The Safe Set Problem on particular graph classes

Roberto Cordone[1], Davide Franchi[2]

[1] Università degli Studi di Milano, Dipartimento di Informatica, Milano, Italy
roberto.cordone@unimi.it
[2] davide9706@gmail.com

Keywords : *Safe Set Problem, Random graphs, Grid graphs.*

1 Introduction

Let $G = (V, E)$ be a connected undirected graph with $n = |V|$ and $m = |E|$. We denote as $G[T]$ the subgraph induced on G by a subset of vertices $T \subseteq V$ and as $\mathcal{C}_G(T)$ the collection of all maximal connected components in $G[T]$. The *Safe Set Problem* (*SSP*) amounts to finding a minimum cardinality nonempty set of vertices $S \subseteq V$ such that each component induced by S has a cardinality not smaller than that of any adjacent component induced by $U = V \setminus S$. We denote the vertices of S and the components of $\mathcal{C}_G(S)$ as *safe*, the vertices of U and the components of $\mathcal{C}_G(U)$ as *unsafe*. The *SSP* can be formulated as

$$\min_{S \subseteq V} z = |S| \tag{1}$$

$$|S_i| \geq |U_j| \qquad S_i \in \mathcal{C}_G(S), U_j \in \mathcal{C}_G(V \setminus S) : E(S_i, U_j) \neq \emptyset \tag{2}$$

$$S \neq \emptyset \tag{3}$$

where $E(S_i, U_j) = \{\{s, u\} \in E : s \in S_i, u \in U_j\}$ is the set of edges between S_i and U_j, and its optimal value is denoted as the *safe number* $s(G)$.

The literature provides a large number of recent theoretical results. The *SSP* is strongly $\mathcal{NP}$-hard on general graphs [6], but also on planar and split graphs, and not approximable within a factor better than 1.3606, unless $\mathcal{P} = \mathcal{NP}$ [1]. It is polynomial on bounded-treewidth graphs and interval graphs: trees admit an $O(n^5)$-time algorithm. It is fixed-parameter tractable with respect to neighbourhood diversity and to the size of the optimal solution [2]. The *Connected Safe Set Problem* (*CSSP*), whose optimum is named the *connected safe number* $cs(G)$, additionally bounds the solution to consist of a single safe component [6]. A number of contributions investigate the relation between $cs(G)$ and $s(G)$. While in general $s(G) \leq cs(G) \leq 2s(G) - 1$, the two values coincide for stars, paths and cycles [6], Cartesian products of complete graphs [10] and a family of connected bipartite graphs characterised in [7]. Special topologies allow not only to solve the *SSP* in polynomial time, but even to find solutions in closed form, that is to determine specific subsets of vertices that provide *a priori* optimal solutions. Ladder, sunlet and wheel graphs have been studied in this regard by [9].

This paper provides new results, based on the matching between lower bounds and feasible solutions, for grids, complete bipartite graphs, and windmill graphs. Moreover, it investigates the asymptotic structure of optimal solutions for random graphs. We prove that, for a sufficiently large density, such instances become intrinsically easy as the size of the graph increases, with an optimum equal to $\lceil n/2 \rceil$, whereas they are generally unfeasible for small densities. This explains the empirical behaviour observed in algorithmic studies [3, 4].

2 Cardinality-separated solutions

We here define a special class of solutions of the *SSP* and prove a simple lower bound on their value. Since in several topologies at least one optimal solution falls within this special class, the lower bound holds for the safe number $s(G)$.

Definition 1 *A solution of the* SSP *is* cardinality-separated *when all unsafe components have cardinality not larger than all safe ones: $S_i \geq U_j$ for all $S_i \in \mathcal{C}_G(S)$ and $U_j \in \mathcal{C}_G(V \setminus S)$. By construction, all such solutions are feasible.*

Remark 1 *All feasible solutions with a single safe or unsafe component are cardinality-separated.*

Proof : If there is only one component of a given kind, the components of the other kind must be adjacent to it, because they cannot be adjacent to each other and the graph is connected. The feasibility constraint implies that feasible solutions are cardinality-separated. $\square$

Lemma 1 *Every cardinality-separated solution of the* SSP *that induces $k_S = |\mathcal{C}_G(S)|$ safe components and $k_U = |\mathcal{C}_G(V \setminus S)|$ unsafe ones has value $z \geq \left\lceil n \cdot \dfrac{k_S}{k_S + k_U} \right\rceil$.*

Proof : Summing inequalities $|S_i| \geq |U_j|$ on all pairs (i, j) implies $\sum\limits_{j=1}^{k_U} \sum\limits_{i=1}^{k_S} |S_i| \geq \sum\limits_{i=1}^{k_S} \sum\limits_{j=1}^{k_U} |U_j|$. Since $z = \sum\limits_{i=1}^{k_S} |S_i| = n - \sum\limits_{j=1}^{k_U} |U_j|$, we conclude that $\sum\limits_{j=1}^{k_U} z \geq \sum\limits_{i=1}^{k_S} (n - z) \Rightarrow k_U z \geq k_S n - k_S z$. $\square$

Corollary 1 *Given a complete graph K_n, any subset of $\lceil n/2 \rceil$ vertices is an optimal solution of the* SSP*.*

Proof : As in a complete graph all safe vertices are adjacent and all unsafe vertices are adjacent, $k_S = k_U = 1$, and $s(G) \geq \lceil n/2 \rceil$. However, any subset S of $\lceil n/2 \rceil$ vertices is a feasible solution, since it is connected and larger than the unsafe set $V \setminus S$. $\square$

3 Random graphs

Corollary 1 admits an asymptotic generalization in probability to random graphs. We remind that in a random graph $G(n, p)$, as defined by Gilbert [8], there are n vertices and every unordered pair of vertices $\{u, v\}$ falls within the edge set E independently with probability p. We also remind that a sequence of events depending on an integer parameter n occurs "with high probability" when its probability converges to 1 as n grows to $+\infty$.

Theorem 1 *Let $G(n, p)$ be a Gilbert random graph with $p = \frac{\log n + c_n}{n}$. When $\lim\limits_{n \to +\infty} c_n = -\infty$, the* SSP *is infeasible with high probability; when $\lim\limits_{n \to +\infty} c_n = +\infty$, any subset of $\lceil n/2 \rceil$ vertices is an optimal solution with high probability.*

Proof : A Gilbert graph with n vertices, $p = \frac{\log n + c_n}{n}$ and $\lim\limits_{n \to +\infty} c_n = -\infty$ is disconnected with high probability [8]. The *SSP* is unfeasible on such graphs.

On the other hand, when $\lim\limits_{n \to +\infty} c_n = -\infty$, the graph is connected with high probability. Given a connected instance, any subset S of $\lceil n/2 \rceil$ vertices induces a graph $G[S]$ that is a Gilbert graph $G(\lceil n/2 \rceil, p)$, as the edges are still independently distributed with probability p. Therefore, $G[S]$ is also connected with high probability. Since S forms a single safe component and every unsafe component has cardinality $|U_j| \leq |V \setminus S| = \lfloor n/2 \rfloor \leq |S|$, S is feasible.

Consider any smaller set of vertices S (with $|S| = n' < \lceil n/2 \rceil$). The unsafe set induces a graph $G[V \setminus S]$ that is a Gilbert graph $G(n - n', p)$, once again connected with high probability. Therefore, S induces on the original graph a single unsafe component of cardinality $|U| = n - n' \geq \lceil n/2 \rceil$, which makes S unfeasible. Consequently, any subset of $\lceil n/2 \rceil$ vertices is an optimal solution with high probability. $\square$

The above finding explains the experimental results of [4], where nearly all random graphs with $n \geq 150$ and $p \in \{0.1, 0.2, 0.3, 0.4\}$ exhibit best known solutions with $z = \lceil n/2 \rceil$. Theorem 1 suggests that the interesting instances of the *SSP* are sparse, but with a topological structure that keeps them connected. In the following we consider some classes of instances with very strong structures.

4 Deterministic graphs with special topologies

A grid graph $L(r, c)$, with rows indexed from 1 to r and columns from 1 to c, is the Cartesian product of two path graphs $P(r)\Box P(c)$. In the following, we first prove a lower bound on $s(G)$, then we build a feasible solution of the *CSSP*, showing that the two bounds have the same asymptotic behaviour.

Proposition 1 *The safe number of a grid graph $G = L(r, c)$ with $c = \Theta(r)$ is $s(G) \in \Omega(r^{\frac{4}{3}})$.*

Proof : Given an optimal solution S of the *SSP*, let $\{U_1, \dots, U_{k_U}\}$ be the unsafe components induced by $G[V \setminus S]$, sorted by nonincreasing cardinalities, so that $|S| \geq |U_1| \geq \dots \geq |U_{k_U}|$. We extend G to $G' = L(r+2, c+2)$, with vertex set V' and edge set E', adding row indices 0 and $r+1$ and column indices 0 and $c+1$. For each unsafe component U_i, we denote the subset of adjacent vertices in V' as its *frontier* $F_i = \{v \in V' \setminus U_i : \exists u \in U_i : \{v, u\} \in E'\}$. Notice that any path from a vertex of U_i to $V' \setminus V$ necessarily intersects F_i, in particular the paths in which all vertices have the same row index (horizontal) or column index (vertical).

We first prove that $|F_i| \geq \sqrt{|U_i|}$. Denoting the vertices as index pairs (ρ, γ), let $r_i^m = \min\limits_{(\rho,\gamma)\in U_i} \rho$, $r_i^M = \max\limits_{(\rho,\gamma)\in U_i} \rho$, $c_i^m = \min\limits_{(\rho,\gamma)\in U_i} \gamma$ and $c_i^M = \max\limits_{(\rho,\gamma)\in U_i} \gamma$ be the extreme values, respectively, of the row and the column indices for the vertices of U_i. Then, let $b_i = c_i^M - c_i^m + 1$ and $h_i = r_i^M - r_i^m + 1$. Since U_i is connected, it contains a path with at least b_i vertices (ρ, γ), one for each different value of γ between c_i^m and c_i^M. Each vertex (ρ, γ) corresponds to at least two different vertices (r', γ) and $(r'', \gamma) \in F_i$, such that $r_i^m \leq r' < \rho < r'' \leq r_i^M$. Therefore, $|F_i| \geq 2b_i$. Analogously, $|F_i| \geq 2h_i$. As a consequence, $|F_i| \geq (b_i + h_i) \geq \sqrt{b_i h_i}$, but $b_i h_i \geq |U_i| \Rightarrow |F_i| \geq \sqrt{|U_i|}$.

The vertices of the frontier subsets are either safe or belong to the sides of the grid: $\bigcup_{i=1}^{k_U} F_i \subseteq S \cup (V' \setminus V \setminus \{(0,0),(0,c+1),(r+1,0),(r+1,c+1)\}) \Rightarrow |\bigcup_{i=1}^{k_U} F_i| \leq |S| + 2(r+c)$. Though in general they overlap, each vertex belongs to at most 4 different subsets: $\sum_{i=1}^{k_U} |F_i| \leq 4|\bigcup_{i=1}^{k_U} F_i|$, which implies that $|S| \geq 1/4 \sum_{i=1}^{k_U} \sqrt{|U_i|} - 2(r+c)$.

Now, let $n_1 = |U_1|$ for the sake of briefness. Since $|U_i| \leq n_1$ for all i, there is always a partition of the unsafe component indices $\{1, \dots, k_U\}$ into ℓ disjoint groups I_l such that $n_1 \leq \sum_{k\in I_l} |U_i| < 2n_1$ for $l = 1, \dots, \ell - 1$ and $\sum_{k\in I_\ell} |U_i| < n_1$. We minorize the total cardinality of each group of components by observing that $\sum_{i\in I_l} \sqrt{|U_i|} \geq \sqrt{\sum_{i\in I_l} |U_i|} \geq \sqrt{n_1}$ for all $l < \ell$. Then, we minorize the number of groups: $\ell \geq \frac{\sum_{i=1}^{k_U} |U_i|}{2n_1}$. Therefore, $\sum_{i=1}^{k_U} \sqrt{|U_i|} \geq \frac{\sum_{i=1}^{k_U} |U_i|}{2n_1} \sqrt{n_1} = \frac{rc - |S|}{2\sqrt{n_1}} \geq \frac{rc}{4\sqrt{n_1}}$, as $|S| \leq \frac{rc}{2}$ in any optimal solution [6]. Finally, $|S| \geq n_1$ implies $|S| \geq \max(n_1, \frac{rc}{16\sqrt{n_1}} - 2(r+c))$. Assuming that $c \in \Theta(r)$, it is $|S| \in \Omega(\max(n_1, \frac{r^2}{\sqrt{n_1}} - r))$. If $n_1 \in \Omega(r^{\frac{4}{3}})$, the thesis trivially holds. If, on the contrary, $n_1 \in O(r^{\frac{4}{3}})$, then $(r^2/\sqrt{n_1} - r) \in \Omega(r^2/r^{\frac{2}{3}} - r) = \Omega(r^{\frac{4}{3}})$, which is again the thesis. □

Proposition 2 *The connected safe number of a grid graph $G = L(r, c)$ with $c \in \Theta(r)$ is $cs(G) \in O(r^{4/3})$.*

Proof : Let the safe set S include all vertices with row and column index multiple of $l = \lceil \sqrt[3]{rc} \rceil$. This forms a connected grid of $|S| = c\left\lfloor \frac{r}{l} \right\rfloor + r\left\lfloor \frac{c}{l} \right\rfloor - \left\lfloor \frac{r}{l} \right\rfloor \left\lfloor \frac{c}{l} \right\rfloor \leq \frac{2rc}{l}$ vertices. The unsafe vertices form components that are squares or rectangles with at most $l - 1$ rows and columns. S is feasible when $c\left\lfloor \frac{r}{l} \right\rfloor + r\left\lfloor \frac{c}{l} \right\rfloor - \left\lfloor \frac{r}{l} \right\rfloor \left\lfloor \frac{c}{l} \right\rfloor \geq (l-1)^2$, that can be strengthened to $(c-1)(\frac{r}{l} - 1) + (r-1)(\frac{c}{l} - 1) \geq l^2$. As r and c increase, the inequality is asymptotically verified. Since $|S| \leq 2rc/\sqrt[3]{rc}$, assuming $c \in \Theta(r)$, this implies $|S| \in O(r^{\frac{4}{3}})$. □

Corollary 2 *Given a grid graph $G = L(r, c)$ with $c \in \Theta(r)$, both $s(G)$ and $cs(G) \in O(r^{4/3})$.*

Bipartite graphs have been discussed in [1], proving the NP-hardness of the *SSP* for planar graphs of degree ≤ 7, and in [5], stating that balanced complete graphs have optimal connected solution of value $cs(K_{n/2-1,n/2+1}) = n/2$. We here consider general complete bipartite graphs.

Proposition 3 *The* SSP *on a complete bipartite graph $K_{a,n-a}$ has an optimal solution of value* $\min(a, n-a)$ *composed by the vertices of the less numerous shore.*

Proof : The best solution including all the vertices on a shore consists exactly of the less numerous shore, and its cost is $\leq \lceil n/2 \rceil$. Any other solution has unsafe vertices on both shores, that are connected to each other and adjacent to all other vertices. Hence, there is a single unsafe component and $z \geq \left\lceil n \frac{k_S}{k_S+1} \right\rceil \geq \lceil n/2 \rceil$. □

A *windmill graph* $Wd(k,h)$ consists of $h \geq 2$ copies of the complete graph K_k and a central vertex linked to all other vertices.

Proposition 4 *The safe number of a windmill graph $Wd(k,h)$ is $s(G) = \left\lceil \frac{hk+1}{h+1} \right\rceil$.*

Proof : Since all vertices are adjacent to the central one, if this is unsafe, there is a single unsafe component, and $k_S \leq n$ safe components (the vertices in each complete subgraph are adjacent to each other). Therefore, $z \geq \left\lceil (hk+1) \frac{k_S}{k_S+1} \right\rceil \geq \lceil (hk+1)/2 \rceil$.

If the central vertex is safe, there is a single safe component and $k_U \leq n$ unsafe ones. Therefore, $z \geq \left\lceil \frac{hk+1}{1+h} \right\rceil = \left\lceil k - \frac{k-1}{1+h} \right\rceil = k - \left\lfloor \frac{k-1}{1+h} \right\rfloor$. This bound can be hit by setting a safe central vertex, $\left\lceil \frac{k-1}{h+1} \right\rceil$ vertices in $(k-1)-(h+1)\left\lfloor \frac{k-1}{h+1} \right\rfloor$ complete subgraphs and $\left\lfloor \frac{k-1}{h+1} \right\rfloor$ vertices in the other complete subgraphs. In fact, $|S| = 1 + \left\lceil \frac{k-1}{h+1} \right\rceil \left[(k-1) - (h+1) \left\lfloor \frac{k-1}{h+1} \right\rfloor \right] + \left\lfloor \frac{k-1}{h+1} \right\rfloor \left[h - (k-1) + (h+1) \left\lfloor \frac{k-1}{h+1} \right\rfloor \right] = 1 + (k-1) - (h+1) \left\lfloor \frac{k-1}{h+1} \right\rfloor + h \left\lfloor \frac{k-1}{h+1} \right\rfloor = k - \left\lfloor \frac{k-1}{h+1} \right\rfloor$, and the unsafe components have cardinality $\leq k - \left\lfloor \frac{k-1}{h+1} \right\rfloor$. □

References

[1] R. Àgueda, N. Cohen, S. Fujita, S. Legay, Y. Manoussakis, Y. Matsui, L. Montero, R. Naserasr, H. Ono, Y. Otachi, T. Sakuma, Z. Tuza, and R. Xu. Safe sets in graphs: Graph classes and structural parameters. *J. Comb. Optim.*, 36(4):1221–1242, 2018.

[2] R. Belmonte, T. Hanaka, I. Katsikarelis, M. Lampis, H. Ono, and Y. Otachi. Parameterized complexity of safe set. *J. Graph Algorithms Appl.*, 24(3):215–245, 2020.

[3] A. Boggio Tomasaz, R. Cordone, and P. Hosteins. A combinatorial branch-and-bound for the Weighted Safe Set Problem. *Networks*, 2022. [accepted].

[4] A. Boggio Tomasaz, R. Cordone, and P. Hosteins. Constructive-destructive heuristics for the safe set problem. *Computers & Operations Research*, 2023. [under review].

[5] S. Fujita and M. Furuya. Safe number and integrity of graphs. *Discrete Applied Mathematics*, 247:398–406, October 2018.

[6] S. Fujita, G. MacGillivray, and T. Sakuma. Safe set problem on graphs. *Discrete Applied Mathematics*, 215:106–111, 2016.

[7] S. Fujita, B. Park, and T. Sakuma. Stable structure on safe set problems in vertex-weighted graphs. *European Journal of Combinatorics*, 91:103211, 2021.

[8] E. N. Gilbert. Random Graphs. *Ann. Math. Stat.*, 30(4):1141–1144, 1959.

[9] R. Iqbal, M. S. Sardar, D. Alrowaili, S. Zafar, and I. Siddiqu. On the connected safe number of some classes of graphs. *Journal of Mathematics*, 2021:9483892, October 2021.

[10] B. Kang, S.-R. Kim, and B. Park. On the safe sets of Cartesian product of two complete graphs. *Ars Combinatoria*, 141:243–257, October 2018.