Covering with Clubs: Complexity and Approximability

Riccardo Dondi [1] Giancarlo Mauri [2] Florian Sikora [3]
Italo Zoppis [2]

[1]Università di Bergamo, [2]Università degli Studi di Milano-Bicocca, [3]Université Paris-Dauphine

IWOCA 2018, 16-19 July, 2018

Outline

Outline

Motivations

- Finding **cohesive subgraphs** in networks/graphs has many applications: social network analysis and bioinformatics
- Several definitions of cohesive subgraph, e.g. complete graph or **clique**
- Combinatorial problems based on clique:
  - Maximum Clique
  - **Minimum Clique Partition**
  - Minimum Clique Cover

Minimum Clique Partition

Minimum Clique Partition:

- Classical problem in TCS
- Partition the vertices of a graph into the **minimum number of cliques**
- **Complexity results** (decision version):
  - NP-complete finding a partition of three (or more) cliques
  - In P finding a partition of two cliques
- **Approximation**: not approximable within factor $O(|V|^{1-\varepsilon})$, for any $\varepsilon > 0$, unless $P = NP$ [Zuckerman 2007]

Relaxed Cliques

A **relaxed clique** is a graph whose vertices satisfy a property which is a **relaxation** of the clique property
Two examples:

- **Diameter** of a clique: 1; of a relaxed clique: $s \geq 1$
- **Degree** in the vertices of a clique: $|V| - 1$; in a relaxed clique: $\geq |V| - s$, with $s \geq 1$

s-club

One of the most applied definition of relaxed clique in the analysis of networks: ***s*-club**

Definition

Given a graph $G = (V, E)$, and a subset $V' \subseteq V$, $G[V']$ is an *s*-club, with $s \geq 1$, if it has diameter at most *s*.

Maximum s-club

Maximum s-club: finding an s-club of maximum cardinality

- NP-hard [Bourjolly et al 2002]
- NP-hard even if the graph has diameter $s + 1$ [Balasundaram et al 2005]
- For every $s \geq 2$, approximable within factor $|V|^{1/2}$ and not approximable within factor $|V|^{1/2-\varepsilon}$, for each $\varepsilon > 0$ [Asahiro et al 2017]

Outline

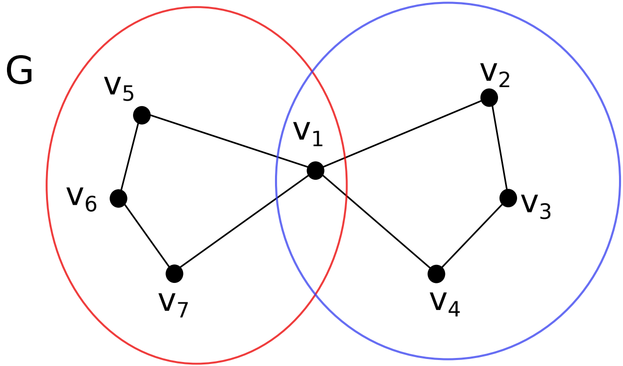
Problem Definition

Problem (Min s-Club Cover)

Input: *a graph* $G = (V, E)$ *and an integer* $s \geq 2$.
Output: *a minimum cardinality collection* $\mathcal{S} = \{V_1, \ldots, V_h\}$ *such that each* $G[V_i]$ *is an s-club, and, for each vertex* $v \in V$*, there exists an s-club* $G[V_j]$ *with* $v \in V_j$.

We focus on the cases $s = 2$ and $s = 3$

Min 2-Club Cover: an Example



Outline

Computational Complexity of Min s-Club Cover

Clique Partition:

- In *P*: covering a graph with two cliques
- *NP*-complete: covering a graph with three cliques

For *s*-clubs:

- Complexity of covering a graph with **two *s*-clubs**?
- Complexity of covering a graph with **three *s*-clubs**?

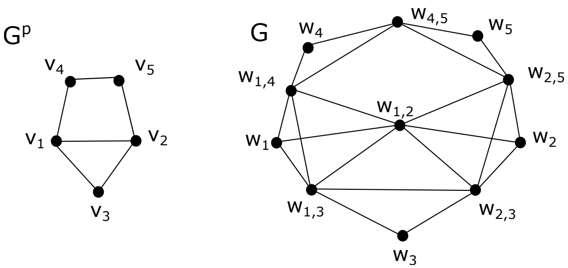
Hardness of Covering with 2-clubs

Theorem

Covering a graph with three 2-clubs is NP-complete.

Proof.

Reduction from Clique Partition.



□

Hardness of Covering with 3-clubs

Theorem

Covering a graph with two 3-clubs is NP-complete.

Proof.

- Reduction from a variant of SAT called 5-Double-Sat
- Goal: **double satisfy** a set of clauses, each one with 5 literals

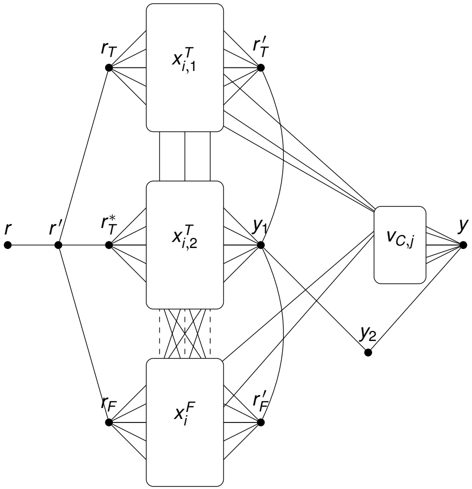
□

Outline of the Reduction

Outline of the reduction

1. 5-Double-Sat is NP-complete: reduction from 3-SAT
2. Reduction from 5-Double-Sat to covering a graph with two 3-clubs

Outline of the Reduction



Outline

Inapproximability of Min s-Club Cover

Theorem

Min 2-Club Cover *is not approximable within factor* $O(|V|^{1/2-\varepsilon})$, *for each* $\varepsilon > 0$, *unless* $P = NP$.
Min 3-Club Cover *is not approximable within factor* $O(|V|^{1-\varepsilon})$, *for each* $\varepsilon > 0$, *unless* $P = NP$.

Proof.

Reduction from Minimum Clique Partition. □

Approximation of Min 2-Club Cover

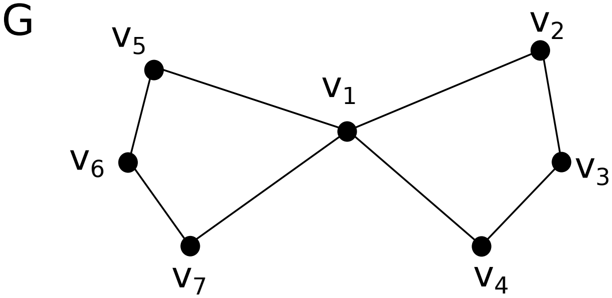
Theorem

Min 2-Club Cover *can be approximated within factor* $O(|V|^{1/2}\log^{3/2}|V|)$*.*

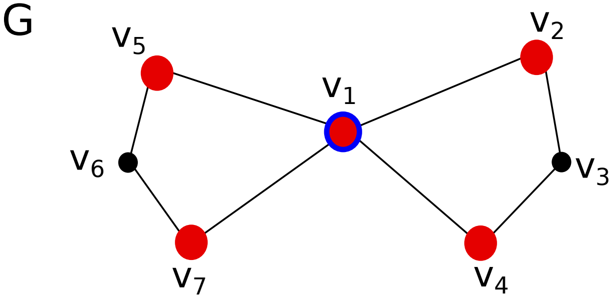
Proof.

Greedy algorithm: at each step, add a **star** the covers the **maximum number of uncovered vertices.** □

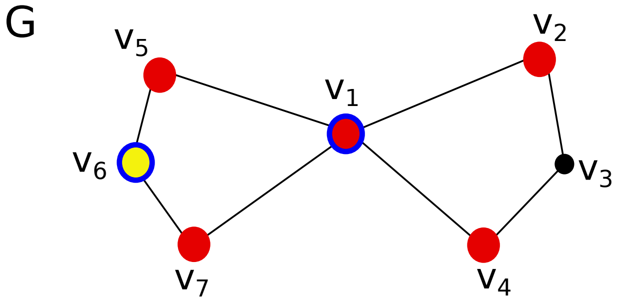
Approximation of Min 2-Club Cover



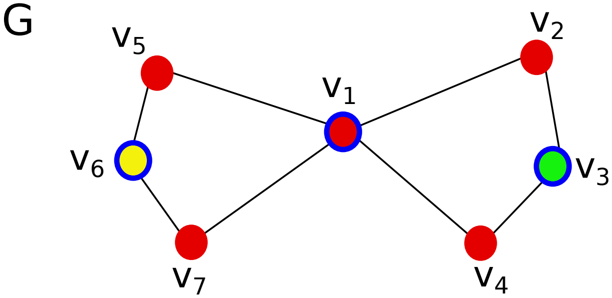
Approximation of Min 2-Club Cover



Approximation of Min 2-Club Cover



Approximation of Min 2-Club Cover



Approximation of Min 2-Club Cover

Theorem

Min 2-Club Cover *can be approximated within factor* $O(|V|^{1/2} \log^{3/2} |V|)$.

Proof.

Analysis of the approximation factor based on:

Lemma (Desormeaux et al 2014)

Let $H = (V_H, E_H)$ *be a* 2*-club, then* H *has a dominating set of size at most* $1 + \sqrt{|V_H| + \ln(|V_H|)}$.

□

Outline

Open Problems

Open problems:

- Complexity of covering a graph with two 2-clubs
- Improve the approximation factor to $O(|V|^{1/2})$
- New variants with other definitions of relaxed clique

Thank you!