9 + 0.5 + 1 = 2.4$;
3. Например $a_{население}(h_{население}) = \frac{1}{h_{население}^2}; a_{климат}(h_{климат}) = \frac{1}{(h_{климат}+1)^2}; a_{университеты}(h_{университеты}) = \frac{1}{(h_{университеты}+1)^2}$. Тогда, полагая h минимально

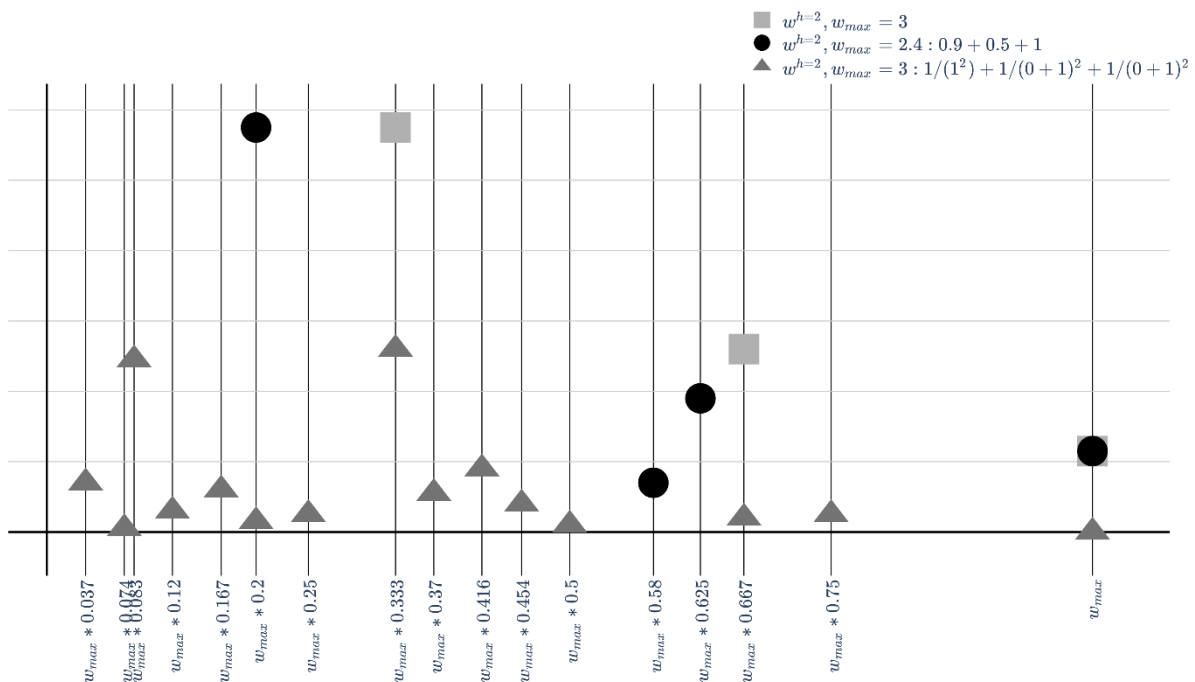


Рисунок 6. Распределение количества связей по силе связывания (величине подбора) при изменении способа расчёта w при $h = 2$

$$\text{возможным: } w_{max} = \frac{1}{1^2} + \frac{1}{(0+1)^2} + \frac{1}{(0+1)^2} = 3;$$

График на рис. 6 показывает полученное число обнаружений по силе связывания в долях максимально возможного значения для каждой функции w_{max} .

Выбор формы графика и, соответственно, коэффициентов для модальностей определяется в первую очередь задачами дальнейшей обработки результатов метода. Например, в связи с тем, что по результатам метода можно получить матрицу сходств между объектами и матрицу расстояний между ними (на основе вычисления коэффициента обратного w - чем больше сила связывания, тем меньше расстояние), то следует применить кластерный анализ. Если используемые методы кластерного анализа предполагают малое число центров кластеризации, то целевая функция должна быть такой, чтобы число центров с максимальной силой связывания было минимизировано. Если ставится задача равномерного распределения, то следует выбрать метод расчета w с использованием взвешенных коэффициентов. Чаще всего на практике встречаются распределения Пуассона той или иной степени, зависящие от параметра h для w^h .

VI. НАПРАВЛЕНИЯ ДАЛЬНЕЙШИХ ИССЛЕДОВАНИЙ

Дальнейшими задачами исследований становятся снижение числа производимых операций при добавлении новых элементов в систему (что может быть выполнено на основе функции хеширования мультимодальных объектов) и возможность минимизации числа находимых связей таким образом, чтобы оставались только лишь значимые (что может быть выполнено через построение остова графа связей мультимодальных объектов классическими алгоритмами).

VII. ЗАКЛЮЧЕНИЕ

В данной статье был показан подход к определению мультимодального объекта для представления гетерогенных реальных сущностей. Также показаны способы представления сложных связей между мультимодальными объектами. В качестве следующего шага также показана эскалация уровней мультимодальных данных, позволяющая найти корреляции между всеми задействованными мультимодальными объектами, независимо от того, были ли они реальными сущностями в начале или они представляют собой связи между ними.

Ключевым направлением данного исследования является повышение эффективности метода за счет улучшения вычисления значений весовых функций сил связывания w .

Увеличение сложности мультимодальных структур данных увеличивает количество обрабатываемых данных. Однако становится возможным представлять сложные гетерогенные данные таким образом, чтобы находить новые связи в множествах и фактическую релевантность найденных связей между объектами

разных типов.

В статье был продемонстрирован пример использования метода представления мультимодальных данных для получения косвенных связей.

БИБЛИОГРАФИЯ

- [1] L. A. Nguyen, "Multimodal logic programming" *Theoretical Computer Science*, vol. 360, no. 1, pp. 247-288, 2006.
- [2] R. Dockins, A. W. Appel and A. Hobor, "Multimodal Separation Logic for Reasoning About Operational Semantics" in *Proceedings of the 24th Conference on the Mathematical Foundations of Programming Semantics (MFPS XXIV)*, 2008.
- [3] B. Karimi и M. Bashiri, «Designing a Multi-commodity multimodal splittable supply chain network by logistic hubs for intelligent manufacturing» в *Procedia Manufacturing*, Columbus, OH, 2018.
- [4] K. Atchaneeyasakul, D. S. Liebeskind, R. Jahan, S. Starkman, L. Sharma, B. Yoo, J. Avelar, N. Rao, J. Hinman, G. Duckwiler, M. Nour, V. Szeder, S. Tateshima, G. Colby, M. B. Hosseini, R. Raychev, D. Kim и J. L. Saver, «Efficient Multimodal MRI Evaluation for Endovascular Thrombectomy of Anterior Circulation Large Vessel Occlusion» *Journal of Stroke and Cerebrovascular Diseases*, т. 29, № 12, p. 105271, 2020.
- [5] J. D. Tward, T. Schlomm, S. Bardot, D. J. Canter, T. Scroggins, S. J. Freedland, L. Lenz, D. D. Flake, T. Cohen, M. K. Brawer, S. Stone и J. Bishoff, «Personalizing Localized Prostate Cancer: Validation of a Combined Clinical Cell-cycle Risk (CCR) Score Threshold for Prognosticating Benefit From Multimodality Therapy» *Clinical Genitourinary Cancer*, 2021.
- [6] J.-E. Peng, «The roles of multimodal pedagogic effects and classroom environment in willingness to communicate in English» *System*, т. 82, pp. 161-173, 2019.
- [7] S. Philippe, A. D. Souchet, P. Lameris, P. Petridis, J. Caporal, G. Coldeboeuf and H. Duzan, "Multimodal teaching, learning and training in virtual reality: a review and case study" *Virtual Reality & Intelligent Hardware*, vol. 2, no. 5, pp. 421-442, 2020.
- [8] L. C. O. Tiong, S. T. Kim and Y. M. Ro, "Multimodal facial biometrics recognition: Dual-stream convolutional neural networks with multi-feature fusion layers" *Image and Vision Computing*, vol. 102, p. 103977, 2020.
- [9] W. Zhang, J. Yu, W. Zhao и C. Ran, «DMRFNet: Deep Multimodal Reasoning and Fusion for Visual Question Answering and explanation generation» *Information Fusion*, т. 72, pp. 70-79, 2021.
- [10] W. Zhang, J. Yu, Y. Wang and W. Wang, "Multimodal deep fusion for image question answering" *Knowledge-Based Systems*, vol. 212, p. 106639, 2021.
- [11] O. Kalyonova and I. Perl, "Revealing of entities interconnections in system dynamics modelling process by applying multimodal data analysis paradigm" in *21st*

Conference of Open Innovations Association (FRUCT), Helsinki, Finland, 2017.

- [12] G. Zhu, J. Wang, Z. Ren, Z. Lin and C. Zhang, "Object-Oriented Dynamics Learning through Multi-Level Abstraction" *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 6989-6998, 04 2020.
- [13] K. Wenzel and H. Reinhardt, "Mathematical computations for linked data applications with openmath" in *CEUR Workshop Proceedings*, 2012.
- [14] J. Guo, J. Xu, Z. He и W. Liao, «Research on risk propagation method of multimodal transport network under uncertainty» *Physica A: Statistical Mechanics and its Applications*, т. 563, p. 125494, 2021.
- [15] *Список городов России с населением более 100 тысяч жителей*, Википедия, 2021.
- [16] *Vuzoteka.ru, Вузы по городам России*, 2021.

Application of multimodal approach for identifying similarities in multi-dimensional datasets with usage example

Olga Perl, Ivan Perl

Abstract – Increasingly, one can find research from various fields devoted to working with different formats of data, tasks, research perspectives or views on the object of research. Often these studies use the term "multimodal", which at the same time varies greatly from region to region. Thus, researchers from different subject areas are faced with the same task: choosing an approach and processing multimodal data. The problem with this task lies in the narrow area of applicability of the proposed solutions. Then the very concept of multimodality needs formalization. At the same time, it becomes necessary to allocate data structures to work with this concept. It is also important to describe the approaches for analyzing data in the selected data structures. This article proposes a definition of the concept of multimodal data, describes 4 structures for working with it, and also proposes the method for identifying the most similar multimodal objects. In addition, the given structures are illustrated by examples. The method for finding similar multimodal objects is supplemented with modification descriptions so that it can be applied to all 4 multimodal data structures. The article also demonstrates the application of the method on a general example with a description of data structures - the study of cities according to the characteristics of the population, climate and the number of universities. The example of the study is for illustrative purposes only, however, it can be used for further research after verification by appropriate specialists. The article provides methods for configuring the method and recommendations for working with them. Calculated object similarity (coherence power) is a way to define a multidimensional metric over a complex data structure. At the end of the article, directions for further research are given, which are already being carried out by the authors at the present time.

Keywords – Revealing entities similarity, Information model, Multimodal data, Multimodal data structure

VIII. BIBLIOGRAPHY

- [1] L. A. Nguyen, "Multimodal logic programming," *Theoretical Computer Science*, vol. 360, no. 1, pp. 247-288, 2006.
- [2] R. Dockins, A. W. Appel and A. Hobor, "Multimodal Separation Logic for Reasoning About Operational Semantics," in *Proceedings of the 24th Conference on the Mathematical Foundations of Programming Semantics (MFPS XXIV)*, 2008.
- [3] B. Karimi and M. Bashiri, "Designing a Multi-commodity multimodal splittable supply chain network by logistic hubs for intelligent manufacturing," in *Procedia Manufacturing*, Columbus, OH, 2018.
- [4] K. Atchaneeyasakul, D. S. Liebeskind, R. Jahan, S. Starkman, L. Sharma, B. Yoo, J. Avelar, N. Rao, J. Hinman, G. Duckwiler, M. Nour, V. Szeder, S. Tateshima, G. Colby, M. B. Hosseini, R. Raychev, D. Kim and J. L. Saver, "Efficient Multimodal MRI Evaluation for Endovascular Thrombectomy of Anterior Circulation Large Vessel Occlusion," *Journal of Stroke and Cerebrovascular Diseases*, vol. 29, no. 12, p. 105271, 2020.
- [5] J. D. Tward, T. Schlomm, S. Bardot, D. J. Canter, T. Scroggins, S. J. Freedland, L. Lenz, D. D. Flake, T. Cohen, M. K. Brawer, S. Stone and J. Bishoff, "{Personalizing Localized Prostate Cancer: Validation of a Combined Clinical Cell-cycle Risk (CCR) Score Threshold for Prognosticating Benefit From Multimodality Therapy," *Clinical Genitourinary Cancer*, 2021.
- [6] J.-E. Peng, "The roles of multimodal pedagogic effects and classroom environment in willingness to communicate in English," *System*, vol. 82, pp. 161-173, 2019.
- [7] S. Philippe, A. D. Souchet, P. Lameris, P. Petridis, J. Caporal, G. Coldeboeuf and H. Duzan, "Multimodal teaching, learning and training in virtual reality: a review and case study," *Virtual Reality & Intelligent Hardware*, vol. 2, no. 5, pp. 421-442, 2020.
- [8] L. C. O. Tiong, S. T. Kim and Y. M. Ro, "Multimodal facial biometrics recognition: Dual-stream convolutional neural networks with multi-feature fusion layers," *Image and Vision Computing*, vol. 102, p. 103977, 2020.
- [9] W. Zhang, J. Yu, W. Zhao and C. Ran, "DMRFNet: Deep Multimodal Reasoning and Fusion for Visual Question Answering and explanation generation," *Information Fusion*, vol. 72, pp. 70-79, 2021.
- [10] W. Zhang, J. Yu, Y. Wang and W. Wang, "Multimodal deep fusion for image question answering," *Knowledge-Based Systems*, vol. 212, p. 106639, 2021.
- [11] O. Kalyonova and I. Perl, "Revealing of entities interconnections in system dynamics modelling process by applying multimodal data analysis paradigm," in *21st Conference of Open Innovations Association (FRUCT)*, Helsinki, Finland, 2017.
- [12] G. Zhu, J. Wang, Z. Ren, Z. Lin and C. Zhang, "Object-Oriented Dynamics Learning through Multi-Level Abstraction," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 6989-6998, 04 2020.
- [13] K. Wenzel and H. Reinhardt, "Mathematical computations for linked data applications with openmath," in *CEUR Workshop Proceedings*, 2012.

- [14] J. Guo, J. Xu, Z. He and W. Liao, "Research on risk propagation method of multimodal transport network under uncertainty," *Physica A: Statistical Mechanics and its Applications*, vol. 563, p. 125494, 2021.
- [15] *Spisok gorodov Rossii s naseleniem bolee 100 tysjach zhitelej*, Vikipedija, 2021.
- [16] Vuzoteka.ru, Vuzy po gorodam Rossii, 2021.