



Развитие матричного представления обобщенных графовых структур в задачах описания и анализа Больших Данных

.....
Приньков Алексей

д.ф.-м.н., профессор Блюмин С.Л., sabl@lipetsk.ru, +7(910)-350-44-40

студент Приньков А.С., aprinkov@gmail.com, +7(910)-250-55-40

Кафедра прикладной математики, Липецкий государственный технический университет

Цель работы

Целью данной работы является рассмотрение практических аспектов графоструктурного моделирования в задачах описания и анализа больших данных, а также развитие матричного представления обобщенных графовых структур; на этом основании будут изложены полученные результаты предварительных исследований и обозначены перспективы дальнейших.

Содержание раздела 1

1. Графоструктурное моделирование
 - 1.1 Графы и их обобщение
 - 1.2 Особенности метаграфов
 - 1.3 Матричное представление
 - 1.4 Ремоделирование и его мотивация
 - 1.5 Список литературы
2. Описание и анализ Больших Данных
 - 2.1 Алгоритм преобразования графа в метаграф
 - 2.2 Сложные системы
 - 2.3 ИИС и ГИИС
 - 2.4 Практические примеры
 - 2.5 Список литературы

Графы

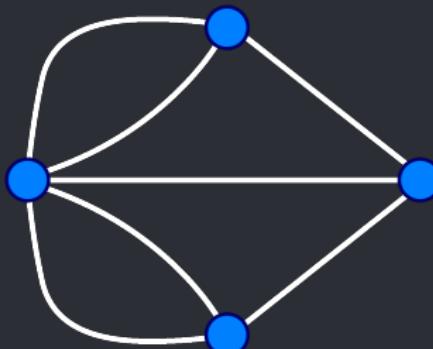
Определение функции B

$$B(V, n) = [V \rightarrow 2^V]^n,$$

где $[S]^n = \{X \mid X \subseteq S \wedge |X| = n\}.$

Формальное определение графа

Граф – это $\langle V, E \rangle$, где $E \subseteq B(V, 2)$.



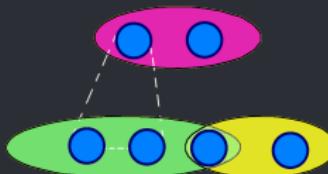
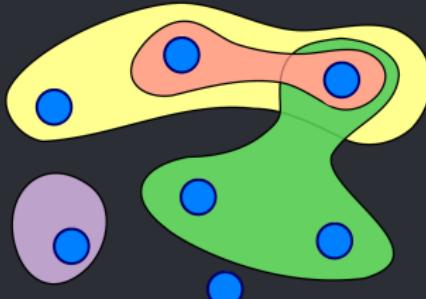
Обобщенные графовые структуры

Гиперграф

Гиперграф определен как $\langle V, HE \rangle$, где $HE \subseteq \bigcup_{i=1}^n B(V, i)$.

Гиперсеть

Гиперсеть – это пара, состоящая из множества гиперребер и множества отображений между ними $HN = \langle \{WS_i\}, \{\Phi_i\} \rangle$,
где WS_i – гиперребро, а $\Phi_i : WS_i \rightarrow WS_{i-1}$.



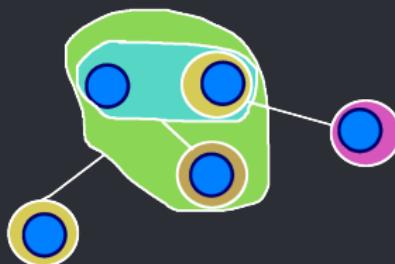
Обобщенные графовые структуры

Метаграф

Метаграф – это $MG = \langle V, ME \rangle$, где V – множество вершин, ME – множество метаребер и $ME \subseteq B(\bigcup_{i=1}^n B(V, i), 2)$.

Эквивалентное определение метаграфа

$MG = \langle V, MV, ME \rangle$, где множество метавершин $MV \subseteq \bigcup_{i=1}^n B(V, i)$, а $ME \subseteq B(MV, 2)$.



Обобщенные графовые структуры

Итергиперграфы

V – некоторое множество, $f : V \rightarrow 2^V$.

$f(V, 0) = V, f(V, 1) = f(V), f(V, 2) = f(f(V)), f(V, p) = f(f(V, p-1)),$
где p – показатель итерации и

$$e(f(V, p)) = p, e(f(V, p)) = e(f(V, p-1)) + 1.$$

Если $|V| = n, |E| = m |f(V, p)| \leq m \leq |f(V, p+1)|,$
то $e(E)$ дробный показатель итерации.

Для графа показатель итерации равен $\frac{n-m}{2^n-n}$.

Для гиперграфа и метаграфа $\frac{n-m}{2^n-n}, \frac{2^n-m}{2^{2^n}-2^n}$ соответственно.

Класс итергиперграфов охватывает все графовые структуры при варьировании степени итерации.

Особенности метаграфов

Характерной особенностью метаграфа является то, что он обобщает приведенные графовые структуры и что в нём реализованы **два вида смежности** вершин.

Причем одна из смежностей должна однозначно определять связи между вершинами.

Метавершинная

- $G\langle V, E \rangle \cong MG\langle V, MV, ME \rangle$, где $ME = \emptyset$, а $MV \subseteq B(V, 2) = E$
- $HG\langle V, HE \rangle \cong MG\langle V, MV, ME \rangle$, где $ME = \emptyset$, а $MV \subseteq \bigcup_{i=1}^n B(V, i) = HE$
- $HN\langle WS, \Phi \rangle \cong MG\langle V, MV, ME \rangle = \langle V, \{WS_i\}, \langle \dots, \langle WS_i, WS_{i-1} \rangle, \dots \rangle \rangle$

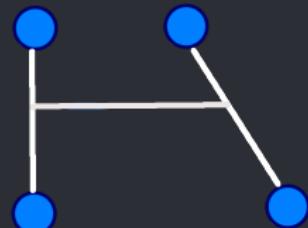
Метареберная

- $G\langle V, E \rangle \cong MG\langle V, MV, ME \rangle$, где $MV \subseteq \bigcup_{i=1}^n B(V, i)$, $ME \subseteq B(MV, 2)$
- $HG\langle V, HE \rangle \cong MG\langle V, MV, ME \rangle$, где $MV \subseteq \bigcup_{i=1}^n B(V, i)$, $ME \subseteq B(MV, 2)$
- $HN\langle WS, \Phi \rangle \cong MG\langle V, MV, ME \rangle = \langle V, \{WS_i\}, \langle \dots, \langle WS_i, WS_{i-1} \rangle, \dots \rangle \rangle$

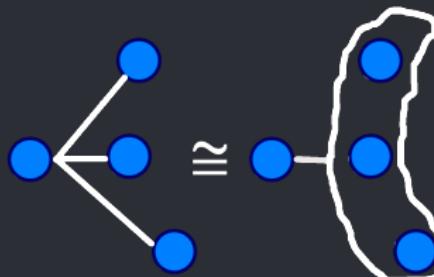
Особенности метаграфов

Вторая функция смежности может использоваться для ...

... рассмотрения связей между связями. Например, между двумя ребрами:



... кластеризации вершин. Например, для кластеризации вершин по признаку совместной смежности другой вершине:



Особенности метаграфов

Метаграф не только обобщает приведенные структуры, но и достаточен по отношению к более абстрактным структурам в задачах моделирования, т.к. ...

... связи могут быть между связями произвольного уровня.

... кластеризация вершин может быть многокритериальной.

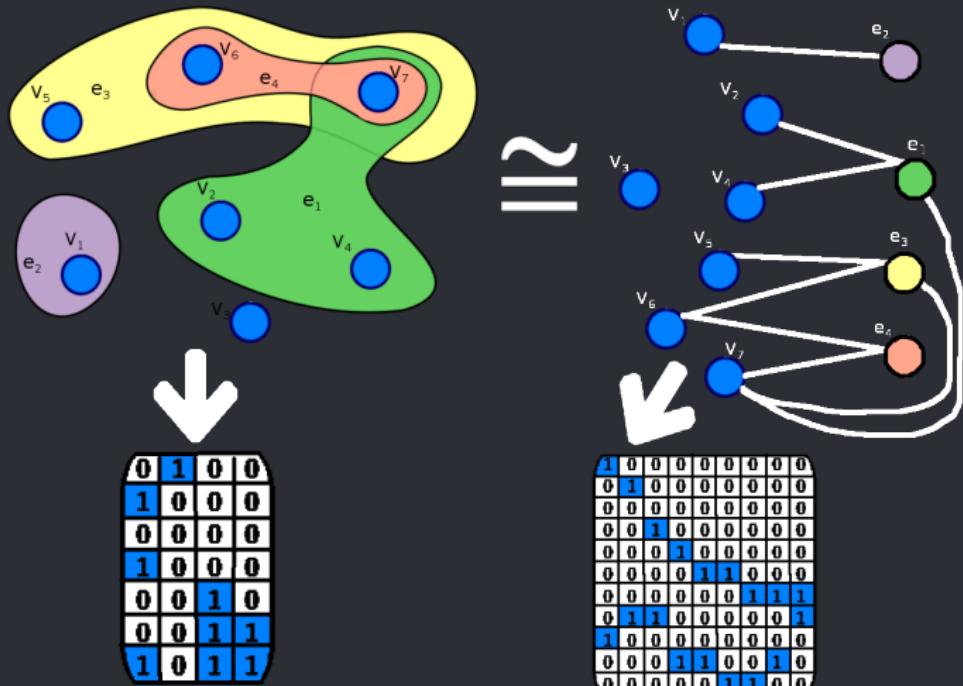
Матричное представление

Графовые структуры удобно представлять с помощью матриц. Для такого представления наиболее популярны матрицы инцидентности (I), смежности (A), валентности (D) и лапласиан (L), которые связаны равенством

$$I \cdot I^\tau = L = D \pm A.$$

Матричное представление

Также любую графовую структуру можно представить в матричном виде с помощью графа Кёнига:



Матричное представление

Матрица смежности по Басу:

$$a_{ij} = \bigcup_k (\alpha_{ij})_k, \text{ где}$$

$$(\alpha_{ij})_k = \begin{cases} \langle mu_k\{v_i\}, mv_k\{v_j\}, \langle me_k \rangle \rangle, & \text{если } v_i \in mu_k \wedge v_j \in mv_k; \\ \emptyset, & \text{иначе.} \end{cases}$$

\emptyset	\emptyset	$\langle \phi, \{x_4\}, e_1 \rangle$	$\langle \phi, \{x_3\}, e_1 \rangle$	\emptyset	$\langle \{x_1\}, \{x_1, x_4\}, \langle e_1, e_5 \rangle \rangle$	$\langle \phi, \{x_1, x_4\}, \langle e_1, e_5 \rangle \rangle$	$\langle \{x_5, \{x_5, x_1, x_7\}, \langle e_1, e_2, e_5 \rangle \rangle, \langle \{x_5, x_7\}, \{x_3, x_4, x_6\}, \langle e_1, e_4, e_5 \rangle \rangle \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	$\langle \phi, \phi, e_2 \rangle$	$\langle \{x_4\}, \{x_5\}, \langle e_2, e_4 \rangle \rangle$	\emptyset	$\langle \{x_4, x_7\}, \{x_5, x_6\}, \langle e_1, e_4, e_5 \rangle \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	$\langle \phi, \phi, e_2 \rangle$	$\langle \{x_5\}, \{x_7\}, \langle e_2, e_5 \rangle \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	$\langle \{x_5\}, \phi, e_4 \rangle$	\emptyset	$\langle \{x_5, x_7\}, \{x_6\}, \langle e_4, e_5 \rangle \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	$\langle \{x_4\}, \phi, e_4 \rangle$	\emptyset	$\langle \{x_4, x_7\}, \{x_6\}, \langle e_4, e_5 \rangle \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	$\langle \{x_7\}, \phi, e_5 \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	$\langle \{x_6\}, \phi, e_5 \rangle$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset

Матричное представление

На основе слайда 6 выводится связь между матричными представлениями разных графовых структур как частных случаев итергиперграфов. Например, для метаграфа матрица инцидентности определяется как

$$I(V, ME) = I(V, E) \cdot I(MV, ME).$$

Матричное представление

Взаимосвязь $L(V, ME)$ и $I(V, E)$, $I(MV, ME)$:

$$L(V, ME) = I(V, ME) \cdot I(MV, ME) =$$

$$= \begin{pmatrix} I_{1,1} & \dots & 0 & | & I(V, MV) & | & 0 \\ \vdots & \ddots & \vdots & | & & | & \\ 0 & \dots & \ddots & | & & | & \\ \hline & & & | & \ddots & \dots & 0 \\ & (I(V, MV))^T & & | & \vdots & \ddots & \vdots & | & I(MV, ME) \\ & & & | & 0 & \dots & \ddots & | & \\ \hline & 0 & & | & & \ddots & \dots & 0 \\ & & (I(MV, ME))^T & | & & \vdots & \ddots & \vdots \\ & & & | & 0 & \dots & I_{|V|+|MV|+|ME|, |V|+|MV|+|ME|} & \end{pmatrix},$$

$$\text{где } I_{i,i} = \sum_{k=1}^{|V+E|} |I_{i,k}| = \sum_{k=1}^{|V+E|} |I_{k,i}|.$$

Какое из представлений выбрать?!

Выбор матричного представления должен осуществляться в зависимости от постановки задач, наличия реализованных алгоритмов для данного представления, а также возможности корректной интерпретации промежуточных и конечного результата.

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Пути реализации матричного представления

Алгоритмический

+

Алгебраический

- Базовые операции линейной алгебры;
- алгоритмические операции;
- логика и алгебра высказываний;
- идемпотентные полукольца.

Преимущества матричного подхода

- Возможность оптимизации вычислений с использованием архитектуры параллельного программирования и распределенного хранения данных;
- строгая алгебраическая формализация;
- применимость средств линейной алгебры и основанных на ней методов прикладной математики;
- вариативность представления в зависимости от условий задачи.

Ремоделирование и его мотивация

Математическое ремоделирование – это подход на основе перехода от математических или имитационных моделей одного или различных типов к моделям некоторого одного класса.

Суть подхода – приведение исходных моделей к единой форме, пригодной или удобной для дальнейшего анализа и исследования.

Особенностью подхода является то, что начальной информацией является некоторая математическая или имитационная модель, а не информация об объекте или процессе моделирования.

Результаты применения гиперграфов

- Сложные системы
 - Машинное обучение
 - Реализация неинтуитивных связей
 - Относительная простота интерпретации
-
- 1 S. Tan, Z. Guan, D. Cai, X. Qin, J. Bu, and C. Chen. Mapping users across networks by manifold alignment on hypergraph. In AAAI, 2014.
 - 2 Huang, J. Scalable Hypergraph Learning and Processing [Text] / Huang J. Zhang R., Xu Yu J. // Data Mining (ICDM), 2015 IEEE International Conference on — Atlantic City, NJ, USA, 14-17 Nov. 2015.
 - 3 Zhou, D. Learning with Hypergraphs: Clustering, Classification, and Embedding [Text] / D. Zhou, J. Huang, B. Schölkopf // AINPS 19, 2007 — p. 1601-1608.
 - 4 Blyumin, S.L. Interval Cyclic Hypergraphs for Modeling of Transportation Systems [Text] / S.L. Blyumin, A.V. Galkin, D.L. Prikhodko, P.V. Saraev, A.S Sysoev // ICTTE, 2014. — Belgrade , Serbia — pp 157-163.

Список литературы

- 1 Емеличев, В.А. Лекции по теории графов [Текст] / В.А. Емеличев, О.И. Мельников, В.И. Сарванов, Р.И. Тышкевич. — М.: Наука, 1990. — 384 с.
- 2 Basu, A. Metagraphs and Their Applications [Text] / A. Basu, R. Blanning. — NY: Springer, 2007. — 172 p.
- 3 Блюмин, С.Л. Итергиперграфы: расширенный класс графовых моделей больших систем [Текст] / С.Л. Блюмин // Труды конференции «Теория активных систем-2011» (ТАС) в рамках Международной научно-практической мультиконференции «Управление большими системами» (УБС-2011). — Москва: ИПУ РАН, 2011. — Т.1. — С. 11-15.
- 4 Блюмин, С.Л. Графоструктурное моделирование. Метаграфы и их матрицы [Текст] / С.Л. Блюмин // Вестник ЛГТУ. — 2015. — № 1(23). — С. 7-13.

Список литературы

- 5 Блюмин, С.Л. Оргиперграфы: матрицы инцидентности и лапласианы [Текст] / С.Л. Блюмин // Вестник ЛГТУ. — 2013. — № 1(21). — С. 15-27.
- 6 Приньков, А.С. Графоструктурное ремоделирование метаграфами сложных систем на примере московского метрополитена [Текст] : HTCS'2017 : мат-лы XII междунар. науч.-практ. конф., 25-27 октября 2017 г. — С. 125-129.
- 7 Приньков, А.С. Разработка программного обеспечения для графоструктурного ремоделирования сложных систем [Текст] : HTCS'2017 : мат-лы XII междунар. науч.-практ. конф., 25-27 октября 2017 г.. — С. 65-69.
- 8 Etsuji, T. The worst-case time complexity for generating all maximal cliques and computational experiments / T. Etsuji, T. Akira, T. Haruhisa // Theoretical Computer Science, volume 363, issue 1, 2006 — pp 28-42.

Содержание раздела 2

1. Графоструктурное моделирование
 - 1.1 Графы и их обобщение
 - 1.2 Особенности метаграфов
 - 1.3 Матричное представление
 - 1.4 Ремоделирование и его мотивация
 - 1.5 Список литературы
2. Описание и анализ Больших Данных
 - 2.1 Алгоритм преобразования графа в метаграф
 - 2.2 Сложные системы
 - 2.3 ИИС и ГИИС
 - 2.4 Практические примеры
 - 2.5 Список литературы

Алгоритм преобразования графа в метаграф

1. Запишем граф в виде матрицы инцидентности $I(V, E)$.
2. Если матрица $I(V, E) = \emptyset$, то алгоритм завершен.
3. Находим строку в $I(V, E)$ с максимальным количеством единиц (вершину с максимальной валентностью).
4. Добавляем метаребро в $I(V, ME)$, содержащее в одной метавершине эту вершину, а в другой метавершине вершины, смежные с данной.
5. Обнуляем в $I(V, ME)$ найденную строку и столбцы, значения которых в этой строке равны единице.
6. Переходим к пункту 2.

Пример использования алгоритма

Пример. Дан граф

$G = \langle V, E \rangle = \langle \{v_1, v_2, v_3, v_4, v_5\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_2, v_4\}, \{v_3, v_4\}, \{v_4, v_5\} \rangle$. Необходимо преобразовать его в метаграф, иными словами, построить изоморфный ему метаграф.

Матрица инцидентности графа $G(V, E)$ и матрица инцидентности метаграфа $MG(V, ME)$ в начале вычислений равны

$$I(V, E) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, I(V, ME) = () .$$

Пример использования алгоритма

Первая итерация.

$$I(V, E) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$I(V, ME) = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Пример использования алгоритма

Вторая итерация.

$$I(V, E) = \left(\begin{array}{c|cccc|cc} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right) = \left(\begin{array}{ccccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right),$$

$$I(V, ME) = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Пример использования алгоритма

Третья итерация.

$$I(V, E) = \left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right) = \left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) = \emptyset,$$

$$I(V, ME) = \left(\begin{array}{ccc} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{array} \right).$$

Сложные системы

Определение сложной системы

Сложная система – система, состоящая из множества взаимодействующих подсистем, вследствие чего сложная система приобретает новые свойства, характерные для неё как для целостной структуры, которые отсутствуют не только у каждой подсистемы в частности, но и у суммы подсистем.

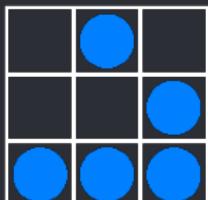
Сложные системы – эмерджентные свойства

Примеры систем с эмерджентными свойствами:

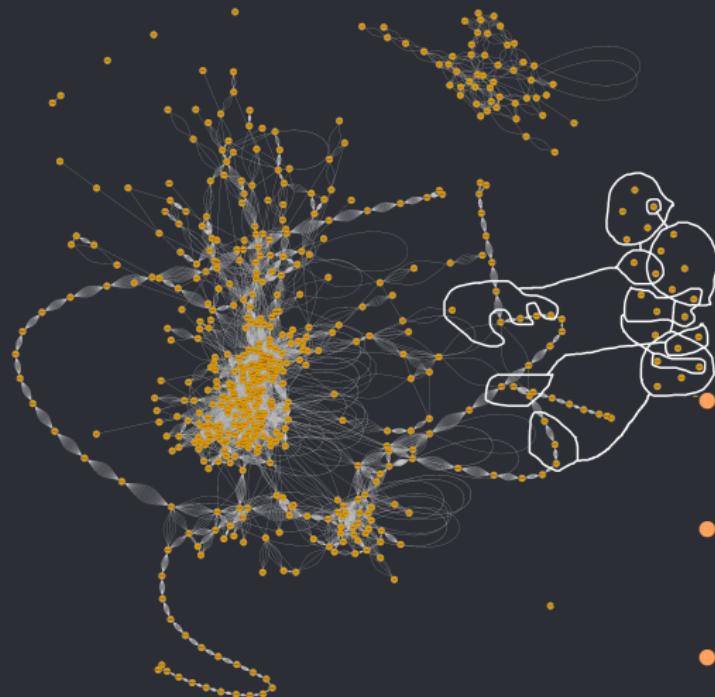
- живые организмы;
- текст на естественном языке;
- интеллект;
- социальная сеть;
- производственное предприятие.

Эмерджентные свойства не восходят к количеству.

Источником эмерджентных свойств является структура системы: при различной структуре у систем, образуемых из одних и тех же элементов, возникают разные свойства и особенности.



Сложные системы – иллюстрация



- Распределенные вычисления;
- не нужно искать клики (!NP);
- относительная простота интерпретации.

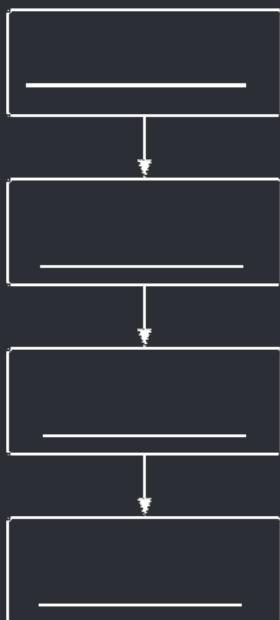
ИИС и ГИИС

Применение графоструктурного подхода к решению задач больших данных можно условно разделить на два вида – это непосредственно моделирование исходных данных графовыми структурами и моделирование информационных систем (ИС), с помощью которых решаются эти задачи.

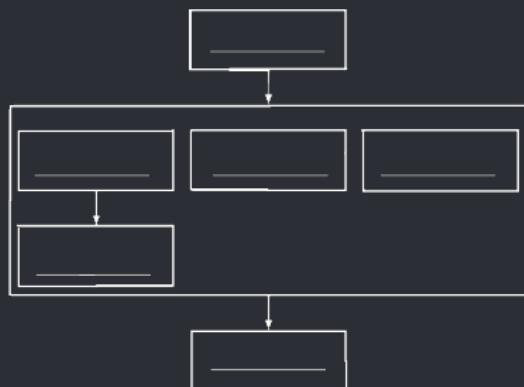
Реализация ГИИС

Программную реализацию простейших видов ГИИС можно найти в библиотеке машинного обучения sklearn.

Pipeline



FeatureUnions



Практические примеры-классификация текста

Исходные данные

PhraseId	SentenceId	Phrase	Sentiment
1	1	A series of escapades demonstrating the adage ...	1
2	1	demonstrating the adage	2
...
156060	8544	avuncular chortles	3

Источник: <https://www.kaggle.com/c/sentiment-analysis-on-movie-reviews>

Практические примеры-классификация текста

Лексический анализатор

$$x' = g(x), g : \mathbb{L} \rightarrow \mathbb{R}^{n_1}.$$

Отображение-трансформатор

$$x'_2 = h(x'_1), h : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2}.$$

Процесс предварительной обработки данных

$$h_n(\dots(h_1(g_1(x)))), g_1 : \mathbb{L} \rightarrow \mathbb{R}^{n_1}, h_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_{i+1}}.$$

Практические примеры-классификация текста

Для параллельной обработки данных необходимо использовать метаграфовую структуру ГИИС:

$$x'_g = Gx = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n-1} \\ g_n \end{pmatrix}^\tau \begin{pmatrix} x & 0 & \dots & 0 & 0 \\ 0 & x & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & x & 0 \\ 0 & 0 & \dots & 0 & x \end{pmatrix} = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_{n-1}(x) \\ g_n(x) \end{pmatrix}^\tau.$$

После чего получится разреженная матрица, которую необходимо преобразовать в векторную форму

$$x'_g = m_g(x'), m_g : \mathbb{R}^{n_1 \times n} \rightarrow \mathbb{R}^{n'},$$

Практические примеры-классификация текста

к которой можно применить отображения-трансформаторы либо последовательно

$$h_n(\dots(h_1(g_1(x_g)))), g_1 : \mathbb{L} \rightarrow \mathbb{R}^{n_1}, h_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_{i+1}},$$

либо параллельно

$$x'_h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n-1} \\ h_n \end{pmatrix}^\tau \cdot m_g \left(\begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n-1} \\ g_n \end{pmatrix}^\tau \cdot \begin{pmatrix} x & 0 & \dots & 0 & 0 \\ 0 & x & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & x & 0 \\ 0 & 0 & \dots & 0 & x \end{pmatrix} \right) = \begin{pmatrix} h_1(x'_g) \\ h_2(x'_g) \\ \vdots \\ h_{n-1}(x'_g) \\ h_n(x'_g) \end{pmatrix}^\tau$$

В случае параллельного применения трансформаторов

$$x'_h = m_h(x'), m_h : \mathbb{R}^{n_2 \times n} \rightarrow \mathbb{R}^{n'}.$$

Практические примеры-классификация текста

Построим ГИИС для предварительной обработки данных, состоящую из трех метавершин. $MG = \left\langle \{h_1, h_2, h_3, g_1, g_2, g_3, g_4\}, \langle \{\{h_2, h_3\}, \{g_1, g_2, g_3, g_4\}\}, \langle \{h_1\}, \{h_2, h_3, g_1, g_2, g_3, g_4\} \rangle \rangle \right\rangle$, где g_1, g_2, g_3, g_4 – лексические анализаторы, а h_1, h_2, h_3 – трансформаторы.

- g_1 = tfidf (N-граммы по словам);
- g_2 = tfidf (N-граммы по буквам);
- g_3, g_4 – «счетчики»;

h_1, h_2, h_3 – трансформаторы для масштабирования входных векторов до единичной нормы с использованием l_1, l_2, \max нормы соответственно.

Результат: 84% train, 67% test.

Список литературы

- 1 Блюмин, С.Л. Графоструктурные тенденции развития ИИС: применение гиперграфов, метаграфов, итерграфов и их матричных представлений [Текст] / С.Л. Блюмин, А.С. Приньков // Проблемы фунд. и прикладной информ. в управ., автомат. и мехат. — Курск, Юго-Зап. гос. ун-т, 2017 — С. 5-13
- 2 Черненький В.М. Метаграфовый подход для описания гибридных интеллектуальных информационных систем [Текст] / В.М. Черненький, Ю.Е. Гапанюк, Г.И. Ревунков, В.И. Терехов, Ю.Т. Каганов // Прикладная информатика. — Москва, — Том 12. №3 (69). 2017. — С 57-79.
- 3 Черненький, В.М. Структура гибридной интеллектуальной информационной системы на основе метаграфов [Текст] / В.М. Черненький, В.И. Терехов, Ю.Е. Гапанюк // Нейрокомпьютеры: разработка, применение , 2016. — С. 3-13.
- 4 Drexl M. On the generalized directed rural postman problem / M. Drexl // Journal of the Operational Research Society — NY: Springer — August 2014, Volume 65, Issue 8, — pp 1143–1154

Список литературы

- 5 Черненький В.М. Представление сложных сетей на основе метаграфов [Текст] / В.М. Черненький, В.И. Терехов, Ю.Е. Гапанюк // Нейроинформатика-2016. XVIII Всероссийская научно-техническая конференция. — Сб. науч. трудов. Ч. 1. М.: НИЯУ МИФИ, 2016. — С. 173 – 178.
- 6 Sikora, F. The shortest way to visit all metro lines in Paris [Text] / F. Sikora // Preprint arXiv:1709.05948
- 7 Etsuji, T. The worst-case time complexity for generating all maximal cliques and computational experiments / T. Etsuji, T. Akira, T. Haruhisa // Theoretical Computer Science, volume 363, issue 1, 2006 — pp 28-42.
- 8 Traud, A. Social structure of Facebook networks [Text] / A. Traud, P. Mucha, M. Porter // Preprint arXiv:1102.2166.

Спасибо за внимание!