The densest k-subgraph problem on clique graphs

Maria Liazi

Department of Informatics and Telecommunications
National and Kapodistrian University of Athens (NKUA)

joint work with
Ioannis Milis, **Fanny Pascual** and **Vassilis Zissimopoulos**

Combinatorics, **G**eometry and **C**omputer **S**cience **C**onference **2007**

- **Densest k-subgraph (DkS)**:

  Input: A graph $G = (V, E)$, $|V| = n$, and an integer k, $k \leq n$

  Output: Find a k vertices subgraph of G with the maximum number of edges

- **DkS is strongly NP-hard** as generalization of the CLIQUE problem.

- **Densest k-subgraph (DkS)**:

  Input: A graph $G = (V, E)$, $|V| = n$, and an integer k, $k \leq n$

  Output: Find a k vertices subgraph of G with the maximum number of edges

- **DkS is strongly NP-hard** as generalization of the CLIQUE problem.

DkS remains **NP-hard** even for:

- bipartite graphs, comparability graphs, chordal graphs
  [Corneil and Perl, Discr. Appl. Math., 1984]
- planar graphs
  [Keil and Brecht, J. Comb. Math. Comb. Comp., 1991]
- bipartite graphs of maximal degree three
  [Feige and Seltser, Weizmann Inst., 1997]

DkS can be solved in **polynomial** time on:

- trees
  [Perl and Shiloach, SIAM J. Alg. Discr. Methods, 1983]
  [Fischetti et al., Networks, 1994]
  [Goldschmidt and Hochbaum, Discr. Appl. Math., 1997]
  [Rader and Woeginger, Oper. Res. Lett., 2002]
- cographs, split graphs, k-trees
  [Corneil and Perl, Discr. Appl. Math., 1984]

DkS remains **NP-hard** even for:

- bipartite graphs, comparability graphs, chordal graphs

  [Corneil and Perl, Discr. Appl. Math., 1984]

- planar graphs

  [Keil and Brecht, J. Comb. Math. Comb. Comp., 1991]

- bipartite graphs of maximal degree three

  [Feige and Seltser, Weizmann Inst., 1997]

DkS can be solved in **polynomial** time on:

- trees

  [Perl and Shiloach, SIAM J. Alg. Discr. Methods, 1983]
  [Fischetti et al., Networks, 1994]
  [Goldschmidt and Hochbaum, Discr. Appl. Math., 1997]
  [Rader and Woeginger, Oper. Res. Lett., 2002]

- cographs, split graphs, k-trees

  [Corneil and Perl, Discr. Appl. Math., 1984]

- Approximation algorithms
  1. [Kortsarz and Peleg, FOCS, 1993]
  2. [Feige and Seltser, Weizmann Inst., 1997]
  3. [Srivastav and Wolf, APPROX, 1998]
  4. [Ye and Zang, T.R. U. Iowa, 1999]
  5. [Asahiro et al., J. of Algorithms, 2000]
  6. [Feige et. al., Algorithmica, 2001]
  7. [Feige and Langberg, J. of Algorithms, 2001]
  8. [Han et al., Math. Progr., 2002]
  9. [Billionnet and Roupin, Cedric-CNAM, 2004]

- Best known approximation ratio $\rightarrow$ $\mathbf{O}(\mathbf{n}^{\frac{1}{3}})$, [6]

- No one of them achieves a constant approximation ratio

- Approximation algorithms

  1. [Kortsarz and Peleg, FOCS, 1993]
  2. [Feige and Seltser, Weizmann Inst., 1997]
  3. [Srivastav and Wolf, APPROX, 1998]
  4. [Ye and Zang, T.R. U. Iowa, 1999]
  5. [Asahiro et al., J. of Algorithms, 2000]
  6. [Feige et. al., Algorithmica, 2001]
  7. [Feige and Langberg, J. of Algorithms, 2001]
  8. [Han et al., Math. Progr., 2002]
  9. [Billionnet and Roupin, Cedric-CNAM, 2004]

- Best known approximation ratio $\rightarrow$ $\mathbf{O(n^{\frac{1}{3}})}$, [6]

- No one of them achieves a constant approximation ratio

- There is **not a PTAS** for the DkS problem

[Khot, FOCS, 2004]

- There is **not known inapproximability result** for some approximation ratio

- A PTAS for dense graphs

  - of minimum degree $\Omega(n)$

  - of $\Omega(n^2)$ edges when k is $\Omega(n)$

[Arora et al., J. Comput. Syst. Sci., 1999]

- There is **not a PTAS** for the DkS problem

[Khot, FOCS, 2004]

- There is **not known inapproximability result** for some approximation ratio

- A PTAS for dense graphs

  - of minimum degree $\Omega(n)$
  - of $\Omega(n^2)$ edges when k is $\Omega(n)$

[Arora et al., J. Comput. Syst. Sci., 1999]

- A **clique** of an undirected graph, $G = (V, E)$, is a subset of its vertices inducing a complete subgraph in G.
- The **intersection graph** of a family, F, of subsets of a set is defined as a graph, $\mathcal{G}$, whose vertices correspond to the subsets in F, and there is an edge between two vertices of $\mathcal{G}$ if the corresponding pair of subsets intersect.
- The **clique graph** of a graph G is defined as the intersection graph of the maximal cliques of G.

All maximal cliques, and hence the clique graph, of a chordal graph can be found in polynomial time.

- **Star (Tree) of cliques** are graphs that having as clique graph a star (tree).

- A **clique** of an undirected graph, $G = (V, E)$, is a subset of its vertices inducing a complete subgraph in G.

- The **intersection graph** of a family, F, of subsets of a set is defined as a graph, $\mathcal{G}$, whose vertices correspond to the subsets in F, and there is an edge between two vertices of $\mathcal{G}$ if the corresponding pair of subsets intersect.

- The **clique graph** of a graph G is defined as the intersection graph of the maximal cliques of G.

All maximal cliques, and hence the clique graph, of a chordal graph can be found in polynomial time.

- **Star (Tree) of cliques** are graphs that having as clique graph a star (tree).

- A **clique** of an undirected graph, $G = (V, E)$, is a subset of its vertices inducing a complete subgraph in G.
- The **intersection graph** of a family, F, of subsets of a set is defined as a graph, $\mathcal{G}$, whose vertices correspond to the subsets in F, and there is an edge between two vertices of $\mathcal{G}$ if the corresponding pair of subsets intersect.
- The **clique graph** of a graph G is defined as the intersection graph of the maximal cliques of G.

All maximal cliques, and hence the clique graph, of a chordal graph can be found in polynomial time.

- **Star (Tree) of cliques** are graphs that having as clique graph a star (tree).

- **S** is a subset of $|S| = k$ vertices
- **E(S)** is the number of edges in the subgraph induced by S
- $\mathbf{S}^*$ is the optimal solution to the DkS problem
- $G = (V, E)$ a star of cliques with **m** maximal cliques
- $\mathbf{C_0}$ is the **central** clique
- and $\mathbf{C_1}, \ldots, \mathbf{C_{m-1}}$ are the **exterior** maximal cliques of the star of cliques

C_0 intersects with each other clique and no other intersection exists

- **S** is a subset of $|S| = k$ vertices
- **E(S)** is the number of edges in the subgraph induced by S
- $\mathbf{S}^*$ is the optimal solution to the DkS problem

- $G = (V, E)$ a star of cliques with **m** maximal cliques
- $\mathbf{C_0}$ is the **central** clique
- and $\mathbf{C_1}, \ldots, \mathbf{C_{m-1}}$ are the **exterior** maximal cliques of the star of cliques

C_0 intersects with each other clique and no other intersection exists

- **S** is a subset of $|S| = k$ vertices
- **E(S)** is the number of edges in the subgraph induced by S
- **S*** is the optimal solution to the DkS problem

- $G = (V, E)$ a star of cliques with **m** maximal cliques
- $\mathbf{C_0}$ is the **central** clique
- and $\mathbf{C_1}, \ldots, \mathbf{C_{m-1}}$ are the **exterior** maximal cliques of the star of cliques

C_0 intersects with each other clique and no other intersection exists

For each exterior clique C_i,

- $\mathbf{a_i = |C_i \cap C_0|}$, i.e., the number of vertices in its intersection with C_0
- $\mathbf{b_i = |C_i| - a_i > 0}$, i.e., the number of its vertices outside C_0
- $\mathbf{C_0' = C_0 \setminus \bigcup_{i=1}^{m-1} C_i}$, i.e., is the clique consisting of the vertices of C_0 not belonging to any other clique

- A clique $C_i,\ 0 \leq i \leq m-1$, is **completely** in a solution S if all its vertices are in S.

- The cliques C_0 and C_0' are **partially** in a solution S if a non-empty subset of their vertices, but not all, are in S.

- An **exterior** clique $C_i,\ 1 \leq i \leq m-1$ is **partially** in S if a non-empty subset of its $C_i \setminus C_0$ vertices, but not all, are in S.

If an exterior clique C_i is partially in S^*, then all its $|C_i \cap C_0| = a_i$ vertices are in S^*.

- A clique **is participating** in a solution S if it is either completely or partially in S.

- A clique $C_i,\ 0 \leq i \leq m-1$, is **completely** in a solution S if all its vertices are in S.

- The cliques C_0 and C_0' are **partially** in a solution S if a non-empty subset of their vertices, but not all, are in S.

- An **exterior** clique $C_i,\ 1 \leq i \leq m-1$ is **partially** in S if a non-empty subset of its $C_i \setminus C_0$ vertices, but not all, are in S.

If an exterior clique C_i is partially in S^*, then all its $|C_i \cap C_0| = a_i$ vertices are in S^*.

- A clique **is participating** in a solution S if it is either completely or partially in S.

- A clique $C_i,\ 0 \leq i \leq m-1$, is **completely** in a solution S if all its vertices are in S.
- The cliques C_0 and C_0' are **partially** in a solution S if a non-empty subset of their vertices, but not all, are in S.
- An **exterior** clique $C_i,\ 1 \leq i \leq m-1$ is **partially** in S if a non-empty subset of its $C_i \setminus C_0$ vertices, but not all, are in S.

If an exterior clique C_i is partially in S^*, then all its $|C_i \cap C_0| = a_i$ vertices are in S^*.

- A clique **is participating** in a solution S if it is either completely or partially in S.

Proposition

At most one of the cliques $C_0', C_1, \ldots, C_{m-1}$ is partially in an optimal solution.

Proposition

(i) If C_0 is the largest clique i.e., $|C_0| > |C_i|$, $1 \leq i \leq m-1$, then C_0 belongs completely to every optimal solution.

(ii) If C_0 is partially in an optimal solution S^*, then $|C_0| \leq |C_i|$ for every clique C_i participating in S^*.

Proposition

At most one of the cliques $C'_0, C_1, \ldots, C_{m-1}$ is partially in an optimal solution.

Proposition

(i) If C_0 is the largest clique i.e., $|C_0| > |C_i|,\ 1 \leq i \leq m-1$, then C_0 belongs completely to every optimal solution.

(ii) If C_0 is partially in an optimal solution S^*, then $|C_0| \leq |C_i|$ for every clique C_i participating in S^*.

Lemma

If clique C_0 is completely in the optimal solution, then there is an $O(nk^2)$ dynamic programming algorithm for the DkS problem on a star of cliques.

$k' = k - |C_0|$ vertices from exterior cliques

For $i = 0, 1, 2, \ldots, k'$ and $j = 2, 3, \ldots, m-1$

$$f(i,j) = \max_{0 \le q \le \min\{i, b_j\}} \{ f(i-q, j-1) + q \cdot a_j + \binom{q}{2} \}$$

Complexity: $O(nk^2)$

Notice that if C_0 is the largest clique, then, C_0 belongs completely to every optimal solution.

Lemma

If clique C_0 is completely in the optimal solution, then there is an $O(nk^2)$ dynamic programming algorithm for the DkS problem on a star of cliques.

$k' = k - |C_0|$ vertices from exterior cliques

For $i = 0, 1, 2, \ldots, k'$ and $j = 2, 3, \ldots, m-1$

$$f(i,j) = \max_{0 \le q \le \min\{i, b_j\}} \left\{ f(i-q, j-1) + q \cdot a_j + \binom{q}{2} \right\}$$

Complexity: $O(nk^2)$

Notice that if C_0 is the largest clique, then, C_0 belongs completely to every optimal solution.

Lemma

If clique C_0 is completely in the optimal solution, then there is an $O(nk^2)$ dynamic programming algorithm for the DkS problem on a star of cliques.

$k' = k - |C_0|$ vertices from exterior cliques

For $i = 0, 1, 2, \ldots, k'$ and $j = 2, 3, \ldots, m-1$

$$f(i,j) = \max_{0 \le q \le \min\{i, b_j\}} \left\{ f(i-q, j-1) + q \cdot a_j + \binom{q}{2} \right\}$$

Complexity: $O(nk^2)$

Notice that if C_0 is the largest clique, then, C_0 belongs completely to every optimal solution.

Lemma

If clique C_0 is completely in the optimal solution, then there is an $O(nk^2)$ dynamic programming algorithm for the DkS problem on a star of cliques.

$k' = k - |C_0|$ vertices from exterior cliques

For $i = 0, 1, 2, \ldots, k'$ and $j = 2, 3, \ldots, m-1$

$$f(i,j) = \max_{0 \le q \le \min\{i, b_j\}} \{f(i-q, j-1) + q \cdot a_j + \binom{q}{2}\}$$

Complexity: $O(nk^2)$

Notice that if C_0 is the largest clique, then, C_0 belongs completely to every optimal solution.

Lemma

If clique C_0 is completely in the optimal solution, then there is an $O(nk^2)$ dynamic programming algorithm for the DkS problem on a star of cliques.

$k' = k - |C_0|$ vertices from exterior cliques

For $i = 0, 1, 2, \ldots, k'$ and $j = 2, 3, \ldots, m-1$

$$f(i,j) = \max_{0 \le q \le \min\{i, b_j\}} \left\{ f(i-q, j-1) + q \cdot a_j + \binom{q}{2} \right\}$$

Complexity: $O(nk^2)$

Notice that if C_0 is the largest clique, then, C_0 belongs completely to every optimal solution.

Proposition

If C_0 is partially in an optimal solution, then the number of exterior cliques of size at least $|C_0|$ is at most $\sqrt{n}$.

$\mathbf{r} = \lfloor \frac{k}{|C_0|} \rfloor$, δ is a fixed number which will be defined later.

Proposition

If C_0 is partially in an optimal solution, then the number of exterior cliques of size at least $|C_0|$ is at most $\sqrt{n}$.

$\mathbf{r} = \lfloor \frac{k}{|C_0|} \rfloor$, δ is a fixed number which will be defined later.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\, 2^r\, n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\,2^r\,n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
    - at most one of the cliques in a set of r cliques is partially in S^*
    - consider all the $2^r - 1$ subsets of this set
    - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\, 2^r\, n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\,2^r\,n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\,2^r\,n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\,2^r\,n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\,2^r\,n^{\frac{r}{2}})$ time.

If $r < \delta$, then we proceed in an **exhaustive** manner

- examine all the possible sets of r cliques out of cliques of size at least $|C_0|$
- for each one of these sets we compute the k vertices that maximize the number of edges
  - at most one of the cliques in a set of r cliques is partially in S^*
  - consider all the $2^r - 1$ subsets of this set
  - the solutions are all the possible k-vertex solutions for the set
- the optimal solution is the one with the maximum number of edges.

Lemma

For the case $r < \delta$, δ be a fixed number, an optimal solution for the DkS problem in a star of cliques can be found in $O(r\,2^r\,n^{\frac{r}{2}})$ time.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

If $r \geq \delta$, then we proceed in a **greedy** manner

- S is obtained by the following algorithm:
  - t is the largest integer number such that $k \geq \sum_{i=1}^{t} |C_i| = k'$, where $C_1 \geq C_2 \geq \ldots \geq C_{m-1}$
  - return all the vertices of the cliques $C_1 \geq C_2 \geq \ldots \geq C_t$ and $k - k'$ vertices of clique C_{t+1}
- all cliques in S are of size at least $|C_0|$ and we need at least r cliques of size $|C_0|$ in order to fill k
  thus, $E(S) \geq rE(C_0)$
- S^* involves exterior cliques of size at least $|C_0|$ and these cliques are selected by S^* due to the edges between their overlaps in C_0
  thus, $E(S^*) \leq E(S) + E(C_0)$

Lemma

For the case $r \geq \delta$, where $\delta = \frac{1-\epsilon}{\epsilon}$, $0 < \epsilon < 1$, there is an $(1-\epsilon)$-approximation algorithm for the DkS problem in a star of cliques.

The complexity of the exhaustive optimal algorithm is exponential in $r \leq \delta = \frac{1-\epsilon}{\epsilon}$, that is exponential in $\frac{1}{\epsilon}$.

The complexity of the greedy approximation algorithm is $O(n\, logn)$.

Theorem

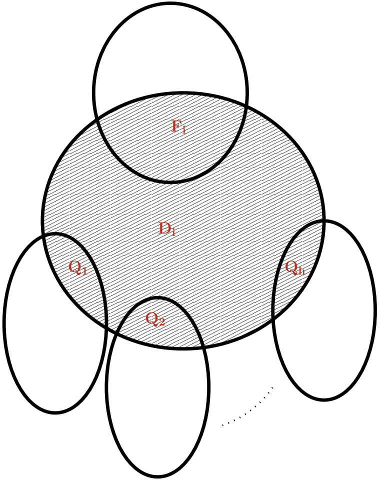
There is a polynomial time approximation scheme for the DkS problem in stars of cliques.

The complexity of the exhaustive optimal algorithm is exponential in $r \leq \delta = \frac{1-\epsilon}{\epsilon}$, that is exponential in $\frac{1}{\epsilon}$.

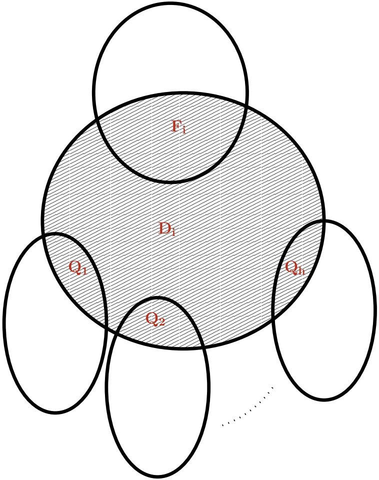
The complexity of the greedy approximation algorithm is $O(n\, logn)$.

Theorem

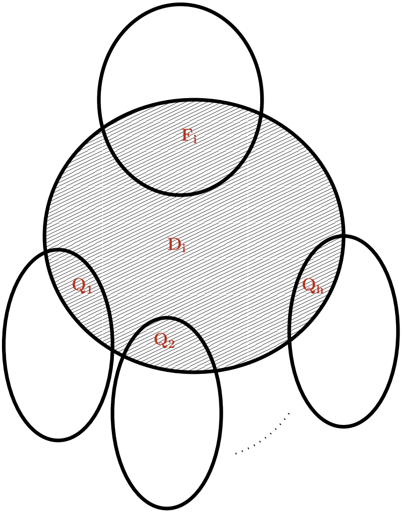
There is a polynomial time approximation scheme for the DkS problem in stars of cliques.



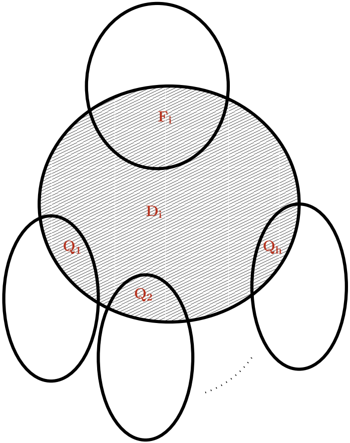
- $\mathbf{C_i}$ a non-leaf clique with $m_i \geq 1$ children
- $\mathbf{Q_h}$ the intersection of C_i with its h^{th} child clique
- $\mathbf{F_i}$ the intersection of the clique C_i with its father clique
- $\mathbf{D_i}$ the vertices of a clique C_i not belonging to any intersection



- $\mathbf{C_i}$ a non-leaf clique with $m_i \geq 1$ children
- $\mathbf{Q_h}$ the intersection of C_i with its h^{th} child clique
- $\mathbf{F_i}$ the intersection of the clique C_i with its father clique
- $\mathbf{D_i}$ the vertices of a clique C_i not belonging to any intersection



$\mathbf{f_i(j, a)}$ optimal solution to the DjS problem on the subtree rooted at clique C_i including *exactly a* vertices from the clique F_i.



$$f_i(j,a) = \begin{cases} \binom{j}{2}, & \text{if } j \le \sum_{h=i_1}^{i_{m_i}} |Q_h| + |D_i| + a \\ \max\limits_{a+b+\sum_{h=i_1}^{i_{m_i}} j_h = j} \{ \sum\limits_{h=i_1}^{i_{m_i}} f_h(j_h, a_h) + \binom{a+b}{2} + (a+b) \sum\limits_{h=i_1}^{i_{m_i}} a_h + \\ \qquad \sum\limits_{\substack{i,j=i_1 \\ i \ne j}}^{i_{m_i}} \frac{a_i \cdot a_j}{2} \}, & \text{otherwise.} \end{cases}$$

- $a \in F_i$, $b \in D_i$, $a_h \ge 1$ vertices of each Q_h

Theorem

There is an $O(nk^{m+1})$ algorithm for the DkS problem on a tree of cliques of maximum degree m.

Corollary

There is an $O(nk^3)$ optimal algorithm for the DkS problem on a path of cliques.

Theorem

There is an $O(nk^{m+1})$ algorithm for the DkS problem on a tree of cliques of maximum degree m.

Corollary

There is an $O(nk^3)$ optimal algorithm for the DkS problem on a path of cliques.

- What is the complexity of the DkS problem on:
  - (Proper) Interval graphs
  - Permutation graphs

- Approximation algorithms for special graph classes:
  - Bipartite graphs
  - Comparability graphs
  - Chordal graphs
  - Planar graphs
  - Bounded degree graphs (even bipartite)
  - Regular graphs

- Better approximation algorithms for arbitrary graphs