

Лекция 5

Алгоритмы на графах.

Введение в графы и базовые алгоритмы на графах

Анализ и разработка алгоритмов



УНИВЕРСИТЕТ ИТМО

- 1 Графы и их приложения
- 2 Дерево и лес
- 3 Представления графов
- 4 Графы реального мира
- 5 Поиск в глубину и его приложения
- 6 Поиск в ширину и его приложения

Графы и их приложения

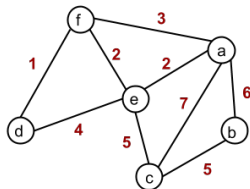
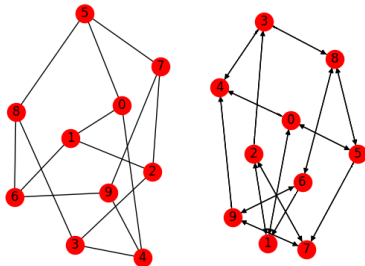
Неориентированным графом называют пару $G = (V, E)$, где $V = \{v_i\}$ — множество вершин (или узлов), и $E = \{e_{ij}\} = \{(v_i, v_j)\}$ — множество пар вершин, называемых *ребрами* (или *связями*).

Число вершин обозначают через $|V|$, а число ребер — через $|E|$.

Оrientированный граф — это граф, в котором ребра имеют направления (ориентации).

Взвешенный граф — это граф, в котором каждому ребру приписаны *веса*.

Простой граф — это граф, в котором возможно только одно ребро между парой вершин. *Мультиграф* — это обобщение простого графа, при котором в графе возможны несколько ребер между парой вершин.



Графы и их приложения

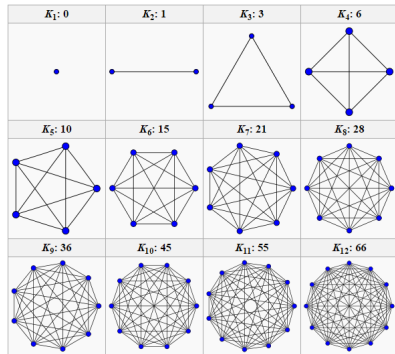
Полный граф — это граф, в котором все вершины соединены ребром.

Путь (цепь) в графе — это последовательность попарно различных ребер, соединяющих попарно различные вершины. *Длиной пути (цепи)* называют количество ребер в пути (цепи).

Вершины v_1 и v_2 в графе называются *связанными*, если существует путь из v_1 в v_2 . В противном случае, эти вершины называются *несвязанными*.

Связный граф — это граф, в котором любая пара вершин связана. В противном случае граф называют *несвязным*.

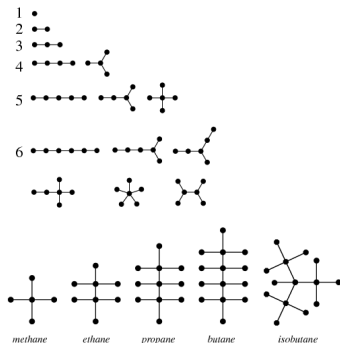
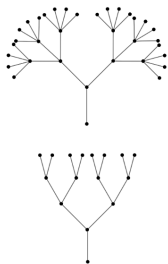
Мы будем обычно рассматривать **простые неориентированные (взвешенные или невзвешенные) графы**.



Демонстрация: Приложения графов в реальной жизни

Дерево — граф специального вида

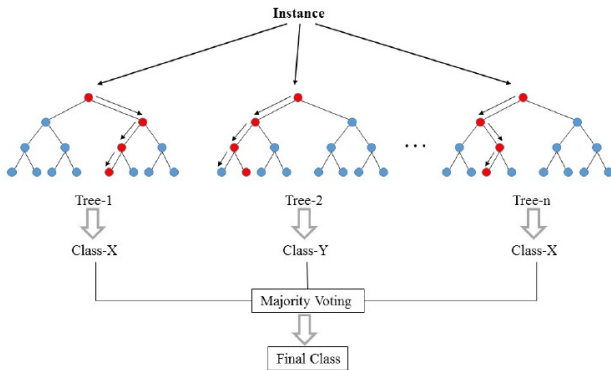
Дерево — это неориентированный граф, в котором между любыми двумя вершинами существует ровно один путь.



Демонстрация: Дерево решений, Фрактальное дерево

Лес — граф специального вида

Лес — это неориентированный граф, в котором между любыми двумя вершинами существует не более одного пути или, эквивалентно, это непересекающееся объединение деревьев.



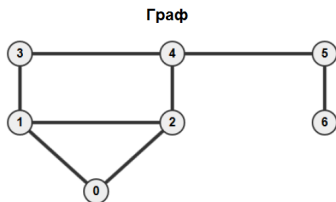
Демонстрация: Классификатор на основе случайного леса

Представления графа: матрица и список смежности

Матрица смежности — это матрица, строки и столбцы которой индексируются вершинами и ячейки которой содержат булево значение (0 или 1), указывающее, являются ли соответствующие вершины смежными (для взвешенных графов вместо 1 стоят соответствующие веса). Матрица смежности (как 2D массив) требует $O(|V|^2)$ памяти.

Список смежности — это множество списков с перечнем смежных вершин. Список смежности (как 1D массив списков) требует $O(|V| + |E|)$ памяти.

Для *разреженного графа*, т.е. графа, в котором большинство пар вершин не связаны ребрами, $|E| \ll |V|^2$, список смежности значительно эффективнее для хранения, чем матрица смежности.



Матрица смежности

	0	1	2	3	4	5	6
0	0	1	1	0	0	0	0
1	1	0	1	1	0	0	0
2	1	1	0	0	1	0	0
3	0	1	0	0	1	0	0
4	0	0	1	1	0	1	0
5	0	0	0	0	1	0	1
6	0	0	0	0	0	1	0

Список смежности

Вершина	Смежные вершины
0	1, 2
1	0, 2, 3
2	0, 1, 4
3	1, 4
4	2, 3, 5
5	4, 6
6	5

Графы (сети) реального мира

Ссылка 1: Stanford Large Network Dataset Collection

Ссылка 2: Network Repository

Data & Network Collections. Find and interactively [VISUALIZE](#) and [EXPLORE](#) hundreds of network data

ANIMAL SOCIAL NETWORKS	816	INTERACTION NETWORKS	29	SCIENTIFIC COMPUTING	11
BIOLOGICAL NETWORKS	37	INFRASTRUCTURE NETWORKS	8	SOCIAL NETWORKS	77
BRAIN NETWORKS	116	LABELED NETWORKS	104	FACEBOOK NETWORKS	114
COLLABORATION NETWORKS	19	MASSIVE NETWORK DATA	21	TECHNOLOGICAL NETWORKS	12
CHEMINFORMATICS	646	MISCELLANEOUS NETWORKS	2668	WEB GRAPHS	33
CITATION NETWORKS	4	POWER NETWORKS	8	DYNAMIC NETWORKS	115
ECOLOGY NETWORKS	6	PROXIMITY NETWORKS	13	TEMPORAL REACHABILITY	38
ECONOMIC NETWORKS	16	GENERATED GRAPHS	221	BHOSLIB	36
EMAIL NETWORKS	6	RECOMMENDATION NETWORKS	36	DIMACS	78
GRAPH 500	8	ROAD NETWORKS	15	DIMACS10	84
HETEROGENEOUS NETWORKS	15	RETWEET NETWORKS	34	NON-RELATIONAL ML DATA	211

Поиск в глубину (DFS) и его приложения

Поиск в глубину (Depth-first search, DFS)

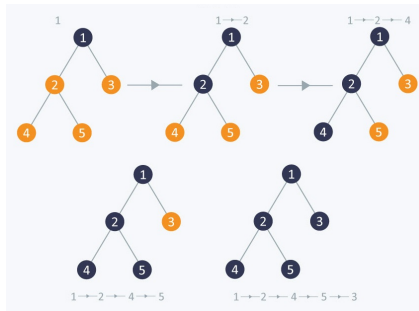
— это алгоритм обхода графа, который начинает обход с выбранной корневой вершины и идет «вглубь» графа, насколько это возможно, перед обходом из новой вершины.

Применение: поиск связных компонент, поиск циклов в графе и т.д.

Временная сложность: $O(|V| + |E|)$.

Компонента связности графа — максимальный связный подграф графа.

Демонстрация: DFS и поиск связных компонент



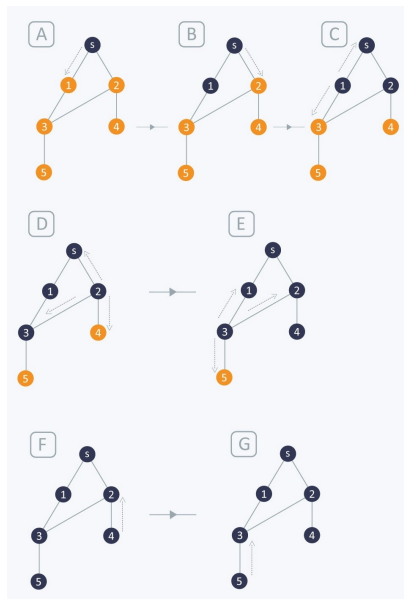
Поиск в ширину (BFS) и его приложения

Поиск в ширину (Breadth-first search, BFS) — это алгоритм обхода графа, в котором обход начинается с выбранной корневой вершины и исследуются все смежные вершины на текущей «глубине», прежде чем перейти к вершинам на следующей «глубине». Этот алгоритм обхода использует **стратегию, противоположную DFS**.

Применение: поиск кратчайшего пути между вершинами

Временная сложность: $O(|V| + |E|)$.

Демонстрация: BFS



Спасибо за внимание!