Finding Maximum Subgraphs with Relatively Large Vertex Connectivity

Oleg A. Prokopyev

Department of Industrial Engineering
University of Pittsburgh

May 19, 2015

- Joint work with:
  - Alexander Veremyev
    University of Florida
  - Eduardo L. Pasiliao
    US Air Force Research Laboratory (AFRL)
  - Vladimir Boginski
    University of Florida
- Reference:
  - "Finding maximum subgraphs with relatively large vertex connectivity," *European Journal of Operational Research*, Vol. 239(2), pp. 349–362 (2014)

Background: Vertex Connectivity

- The *vertex connectivity* of a graph G, referred to as $\kappa(G)$, is defined as the minimum number of vertices of G whose removal results in a disconnected graph or a trivial graph (i.e., consisting of exactly one vertex)

- Graph G is *k-vertex-connected* if its vertex connectivity is at least k, i.e., $\kappa(G) \geq k$
- Vertex connectivity and k-vertex connectivity of a given graph can be verified in polynomial time

- Vertex connectivity is among the fundamental graph properties and there is a considerable body of work on this topic

Background: Clique Properties

- A clique is a very intuitive and simple concept of a cohesive subgraph with numerous important applications
- Cliques possess a number of “ideal” cohesiveness properties:
  - each vertex is connected to all other vertices
  - a clique has maximum possible edge density
  - a clique has maximum edge and vertex connectivity
  - distance between any pair of vertices is one, etc

Background: Clique Properties

- A clique is a very intuitive and simple concept of a cohesive subgraph with numerous important applications
- Cliques possess a number of “ideal” cohesiveness properties:
  - each vertex is connected to all other vertices
  - a clique has maximum possible edge density
  - a clique has maximum edge and vertex connectivity
  - distance between any pair of vertices is one, etc
- In many practical scenarios cliques are overly restrictive
  - e.g., some links in the graph may be missing due to noisy observations or experimental errors

Background: Taxonomy of Clique Relaxations

- Existing clique relaxation models are based on relaxing some of the *elementary clique-defining properties*, namely, distance, diameter, domination, degree, density and connectivity
  - Pattillo, Youssef, Butenko, “On clique relaxation models in network analysis,” *European Journal of Operational Research*, 2013

Background: Taxonomy of Clique Relaxations

- Existing clique relaxation models are based on relaxing some of the *elementary clique-defining properties*, namely, distance, diameter, domination, degree, density and connectivity
  - Pattillo, Youssef, Butenko, “On clique relaxation models in network analysis,” *European Journal of Operational Research*, 2013
- Relaxations are further classified into *absolute* and *relative* ones

Background: Absolute Clique Relaxations

- One example of an absolute relaxation is an *s-club*
  - a subset $S \subseteq V$ such that the subgraph $G[S]$ induced by S in G has diameter at most $s \in \mathbb{Z}_{>0}$, i.e., $diam(G[S]) \leq s$
  - Clearly, requiring $s = 1$ results in a clique
  - The problem of finding maximum s-clubs is known to be *NP*-hard for any fixed $s \geq 2$ (Balasundaram, et al., 2005)

Background: Absolute Clique Relaxations

- One example of an absolute relaxation is an *s-club*
  - a subset $S \subseteq V$ such that the subgraph $G[S]$ induced by S in G has diameter at most $s \in \mathbb{Z}_{>0}$, i.e., $diam(G[S]) \leq s$
  - Clearly, requiring $s = 1$ results in a clique
  - The problem of finding maximum s-clubs is known to be *NP*-hard for any fixed $s \geq 2$ (Balasundaram, et al., 2005)
- Another example is a *k-block*
  - a subset $S \subseteq V$ such that the subgraph $G[S]$ induced by S in G has vertex connectivity at least k, i.e., $\kappa(G[S]) \geq k$
  - Clearly, requiring $k = n - 1$ results in a clique
  - In contrast to the above model, finding k-connected components can be performed in polynomial time (assuming fixed k)
    - e.g., finding a maximum 1-block corresponds to finding the largest connected component of a graph

Background: Relative Clique Relaxations

- A classical example of a relative relaxation is a γ-*quasi-clique*
  - a subset $S \subseteq V$ such that the subgraph $G[S]$ induced by S in G has edge density at least γ, i.e., $\rho(G[S]) = |(S \times S) \cap E| / \binom{|S|}{2} \geq \gamma$, where $\gamma \in [0, 1]$ is a fixed constant parameter
  - Clearly, requiring $\gamma = 1$ results in a clique
  - The problem of finding maximum γ-quasi-cliques is known to be *NP*-hard for any fixed $\gamma = p/q$, where $p,\ q \in \mathbb{Z}_{>0}$ and $p \leq q$ (Patillo et al., 2013)

Background: Relative Clique Relaxations

- A classical example of a relative relaxation is a γ-*quasi-clique*
  - a subset $S \subseteq V$ such that the subgraph $G[S]$ induced by S in G has edge density at least γ, i.e., $\rho(G[S]) = |(S \times S) \cap E| / \binom{|S|}{2} \geq \gamma$, where $\gamma \in [0, 1]$ is a fixed constant parameter
  - Clearly, requiring $\gamma = 1$ results in a clique
  - The problem of finding maximum γ-quasi-cliques is known to be NP-hard for any fixed $\gamma = p/q$, where $p,\ q \in \mathbb{Z}_{>0}$ and $p \leq q$ (Patillo et al., 2013)
- An alternative (degree-based) definition of a γ-*quasi-clique*
  - a subset $S \subseteq V$ such that in the subgraph $G[S]$ induced by S in G degree of every node is at least $\gamma \cdot (|S| - 1)$, i.e., $\deg_{G[S]}(i) \geq \gamma \cdot (|S| - 1)$ for any $i \in S$, where $\gamma \in [0, 1]$ is a fixed constant parameter

Background: Relative Clique Relaxations

- A classical example of a relative relaxation is a γ-*quasi-clique*
  - a subset $S \subseteq V$ such that the subgraph $G[S]$ induced by S in G has edge density at least γ, i.e., $\rho(G[S]) = |(S \times S) \cap E|/\binom{|S|}{2} \geq \gamma$, where $\gamma \in [0, 1]$ is a fixed constant parameter
  - Clearly, requiring $\gamma = 1$ results in a clique
  - The problem of finding maximum γ-quasi-cliques is known to be *NP*-hard for any fixed $\gamma = p/q$, where $p,\ q \in \mathbb{Z}_{>0}$ and $p \leq q$ (Patillo et al., 2013)
- An alternative (degree-based) definition of a γ-*quasi-clique*
  - a subset $S \subseteq V$ such that in the subgraph $G[S]$ induced by S in G degree of every node is at least $\gamma \cdot (|S| - 1)$, i.e., $\deg_{G[S]}(i) \geq \gamma \cdot (|S| - 1)$ for any $i \in S$, where $\gamma \in [0, 1]$ is a fixed constant parameter
- Note that a γ-quasi-clique may be a *disconnected graph*, which is often mentioned as the key disadvantage of this relative clique relaxation model

γ-Relative-Vertex-Connected Subgraph

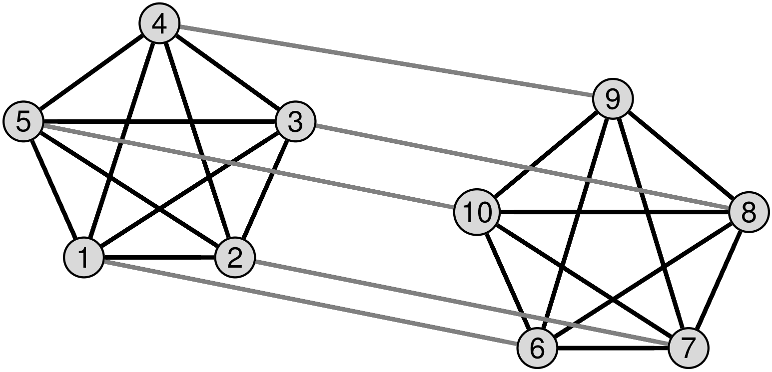
Definition

Given $G = (V, E)$ and a fixed $\gamma \in [0, 1]$, a subgraph $G[S]$, $S \subseteq V$, is called γ*-relative-vertex-connected* (or relative γ-vertex-connected) if the minimum number of vertices, whose removal disconnects $G[S]$ (or results in a trivial subgraph with exactly one vertex), is at least $\gamma(|S| - 1)$

γ-Relative-Vertex-Connected Subgraph

Definition

Given $G = (V, E)$ and a fixed $\gamma \in [0, 1]$, a subgraph $G[S]$, $S \subseteq V$, is called γ*-relative-vertex-connected* (or relative γ-vertex-connected) if the minimum number of vertices, whose removal disconnects $G[S]$ (or results in a trivial subgraph with exactly one vertex), is at least $\gamma(|S| - 1)$



γ-Relative-Vertex-Connected Subgraph

Definition

Given $G = (V, E)$ and a fixed $\gamma \in [0, 1]$, a subgraph $G[S]$, $S \subseteq V$, is called γ*-relative-vertex-connected* (or relative γ-vertex-connected) if the minimum number of vertices, whose removal disconnects $G[S]$ (or results in a trivial subgraph with exactly one vertex), is at least $\gamma(|S| - 1)$

Lemma

Graph $K_{(n,n)}$ is $\frac{1}{2}$-relative-vertex-connected

γ-Relative-Vertex-Connected Subgraph

Definition

Given $G = (V, E)$ and a fixed $\gamma \in [0, 1]$, a subgraph $G[S]$, $S \subseteq V$, is called γ*-relative-vertex-connected* (or relative γ-vertex-connected) if the minimum number of vertices, whose removal disconnects $G[S]$ (or results in a trivial subgraph with exactly one vertex), is at least $\gamma(|S| - 1)$

	max	max $-s$	$\gamma \cdot$ max	k
Edge density	$\lvert S\rvert(\lvert S\rvert - 1)/2$	s-defective clique	γ-quasi-clique	–
Min degree Connectivity	$\lvert S\rvert - 1$ $\lvert S\rvert - 1$	s-plex s-bundle	γ-quasi-clique γ-relative-connected	k-core k-block

- s-defective clique is a subset $S \subseteq V$ such that $G[S]$ contains at least $|S|(|S| - 1)/2 - s$ edges
- s-plex is a subset $S \subseteq V$ such that $\delta(G[S]) \geq |S| - s$
- s-bundle is a subset $S \subseteq V$ such that $\kappa(G[S]) \geq |S| - s$
- k-core is a subset $S \subseteq V$ such that $\delta(G[S]) \geq k$

f-Vertex-Connected Subgraph

Definition

Given $G = (V, E)$ and a function $f(\cdot)$ such that $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, a subgraph $G[S]$, $S \subseteq V$, is called f-vertex-connected if the minimum number of vertices, whose removal disconnects $G[S]$ (or results in a trivial graph with exactly one vertex) is at least $f(|S|)$

	max	max $-s$	$\gamma \cdot$ max	k
Edge density	$\|S\|(\|S\|-1)/2$	s-defective clique	γ-quasi-clique	–
Min degree	$\|S\|-1$	s-plex	γ-quasi-clique	k-core
Connectivity	$\|S\|-1$	s-bundle	γ-relative-connected	k-block

f-Vertex-Connected Subgraph

Definition

Given $G = (V, E)$ and a function $f(\cdot)$ such that $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, a subgraph $G[S]$, $S \subseteq V$, is called f-vertex-connected if the minimum number of vertices, whose removal disconnects $G[S]$ (or results in a trivial graph with exactly one vertex) is at least $f(|S|)$

	max	$\max - s$	$\gamma \cdot \max$	k
Edge density	$\lvert S\rvert(\lvert S\rvert - 1)/2$	s-defective clique	γ-quasi-clique	–
Min degree	$\lvert S\rvert - 1$	s-plex	γ-quasi-clique	k-core
Connectivity	$\lvert S\rvert - 1$	s-bundle	γ-relative-connected	k-block

- clique: $\gamma = 1$, or $f(|S|) = |S| - 1$
- k-block: $f(|S|) = k$
- s-bundle: $f(|S|) = |S| - s$
- γ-relative-vertex-connected: $f(|S|) = \gamma(|S| - 1)$

Optimization and Decision Problems

- We consider the problem of finding a maximum (in terms of cardinality $|S|$, $S \subseteq V$) subgraph $G[S]$ that is γ-relative-vertex-connected
  - $\gamma = 0$ corresponds to a polynomially solvable case as any graph G is 0-relative-vertex-connected
  - $\gamma = 1$ reduces to the classical maximum clique problem
- We refer to the decision version of this problem as the γ-RELATIVE-VERTEX-CONNECTED subgraph problem

Optimization and Decision Problems

- We consider the problem of finding a maximum (in terms of cardinality $|S|$, $S \subseteq V$) subgraph $G[S]$ that is γ-relative-vertex-connected
  - $\gamma = 0$ corresponds to a polynomially solvable case as any graph G is 0-relative-vertex-connected
  - $\gamma = 1$ reduces to the classical maximum clique problem
- We refer to the decision version of this problem as the γ-RELATIVE-VERTEX-CONNECTED subgraph problem
- Similarly, for a fixed function $f(\cdot)$ we consider the problem of finding a maximum (in terms of cardinality, $|S|$) f-vertex-connected subgraph $G[S]$, $S \subseteq V$
- Its decision version is referred to as f-VERTEX-CONNECTED subgraph problem

Computational Complexity

Proposition

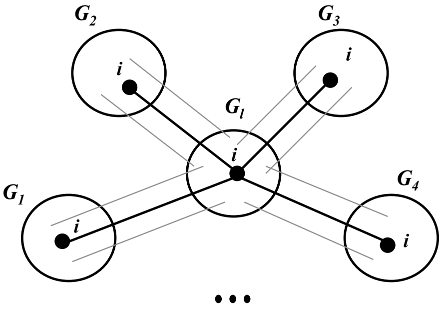
The $\frac{1}{\ell}$-RELATIVE-VERTEX-CONNECTED subgraph problem is NP-complete for any fixed positive integer ℓ

Computational Complexity

Proposition

The $\frac{1}{\ell}$-RELATIVE-VERTEX-CONNECTED subgraph problem is NP-complete for any fixed positive integer ℓ

- Key idea of the proof:

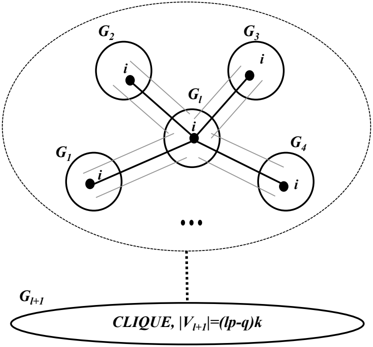


Computational Complexity

Proposition

The $\frac{p}{q}$-RELATIVE-VERTEX-CONNECTED subgraph problem is NP-complete for any fixed positive integers p and q such that $\frac{p}{q} \in (0,1]$

- Key idea of the proof:



Computational Complexity

Proposition

The f-VERTEX-CONNECTED subgraph problem is NP-complete for

(i) $f(|S|) = |S|^{1-\alpha} - 1$ *and any fixed* α *such that* $\alpha \in [0, \frac{1}{2})$

(ii) $f(|S|) = \gamma(|S| - |S|^{\alpha})$ *and any fixed* α *and* γ *such that* $\alpha \in [0, 1)$ *and* $\gamma = \frac{p}{q} \in (0, 1]$, *where p and q are positive integers*

- Recall that the problem of finding maximum
  - *k*-block, i.e., $f(|S|) = k$, is polynomially solvable
  - clique, i.e., $\gamma = 1$, or $f(|S|) = |S| - 1$, is *NP*-hard
  - *s*-bundle, i.e., $f(|S|) = |S| - s$, is *NP*-hard
  - γ-relative-vertex-connected subgraph, i.e., $f(|S|) = \gamma(|S| - 1)$, is *NP*-hard

Flow-based MIP Model

- There should be at least $\gamma(|S|-1)$ vertex-disjoint paths between any pair of vertices in $G[S]$

$$(\gamma\textbf{-CP}): \quad \max \sum_{k=1}^{|V|} x_k$$

subject to

$$\sum_{j:(s,j)\in E} u_{sj}^{st} - \sum_{i:(i,s)\in E} u_{is}^{st} \geq \gamma \left(\sum_{k=1}^{|V|} x_k - 1 \right) + (\gamma|V|-1)(x_s + x_t - 2) \quad \forall\, s < t$$

$$\sum_{i:(i,t)\in E} u_{it}^{st} - \sum_{j:(t,j)\in E} u_{tj}^{st} \geq \gamma \left(\sum_{k=1}^{|V|} x_k - 1 \right) + (\gamma|V|-1)\,(x_s + x_t - 2) \quad \forall\, s < t$$

$$\sum_{j:(k,j)\in E} u_{kj}^{st} \leq x_k \qquad \forall s,t,\ s<t,\ \forall k \in V \setminus \{s,t\}$$

$$\sum_{j:(i,j)\in E} \left(u_{ij}^{st} - u_{ji}^{st} \right) = 0 \qquad \forall s,t,\ s<t,\ \forall i \in V \setminus \{s,t\}$$

$$x_k \in \{0,1\},\ 0 \leq u_{ij}^{st} \leq 1 \qquad \forall s,t,k,\ s<t,\ \forall (i,j) \in E$$

Flow-based MIP Model: *f*-vertex-connected version

- We use value disjunctions to represent function $f(\cdot)$
- Given $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, define $c_k = \lceil f(k) \rceil$, $k = 1, \ldots, |V|$
- Let $\bar{c} = \max_{1 \le k \le |V|} c_k$
- The modified model is referred to as ***f*-CP**, with the "key" constraints given by:

$$\sum_{j:(s,j)\in E} u_{sj}^{st} - \sum_{i:(i,s)\in E} u_{is}^{st} \ge \sum_{k=1}^{|V|} c_k z_k + \bar{c}(x_s + x_t - 2) \qquad \forall s,t \in V,\ s<t$$

$$\sum_{i:(i,t)\in E} u_{it}^{st} - \sum_{j:(t,j)\in E} u_{tj}^{st} \ge \sum_{k=1}^{|V|} c_k z_k + \bar{c}(x_s + x_t - 2) \qquad \forall s,t \in V,\ s<t$$

$$\sum_{k=1}^{|V|} x_k = \sum_{k=1}^{|V|} k z_k, \qquad \sum_{k=1}^{|V|} z_k = 1$$

Flow-based MIP Model: f-vertex-connected version

- We use value disjunctions to represent function $f(\cdot)$
- Given $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, define $c_k = \lceil f(k) \rceil$, $k = 1, \ldots, |V|$
- Let $\bar{c} = \max_{1 \le k \le |V|} c_k$
- The modified model is referred to as *f*-**CP**, with the "key" constraints given by:

$$\sum_{j:(s,j)\in E} u_{sj}^{st} - \sum_{i:(i,s)\in E} u_{is}^{st} \ge \sum_{k=1}^{|V|} c_k z_k + \bar{c}(x_s + x_t - 2) \qquad \forall s,t \in V,\ s < t$$

$$\sum_{i:(i,t)\in E} u_{it}^{st} - \sum_{j:(t,j)\in E} u_{tj}^{st} \ge \sum_{k=1}^{|V|} c_k z_k + \bar{c}(x_s + x_t - 2) \qquad \forall s,t \in V,\ s < t$$

$$\sum_{k=1}^{|V|} x_k = \sum_{k=1}^{|V|} k z_k, \qquad \sum_{k=1}^{|V|} z_k = 1$$

- The number of binary and continuous variables in the proposed formulations is $O(|V|)$ and $O(|V|^2|E|)$, respectively; the number of constraints is $O(|V|^3)$

Exact Iterative Algorithm (**EIA**): Idea

- Simple observation:
  - Let $L \in \{1, \ldots, |V|\}$
  - Given a subset $S_L \subseteq V$, if a subgraph $G[S_L]$ is L-vertex-connected, i.e., $\kappa(G[S_L]) \geq L$, and $f(|S_L|) \leq L$, then $G[S_L]$ is also f-vertex-connected
  - Conversely, if a subgraph $G[S]$ is f-vertex-connected, then there exists a non-negative integer L such that the vertex connectivity of $G[S]$ is at least L and $f(|S|) \leq L$
- Therefore, one should simply find the largest subgraph $G[S_L]$ such that its vertex connectivity is L and $f(|S_L|) \leq L$

Exact Iterative Algorithm (**EIA**): f-**CP**(L)

$$f\text{-}\mathbf{CP}(\mathrm{L}): \quad \max \sum_{k=1}^{|V|} x_k$$

subject to

$$\sum_{j:(s,j)\in E} u_{sj}^{st} - \sum_{i:(i,s)\in E} u_{is}^{st} \geq L(x_s + x_t - 1) \qquad \forall s,t \in V,\ s<t$$

$$\sum_{i:(i,t)\in E} u_{it}^{st} - \sum_{j:(t,j)\in E} u_{tj}^{st} \geq L(x_s + x_t - 1) \qquad \forall s,t \in V,\ s<t$$

$$\sum_{k=1}^{|V|} x_k \leq \max\left\{i \,\middle|\, f(i) \leq L,\ i = 1,\ldots,|V|\right\}$$

$$x_k \in \{0,1\},\ 0 \leq u_{ij}^{st} \leq 1, \qquad \forall s,t,k \in V,\ s<t\ \forall (i,j)\in E$$

- Model f-**CP**(L) is simpler and easier to solve than f-**CP** due to the simpler structure of the constraint right-hand sides

Exact Iterative Algorithm (**EIA**)

Proposition

*Given a graph $G = (V, E)$ and a function $f(\cdot)$ such that $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, let $S_L \subseteq V$ be an optimal solution of f-**CP**(L) , where $L \in \{1, \ldots, |V|\}$. Denote by S^* a subset of V that induces a maximum f-vertex-connected subgraph of G, i.e., S^* is an optimal solution of f-**CP**. Then:*

$$|S^*| = \max_{L \in \{1,\ldots,|V|\}} \left\{|S_L|\right\}$$

Exact Iterative Algorithm (**EIA**)

Proposition

*Given a graph $G = (V, E)$ and a function $f(\cdot)$ such that $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, let $S_L \subseteq V$ be an optimal solution of f-**CP**(L) , where $L \in \{1, \ldots, |V|\}$. Denote by S^* a subset of V that induces a maximum f-vertex-connected subgraph of G, i.e., S^* is an optimal solution of f-**CP**. Then:*

$$|S^*| = \max_{L \in \{1,\ldots,|V|\}} \big\{|S_L|\big\}$$

- We iteratively decrease L starting with $L = \max_{v \in V}\{deg_G(v)\}$

Exact Iterative Algorithm (**EIA**)

Proposition

Given a graph $G = (V, E)$ *and a function* $f(\cdot)$ *such that* $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$*, let* $S_L \subseteq V$ *be an optimal solution of* f*-**CP**(L) , where* $L \in \{1, \ldots, |V|\}$*. Denote by* S^* *a subset of* V *that induces a maximum* f*-vertex-connected subgraph of G, i.e.,* S^* *is an optimal solution of* f*-**CP**. Then:*

$$|S^*| = \max_{L \in \{1,\ldots,|V|\}} \{|S_L|\}$$

- We iteratively decrease L starting with $L = \max_{v \in V}\{deg_G(v)\}$
- We should consider all $L \in \{1, \ldots, |V|\}$ for general $f(\cdot)$

Exact Iterative Algorithm (**EIA**)

Proposition

*Given a graph $G = (V, E)$ and a function $f(\cdot)$ such that $f : \mathbb{Z}_{>0} \to \mathbb{R}_+$, let $S_L \subseteq V$ be an optimal solution of f-**CP**(L) , where $L \in \{1, \ldots, |V|\}$. Denote by S^* a subset of V that induces a maximum f-vertex-connected subgraph of G, i.e., S^* is an optimal solution of f-**CP**. Then:*

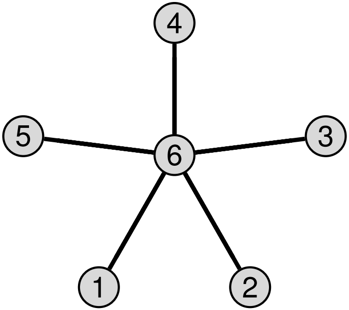
$$|S^*| = \max_{L \in \{1,\ldots,|V|\}} \{|S_L|\}$$

- We iteratively decrease L starting with $L = \max_{v \in V}\{deg_G(v)\}$
- We should consider all $L \in \{1, \ldots, |V|\}$ for general $f(\cdot)$
- If $f(\cdot)$ is non-decreasing, then we stop when the right-hand side of the "budget" constraint $\max\left\{i \mid f(i) \le L,\ 1 \le i \le |V|\right\}$ is equal or smaller than the size of the best solution identified at previous iterations

Special Case: γ-relative-robust and *f*-robust 2-clubs

Special Case: γ-relative-robust and f-robust 2-clubs

- 2-clubs ensure that there exists a very “short” path (of length at most 2) between any pair of vertices
- However, unlike cliques such graphs are potentially still vulnerable to specific vertex-targeted attacks (or errors)
- Example of a vulnerable 2-club:



Special Case: γ-relative-robust and f-robust 2-clubs

- R-robust 2-club proposed (Veremyev and Boginski, 2012)
  - a subset $S \subseteq V$ such that for any pair of vertices v and v' in S there exist at least R vertex-disjoint paths of length at most 2 connecting them in $G[S]$

Special Case: γ-relative-robust and f-robust 2-clubs

- R-robust 2-club proposed (Veremyev and Boginski, 2012)
  - a subset $S \subseteq V$ such that for any pair of vertices v and v' in S there exist at least R vertex-disjoint paths of length at most 2 connecting them in $G[S]$
- However, parameter R is fixed and does not depend on the size of the subgraph
- Thus, it is natural to consider more "robust" (i.e., vertex-attack tolerant) subgraphs that require a larger number of vertex-disjoint "short" paths as their sizes grow

Special Case: γ-relative-robust and f-robust 2-clubs

- γ-relative-robust 2-club
  - Given a graph $G = (V, E)$ and a fixed parameter $\gamma \in (0, 1]$, a subset S, $S \subseteq V$, is called a γ-relative-robust 2-club if for any pair of vertices v and v' in S there exist at least $\gamma(|S| - 1)$ vertex-disjoint paths of length at most 2 connecting them in $G[S]$

Special Case: γ-relative-robust and f-robust 2-clubs

- γ-relative-robust 2-club
  - Given a graph $G = (V, E)$ and a fixed parameter $\gamma \in (0, 1]$, a subset S, $S \subseteq V$, is called a γ-relative-robust 2-club if for any pair of vertices v and v' in S there exist at least $\gamma(|S| - 1)$ vertex-disjoint paths of length at most 2 connecting them in $G[S]$
- Observe that 1-relative-robust 2-club is a clique

Special Case: γ-relative-robust and f-robust 2-clubs

- γ-relative-robust 2-club
  - Given a graph $G = (V, E)$ and a fixed parameter $\gamma \in (0, 1]$, a subset S, $S \subseteq V$, is called a γ-relative-robust 2-club if for any pair of vertices v and v' in S there exist at least $\gamma(|S| - 1)$ vertex-disjoint paths of length at most 2 connecting them in $G[S]$
- Observe that 1-relative-robust 2-club is a clique

Proposition

For any fixed $\gamma \in (0, 1]$, if $S \subseteq V$ is a γ-relative-robust 2-club, then $G[S]$ is also a γ-relative-vertex-connected subgraph

Special Case: γ-relative-robust and f-robust 2-clubs

- γ-relative-robust 2-club
  - Given a graph $G = (V, E)$ and a fixed parameter $\gamma \in (0, 1]$, a subset S, $S \subseteq V$, is called a γ-relative-robust 2-club if for any pair of vertices v and v' in S there exist at least $\gamma(|S| - 1)$ vertex-disjoint paths of length at most 2 connecting them in $G[S]$
- Observe that 1-relative-robust 2-club is a clique

Proposition

For any fixed $\gamma \in (0, 1]$, if $S \subseteq V$ is a γ-relative-robust 2-club, then $G[S]$ is also a γ-relative-vertex-connected subgraph

Proposition

For any fixed $\gamma \in (1/2, 1]$ and $S \subseteq V$, if $G[S]$ is a γ-relative-vertex-connected subgraph then S is also a $2(\gamma - \frac{1}{2})$-relative-robust 2-club

Special Case: γ-relative-robust and f-robust 2-clubs

Proposition

The $\frac{p}{q}$-RELATIVE-ROBUST 2-CLUB problem is NP-complete for any fixed positive integers p and q such that $\frac{p}{q} \in (0, 1]$

Special Case: γ-relative-robust and f-robust 2-clubs

- f-robust 2-club
  - Given a graph $G = (V, E)$ and a function $f(\cdot)$ such that $f : \mathbb{Z}_{>0} \to [1, +\infty)$, a subset S, $S \subseteq V$, is called an f-robust 2-club if for any pair of vertices v and v' in S there exist at least $f(|S|)$ vertex-disjoint paths of length at most 2 connecting them in $G[S]$
- Observe that
  - If $f(|S|) = R$ (i.e., constant) then we have R-robust 2-clubs
  - If $f(|S|) = \gamma(|S| - 1)$ then we have γ-relative-robust 2-clubs

γ-relative-robust and f-robust 2-clubs: MIPs and EIA

$$(\gamma\text{-}\mathbf{RCP}): \quad \max \sum_{k=1}^{|V|} x_k$$

subject to

$$\mathbb{1}_{\{(i,j)\in E\}} + \sum_{k:\{(i,k),(k,j)\}\subseteq E} x_k \geq \gamma\left(\sum_{k=1}^{|V|} x_k - 1\right) + (\gamma|V|-1)(x_i + x_j - 2) \qquad \forall i,j \in V$$

$$x_k \in \{0,1\} \qquad \forall k \in V$$

γ-relative-robust and f-robust 2-clubs: MIPs and EIA

$$(\gamma\text{-}\mathbf{RCP}): \quad \max \sum_{k=1}^{|V|} x_k$$

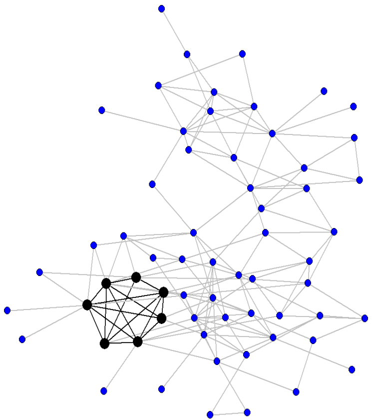
subject to

$$\mathbb{1}_{\{(i,j)\in E\}} + \sum_{k:\{(i,k),(k,j)\}\subseteq E} x_k \geq \gamma \left(\sum_{k=1}^{|V|} x_k - 1 \right) + (\gamma|V| - 1)(x_i + x_j - 2) \qquad \forall i, j \in V$$

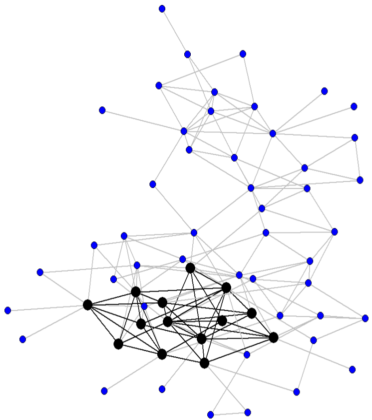
$$x_k \in \{0, 1\} \qquad \forall k \in V$$

- Similar to the general case, we can derive an MIP formulation (referred to as f-**RCP**) for finding maximum f-robust 2-clubs using the value disjunction idea
- Furthermore, **EIA** (with minor modifications) can be used to solve f-**RCP**

Illustrative Examples: Dolphins ($|V| = 62, |E| = 159$)

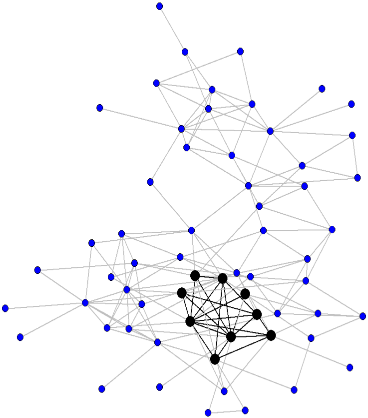


(a) 0.3-relative-robust 2-club, $|S| = 7$

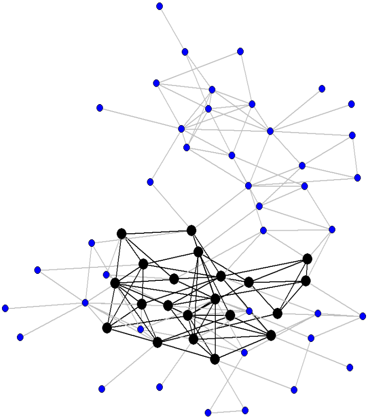


(b) 0.3-relative-vertex-connected, $|S| = 14$

Illustrative Examples: Dolphins ($|V| = 62, |E| = 159$)



(a) 0.2-relative-robust 2-club, $|S| = 9$



(b) 0.2-relative-vertex-connected, $|S| = 21$

Illustrative Results: Sizes of the Maximum Subgraphs

γ-relative-vertex-connected subgraphs

Network Parameters				max	$\gamma = \dots$					
Name	Type	$\|V\|$	$\|E\|$	clique	0.6	0.5	0.4	0.3	0.2	0.1
karate	Social	34	78	5	6	9	10	11	16	21
chesapeake	Food	39	170	5	11	13	16	21	26	33
huck	Book	74	301	11	11	13	16	17	21	31
miles250	Geo	128	387	8	11	13	16	18	26	41

γ-relative-robust 2-clubs

Network Parameters				$\gamma = \dots$					
Name	Type	$\|V\|$	$\|E\|$	0.6	0.5	0.4	0.3	0.2	0.1
karate	Social	34	78	6	6	6	7	11	12
chesapeake	Food	39	170	7	9	11	12	14	21
huck	Book	74	301	11	11	13	14	17	21
miles250	Geo	128	387	10	11	11	13	14	16
miles500	Geo	128	1170	26	27	28	31	32	36
USAir97	Transport	332	2126	34	36	39	43	49	55
NetScience	Collaboration	1589	2742	20	20	20	20	20	21
h.pylori	Biological	1570	1399	4	5	6	7	7	11
s.cerevisae	Biological	2112	2203	7	7	7	7	11	12

Illustrative Results: **EIA** Running Times (in seconds)

γ-relative-vertex-connected subgraphs

Graph	$\lvert V\rvert$	$\lvert E\rvert$	max clique	$\gamma = \dots$ 0.6	0.5	0.4	0.3	0.2	0.1
karate	34	78	0.1	0.8	0.2	0.2	3.2	3.2	14.9
chesapeake	39	170	0.1	5.9	8.3	29.7	22.0	4.6	19.4
huck	74	301	51.3	182.2	400.7	365.5	1561.7	2929.8	2406.9
miles250	128	387	1.5	7.7	34.6	30.1	64.9	166.2	7209.2

γ-relative-robust 2-clubs

Name	$\lvert V\rvert$	$\lvert E\rvert$	$\gamma = \dots$ 0.6	0.5	0.4	0.3	0.2	0.1
karate	34	78	0.1	0.1	0.1	0.1	0.1	0.1
chesapeake	39	170	0.1	0.1	0.1	0.1	0.1	0.2
huck	74	301	0.1	0.2	0.3	0.3	0.3	0.5
miles250	128	387	0.1	0.1	0.1	0.3	0.5	0.7
miles500	128	1170	0.4	0.5	0.9	1.4	2.4	4.1
USAir97	332	2126	2.7	2.8	3.4	4.4	5.8	9.6
NetScience	1589	2742	0.5	0.6	1.2	2.2	56.7	761.9
h.pylori	1570	1399	20.8	25.9	27.4	23.9	30.8	92.4
s.cerevisae	2112	2203	0.1	0.1	1.3	57.7	57.9	73.3

Comparisons of **EIA** vs. MIPs

- The numbers are given in ratios (MIP solution time)/(EIA solution time). If γ-**CP** does not find an optimal solution within the allotted time limit of 50000 seconds, the ratio 50000/(EIA time) is reported and "$>$" is placed in front of it.

Graph	$\|V\|$	$\|E\|$	γ 0.6	0.5	0.4	0.3	0.2	0.1	Average Ratio	# Iter
						EIA vs. γ-**CP**				
karate	34	78	21.85	120.4	1.02	8.13	34.22	21.56	23.46	1.44
chesapeake	39	170	18.22	30.75	8.22	3.79	67.11	30.75	19.1	1.67
dolphins	62	159	34.67	21.05	12.08	12.05	35.16	5.74	16.63	1.11
kreb	62	153	3.12	36.1	17.28	15.34	33.54	13.67	18.08	2.11
						EIA vs. γ-**RCP**				
jazz	198	2742	0.94	170.6	540.38	1876.82	2040.91	$>$2183.41	$>$757.05	13.89
USAir97	332	2126	2.99	9.9	14.99	13.3	45.78	1184.79	141.97	10.78
HarvardWeb	500	2043	0.63	0.76	0.52	57.43	188.21	2016.73	251.87	9.89
emails	1133	5451	0.01	0.48	4.12	24.49	$>$90.06	$>$38.14	$>$17.78	5.33

Concluding Remarks

- We have proposed a family of flexible and intuitive clique relaxation models based on the concept of relative vertex connectivity and its functional generalizations
- Our work has been mostly focused on graph-theoretical modeling and computational complexity issues as well as MIP-based exact solution approaches
  - The computational results demonstrate that the proposed approach can handle reasonably large instances of sparse graphs as long as vertex connectivity with "short" paths is required
  - Development of more advanced solution methods for large-scale and dense graph problems with general vertex connectivity metrics is a promising direction of future research
- The overall framework can be naturally extended to consider relative edge connectivity