Различные алгоритмы выделения остовного дерева и проверки гиперграфа на связность

Сыров Денис Игоревич, гр. 522

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: д. ф.-м. н., профессор Ю. А. Сушков
Рецензент: асп. Г. С. Тамазян

Санкт-Петербург
2012г.

Постановка задачи

Существует подход к моделированию систем, описанный в работах Сушкова Ю.А., в котором в качестве моделей выступают q-униформные гиперграфы. Ключевым понятием при анализе модели является связность гиперграфа. При построении теории связности гиперграфов одними из основных понятий являются лес, дерево и цикл. Описанный Сушковым Ю.А. подход замечателен тем, что он является обобщением понятия связности графов на случай гиперграфов. В работе рассматривалось $q = 3$.

Цель работы

- *Найти полиномиальные алгоритмы, которые позволяли бы определять, является ли гиперграф лесом или деревом, и выделять компоненты связности гиперграфа.*
- *Теория должна работать в случае графов.*

Необходимые определения

Определение

Гиперграф G – это пара (V, E), где V – множество вершин, а $E \subseteq 2^V$.
Если $\forall e \in E$ выполняется $|\Gamma e| = q$, то G называют q-униформным.

Тут и далее ΓS обозначает множество вершин, инцидентных множеству ребер S.
Для определения понятий дерева и леса очень важным является следующее соотношение:

$$\forall S \subseteq E : |S| \leq |\Gamma S| - q + 1 \qquad (1)$$

Определение

Лес – это гиперграф $W = (V, E)$, множество ребер которого удовлетворяет неравенству (1).

Необходимые определения

Определение

Каркас гиперграфа G – это подгиперграф W такой, что он является максимальным по включению лесом.

Вообше говоря, у гиперграфа, как и у графа, может существовать несколько каркасов.

Определение

Дерево – это гиперграф $T = (V, E)$, который является лесом и для множества ребер E которого выполняется равенство $|E| = |\Gamma E| - q + 1$.

Определение

Цикл – это гиперграф $C = (V, E)$, все подгиперграфы которого являются лесами, а сам C не лес.

Необходимые утверждения

Теорема (Ю.А. Сушков, 2002)

Всякий лес W можно представить как $W = T_1 + ... + T_k$, где $\forall i \in 1..k$ выполняется: T_i – дерево и T_i максимальны по включению.

- T_i из теоремы 1 называются компонентами связности леса W.

Теорема (Ю.А. Сушков, 2002)

1) Если в гиперграфе G зафиксировать каркас W, то $G = D_1 + ... + D_m$, где $D_i = T_i + H_i$, T_i – компоненты связности каркаса W, а H_i такие, что $\forall h \in H_i$ выполняется $|T_i + h| > |\Gamma(T_i + h)| - q + 1$ и $|T_j + h| \leq |\Gamma(T_j + h)| - q + 1$ при $i \neq j$.
2) Разбиение гиперграфа $G = D_1 + ... + D_m$ не зависит от выбора каркаса.

- D_i из теоремы 2 называются компонентыми связности гиперграфа G.

Связи между ребрами гиперграфов устроены сложней, чем у графов

Пример

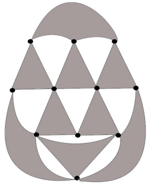

Рис. 1
Дерево

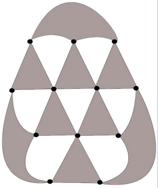

Рис. 2
Лес

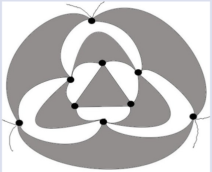

Рис. 3
Дерево

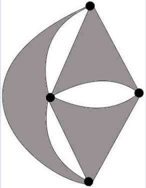

Рис. 4
Цикл

Уже из приведенных примеров видно, что свойства ребер иметь общие вершины и быть в одной компоненте связности для гиперграфов – не одно и то же. В связи с этим компоненты связности в гиперграфах устроены сложнее, чем в графах.

Проверка на свойство быть лесом

То, что возможна эффективная проверка гиперграфа на свойство быть лесом, обосновывается следующей теоремой:

Теорема (А.Ш. Абакаров, Ю.А. Сушков, 2006)

3-граф $G = (V, E)$ – лес тогда и только тогда, когда в его Кениговом представлении G_k при удалении любой пары вершин p, q, где p, q соответствуют вершинам G, в $G_k - \{p, q\}$ существует полное паросочетание.

- Полученный результат применим к графам.
- Полиномиальная сложность алгоритма.

В дипломной работе получен следующий вероятностный аналог теоремы Абакарова-Сушкова:

Теорема

Рассмотрим 3-граф $G = (V, E)$, A – его матрица смежности, A' получена из A заменой ненулевых элементов на независимые равномерно распределенные на $[0, 1]$ случайные величины. Гиперграф G является лесом тогда и только тогда, когда для всякой пары столбцов c_1, c_2 почти наверное вектор-строки матрицы $A' - c_1 \cup c_2$ являются линейно независимыми.

Тут $A' - c_1 \cup c_2$ обозначает матрицу A', из которой удалили столбцы c_1 и c_2.

- Полученные результаты работают в случае графов.
- Алгоритм прост в реализации.
- Полиномиальная сложность.

В качестве частного случая предыдущей теоремы сформулирована и доказана следующая теорема:

Теорема

Рассмотрим 3-граф $G = (V, E)$, A – его матрица смежности, A' получена из A заменой ненулевых элементов на независимые равномерно распределенные на $[0, 1]$ случайные величины. Гиперграф G является деревом тогда и только тогда, когда для всякой пары столбцов c_1, c_2 почти наверное определитель матрицы $A' - c_1 \cup c_2$ отличен от 0.

- Этот случай интересен тем, что когда случайные величины являются дискретными, то возможна оценка вероятности ошибки алгоритма.

Теорема (Д. Шварц, 1980)

Вероятность обращения в ноль ненулевого многочлена $p(x_1, ..., x_j)$ степени k при подстановке независимых равномерно распределенных случайных величин, которые могут принимать s значений, оценивается сверху как $P[p(x_1, ..., x_j) = 0] \leq \frac{k}{s}$.

Раскраска гиперграфа

После того как выделен каркас, нужно разбить его на компоненты связности.
В теории графов для выделения компонент связности используют алгоритм раскраски.

- В случае гиперграфов не удается выделить всю компоненту связности, просто переходя по общим вершинам.
- Компоненты связности устроены сложно, поэтому нужно запоминать их структуру при раскрашивании.
- Можно использовать различные цвета при раскрашивании текущей компоненты.
- После того как компоненты выделена, можно забыть про её сложность и структуру, просто окрасив её в уникальный цвет.

Раскраска леса

Как покрасить все, что нужно, и ничего лишнего?

- Нужно следить за степенью свободы выделенной компоненты, то есть за разностью $|S| - |\Gamma S|$
- Ключевое соотношение при этом также $|S| \leq |\Gamma S| - q + 1$.

На основании описанных правил получен алгоритм раскраски, который выделяет компоненты связности в каркасе.

Выделение компонент связности

После выделения компонент связности каркаса нужно распределить по компонентам связности ребра, не вошедшие в каркас.

- По теореме о разбиении гиперграфа на компоненты связности найдется компонента связности каркаса, содержащая все вершины такого ребра.
- Для того, чтобы её найти, достаточно одного просмотра всех ребер.
- Таким образом, нахождение компонент связности самого гиперграфа имеет полиномиальную сложность.

Заключение

- В результате проведенной работы найдены алгоритмы, проверяющие гиперграф на то, что он является деревом.
- Получены алгоритмы проверки того, что гиперграф является лесом.
- Для проверки гипотез и утверждений был использован пакет R.
- Построен алгоритм раскраски, который позволяет выделять компоненты связности произвольного гиперграфа.
- Все найденные алгоритмы имеют полиномиальную сложность.
- Результаты применимы к теории графов.