Maximum clique: limits of approximation

Sapienza Università di Roma

Outline

Notation

The MCP is not approximable

Special cases

Example

Sources

Typical mistakes

Learning objectives

At the end of this chapter you will be able to:

- build the strong product of two graphs and prove $\omega(G \boxtimes H) = \omega(G)\,\omega(H)$;
- prove that a constant-ratio algorithm for the MCP would give one for every constant ratio;
- state the results of Håstad and Zuckerman, and conclude that the MCP and the MSSP are not approximable;
- solve the MCP exactly on planar graphs and on graphs of bounded degree, and approximate the MSSP on graphs of bounded degree.

Prerequisites

- From this module, the chapter «Principles of approximation algorithms»: performance ratio for a maximization problem, α-approximation algorithm, PTAS.
- From the module *Combinatorial Optimization Problems*, the chapter «Clique problems»: the definitions of the MCP and of the MSSP, the clique number $\omega(G)$, the complement graph and $\omega(G) = \alpha(\overline{G})$.
- From this module, the chapter «Vertex coloring: limits of approximation»: planar graphs and the bound $m \leq 3n - 6$.

The problem of this chapter

- The problem of this chapter is the maximum clique problem (MCP)
- Its definition and its assumptions are on its page
- The question of the chapter: how well can a **polynomial** algorithm approximate the clique number $\omega(G)$?

Notation

Introduction: Notation

- In this section we fix the notation on graphs and on the optimal values of the MCP and of the MSSP
- We recall how the performance ratio is measured for a maximization problem
- We see which ratio is trivial: the rest of the chapter compares everything with it

Graphs and optimal values

- As on the problem page, $G = (V, E)$ is a simple graph with n vertices and m edges
- $N(v)$ is the set of the neighbours of v, and $\Delta(G)$ the maximum degree
- The optimal value of the MCP is the **clique number**:

$$z(I) = \omega(G)$$

- The optimal value of the MSSP is the **stability number** $\alpha(G)$

Recall: the performance ratio

- As in the chapter «Principles of approximation algorithms», $\mathscr{A}(I)$ is the number of vertices of the clique returned by the algorithm $\mathscr{A}$
- The MCP is a maximization problem, hence

$$\mathscr{A}(I) \leq \omega(G) \qquad \text{and} \qquad \frac{\omega(G)}{\mathscr{A}(I)} \geq 1$$

for every instance I

The trivial ratio

- A single vertex is a clique, so the ratio n is trivial
- This chapter shows that a polynomial algorithm **cannot do much better** on all graphs
- and then the **classes of graphs** on which it can

Summary: Notation

- $G = (V, E)$ with n vertices, m edges, neighbourhoods $N(v)$ and maximum degree $\Delta(G)$
- $z(I) = \omega(G)$ for the MCP, $\alpha(G)$ for the MSSP
- For a maximization problem the ratio is $\frac{\omega(G)}{\mathscr{A}(I)} \geq 1$
- One vertex already gives the ratio n

The MCP is not approximable

Introduction: The MCP is not approximable

- We define the **strong product** of two graphs: it multiplies the clique numbers
- With it, a constant-ratio algorithm for the MCP would give **every** constant ratio
- Then we state the results of Håstad and Zuckerman: unless $P = NP$, not even the ratio $n^{1-\varepsilon}$ is possible
- Together they say that the MCP is **not approximable**

We start from a construction on graphs:
the strong product, which turns two graphs into one.

Definition 1: strong product

The ***strong product*** *of two graphs* $G = (V, E)$ *and* $H = (W, F)$ *is the graph* $G \boxtimes H$ *with vertex set* $V \times W$*, in which two distinct vertices* (u, x) *and* (v, y) *are adjacent if and only if*

$$\begin{aligned} &\big(u = v \quad \text{or} \quad \{u, v\} \in E\big) \\ \text{and} \quad &\big(x = y \quad \text{or} \quad \{x, y\} \in F\big). \end{aligned}$$

Size of the product

- $G \boxtimes H$ has $|V|\,|W|$ vertices
- It is built in time polynomial in the size of G and H
- In particular $G \boxtimes G$ has n^2 vertices

Projection and fibres

For a set $K \subseteq V \times W$:

- the **projection** is

$$p(K) = \{u \in V : (u, x) \in K \text{ for some } x\}$$

- the **fibre** of $u \in p(K)$ is

$$K_u = \{x \in W : (u, x) \in K\}$$

- Projection and fibres take a set of the product apart, one coordinate at a time
- We now see that they take a clique of the product apart into cliques of the two graphs
- This is the key fact of the whole chapter

Proposition 1

For every two graphs G and H:

1. *if $K \subseteq V \times W$ is a clique of $G \boxtimes H$, then $p(K)$ is a clique of G, every fibre K_u is a clique of H, and $|K| \leq |p(K)| \cdot \max_u |K_u|$;*
2. $\omega(G \boxtimes H) = \omega(G)\,\omega(H)$.

Proof (1/3)

Part 1: the projection is a clique.

- Let $u \neq v$ be in $p(K)$, with $(u, x), (v, y) \in K$
- These two vertices are adjacent and $u \neq v$, so $\{u, v\} \in E$
- Hence $p(K)$ is a clique

Part 1: every fibre is a clique.

- Let $x \neq y$ be in K_u: (u, x) and (u, y) are adjacent, so $\{x, y\} \in F$
- Hence K_u is a clique

Proof (2/3)

Part 1: the bound on $|K|$.

- The fibres split K, hence

$$|K| \;=\; \sum_{u \in p(K)} |K_u| \;\leq\; |p(K)| \cdot \max_u |K_u|.$$

Proof (3/3)

Part 2.

- If K_G and K_H are maximum cliques of G and H, every two distinct vertices of $K_G \times K_H$ satisfy the definition, so

$$\omega(G \boxtimes H) \geq \omega(G)\,\omega(H)$$

- Conversely, by Part 1 every clique K of $G \boxtimes H$ has

$$|K| \leq |p(K)| \cdot \max_u |K_u| \leq \omega(G)\,\omega(H)$$

□

Strong or lexicographic product

- The *lexicographic* product also satisfies

$$\omega(G[H]) = \omega(G)\,\omega(H)$$

- We use the strong product, whose definition is **symmetric** in the two graphs

Now we use the strong product of G with itself
to improve the ratio of an approximation algorithm.

From a clique of $G \boxtimes G$ back to G

- Part 1 of Proposition 1 is **constructive**
- From a clique K of $G \boxtimes G$, either $p(K)$ or the largest fibre is a clique of G with at least $\sqrt{|K|}$ vertices
- Indeed, the product of their sizes is at least $|K|$
- Both are computed in polynomial time

Theorem 1

If the MCP has a polynomial α-approximation algorithm for some constant $\alpha \geq 1$, then it has a polynomial β-approximation algorithm for every constant $\beta > 1$.

- The idea of the proof: run the algorithm on $G \boxtimes G$, where the optimum is $\omega(G)^2$
- Go back to G with the square root: the ratio becomes $\sqrt{\alpha}$
- Repeat a constant number of times

Proof (1/3)

Let $\mathscr{A}$ be a polynomial α-approximation algorithm, and define the algorithm $\mathscr{B}$:

1. build $G \boxtimes G$ and run $\mathscr{A}$ on it, obtaining a clique K;
2. return the larger of $p(K)$ and of the largest fibre of K, a clique of G with at least $\sqrt{|K|}$ vertices.

Proof (2/3)

- By Proposition 1, $\omega(G \boxtimes G) = \omega(G)^2$, so

$$|K| \geq \frac{\omega(G)^2}{\alpha}$$

- and

$$\mathscr{B}(I) \;\geq\; \sqrt{|K|} \;\geq\; \frac{\omega(G)}{\sqrt{\alpha}}$$

- $\mathscr{B}$ is a polynomial $\sqrt{\alpha}$-approximation algorithm

Proof (3/3)

- Applying the construction t times gives the ratio $\alpha^{1/2^t}$
- $\mathscr{A}$ is run on a graph with n^{2^t} vertices
- Choose the constant t with

$$2^t \geq \frac{\ln \alpha}{\ln \beta}$$

- Then $\alpha^{1/2^t} \leq \beta$, and n^{2^t} is a polynomial

□

The product amplifies the ratio

- In other words, the MCP has a constant-ratio algorithm **if and only if** it has a PTAS
- The strong product *amplifies* any constant ratio
- The running time grows like a power n^{2^t}, which is polynomial only because t is a constant

This is the positive side of the amplification.
The negative side comes from the theory of complexity.

The results of Håstad and Zuckerman

- The negative side is known, and we state it **without proof**
- It uses probabilistically checkable proofs

Theorem 2: Håstad, 1999; Zuckerman, 2007

For every $\varepsilon > 0$, if $P \neq NP$, no polynomial algorithm $\mathscr{A}$ for the MCP satisfies, for every graph G with n vertices,

$$\frac{\omega(G)}{\mathscr{A}(I)} \leq n^{1-\varepsilon}.$$

Two assumptions

- **Håstad (1999)** proved the theorem under the assumption that NP problems cannot be solved in randomized polynomial time ($NP \neq ZPP$)
- **Zuckerman (2007)** proved it under $P \neq NP$

No constant ratio and no PTAS

- As for the VCP, the theorem excludes every constant ratio
- An α-approximation algorithm has ratio at most $\alpha < n^{1/2}$ on the graphs with $n > \alpha^2$ vertices
- The smaller graphs are solved by enumeration in constant time
- Combined with Theorem 1: unless $P = NP$, the MCP has **neither** a constant ratio **nor** a PTAS

The MCP is not approximable

- The MCP is not approximable: unless $P = NP$, no polynomial algorithm has a constant performance guarantee
- not even a guarantee of the form $n^{1-\varepsilon}$
- Positive results exist only on **special classes of graphs**

Summary: The MCP is not approximable (1/2)

Strong product:

- $G \boxtimes H$ has $|V|\,|W|$ vertices and $\omega(G \boxtimes H) = \omega(G)\,\omega(H)$
- A clique of the product splits into its projection and its fibres, cliques of the two graphs

Amplification:

- From a clique K of $G \boxtimes G$ we get a clique of G with at least $\sqrt{|K|}$ vertices
- A constant ratio α gives $\sqrt{\alpha}$, then every constant $\beta > 1$

Summary: The MCP is not approximable (2/2)

Håstad and Zuckerman:

- Unless $P = NP$, no polynomial algorithm has ratio $n^{1-\varepsilon}$ on every graph
- Hence the MCP has neither a constant ratio nor a PTAS

Next:

- Positive results only on special classes of graphs

Special cases

Introduction: Special cases

- On all graphs the MCP is not approximable, but some classes of graphs are easy
- On **planar graphs** and on graphs of **bounded maximum degree** the MCP is solved exactly by enumeration
- Through the complement graph, the negative result passes to the **MSSP**, but the special cases do not
- On graphs of bounded degree the MSSP has a constant ratio

First class: the planar graphs,
which have no large cliques.

Proposition 2

Every planar graph has $\omega(G) \leq 4$, and on planar graphs the MCP is solved exactly in $O(n^4)$ time.

Proof

- A subgraph of a planar graph is planar
- The complete graph with 5 vertices has 10 edges, more than $3 \cdot 5 - 6 = 9$, so it is not planar
- Hence a planar graph has no clique with 5 vertices
- The subsets of at most 4 vertices are $O(n^4)$, and each one is checked in constant time

□

Second class: the graphs of bounded maximum degree, where every clique stays inside one neighbourhood.

Proposition 3

Every graph has $\omega(G) \leq \Delta(G) + 1$, *and the MCP is solved exactly in* $O(n\, 2^{\Delta}\, \Delta^2)$ *time, with* $\Delta = \Delta(G)$.

Proof

- A clique that contains the vertex v is contained in $\{v\} \cup N(v)$, which has at most $\Delta + 1$ vertices
- For every vertex v, enumerate the $2^{\deg(v)} \leq 2^{\Delta}$ subsets $S \subseteq N(v)$
- Keep the largest $\{v\} \cup S$ that is a clique
- Checking the pairs of S takes $O(\Delta^2)$ time
- Every maximum clique is found from any of its vertices



When is the time polynomial?

- The time $O(n\,2^{\Delta}\,\Delta^2)$ is polynomial for every fixed Δ
- and even when $\Delta \leq \log_2 n$, since then

$$2^{\Delta} \leq n$$

What happens to the stable sets?
The complement graph turns one problem into the other.

The maximum stable set problem

- By $\omega(G) = \alpha(\overline{G})$ (chapter «Clique problems» of the module *Combinatorial Optimization Problems*), a stable set of G is a clique of $\overline{G}$
- $\overline{G}$ is built in $O(n^2)$ time and has the same n
- An algorithm with ratio $\rho(n)$ for one of the two problems, applied to the complement graph, has the same ratio $\rho(n)$ for the other one

Corollary 1

For every $\varepsilon > 0$, if $P \neq NP$, no polynomial algorithm for the MSSP has ratio $\frac{\alpha(G)}{\mathscr{A}(I)} \leq n^{1-\varepsilon}$ on every graph. In particular, the MSSP has no constant ratio.

The special cases do not transfer

- The special cases do *not* transfer to the MSSP
- The complement of a planar graph, or of a graph of bounded degree, is **dense**
- On those classes the MSSP remains NP-hard, but it is **approximable**

Proposition 4

On every graph, a ***maximal*** *stable set S (no vertex can be added to it) has $|S| \geq \frac{n}{\Delta(G)+1}$. Hence the greedy algorithm that adds vertices while possible satisfies*

$$\frac{\alpha(G)}{|S|} \leq \Delta(G) + 1.$$

Proof

- Every vertex outside S has a neighbour in S, otherwise it could be added
- Hence

$$V = \bigcup_{v \in S} \left(\{v\} \cup N(v) \right) \qquad \text{and} \qquad n \leq |S| \left(\Delta(G) + 1 \right)$$

- Since $\alpha(G) \leq n$, the ratio is at most

$$\frac{\alpha(G)}{|S|} \leq \frac{n}{|S|} \leq \Delta(G) + 1$$



Bounded degree and planar graphs for the MSSP

- On graphs of maximum degree at most a fixed Δ this is a **constant ratio**
- On planar graphs the MSSP even has a **PTAS** (Baker, 1994)

Summary: Special cases (1/2)

Planar graphs:

- K_5 is not planar, so $\omega(G) \leq 4$, and the MCP is solved exactly in $O(n^4)$ time

Bounded maximum degree:

- $\omega(G) \leq \Delta(G) + 1$: every clique lies in $\{v\} \cup N(v)$
- Enumeration in $O(n\, 2^{\Delta}\, \Delta^2)$ time, polynomial for fixed Δ or $\Delta \leq \log_2 n$

Summary: Special cases (2/2)

The MSSP:

- Through $\overline{G}$, the MSSP has no ratio $n^{1-\varepsilon}$, unless $P = NP$
- The special cases do not transfer: the complements are dense
- A maximal stable set has ratio at most $\Delta(G) + 1$; on planar graphs a PTAS exists

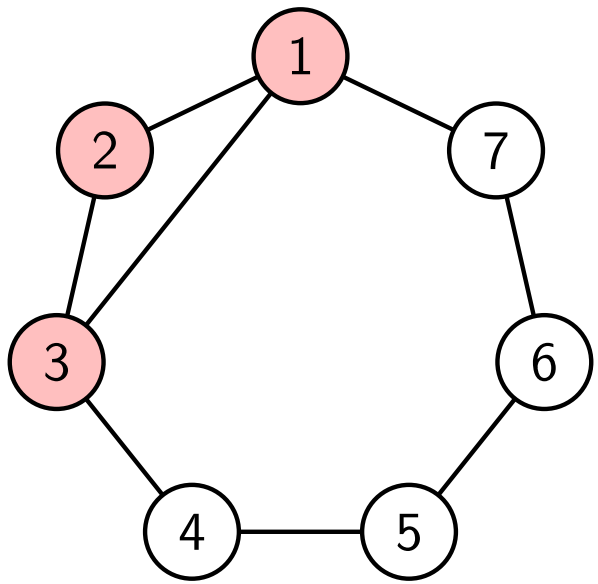
Example

Introduction: Example

- We apply the results of the chapter to the reference instance of the MCP
- We see the strong product, a clique of the product taken apart, the enumeration for bounded degree and the bound for planar graphs
- All the numbers are small enough to be checked by hand

Example (1/4)

- The reference instance of the MCP is the cycle $1, 2, \ldots, 7, 1$ with the chord $\{1, 3\}$
- $n = 7$, $m = 8$, $\Delta(G) = 3$ and $\omega(G) = 3$, with the unique maximum clique $\{1, 2, 3\}$ (in red)



Example (2/4)

Strong product.

- $G \boxtimes G$ has $7^2 = 49$ vertices and 240 edges
- Its clique number is

$$\omega(G \boxtimes G) = 3^2 = 9$$

attained by $\{1, 2, 3\} \times \{1, 2, 3\}$

Example (3/4)

- Suppose an algorithm returns on $G \boxtimes G$ the clique

$$K = \{1, 2, 3\} \times \{4, 5\}, \qquad |K| = 6$$

- Its projection $p(K) = \{1, 2, 3\}$ and its fibres $K_u = \{4, 5\}$ are cliques of G
- The larger one has $3 \geq \sqrt{6}$ vertices

Example (4/4)

Bounded degree.

- The enumeration of Proposition 3 examines

$$\sum_{v} 2^{\deg(v)} = 2^3 + 2^2 + 2^3 + 4 \cdot 2^2 = 36$$

  subsets

- It finds $\{1, 2, 3\}$ from the vertices 1, 2 and 3

Planarity.

- The graph is planar, as the figure shows, and $\omega(G) = 3 \leq 4$

Summary: Example

- On the cycle with a chord, $\omega(G) = 3$ and $\omega(G \boxtimes G) = 9$
- A clique of 6 vertices of the product gives back a clique of G with $3 \geq \sqrt{6}$ vertices
- The enumeration for bounded degree examines 36 subsets and finds $\{1, 2, 3\}$
- The graph is planar, and $\omega(G) = 3 \leq 4$

Sources

Introduction: Sources

- The results of this chapter come from a textbook and from three research papers
- They are the places where to read the full proofs, in particular the one of the theorem of Håstad and Zuckerman

Sources (1/2)

- M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman, 1979 (the self-improvement of the MCP by graph products).

- J. Håstad, Clique is hard to approximate within $n^{1-\varepsilon}$, *Acta Mathematica* 182(1):105–142, 1999.

Sources (2/2)

- D. Zuckerman, Linear degree extractors and the inapproximability of max clique and chromatic number, *Theory of Computing* 3:103–128, 2007.
- B.S. Baker, Approximation algorithms for NP-complete problems on planar graphs, *Journal of the ACM* 41(1):153–180, 1994.

Summary: Sources

- Garey and Johnson (1979): the self-improvement of the MCP by graph products
- Håstad (1999) and Zuckerman (2007): the ratio $n^{1-\varepsilon}$ is impossible
- Baker (1994): a PTAS for the MSSP on planar graphs

Typical mistakes

Introduction: Typical mistakes

- Four mistakes that come up often on the approximation of cliques and stable sets
- They concern the direction of the ratio, the number of products, and the passage from the MCP to the MSSP
- Checking them is a quick way to avoid losing points

Observation 1: typical mistakes (1/2)

- ***Inverting the ratio.*** *The MCP is a maximization problem: the ratio is* $\frac{\omega(G)}{\mathscr{A}(I)}$*, at least* 1*.*
- ***Iterating the product a number of times that grows with*** n*.* *The graph after* t *products has* n^{2^t} *vertices: polynomial only for a constant* t*, so the amplification gives every constant ratio, not the ratio* 1*.*

Observation: typical mistakes (2/2)

- ***Transferring the special cases to the MSSP.*** *The complement of a planar graph is not planar: the MCP is easy on planar graphs, the MSSP is NP-hard there.*
- ***Confusing the bound*** $\Delta(G)+1$ ***of the MCP with that of the MSSP.*** *For the MCP,* $\omega(G) \leq \Delta(G)+1$ *bounds the optimum and makes enumeration fast; for the MSSP,* $\Delta(G)+1$ *bounds the ratio of a maximal stable set.*

Summary: Typical mistakes

- The ratio of the MCP is $\frac{\omega(G)}{\mathscr{A}(I)} \geq 1$
- The number t of products must be a constant
- Planar graphs are easy for the MCP, not for the MSSP
- $\Delta(G) + 1$ bounds the optimum of the MCP, and the ratio of a maximal stable set for the MSSP

Summary (1/3)

- The strong product squares the clique number, and a clique of the product gives back a clique of G of at least the square root of its size: a constant ratio for the MCP would give every constant ratio.
- By the results of Håstad and Zuckerman, unless $P = NP$ not even the ratio $n^{1-\varepsilon}$ is possible: the MCP is not approximable, and neither is the MSSP, through the complement graph.
- On planar graphs and on graphs of bounded degree the MCP is solved exactly by enumeration; on graphs of bounded degree a maximal stable set has ratio at most $\Delta(G) + 1$ for the MSSP.

Summary (2/3)

problem and graphs	algorithm	result
MCP, all graphs	any polynomial	no constant ratio, unless $P = NP$
MCP, planar	enumeration	exact, $\omega(G) \leq 4$
MCP, maximum degree $\leq \Delta$	enumeration	exact, $O(n\,2^{\Delta}\Delta^2)$
MSSP, all graphs	any polynomial	no constant ratio, unless $P = NP$
MSSP, maximum degree $\leq \Delta$	maximal stable set	ratio $\leq \Delta + 1$

Summary (3/3)

Dashboard of the reference instance.

Instance	Upper bound	Value found	Optimal value
cycle with a chord	$\Delta(G)+1=4$	3, enumeration	3