Advanced Techniques for Combinatorial Algorithms: Randomized Algorithms

Gianluca Della Vedova

Univ. Milano–Bicocca
https://gianluca.dellavedova.org

May 18, 2020

Karp-Rabin

Find all occurrences of Pattern in a Text

- hash function $H : \Sigma^{|P|} \mapsto \mathbb{N}$
- hash conflicts = false positives
- speedup: simple hash function
- $H(S) = \sum_{i=1}^{|S|} 2^{|S|-i} H(S[i]) mod p$, with p a large random prime
- $|P|$-long sliding window on T
- $H(T[i+1 : i+|P|]) = \left(\left(H(T[i : i+|P|-1]) - T[i] \right) / 2 + 2^{|P|} T[i+|P|] \right)$

Karp-Rabin: false positives

Kinds of error

- False positive (FP): reported false occurrence
- False negative (FN): occurrence not found
- No False negative in Karp-Rabin

Decreasing error probability

Choosing k random primes (independently, without repetitions).

Las Vegas vs. Monte Carlo

Classifying randomized algorithms

- Las Vegas:
  - Always correct
  - Sometimes not fast
  - Example: Quicksort with random pivot
- Monte Carlo:
  - Always fast
  - Sometimes not correct
  - Karp-Rabin

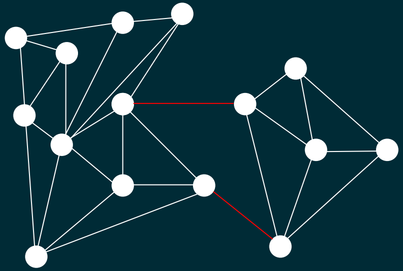
Min Cut

Instance

Undirected graph $G = \langle V, E \rangle$.

Question

Find a smallest subset $C \subset E$ of edges such that its removal from G results in a disconnected graph.



Min Cut: Karger's algorithm

Algorithm 1: Karger's Algorithm

Input: undirected unweighted graph $G = \langle V, E \rangle$ **while** G *has at least two vertices* **do**
- Pick a random edge e;
- Contract the edge into a single vertex, allowing parallel edges;

Output E

Observation

At the end of the algorithm, E is a cut.

Min Cut: Analysis

- C a minimum cut
- Probability that random edge $\notin C$: $1 - \frac{|C|}{|E|} \geq 1 - \frac{2}{n}$ (each node has degree $\geq |C|$)
- Iteration i: probability $p_i \geq \left(1 - \frac{2}{n}\right) p_{i-1}$
- Probability that the algorithm computes a min cut is at least $\binom{n}{2}^{-1}$

Observation

Repeat the algorithm.

The probabilistic method

Shows that an object exists

If the probability that an object exists is > 0

Proposition

Let K_n be the complete graph with n vertices. If $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ then we can color the edges of K_n so that we have no monochromatic K_k subgraph.

The probabilistic method

Proposition

Let K_n be the complete graph with n vertices. If $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ then we can color the edges of K_n so that we have no monochromatic K_k subgraph.

Proof

1. There are $2^{\binom{n}{2}}$ colorings of K_n.
2. Color each edge at random
3. There are $\binom{n}{k}$ k-cliques
4. Let A_i be the event that k-clique i is monochromatic
5. $P[A_i] = 2^{1-\binom{k}{2}}$
6. $P[\bigcup A_i] \leq \sum P[A_i] = \binom{n}{k}2^{1-\binom{k}{2}} < 1$
7. $P[\bigcap \bar{A}_i] = 1 - P[\bigcup A_i] > 0$

The expectation method

Shows that an object exists

Let $E[X]$ be the expected value of event X. Then there exists an event with value $\leq E[X]$ and an event with value $\geq E[X]$

Max Cut

- Let $G = (V, E)$ be an undirected graph, with $|V| = n$, $|E| = m$.
- The cut associated with the bipartition (V_1, V_2) of V is $E \cap V_1 \times V_2$.
- There exists a cut with at least $m/2$ edges.

The expectation method

Random cut

- For each vertex v, assign v to a side with $p = 1/2$

Proof

- $X_i = 1$ if edge i in the cut C, else 0
- Probability of each edge in the random cut C: $1/2$. $E[X_i] = 1/2$
- $E[\sum X_i] = \sum E[X_i] = m/2$

Question

The algorithm?

The expectation method

Random cut

$q = P[|C| \geq m/2]$, for random cut C. Then $q \geq \frac{1}{m/2+1}$

Proof

$$\frac{m}{2} = E[C] = \sum_{i \leq m/2-1} i \cdot P[|C| = i] + \sum_{i \geq m/2} i \cdot P[|C| = i] \leq$$

$$\leq (m/2 - 1)(1 - q) + qm \Rightarrow q \geq \frac{1}{m/2 + 1}$$

Question

How many random cuts are needed?

The sample and modify method

Independent set

G: undirected graph. Average degree $d = 2m/n$

Algorithm 2: Random Independent Set

$S \leftarrow$ a random sample of V, each vertex is picked with $p = 1/d$;
$I \leftarrow S$;
while *there exists an edge e in* $G|I$ **do**
    Remove an endpoint of e from I

The sample and modify method

Bounding $|I|$

- $E(G) = \frac{nd}{2}$
- $E[|S|] = \frac{n}{d}$
- $E[E(G|S)] = \frac{nd}{2}\frac{1}{d}\frac{1}{d} = \frac{n}{2d}$
- At most $E(G|S)$ are removed in the second step, hence $E[|I|] = E[|S|] - E[E(G|S)] = \frac{n}{2d}$