Connected Surveillance Game

Frédéric Giroire, Dorian Mazauric, Nicolas Nisse, Stéphane Pérennes, Ronan Soares

To cite this version:

Frédéric Giroire, Dorian Mazauric, Nicolas Nisse, Stéphane Pérennes, Ronan Soares. Connected Surveillance Game. 20th Colloquium on Structural Information and Communication Complexity (SIROCCO), Jul 2013, Ischia, Italy. pp.68-79. hal-00845531

HAL Id: hal-00845531

https://hal.archives-ouvertes.fr/hal-00845531

Submitted on 17 Jul 2013

Connected Surveillance Game*

F. Giroire[1], D. Mazauric[2], N. Nisse[1], S. Pérennes[1], and R. Soares[1,3]

[1] COATI, INRIA, I3S(CNRS/UNS), Sophia Antipolis, France
[2] ACRO, Laboratoire d'Informatique Fondamentale de Marseille, France
[3] ParGO Research Group, UFC, Fortaleza, Brazil

Abstract. The *surveillance game* [Fomin *et al.*, 2012] models the problem of web-page prefetching as a pursuit evasion game played on a graph. This two-player game is played turn-by-turn. The first player, called the *observer*, can mark a fixed amount of vertices at each turn. The second one controls a *surfer* that stands at vertices of the graph and can slide along edges. The surfer starts at some initially marked vertex of the graph, her objective is to reach an unmarked node The *surveillance number* $\mathrm{sn}(G)$ of a graph G is the minimum amount of nodes that the observer has to mark at each turn ensuring it wins against any surfer in G. Fomin *et al.* also defined the *connected surveillance game* where the marked nodes must always induce a connected subgraph. They ask if there is a constant $c > 0$ such that $\frac{\mathrm{csn}(G)}{\mathrm{sn}(G)} \leq c$ for any graph G. It has been shown that there are graphs G for which $\mathrm{csn}(G) = \mathrm{sn}(G) + 1$. In this paper, we investigate this question.
We present a family of graphs G such that $\mathrm{csn}(G) > \mathrm{sn}(G)+1$. Moreover, we prove that $\mathrm{csn}(G) \leq \mathrm{sn}(G)\sqrt{n}$ for any n-node graph G. While the gap between these bounds remains huge, it seems difficult to reduce it. We then define the *online surveillance game* where the observer has no *a priori* knowledge of the graph topology and discovers it little-by-little. Unfortunately, we show that no algorithm for solving the online surveillance game has competitive ratio better than $\Omega(\Delta)$.
Keywords: Surveillance game, Cops and robber games, Cost of connectivity, Online strategy, Competitive ratio, Prefetching.

1 Introduction

In this paper, we study two variants of the *surveillance game* introduced in [1]. This two-player game involves one Player moving a mobile agent, called *surfer*, along the edges of a graph, while a second Player, called *observer*, marks the vertices of the graph. The surfer wins if it manages to reach an unmarked vertex. The observer wins otherwise.

Surveillance game. More formally, let $G = (V, E)$ be an undirected simple n-node graph, $v_0 \in V$, and $k \in \mathbb{N}^*$. Initially, the surfer stands at v_0 which is marked and all other nodes are not marked. Then, turn-by-turn, the observer

---

* This work has been partially supported by European Project FP7 EULER, ANR CEDRE, ANR AGAPE, Associated Team AlDyNet, and project ECOS-Sud Chile.

first marks k unmarked vertices and then the surfer may move to a neighbor of her current position. Once a node has been marked, it remains marked.The surfer wins if, at some step, she reaches an unmarked vertex; and the observer wins otherwise. Note that the game lasts at most $\lceil \frac{n}{k} \rceil$ turns. When the game is played on a directed graph, the surfer has to follow arcs when it moves [1]. A *k-strategy for the observer from v_0*, or simply a *k-strategy from v_0*, is a function $\sigma : V \times 2^V \to 2^V$ that assigns the set $\sigma(v, M) \subseteq V$ of vertices, $|\sigma(v, M)| \leq k$, that the observer should mark in the *configuration* (v, M), where $M \subseteq V$, $v_0 \in M$, is the set of already marked vertices and $v \in M$ is the current position of the surfer. We emphasis that σ depends implicitly on the graph G, i.e., it is based on the full knowledge of G. A k-strategy from v_0 is *winning* if it allows the observer to win whatever be the sequence of moves of the surfer starting in v_0. The *surveillance number* of a graph G with initial node v_0, denoted by $\text{sn}(G, v_0)$, is the smallest k such that there exists a winning k-strategy starting from v_0.

Let us define some notations used in the paper. Let Δ be the maximum degree of the nodes in G and, for any $v \in V$, let $N(v)$ be the set of neighbors of v. More generally, the neighborhood $N(F)$ of a set $F \subseteq V$ is the subset of vertices of V which have a neighbor in F. Moreover, we define the closed neighborhood of a set F as $N[F] = N(F) \cup F$.

As an example, let us consider the following *basic strategy*: let $\sigma_{\mathcal{B}}$ be the strategy defined by $\sigma_{\mathcal{B}}(v, M) = N(v) \setminus M$ for any $M \subseteq V$, $v_0 \in M$, and $v \in M$. Intuitively, the basic strategy $\sigma_{\mathcal{B}}$ asks the observer to mark all unmarked neighbors of the current position of the surfer. It is straightforward, and it was already shown in [1], that $\sigma_{\mathcal{B}}$ is a winning strategy for any $v_0 \in V$ and it easily implies that $\text{sn}(G, v_0) \leq \max\{|N(v_0)|, \Delta - 1\}$.

Web-page prefetching, connected and online variants. The surveillance game has been introduced because it models the web-page prefetching problem. This problem can be stated as follows. A web-surfer is following the hyperlinks in the digraph of the web. The web-browser aims at downloading the web-pages before the web-surfer accesses it. The number of web-pages that the browser may download before the web-surfer accesses another web-page is limited due to bandwidth constraints. Therefore, designing efficient strategies for the surveillance game would allow to preserve bandwidth while, at the same time, avoiding the waiting time for the download of the web-page the web-surfer wants to access.

By nature of the web-page prefetching problem, in particular because of the huge size of the web digraph, it is not realistic to assume that a strategy may mark any node of the network, even nodes that are "far" from the current position of the surfer. For this reason, [1] defines the *connected* variant of the surveillance game. A strategy σ is said *connected* if $\sigma(v, M) \cup M$ induces a connected subgraph of G for any M, $v_0 \in M \subseteq V(G)$. Note that the basic strategy $\sigma_{\mathcal{B}}$ is connected. The *connected surveillance number* of a graph G with initial node v_0, denoted by $\text{csn}(G, v_0)$, is the smallest k such that there exists a winning connected k-strategy starting from v_0. By definition, $\text{csn}(G, v_0) \geq \text{sn}(G, v_0)$ for any graph G and $v_0 \in V(G)$. In [1], it is shown that there are graphs G and $v_0 \in V(G)$ such that $\text{csn}(G, v_0) = \text{sn}(G, v_0) + 1$. Only the trivial upper bound $\mathrm{csn}(G, v_0) \leq \Delta\, \mathrm{sn}(G, v_0)$ is known and a natural question is how big the gap between $\mathrm{csn}(G, v_0)$ and $\mathrm{sn}(G, v_0)$ may be [1]. This paper provides a partial answer to this question.

Still the connected surveillance game seems unrealistic since the web-browser cannot be asked to have the full knowledge of the web digraph. For this reason, we define the *online surveillance game*. In this game, the observer discovers the considered graph while marking its nodes. That is, initially, the observer only knows the starting node v_0 and its neighbors. After the observer has marked the subset M of nodes, it knows M and the vertices that have a neighbor in M and the next set of vertices to be marked depends only on this knowledge, i.e., the nodes at distance at least two from M are unknown. In other words, an *online strategy* is based on the current position of the surfer, the set of already marked nodes and knowing only the subgraph H of the marked nodes and their neighbors (a more formal definition is postponed to Section 3). By definition, the next nodes marked by such a strategy must be known, i.e., adjacent to an already marked vertex. Therefore, an online strategy is connected. We are interested in the competitive ratio of winning online strategies. The competitive ratio $\rho(\mathcal{S})$ of a winning online strategy $\mathcal{S}$ is defined as $\rho(\mathcal{S}) = \max_{G, v_0 \in V(G)} \frac{\mathcal{S}(G, v_0)}{\mathrm{sn}(G, v_0)}$, where $\mathcal{S}(G, v_0)$ denotes the maximum number of vertices marked by $\mathcal{S}$ in G at each turn, when the surfer starts in v_0. Note that, because any online winning strategy $\mathcal{S}$ is connected, $\mathrm{csn}(G, v_0) \leq \rho(\mathcal{S})\, \mathrm{sn}(G, v_0)$ for any graph G and $v_0 \in V(G)$.

1.1 Related work

The surveillance game has mainly been studied in the computational complexity point of view. It is shown that the problem of computing the surveillance number is NP-hard in split graphs [1]. Moreover, deciding whether the surveillance number is at most 2 is NP-hard in chordal graphs and deciding whether the surveillance number is at most 4 is PSPACE-complete. Polynomial-time algorithms that compute the surveillance number in trees and interval graphs are designed in [1]. All previous results also hold for the connected surveillance number. Finally, it is shown that, for any graph G and $v_0 \in V(G)$, $\max\lceil \frac{|N[S]|-1}{|S|} \rceil \leq \mathrm{sn}(G, v_0) \leq \mathrm{csn}(G, v_0)$ where the maximum is taken over every subset $S \subseteq V(G)$ inducing a connected subgraph with $v_0 \in S$. Moreover, both previous inequalities turn into an equality in case of trees. [1] asks for an example where the inequalities are strict.

In the literature, there are mainly three types of prefetching: server based hints prefetching [2–4], local prefetching [5] and proxy based prefetching [6]. In local prefetching, the client has no aid from the server when deciding which documents to prefetch. In the server based hints prefetching, the server can aid the client to decide which pages to prefetch. Lastly, in the proxy based prefetching, a proxy that connects its clients with the server decides which pages to prefetch. Moreover, some studies consider that the prefetching mechanism has perfect knowledge of the web-surfer's behaviour [7, 8]. In these studies, the objective is to minimize the waiting time of the web-surfer with a given bandwidth, by designing good prediction strategies for which pages to prefetch. In the context of prefetching web-pages, the surveillance game is a model to study a local prefetching scheme to guarantee that a websurfer never has to wait a web-page to be downloaded, whilst minimizing the bandwidth necessary to achieve this.

1.2 Our results

In this paper, we study both the connected and online variants of the surveillance game. First, we try to evaluate the gap between non-connected and connected surveillance number of graphs. We give a new upper bound, independent from the maximum degree, for the ratio csn / sn. More precisely, we show that, for any n-node graph G and any $v_0 \in V(G)$, $\mathrm{csn}(G, v_0) \leq \mathrm{sn}(G, v_0)\sqrt{n}$. Then, we describe a family of graphs G such that $\mathrm{csn}(G, v_0) = \mathrm{sn}(G, v_0) + 2$. Note that, contrary to the simple example that shows that connected and not connected surveillance number may differ by one, a larger difference seems much more difficult to obtain.

As mentioned above, the online variant of the surveillance game is a more constraint version of the connected game. We prove that any online strategy has competitive ratio at least $\Omega(\Delta)$. More formally, we describe a familly of trees with constant surveillance number such that, for any online winning strategy, there is a step when the strategy has to mark at least $\frac{\Delta}{4}$ vertices.Unfortunately, this shows that the best (up to constant ratio) online strategy is the basic one.

2 Cost of connectedness

In this section, we investigate the cost of the connectivity constraint. We first prove the first non-trivial upper bound for the ratio csn / sn. More precisely, we show that for any n-node graph G, $\mathrm{csn}(G, v_0) \leq \mathrm{sn}(G, v_0)\sqrt{n}$. Then, we improve the lower bound of [1]. That is, we show a family of graphs where $\mathrm{csn}(G, v_0) > \mathrm{sn}(G, v_0) + 1$. Finally, we disprove a conjecture in [1].

2.1 Upper bound

In this section, we give the first non-trivial upper bound (independent from the degree) of the cost of the connectivity in the surveillance game.

Theorem 1. *Let G be any connected n-node graph and $v_0 \in V(G)$, then*

$$\mathrm{csn}(G, v_0) \leq \mathrm{sn}(G, v_0)\sqrt{n}.$$

Proof. $\mathrm{sn}(G, v_0) = 1$ if and only if G is a path with v_0 as an end. In this case, $\mathrm{csn}(G, v_0) = \mathrm{sn}(G, v_0)$ and the result holds.

Let us assume that $k = \mathrm{sn}(G, v_0) > 1$. We describe a connected strategy σ marking at most $k\sqrt{n}$ nodes per turn. Let $M^0 = \{v_0\}$ and let M^t be the set of vertices marked after $t \geq 1$ turns. Assume moreover that M^t induces a connected graph of G containing v_0. Finally, let v_t be the vertex occupied by the surfer after turn t. The set $\sigma(v_t, M^t)$ of nodes marked by the observer at step $t+1$ is defined as follows. If $|V(G) \setminus M^t| \leq k\sqrt{n}$, then let $\sigma(v_t, M^t) = V(G) \setminus M^t$. Otherwise, let $H \subseteq V(G) \setminus M^t$ such that $|H| = k\sqrt{n}$, $H \cup M^t$ induces a connected subgraph and $|H \cap N(v_t)|$ is maximum. Then, $\sigma(v_t, M^t) = H$, i.e., the strategy marks $k\sqrt{n}$ new nodes in a connected way and, moreover, mark as many unmarked nodes as possible among the neighbors of v_t. In particular, if $|N(v_t) \setminus M^t| \leq k\sqrt{n}$, then all neighbors of v_t are marked after turn $t+1$.

By definition, σ is connected and marks at most $k\sqrt{n}$ nodes per turn. We need to show σ is winning.

For purpose of contradiction, let us assume that the surfer wins against σ by following the path $P = (v_0, \ldots, v_t, v_{t+1})$. At its $t+1^{th}$ turn, the surfer moves from a marked vertex v_t to an unmarked vertex v_{t+1}.

Therefore, $n > tk\sqrt{n}$, otherwise the observer marking $k\sqrt{n}$ nodes at each turn would have already marked every vertex on the graph by the end of turn t. Moreover, by definition of sigma, $|N(v_t) \setminus M^t| > k\sqrt{n}$

Since, $\mathrm{sn}(G, v_0) = k$, let $\mathcal{S}$ be any k-winning (non necessarily connected) strategy for the observer. Assume that the observer follows $\mathcal{S}$ against the surfer following $P \setminus \{v_{t+1}\}$. Since, $\mathcal{S}$ is winning, all vertices of $N(v_t)$ must be marked after turn t, otherwise the surfer would win by moving to an unmarked neighbor of v_t. Therefore, since $\mathcal{S}$ can mark at most k vertices each turn, $|N(v_t)| \leq kt$.

Taking both inequalities, we have that $k\sqrt[2]{n} < |N(v_t)| \leq kt$. Hence, $\sqrt[2]{n} < t$. Therefore, $n > tk\sqrt[2]{n} > nk$, a contradiction. $\square$

2.2 Lower Bound

This section is devoted to proving the following theorem.

Theorem 2. *There exists a family of graphs G and $v_0 \in V(G)$ such that*

$$\mathrm{csn}(G, v_0) > \mathrm{sn}(G, v_0) + 1.$$

We use the following result proved in [1]. For any graph $G = (V, E)$ and any vertex $v_0 \in V$, a k-strategy for G with initial vertex v_0 is winning if and only if it is winning against a surfer that is constrained to follow induced paths on G. In other words, the walk of the surfer is contrained to be an induced path.

In the following theorem, by *adding a path $P = (v_1, \cdots, v_r)$ between two vertices u and v of G*, we mean that the induced path P is added as an induced subgraph of G and the edges $\{u, v_1\}$ and $\{v_r, v\}$ are added.

Let x, α, β and γ be four strictly positive integers satisfying the following:

$$(1)\ \max\{\beta, \frac{\beta}{2} + \gamma + 1\} < \alpha < \min\{\beta + \gamma + 1, 2\gamma + 2\} \qquad (2)\ \beta < 2\gamma + 2$$

$$(3)\ 3x \geq \alpha + \beta + 2\gamma + 12 \qquad (4)\ x > \frac{4}{5}(\alpha + \beta + \gamma) + 10 \qquad (5)\ 2\alpha \geq 73 + \beta + 2\gamma.$$

For instance, $x = 250$, $\alpha = 146$, $\beta = \gamma = 73$ satisfy all the above inequalities.

For proving the main theorem in this section we mainly rely in the family of graphs built in the following the procedure described below.

Let $\mathcal{G} = (V, E)$ be a graph with 10 isolated vertices $\{v_0, w_0, w_1, w_2, w'_0, w'_1, w'_2, s_0, s_1, s_2\}$. Then, for all $i \in \{0, 1, 2\}$ do the following:

1. $4x - 9$ vertices of degree one are added and made adjacent to s_i;
2. $3x-2$ vertices of degree one are added and made adjacent to w_i, respectively $3x - 2$ neighbors of degree one are added to w'_i;
3. two disjoint paths $A^i = (a^i_1, \cdots, a^i_\alpha)$ and $A'^i = (a'^i_1, \cdots, a'^i_\alpha)$ are added between v_0 and s_i;
4. a path $B^i = (b^i_1, \cdots, b^i_\beta)$ is added between v_0 and w_i, and a path $B'^i = (b'^i_1, \cdots, b'^i_\beta)$ is added between v_0 and w'_i;
5. for any $j \in \{i, i+1 \mod 3\}$ a path $C^{i,j} = (c^{i,j}_1, \cdots, c^{i,j}_\gamma)$ is added between s_j and w_i, and a path $C'^{i,j} = (c'^{i,j}_1, \cdots, c'^{i,j}_\gamma)$ is added between s_j and w'_i;
6. for any $1 \leq j \leq \alpha$, $3x-1$ vertices of degree one are added and made adjacent to a^i_j, respectively $3x - 1$ neighbors of degree one are added to a'^i_j;
7. for any $1 \leq j \leq \beta$, $3x-1$ vertices of degree one are added and made adjacent to b^i_j, respectively $3x - 1$ neighbors of degree one are added to b'^i_j;
8. for any $1 \leq j \leq \gamma$, $\ell \in \{i, i+1 \mod 3\}$, $3x - 1$ vertices of degree one are added and made adjacent to $c^{i,\ell}_j$, respectively $3x-1$ neighbors of degree one are added to $c'^{i,\ell}_j$.

The shape of $\mathcal{G}$ is depicted in Figure 1. $\mathcal{G}$ has $(30 + 18(\alpha + \beta) + 36\gamma)x - 29$ vertices. For any $i \in \{0, 1, 2\}$, the node s_i has $4x - 3$ neighbors, v_0 has 12 neighbors, and any other non-leaf node has degree $3x + 1$.

Claim. [9] If $\max\{\beta, \frac{\beta}{2} + \gamma + 1\} < \alpha < \min\{\beta + \gamma + 1, 2\gamma + 2\}$ and $\beta < 2\gamma + 2$, the unique (up to symmetries) minimum Steiner-tree for $S = N[v_0] \cup \{s_0, s_1, s_2\}$ in $\mathcal{G}$ has $15 + \alpha + \beta + 2\gamma$ vertices and consists of the vertices of the paths $A^0, B^1, C^{1,1}, C^{1,2}$ and the vertices in $S \cup \{w_1\}$.

In Fig. 1, the scheme of a minimum Steiner-tree for $S = N[v_0] \cup \{s_0, s_1, s_2\}$ is depicted with dashed lines.

For any $i \in \{0, 1, 2\}$, let $\mathcal{A}_i = N[v_0] \cup N[A^i] \cup N[s_i]$ (resp., $\mathcal{A}'_i = N[v_0] \cup N[A'^i] \cup N[s_i]$). Note that $|\mathcal{A}_i| = |\mathcal{A}'_i| = (3\alpha + 4)x + 9$ and that the $\mathcal{A}_i$ and $\mathcal{A}_j$, $i \neq j$, pairwise intersect only in $N[v_0]$.

For any $i \in \{0, 1, 2\}$, let $\mathcal{B}_i = N[v_0] \cup N[B^i] \cup N[w_i] \cup N[C^{i,i}] \cup N[C^{i,i+1 \mod 3}] \cup N[s_i] \cup N[s_{i+1 \mod 3}]$ and $\mathcal{B}'_i$ is defined similarly. $|\mathcal{B}_i| = |\mathcal{B}'_i| = (3\beta + 6\gamma + 11)x + 5$. Finally, for any $i \in \{0, 1, 2\}$ and $j \in \{i, i+1 \mod 3\}$, let $\mathcal{B}_{i,j} = N[v_0] \cup N[B^i] \cup N[w_i] \cup N[C^{i,j}] \cup N[s_j]$ and $\mathcal{B}'_{i,j} = N[v_0] \cup N[B'^i] \cup N[w'_i] \cup N[C'^{i,j}] \cup N[s_j]$.

Lemma 1. *For any $i \in \{0, 1, 2\}$ and $j \in \{i, i+1 \mod 3\}$, during its first step, any winning $(3x + y)$-strategy for $\mathcal{G}$ must mark at least*

- *$x + 8 - y(\alpha + 1)$ nodes in $\mathcal{A}_i$ (resp., in $\mathcal{A}'_i$), and*
- *$x + 8 - y(\beta + \gamma + 2)$ nodes in $\mathcal{B}_{i,j}$ (resp., in $\mathcal{B}'_{i,j}$), and*
- *$2x + 4 - y(\beta + 2\gamma + 3)$ nodes in $\mathcal{B}_i$ (resp., in $\mathcal{B}'_i$).*

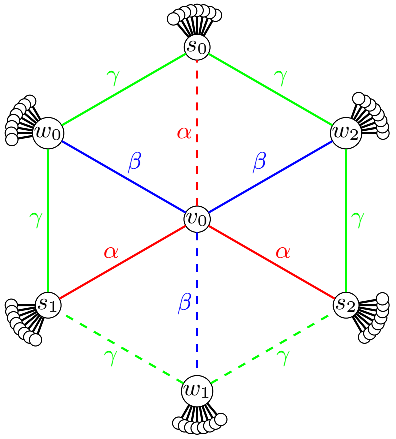

Fig. 1. Graph Family Scheme. Here we show only one "layer" of the graph.

The proof can be found in [9].

Lemma 2. $\mathrm{sn}(\mathcal{G}, v_0) \leq 3x$.

Proof. To show that $\mathrm{sn}(\mathcal{G}, v_0) \leq 3x$, consider the following strategy for the observer. For any $i \in \{0,1,2\}$, in the first step, it marks $x-4$ one-degree neighbots of s_i and the 12 neighbors of v_0. Then, at subsequent step, marks all unmarked neighbors of the current position of the surfer. It is easy to see (see details in [9]), by induction on the number of steps that, each time that the surfer arrives at a new node, this node is marked and has at most $3x$ unmarked neighbors. □

Lemma 3. $\mathrm{csn}(\mathcal{G}, v_0) > 3x + 1$.

Proof. For purpose of contradiction, let us assume that there is a winning connected $3x+1$-strategy. Let F be the set of vertices marked by this strategy during the first step. Clearly, $N(v_0) \subseteq F$ and $|F| \leq 3x+1$.

For any $0 \leq i \leq 2$, let $f_i = |F \cap N[s_i]|$ and let $f_{min} = \min_i f_i$. Without loss of generality, $f_{min} = f_0$. We first show that $f_{min} > 3$.

By Lemma 1, for any $i \in \{0,1,2\}$, $|F \cap (\mathcal{A}_i \setminus N[v_0])| \geq x - 5 - \alpha$ and, for any $i \in \{0,2\}$, $|F \cap (\mathcal{B}_{i,0} \setminus N[v_0])| \geq x - 6 - (\beta + \gamma)$ and $|F \cap (\mathcal{B}'_{i,0} \setminus N[v_0])| \geq x - 6 - (\beta + \gamma)$. Therefore,

$$\begin{aligned} 3x+1 &\geq |F \cap (\mathcal{A}_0 \cup \mathcal{A}'_0 \cup \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{B}_{0,0} \cup \mathcal{B}_{2,0} \cup \mathcal{B}'_{0,0} \cup \mathcal{B}'_{2,0})| \\ &\geq 12 + 4(x - 5 - \alpha) + 4(x - 6 - (\beta+\gamma)) - 5|F \cap N[s_0]| \\ &\geq 8x - 4(\alpha+\beta+\gamma) - 32 - 5f_{min} \end{aligned}$$

Hence, $5f_{min} \geq 5x - 4(\alpha+\beta+\gamma) - 33$, and $f_{min} \geq x - \frac{4}{5}(\alpha+\beta+\gamma) - 7 > 3$ by the above inequality.

Therefore, by definition of f_{min}, $|F \cap N[s_i]| \geq 4$ for any $i \in \{0,1,2\}$. By connectivity of the strategy, $s_i \in F \cap N[s_i]$ for any $i \in \{0,1,2\}$. Hence, F must contain a subset of vertices inducing a subtree spanning $S = N[v_0] \cup \{s_0, s_1, s_2\}$. Let T be an inclusion-minimal subset of F that induces a subtree spanning S. By Claim 2.2, $|T| \geq \alpha + \beta + 2\gamma + 15$. Let $T' = T \setminus (N[v_0] \cup \bigcup_{0 \leq i \leq 2} N[s_i])$. Then, $|T'| \geq \alpha + \beta + 2\gamma - 4$. Moreover, because of the symmetries, we may assume w.l.o.g., that $T' \subseteq \bigcup_{0 \leq i \leq 2} (\mathcal{A}_i \cup \mathcal{B}_i)$.

By Lemma 1 and because $N(v_0) \subseteq F$, for any $0 \leq i \leq 2$, $|F \cap (\mathcal{A}'_i \cup \mathcal{B}'_{i+1 \mod 3})| \geq x + 8 - (\alpha+1) + 2x + 4 - (\beta + 2\gamma + 3) - 12 = 3x - (\alpha + \beta + 2\gamma) - 4$. Hence, $|T'| + |F \cap (\mathcal{A}'_i \cup \mathcal{B}'_{i+1 \mod 3})| \geq 3x - 8$. Let $W_i = F \setminus (\mathcal{A}'_i \cup \mathcal{B}'_{i+1 \mod 3} \cup T')$. Since $|F| \leq 3x + 1$, it follows that $|W_i| \leq 9$.

Let $f_{max} = \max_i f_i$ and assume w.l.o.g. that $f_{max} = f_2$. Since $\sum_{0 \leq i \leq 2} f_i \leq |F \setminus T'|$, we get that $f_0 + f_1 \leq \lfloor \frac{2}{3}(5 + 3x - (\alpha + \beta + 2\gamma) \rfloor$.

To conclude, $|F \cap \mathcal{B}'_0| = |N(v_0)| + f_0 + f_1 + |W_0| \leq 21 + \lfloor \frac{2}{3}(5 + 3x - (\alpha + \beta + 2\gamma)) \rfloor$. On the other hand, Lemma 1 implies that $|F \cap \mathcal{B}'_0| \geq 2x + 1 - (\beta + 2\gamma)$. Therefore, $22 + \frac{2}{3}(5 + 3x - (\alpha + \beta + 2\gamma) > 2x + 1 - (\beta + 2\gamma)$ and it follows $73 > 2\alpha - \beta - 2\gamma$. This contradicts the inequalities. □

Lemmas 2 and 3 are sufficient to prove Theorem 2. More precisely, it shows that there exists a family of graphs G and $v_0 \in V(G)$ such that $\text{csn}(G, v_0) \geq \text{sn}(G, v_0) + 2$. However, the family of graphs we described does not allow to increase further the cost of connectivity. Indeed, $\text{csn}(\mathcal{G}, v_0) \leq 3x + 2$ [9].

To conclude this section, we answer negatively a question in [1]. We show that there is a graph $\mathcal{G}$ such that $\text{sn}(\mathcal{G}, s) = k$ and $\max_{S \subseteq V(\mathcal{G})} \lceil \{ \frac{|N[S]| - 1}{|S|} \} \rceil < k$ [9].

3 Online Surveillance Number

In this section, we study the online variant of the surveillance game motivated by the web-page prefetching problem where the observer (the web-browser) discovers new nodes through hyperlinks in already marked nodes. In this variant, the observer does not know *a priori* the graph in which it is playing. That is, initially, the observer only knows v_0 and the identifiers of its neighbors. Then, when a new node is marked, the observer discovers all its neighbors that are not yet marked. Note that the degree of a node is not known before it is marked.

Another property of an online strategy that must be defined concerns the moment when the observer discovers the unmarked neighbors of a node that it has decided to mark. There are two natural models. Assume that the set M of nodes have been marked and this is the turn of the observer, and let $N(M)$ be the set of nodes with a neighbor in M. Either it first chooses the k nodes that will be marked among the set $N(M) \setminus M$ of the unmarked neighbors of the nodes that were already marked and then the observer marks each of these k nodes and discover their unknown neighbors simultaneously. Or, the observer first chooses one node x in $N(M) \setminus M$, marks it and discovers its unmarked neighbors, then it chooses a new node to be marked in $N(M \cup \{x\}) \setminus (M \cup \{x\})$ and so on until the observer finishes its turn after marking k nodes. We choose to consider the second model because it is less restricted, i.e., the observer has more power, and, even in this case, our result is pessimistic since we show that the basic strategy is the best one with respect to the competitive ratio.

Formal definition of online strategy. Now we are ready to formally define an online strategy. Let $k \geq 1$, let $G = (V, E)$ be a graph, $v_0 \in V$, and let $\mathcal{G}$ be the set of subgraphs of G.

Given $M \subseteq V$ be a subset of nodes inducing a connected subgraph containing v_0 in G. Let $G_M \in \mathcal{G}$ be the subgraph of G known by the observer when M is the set of marked nodes. That is, $G_M = (M \cup N(M), E_M)$ where $E_M = \{(u,v) \in E \mid u \in M\}$. For any $u, v \in N(M) \setminus M$, let us set $u \sim_M v$ if and only if $N(u) \cap M = N(v) \cap M$. Let χ_M be the set of equivalent classes, called *modules*, of $N(M) \setminus M$ with respect to $\sim_M$. The intuition is that two nodes in the same module of χ_M are known by the observer but cannot be distinguished. For instance, $\chi_{\{v_0\}} = \{N(v_0)\}$.

A *k-online strategy for the observer starting from v_0* is a function $\sigma : \mathcal{G} \times V \times 2^V \times \{1, \cdots, k\} \to 2^V$ such that, for any subset $M \subseteq V$ of nodes inducing a connected subgraph containing v_0 in G, for any $v \in M$, and for any $1 \leq i \leq k$, then $\sigma(G_M, v, M, i) \in \chi_M$. This means that, if M is the set of nodes already marked and thus the observer only knows the subgraph G_M, if v is the position of the surfer and it remains $k - i + 1$ nodes to be marked by the observer before the surfer moves, then the observer will mark one node in $\sigma(G_M, v, M, i)$.

More precisely, we say that the observer *follows* the k-online strategy σ if the game proceeds as follows. Let $M = M^0$ be the set of marked nodes just after the surfer has moved to $v \in M$. Initially, $M^0 = \{v_0\}$ and $v = v_0$. Then, the strategy proceeds sequentially in k steps for $i = 1$ to k. First, the observer marks an arbitrary node $x_1 \in \sigma(G_{M^0}, v, M^0, 1)$. Let $M^1 = M^0 \cup \{x_1\}$. Sequentially, after having marked $1 < i < k$ nodes at this turn, the observer marks one arbitrary node $x_{i+1} \in \sigma(G_{M^i}, v, M^i, i+1)$ and $M^{i+1} = M^i \cup \{x_{i+1}\}$. When the observer has marked k nodes, that is after choosing $x_k \in \sigma(G_{M^{k-1}}, v, M^{k-1}, k)$, it is the turn of the surfer, when it may move to a node adjacent to its current position and then a new turn for the observer starts. Note that because we are interested in the worst case for the observer, each marked node $x_i \in \sigma(G_{M^{i-1}}, v, M^{i-1}, i)$ is chosen by an adversary.

The *online surveillance number* of a graph G with initial node v_0, denoted by $\mathrm{on}(G, v_0)$, is the smallest k such that there exists a winning k-online strategy starting from v_0. In other words, there is a winning k-online strategy σ starting from v_0 such that an observer following σ wins whatever be the trajectory of the surfer and the choices done by the adversary at each step. Note that, since we consider the worst scenario for the observer, we may assume that the surfer has full knowledge of G.

Theorem 3. *There exists an infinite family of rooted trees such that, for any T with root $v_0 \in V(T)$ in this family, $\mathrm{sn}(T, v_0) = 2$ and $\mathrm{on}(T, v_0) = \Omega(\Delta)$ where Δ is the maximum degree of T.*

Proof. We first define the family $(T_k)_{k \geq 1}$ of rooted trees as follows.

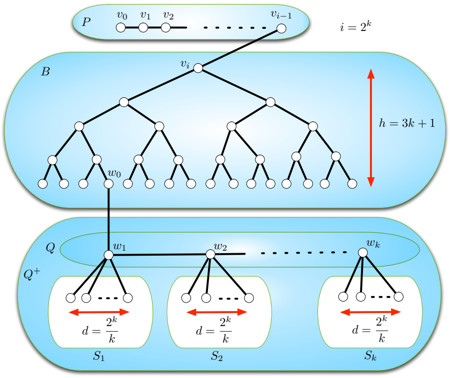

Fig. 2. Tree T_k described in the proof of Theorem 3.

Let $k \geq 4$ be a power of two and let $i = 2^k$ and $d = \frac{2^k}{k}$.

Let us consider a path $P = (v_0, v_1, \ldots, v_{i-1})$ with i nodes Let B be a complete binary tree of height $h = 3k+1$ and rooted at some vertex v_i, i.e., B has $2^{h+1}-1$ vertices. Let w_0 be any leaf of B. Finally, let $Q = (w_1, \ldots, w_k)$ be a path on k nodes. Note that, P, B and Q depend on k.

The tree T_k is obtained from P, B and Q by adding an edge between v_{i-1} and v_i, an edge between w_0 and w_1. Finally, for any $1 \leq j \leq k$, let us add an independent set, S_j, with d vertices and an edge between each vertex of S_j and w_j (i.e., each node in S_j is a leaf). T_k is then rooted in v_0.

Let Q^+ denote the union of vertices of Q and $\bigcup_{j=1}^{k} S_j$. The maximum degree Δ of T_k is reached by any node w_j, $1 \leq j < k$, and $\Delta = d + 2 = \frac{2^k}{k} + 2$.

Clearly, $\mathrm{sn}(T_k, v_0) > 1$. We show that $\mathrm{sn}(T_k, v_0) = 2$.

Consider the following (offline) strategy for the observer. At each turn $j \leq i$, the surfer marks the vertex v_j and one unmarked vertex of Q^+ that is closest to the surfer. For each turn $j > i$ and while the surfer does not occupy a node in $Q^+ \cup \{w_0\}$, the observer marks the neighbors of the current position of the surfer if they are not already marked. Finally, if the surfer occupies a node in $Q^+ \cup \{w_0\}$, the observer marks two unmarked nodes of Q^+ that are closest to the surfer. It is easy to see, by induction on the number of steps that, each time that the surfer arrives at a new node, this node is marked and has at most 2 unmarked neighbors. Hence, $\mathrm{sn}(T_k, v_0) = 2$.

Now it remains to show that $\mathrm{on}(T_k, v_0) = \Omega(\Delta)$. Let γ be any online strategy for T_k and marking at most $\frac{d}{4} = \frac{2^{k-2}}{k}$ nodes per turn. We show that γ fails.

For this purpose, we model the fact that the observer does not know the graph by "building" the tree during the game. More precisely, each time the observer marks a node v, then the adversary may add new nodes adjacent to v or decide that v is a leaf. Of course, the adversary must satisfy the constraint that eventually the graph is T_k. Initially, the observer only knows v_0 that has one neighbor v_1. Now, for any $1 \leq j < i$, when the observer marks the node v_j of

P, then the adversary "adds" a new node v_{j+1} adjacent to v_j, i.e., the observer discovers its single unmarked neighbor v_{j+1}. Now, let v be any node of B. Recall that h is the height of B. When the observer marks v, there are three cases to be considered: if v is at distance at most $h-1$ from v_i, then the adversary adds two new nodes adjacent to v; if v is at distance h from v_i and not all nodes of B have been marked then the adversary decides that v is a leaf; finally, if all nodes of B have been marked (v is the last marked node of B, i.e., B is a complete binary tree of height h), the adversary decides that $v = w_0$ and add one new neighbor w_1 adjacent to it. Note that we can ensure that the last node of B to be marked is at distance h of v_i by connectivity of any online strategy.

Now, let consider the following execution of the game. During the first i steps, the surfer goes from v_0 to v_i. Just after the surfer arrives in v_i, the observer has marked at most $(di)/4$ nodes and all nodes of $P \cup \{v_i\}$ must be marked since otherwise the surfer would have won. Therefore, at most $i(d/4-1)+1 = 2^{2k-2}/k - 2^k + 1$ nodes of B are marked when it is the turn of the surfer at v_i. Since B has $2^{h+1}-1 = 2^{3k+2}-1$ nodes, at least one node of B is not marked.

From v_i, the surfer always goes toward w_0. Note that the observer may guess this strategy but it does not know where is w_0 while all nodes of B have not been marked.

Then let $0 \leq t \leq h$ and let $v'_t \in V(B)$ be the position of the surfer at step $i+t$ and B^t the subtree of B rooted at v'_t. Note that, at step i, $v'_0 = v_i$ and $B^0 = B$. Let B^t_l and B^t_r be the subtrees of B rooted at the children of v'_t. W.l.o.g., let us assume that the number of marked nodes in B^t_l is at most the number of marked nodes in B^t_r, when it is the turn of the surfer standing at v'_t. Then, the surfer moves to the root of B^t_l. That is, v'_{t+1} is the child of v_t whose subtree contains the minimum number of marked nodes.

Let m_t be the number of marks in the subtree of B rooted at v'_t when it is the turn of the surfer at v'_t. Since, at beginning of step i there are at most $2^{2k-2}/k - 2^k + 1$ nodes of B that are marked and $k \geq 4$, $m_0 \leq 2^{2k-2}/k - 2^k + 1 \leq 2^{2k-2}/k$. Note that, for any $t > 0$, $m_t \leq (m_{t-1} - 1 + \frac{d}{4})/2 \leq (m_{t-1} + \frac{d}{4})/2$. Simply expanding this expression we get that, for any $t > 0$,

$$m_t \leq \frac{m_0}{2^t} + \frac{2^k}{k} \sum_{j=3}^{t+2} 2^{-j} \leq \frac{2^{2k-(t+2)}}{k} + \frac{2^k}{k} \sum_{j=3}^{t+2} 2^{-j}.$$

Therefore, for any $t \geq 2k$:

$$m_t \leq \frac{1}{k} + \frac{2^k}{k} \sum_{j=3}^{t+2} 2^{-j} \leq \frac{2^k+1}{k}.$$

In particular, at step $i+2k$ (when it is the turn of the surfer), the surfer is at v'_{2k} which is at distance $k+1$ from w_0. Hence, $|B^{2k}| \geq 2^{k+1}-1$ and at most $\frac{2^k+1}{k} < 2^{k+1}-1$ of its nodes are marked. Hence, w_0 neither no nodes in Q^+ are marked.

From this step, the surfer directly goes to w_k unless she meets an unmarked node, in which case, she goes to it and wins. When the surfer is at w_k and it is her turn, the observer may have marked at most $(2k+2)\frac{d}{4} \leq \frac{kd}{2} + \frac{d}{2} \leq 2^{k-1} + \frac{2^{k-1}}{k}$ nodes in Q^+. Since $|Q^+| = (d+1)k = 2^k + k$ and $k \geq 4$, at least one neighbor of w_k is not marked yet and the surfer wins. $\square$

Theorem 3 implies that, for any online strategy $\mathcal{S}$, $\rho(\mathcal{S}) = \Omega(\Delta)$. Recall that the basic strategy $\mathcal{B}$, that marks all unmarked neighbors of the surfer at each step, is an online strategy. $\mathcal{B}$ has trivially competitive ratio $\rho(\mathcal{B}) = O(\Delta)$. Hence, no online winning strategy has better competitive ratio than the basic strategy up to a constant factor. In other words:

Corollary 1. *The best competitive ratio of online winning strategies is $\Theta(\Delta)$, with Δ the maximum degree.*

As mentioned in the introduction, any online strategy is connected and therefore, for any graph G and $v_0 \in V(G)$, $\mathrm{csn}(G, v_0) \leq \mathrm{on}(G, v_0)$. Moreover, we recall that, for any tree T and for any $v_0 \in V(T)$, $\mathrm{csn}(T, v_0) = \mathrm{sn}(T, v_0)$ [1]. Hence, there might be an arbitrary gap between $\mathrm{csn}(G, v_0)$ and $\mathrm{on}(G, v_0)$.

4 Conclusion

Despite our results, the main question remains open. Can the difference or the ratio between the connected surveillance number of a graph and its surveillance number be bounded by some constant?

References

1. Fomin, F.V., Giroire, F., Jean-Marie, A., Mazauric, D., Nisse, N.: To satisfy impatient web surfers is hard. In: FUN. (2012) 166–176
2. Aumann, Y., Etzioni, O., Feldman, R., Perkowitz, M.: Predicting event sequences: Data mining for prefetching web-pages (1998)
3. Albrecht, D., Zukerman, I., Nicholson, A.: Pre-sending documents on the www: a comparative study. In: 16th Int. Joint Conf. on Artificial Intelligence. (1999) 1274–1279
4. Mogul, J.C.: Hinted caching in the web. In: 7th workshop on ACM SIGOPS European workshop: Systems support for worldwide applications. (1996) 103–108
5. Wang, Z., Lin, F.X., Zhong, L., Chishtie, M.: How far can client-only solutions go for mobile browser speed? In: 21st Int. Conf. on World Wide Web. (2012) 31–40
6. Fan, L., Cao, P., Lin, W., Jacobson, Q.: Web prefetching between low-bandwidth clients and proxies: potential and performance. In: ACM SIGMETRICS Int. Conf. on Measurement and Modeling of Computer Systems. (1999) 178–187
7. Padmanabhan, V.N., Mogul, J.C.: Using predictive prefetching to improve world wide web latency. SIGCOMM Comput. Commun. Rev. **26**(3) (1996) 22–36
8. Kroeger, T.M., Long, D.D.E., Mogul, J.C.: Exploring the bounds of web latency reduction from caching and prefetching. In: USENIX Symposium on Internet Technologies and Systems. (1997) 2–2
9. Giroire, F., Mazauric, D., Nisse, N., Pérennes, S., Pardo Soares, R.: Connected Surveillance Game. Rapport de recherche RR-8297, INRIA (2013)