Color-Coding Method

Noga Alon | Raphael Yuster | Uri Zwick

Darshak Patel

COMP 5112/COMP4900G: Algorithms for Data Science

27-Nov-2024

Introduction

- **Purpose**: Introduce a technique, "Color-Coding" for finding simple paths and cycles and subgraph for specified length in polynomial time.

- **Key Methods to discussed today**:
  - Random Coloring

- A novel randomized method, for finding simple paths and cycles of a specified length k, and other small subgraphs, within a given graph G = (V, E).

Random Coloring

- Let G = (V, E) be a graph (directed or undirected).
- Choose a random coloring of the vertices of G with k colors.
- A **colorful path** is defined as a path where all vertices have distinct colors.
- *This looks very simple.*
- Each simple path of length k-1, has a chance to become colorful.

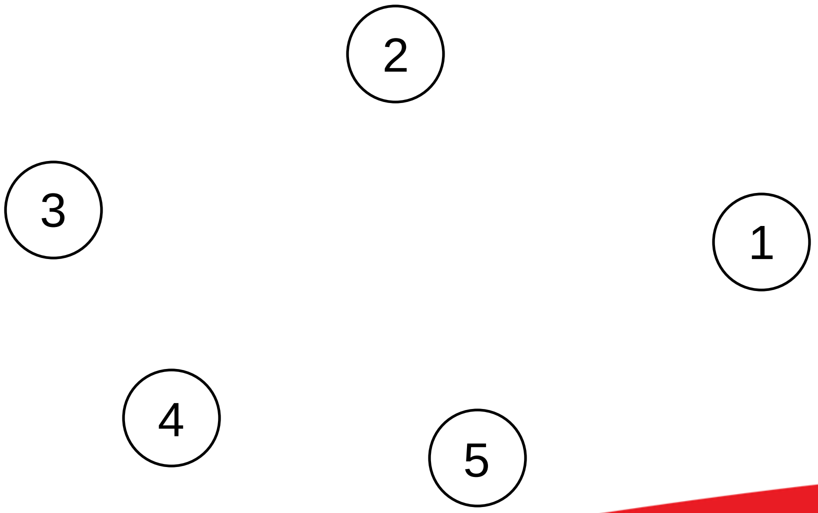
$$\frac{k!}{k^k} > e^{-k}$$

- How much time is needed to find a colorful path of length k-1 in G, if one exists, or all pairs of vertices connected by colorful paths of length k-1 in G?

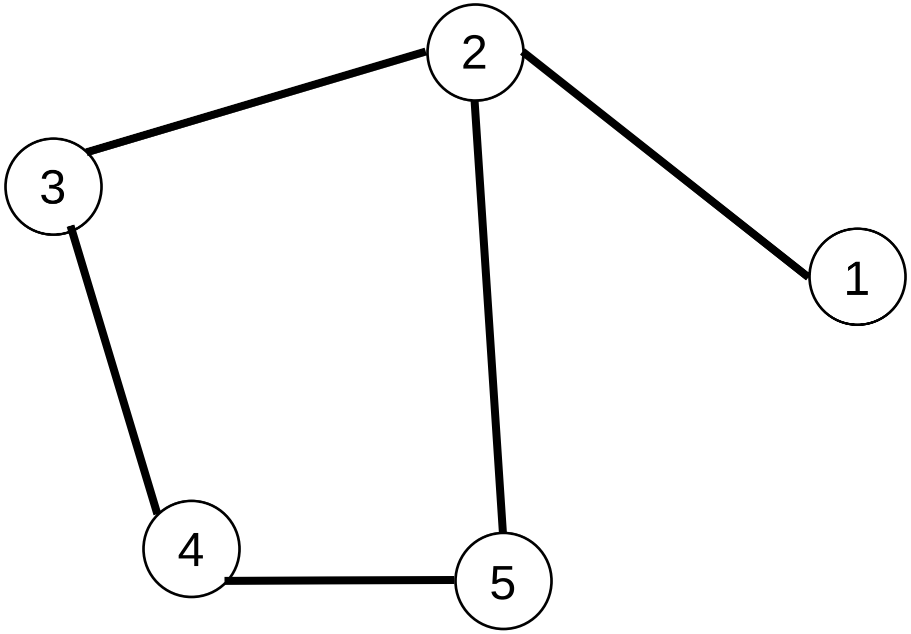
Lemma

Lemma 3.1 *Let $G = (V, E)$ be a directed or undirected graph and let $c : V \to \{1, \ldots, k\}$ be a coloring of its vertices with k colors. A colorful path of length $k - 1$ in G, if one exists, can be found in $2^{O(k)} \cdot E$ worst-case time.*

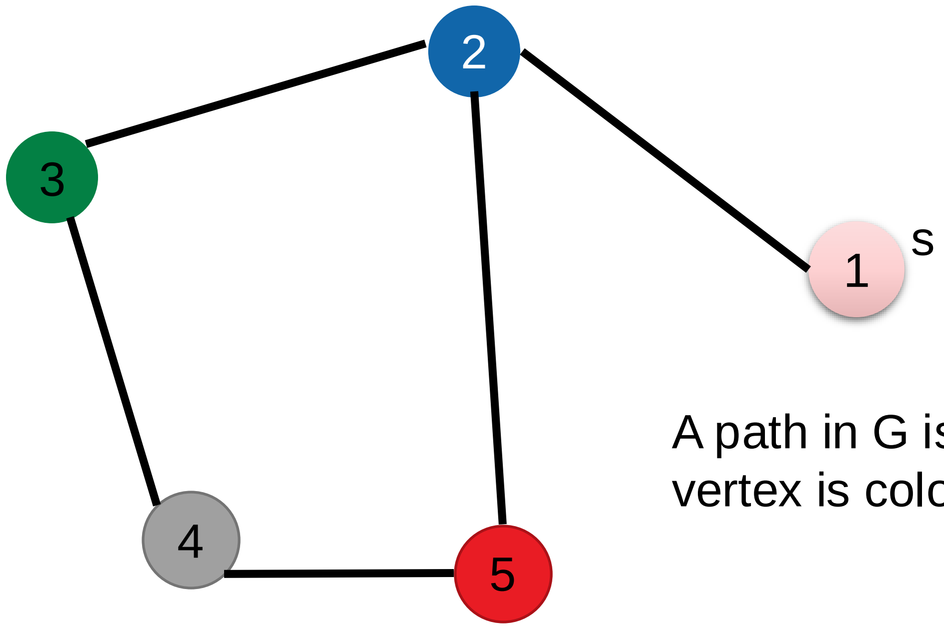
Graph G = (V, E) , Here V=5



Graph G = (V, E) , Here V=5, E=5



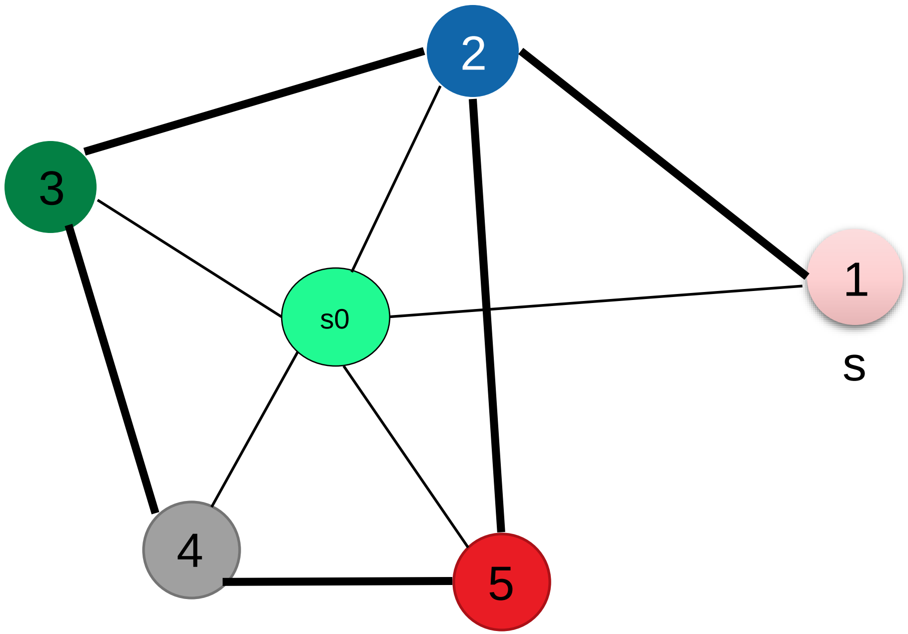
Choose a random coloring of the vertices of G with k colors, c : V -> {1,2,..,k}; k=5 and a vertex s belongs to V.



Finds a colorful path of length k -1 that starts at s, if one exists.

A path in G is said to be colorful if each vertex is colored by a distinct color.

To find a colorful path of length k -1 in G that starts somewhere, we just add a new vertex s0 to V.



We just add a new vertex s0 to V.

Color it with a new color 0 and connect it with edges to all the vertices of V.

We now look for a colorful path of length k that starts at s0.

Lemma Proof

- Example with V = 5 and E = 5
- Let us now apply the algorithm to a specific graph with 5 vertices and 5 edges.
- Consider the graph G = (V, E), where:
  - V = {v1, v2, v3, v4, v5},
  - E = {(v1, v2), (v2, v3), (v3, v4), (v4, v5), (v5, v1)}.

Proof (cont.)

- The graph is a cycle of 5 vertices, and we assign the following coloring:
- $c(v1) = 1$, $c(v2) = 2$, $c(v3) = 3$, $c(v4) = 4$, $c(v5) = 5$.
- We are looking for a colorful path of length 4,
- i.e., a path that uses all 5 colors.
- We can apply the dynamic programming approach described above.

Proof (cont.)

- Step-by-Step Application of the Algorithm
- 1. **Initialization**: Initially, we know that the path starting at v1 uses only the color 1:

$$C(v1, 0) = \{1\},\ C(v2, 0) = \emptyset,\ C(v3, 0) = \emptyset,\ C(v4, 0) = \emptyset,\ C(v5, 0) = \emptyset.$$

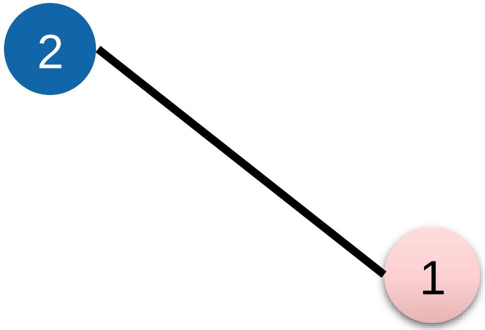
Proof (cont.)

- 2. **Paths of Length 1**: From v1,
- We can extend the path to v2, adding color 2:

$$C(v1, 1) = \{1\}, C(v2, 1) = \{1, 2\}.$$

- For other vertices:

$$C(v3, 1) = \emptyset, C(v4, 1) = \emptyset, C(v5, 1) = \emptyset.$$



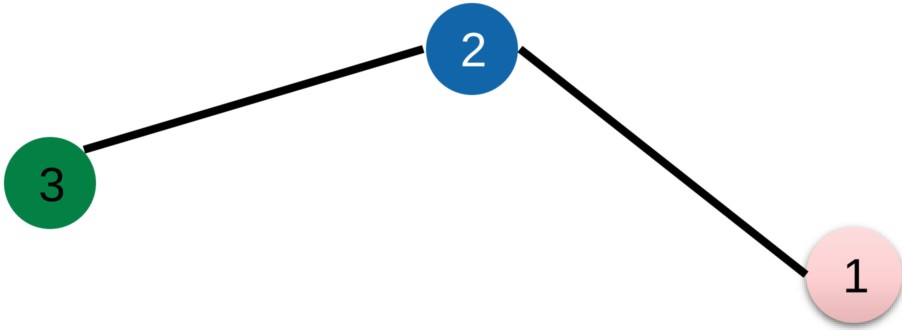
Proof (cont.)

- 3. **Paths of Length 2**: From v2,
- we extend the path to v3, adding color3:

  $$C(v3, 2) = \{1, 2, 3\}.$$

- For other vertices:

  $$C(v4, 2) = \emptyset, C(v5, 2) = \emptyset.$$



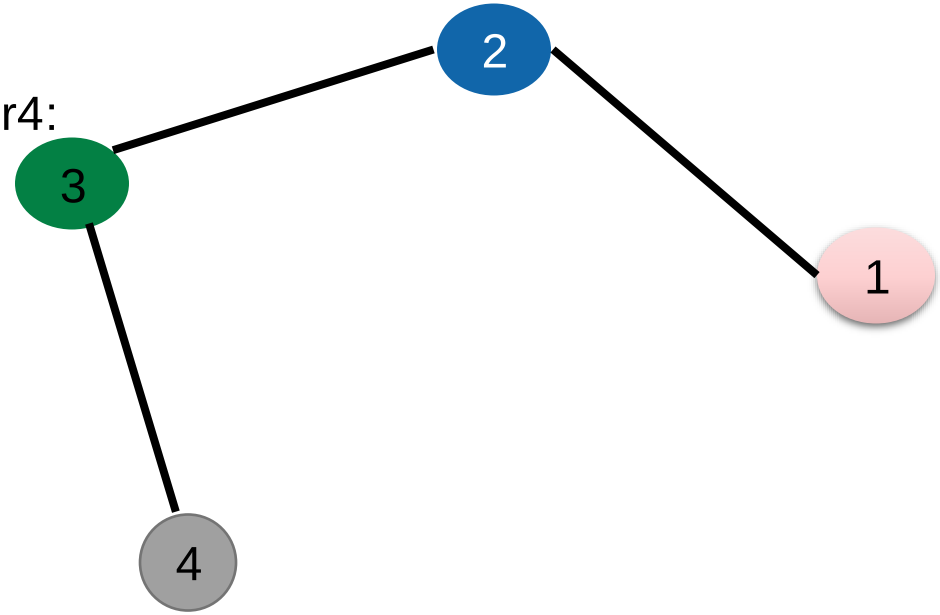
Proof (cont.)

- 4. **Paths of Length 3**: From v3,
- we extend the path to v4, adding color4:

  C(v4, 3) = {1, 2, 3, 4}.

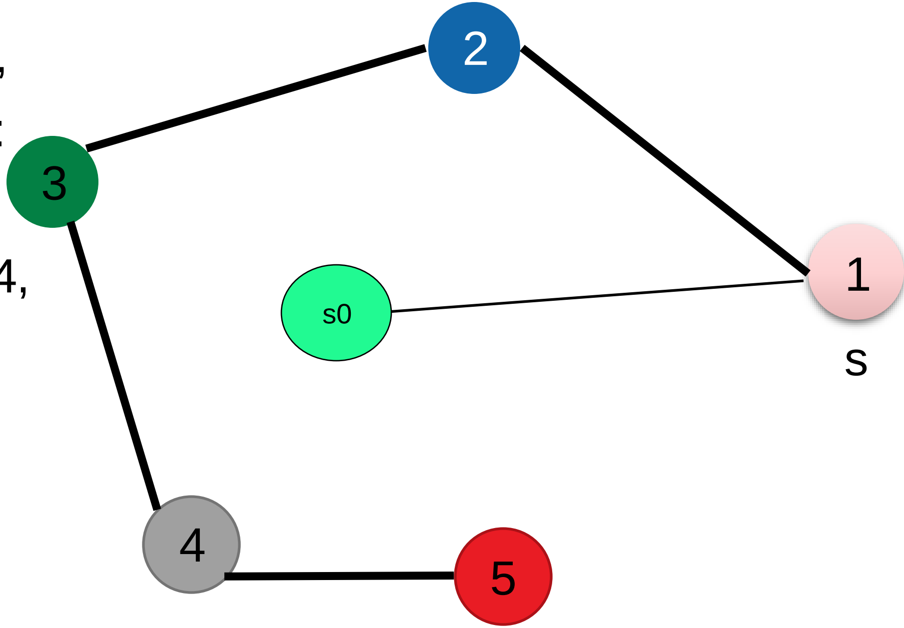
- For other vertices:

  C(v5, 3) = ∅.



Proof (cont.)

- 5. **Paths of Length 4**: Finally, from v4,
- we extend the path to v5, adding color 5:
- C(v5, 4) = {1, 2, 3, 4, 5}.
- We have found a colorful path of length 4,
- which uses all the colors 1, 2, 3, 4, 5.
- The path is : v1 → v2 → v3 → v4 → v5.



Proof (cont.)

- The graph G contains a colorful path of length k − 1 with respect to the coloring c if and only if the final collection corresponding to paths of length k − 1 for at least one vertex is non-empty.
- The number of operations performed by the algorithm is at most

$$O\left(\sum_{i=0}^{k-1} k^i \cdot |E|\right),$$

which is clearly $O(k2^k \cdot E)$. $\square$

Class Notes: *13.4.7 Color Coding*

- **Consider the longest path problem in an undirected graph.**
- **Input:** An undirected simple graph $G = (V, E)$ on n vertices and a positive integer $k \leq n$.
- **Output:** Does there exists a simple path consisting of at least k vertices in G.

Color coding for finding long paths

- Step 1: Repeat Steps 2 and 3 for $O(k^k \ln n)$ times.
- Step 2: Color the vertices of G uniformly at random independently from $\{1, \ldots, k\}$.
- Step 3: Check if there exists a path $\pi = (v_1, .., v_k)$ such that each vertex v_i is colored i. If such a path exists, output TRUE and terminate.
- Step 4: Output G contains no simple path of length k.

Lemma 13.4.29

If there exists a simple path $\boldsymbol{\pi = (v_1, ..., v_k)}$ in G, the probability that for all $\boldsymbol{i \in \{1, ..., k\}}$, the vertex $\boldsymbol{v_i}$ is assigned the color $\boldsymbol{i}$ is $\boldsymbol{\frac{1}{k^k}}$.

Proof:

- Since there are k colors, the probability of assigning a specific color i to vertex v_i is $\frac{1}{k}$.
- The coloring of each vertex is independent of the others.
- Therefore, the probability that v_1 is assigned color 1, v_2 is assigned color 2, and so on, up to v_k being assigned color k, is given by the product of individual probabilities:
- $\Pr(All\ v_i\ are\ assigned\ correct\ colors) = \frac{1}{k} . \frac{1}{k} . \frac{1}{k} \ldots \frac{1}{k} = \prod_{i=1}^{k} \frac{1}{k} = (\frac{1}{k})^k = \frac{1}{k^k}$

Questions ?

Carleton University 80 1942-2022

Refrences

- Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. J. ACM, 42(4):844–856, 1995
- Maheshwari, A. (n.d.). *Design and analysis of algorithms: Lecture notes*. Carleton University. Retrieved from https://people.scs.carleton.ca/~maheshwa/Notes/DAA/notes.pdf