Induced paths in twin-free graphs

David Auger
Télécom ParisTech, 46 rue Barrault,
75634 Paris Cedex 13, France
auger@enst.fr

Submitted: Feb 19, 2008; Accepted: May 27, 2008; Published: Jun 6, 2008
Mathematics Subject Classification: 05C12

Abstract

Let $G = (V, E)$ be a simple, undirected graph. Given an integer $r \geq 1$, we say that G is *r-twin-free* (or *r-identifiable*) if the balls $B(v, r)$ for $v \in V$ are all different, where $B(v, r)$ denotes the set of all vertices which can be linked to v by a path with at most r edges. These graphs are precisely the ones which admit r-identifying codes. We show that if a graph G is r-twin-free, then it contains a path on $2r + 1$ vertices as an induced sugbraph, i.e. a chordless path.

keywords: graph theory; identifying codes; twin-free graphs; induced path; radius

1 Notation and definitions

Let $G = (V, E)$ be a simple, undirected graph. We will denote an edge $\{x, y\} \in E$ simply by xy. A *path* in G is a sequence $P = v_0 v_1 \cdots v_k$ of vertices such that for all $0 \leq i \leq k-1$ we have $v_i v_{i+1} \in E$; if $v_0 = x$ and $v_k = y$, we say that P is a path *between* x and y.

The *length* of a path $P = v_0 v_1 \cdots v_k$ is the number of edges between consecutive vertices, i.e. k. If $x, y \in V$, we define the distance $d(x, y)$ to be the minimum length of a path between x and y. Then a *shortest path* between x and y is a path between x and y of length precisely $d(x, y)$. If $r \geq 0$, $B(x, r)$ will denote the *ball* of centre x and radius r, which is the set of all vertices v of G such that $d(x, v) \leq r$.

If $P = v_0 \cdots v_k$ is a path in G, a *chord* in P is any edge $v_i v_j \in E$ with $|i - j| \neq 1$. A path is *chordless* if it has no chord; in this case there is an edge between two vertices of the path v_i and v_j if and only if i and j are consecutive, i.e. $|i - j| = 1$. It is straightforward to see that any shortest path is chordless.

If $x \in V$, we define the *eccentricity* of x by

$$\mathrm{ecc}(x) = \max_{v \in V} d(x, v).$$

The *diameter* of G is the maximum eccentricity of a vertex in G, whereas the *radius* $\mathrm{rad}(G)$ of G is the minimum eccentricity of a vertex in G. A vertex x such that $\mathrm{ecc}(x) = \mathrm{rad}(G)$ is a *centre* of G. So G has radius $t \geq 1$ and x is a centre of G if and only if $B(x, t) = V$ whereas $B(v, t-1) \neq V$ for all $v \in V$.

If $W \subset V$, the *sugbraph* of G *induced* by W is the graph whose set of vertices is W and whose edges are all the edges $xy \in E$ such that x and y are in W. We denote this graph by $G[W]$; if $W = V \setminus \{v\}$, we simply write $G[V - v]$. An *induced path* in G is a subset P of V such that $G[P]$ is a path; equivalently, the vertices in P define a chordless path in G. All these terminology and notation being standard, we refer to [3] for further explanation.

Two distinct vertices x and y are called *r-twins* if $B(x, r) = B(y, r)$. If there are no r-twins in G, we say that G is *r-twin-free*.

2 Motivations and main results

The notion of identifying code in a graph was introduced by Karpovsky, Chakrabarty and Levitin in [5]. For $r \geq 1$, an r-identifying code in $G = (V, E)$ is a subset $\mathcal{C}$ of V such that the sets

$$I_{\mathcal{C}}(v) = B(v, r) \cap \mathcal{C} \text{ for } v \in V$$

are all distinct and non-empty. The original motivation for identifying codes was the fault diagnosis in multiprocessor systems; we refer to [1], [5] or [7] for further explanation and applications. The interested reader can also find a nearly exhaustive bibliography in [6].

Given a graph $G = (V, E)$, it is easily seen that there exists an r-identifying code in G if and only if V itself is an r-identifying code, which precisely means that G is r-twin-free. Different structural properties which are worth investigating arise when considering a connected r-twin-free graph with $r \geq 1$. For instance, it has been proved in [2] that an r-twin-free graph always contains a path, not necessarily induced, on $2r + 1$ vertices. In the same article, the authors conjectured that we can always find such a path as an induced subgraph of G. We prove this conjecture as a corollary from Theorem 1.

Let us denote by $p(G)$ the maximum number of vertices of an induced path in G. We prove the following theorem and corollary, which we formulate for connected graphs without loss of generality.

Theorem 1. *Let $G = (V, E)$ be a connected graph with at least two vertices, and with a centre $c \in V$ such that no neighbour of c is a centre. Then*

$$p(G) \geq 2\ rad(G) + 1.$$

This implies:

Corollary 2. *Let G be a connected graph with at least two vertices, and $r \geq 1$. If G is r-twin-free then*

$$p(G) \geq 2r + 1.$$

3 Proof of the theorem

A different proof for Corollary 2 can be found in [1]. The one we present here is much shorter and is based on the article by Erdős, Saks and Sós [4] where the following theorem can be found. The authors give credit to Fan Chung for the proof.

Theorem 3. *(Chung) For every connected graph $G = (V, E)$ we have*

$$p(G) \geq 2\ rad(G) - 1.$$

We require the following lemma, inspired by [4], in order to prove Theorem 1.

Lemma 4. *Let $t \geq 2$ and $G = (V, E)$ be a graph such that there are in G two vertices v_0 and v_t with $d(v_0, v_t) = t$, a shortest path $v_0v_1v_2 \cdots v_t$ between v_0 and v_t, and a vertex w such that $d(v_0, w) \leq t-1$ and $d(v_2, w) \geq t$ (see fig. 1). Then there exists an induced path on $2t-1$ vertices in G.*

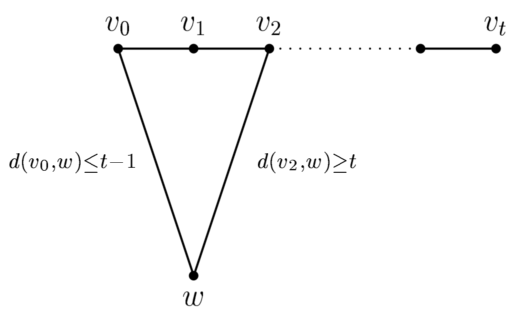

Figure 1: The path $v_0 \cdots v_t$ and w in Lemma 4.

Proof. In the case $t = 2$, the shortest path $v_0v_1v_2$ itself is an induced path on $2t - 1 = 3$ vertices; so we suppose now that $t \geq 3$. First observe that since $d(v_2, w) \geq t$ we have $w \neq v_i$ for all $i \in \{0, 1, \cdots, t\}$. Consider a shortest path P between v_0 and w, and let $u \in P$, distinct from v_0. Let $i \geq 2$; we show that $d(u, v_i) \geq 2$. First we have

$$d(v_0, v_i) = i \leq d(v_0, u) + d(u, v_i)$$

and second

$$t \leq d(v_2, w) \leq d(v_2, v_i) + d(v_i, u) + d(u, w)$$

with $d(v_2, v_i) = i - 2$ because $i \geq 2$. Summing these two inequalities we get

$$t + i \leq d(v_0, u) + d(u, w) + 2d(v_i, u) + i - 2$$

and since

$$d(v_0, u) + d(u, w) = d(v_0, w)$$

we deduce

$$t + 2 \leq d(v_0, w) + 2d(u, v_i).$$

But we have $d(v_0, w) \leq t-1$ and so

$$d(u, v_i) \geq \frac{3}{2}.$$

Let us note that since $d(v_2, w) \geq t$, we have $d(v_0, w) \geq t-2$ and so P consists of v_0 and at least $t-2 \geq 1$ other vertices, i.e. at least $t-1$ vertices. We proved that u satisfies $d(u, v_i) \geq 2$ for $i \geq 2$, so u is distinct from all the v_i's and furthermore can be adjacent only to v_1 or v_0 (see fig. 2).

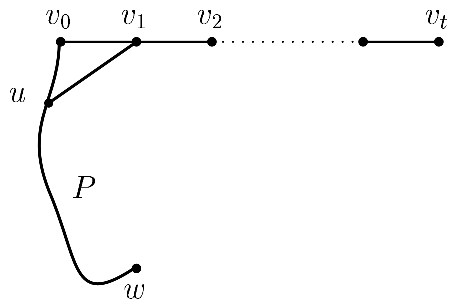

Figure 2: The vertex u can only be adjacent to v_1 or v_0 in Lemma 4.

Now consider two cases:

- if no vertex $u \in P \setminus \{v_0\}$ is adjacent to v_1, then P extended by $v_1 \cdots v_t$ is an induced path of G on at least $(t-1)+t = 2t-1$ vertices;
- if there is a vertex $u \in P \setminus \{v_0\}$ adjacent to v_1, then

$$t \leq d(v_2, w) \leq d(v_2, v_1) + d(v_1, u) + d(u, w)$$

and so $d(u, w) \geq t-2$. Since we have $d(v_0, u) + d(u, w) = d(v_0, w) \leq t-1$, it follows that we must have $d(v_0, u) = 1$ and $d(u, w) = t-2$. The path $w \cdots u v_1 \cdots v_t$ is then an induced path of G on $2t-1$ vertices.

□

For sake of completeness, we rephrase the end of the proof of Theorem 3 in [4]. Consider a connected graph G of radius $t \geq 1$; if $t = 1$, then the result is trivial. Suppose now that $t \geq 2$; we show that the vertices $v_0, v_1, \cdots v_t$ and w as in Lemma 4 exist. To see this, consider the collection of connected induced subgraphs H of G whose radius is at least t, and choose one with the smallest possible number of vertices. Let V_H be the vertex-set of H.

There exists in H a vertex v_t which is not a cutvertex; by minimality of H, the connected induced subgraph $H[V_H - v_t]$ of H must have radius at most $t-1$. If we consider a centre v_0 of $H[V_H - v_t]$, we must have $d(v_0, w) \leq t-1$ for all the vertices $w \neq v_t$ in H; but since H has radius at least t we also have $d(v_0, v_t) = t$. Let $v_0 v_1 v_2 \cdots v_t$ be a shortest path between v_0 and v_t. Since H has radius t, there exists a vertex w such that $d(v_2, w) \geq t$, and we have $d(v_0, w) \leq t-1$ because w cannot be v_t. So we can choose this w and apply Lemma 4.

Proof of Theorem 1. Let $G = (V, E)$ be a graph of radius $t \geq 1$ with a centre $c \in V$ such that no neighbour of c is a centre. We will apply Lemma 4 with $t+1$ instead of t; to do this, we have to find vertices $v_0, v_1, \cdots v_{t+1}$ and w; so let us denote the center c by v_1. We define $N(v_1)$ to be the set of neighbours of v_1. We can choose a vertex v_0 in $N(v_1)$ such that $B(v_0, t)$ is not strictly contained in another $B(x, t)$ for $x \in N(v_1)$: take for instance $v_0 \in N(v_1)$ such that $B(v_0, t)$ is of maximal cardinality. Since v_0 is not a centre, there exists a vertex $v_{t+1} \in V$ such that $d(v_0, v_{t+1}) = t+1$. Then we must have $d(v_1, v_{t+1}) \geq t$, and so $d(v_1, v_{t+1}) = t$ because v_1 is a centre. Consider a shortest path $v_1 v_2 \cdots v_{t+1}$ between v_1 and v_{t+1}; then $v_0 v_1 v_2 \cdots v_{t+1}$ is a shortest path between v_0 and v_{t+1}. Now, if we show that there exists a vertex w such that $d(v_2, w) \geq t+1$ and $d(v_0, w) \leq t$, we can apply Lemma 4. So, assume that such a vertex w does not exist: this means that all the vertices w with $d(v_0, w) \leq t$ must satisfy $d(v_2, w) \leq t$, and so $B(v_0, t) \subset B(v_2, t)$. By maximality of $B(v_0, t)$, we must then have $B(v_0, t) = B(v_2, t)$; but this is impossible, since we have $v_{t+1} \in B(v_2, t) \setminus B(v_0, t)$. This contradiction shows that we can apply Lemma 4, and so there exists in G an induced path on $2(t+1) - 1 = 2t + 1$ vertices; thus we have

$$p(G) \geq 2\text{rad}(G) + 1.$$

□

Proof of Corollary 2. Let G be a graph, x a center of G and y a neighbour of x. Then by definition $B(x, \text{rad}(G)) = V$, and for all $z \in V$ we have

$$d(y, z) \leq d(z, x) + d(x, y) \leq \text{rad}(G) + 1.$$

So

$$B(x, r) = B(y, r) = V$$

for all $r \geq \text{rad}(G) + 1$. Suppose now that G is r-twin-free; then we must have $\text{rad}(G) \geq r$.

Now, either $\text{rad}(G) \geq r+1$ and we can apply Theorem 3, or $\text{rad}(G) = r$. But in the latter case, centers are r-twins so there can only be one in G; in particular we can apply Theorem 1 and so

$$p(G) \geq 2\text{rad}(G) + 1 = 2r + 1.$$

□

4 Conclusion and perspectives

For $n \geq 1$, we denote by P_n the path on n vertices, i.e. the graph consisting of n vertices $v_0, v_1, \cdots, v_{n-1}$ and the $n-1$ edges $v_i v_{i+1}$ for $0 \leq i \leq n-1$. As the path P_{2r+1} on $2r+1$ vertices is itself r-twin-free, the previous results show that P_{2r+1} is the only minimal r-twin-free graph for the induced subgraph relationship. Indeed, we have:

An r-twin-free graph contains a path P_{2r+1} as an induced sugbraph, and P_{2r+1} is r-twin-free.

One could wonder how these results could be extended to different cases. For instance, we have:

An r-twin-free and 2-connected graph G contains a cycle with at least $2r+2$ vertices as a subgraph; and the cycle C_k on k vertices is r-twin-free if and only if $k \geq 2r+2$ (and is, of course, 2-connected).

Let us recall that a graph G is 2-connected if and only if for every pair (x, y) of distinct vertices, there exist at least two paths P_1 and P_2 between x and y in G, such that there are no common vertices to P_1 and P_2 except x and y (see [3], pp. 55-57 for more details). Since an r-twin-free graph has a diameter at least $r+1$, the result above easily follows. This shows that the cycles C_k with $k \geq 2r+2$ are the minimal graphs for the subgraph relationship in the class of 2-connected, r-twin-free graphs. But in this case, the result cannot be extended to the induced subgraph relationship. Indeed, for $r \geq 1$ consider the Cartesian product of a path P_{2r+1} with K_2 (see fig. 3). One can check that this graph is 2-connected, r-twin-free and does not contain a cycle with more than $2r+2$ vertices as an induced subgraph. For $r = 1$, see the counterexample on fig. 4

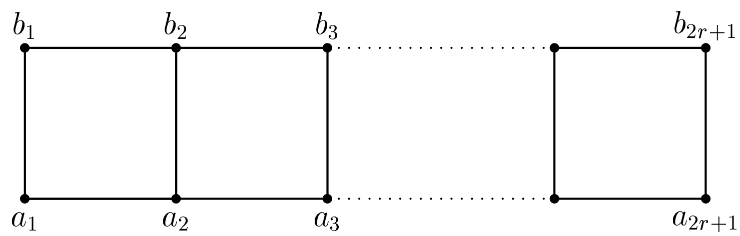

Figure 3: A 2-connected, r-twin-free graph which does not contain a cycle C_k with $k \geq 2r+2$ as an induced subgraph ($r \geq 2$).

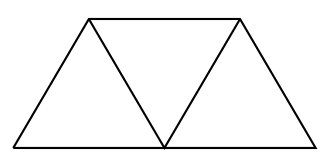

Figure 4: A 2-connected, 1-twin-free graph which does not contain a cycle C_k with $k \geq 4$ as an induced subgraph.

As a conclusion, we leave open the same problem in the class of k-connected graphs with $k \geq 3$:

What are the minimal elements of the class of 3-connected, r-twin-free graphs, for the subgraph relationship, or the induced subgraph relationship?

A first step would be to determine the smallest cardinality for a k-connected r-twin-free graph.

References

[1] D. Auger, *Problèmes d'identification métrique dans les graphes*, ENST Technical Report, ISSN 0751-1345 ENST D013, 2007.

[2] I. Charon, I. Honkala, O. Hudry, A. Lobstein, *Structural Properties of Twin-Free Graphs*, Electronic Journal of Combinatorics, Vol. 14(1), R16, 2007.

[3] R. Diestel, *Graph Theory*, Springer-Verlag, third edition, 2005.

[4] P. Erdős, M. Saks, V. T. Sós, *Maximum Induced Trees in Graphs*, Journal of Combinatorial Theory, Series B, vol. 41, pp. 61-79, 1986.

[5] M. G. Karpovsky, K. Chakrabarty, L. B. Levitin, *On a new class of codes for identifying vertices in graphs*, IEEE Transactions on Information Theory, vol 44, pp. 599-611, 1998.

[6] A. Lobstein, *Bibliography on identifying and locating-dominating codes in graphs*, http://www.infres.enst.fr/∼lobstein/debutBIBidetlocdom.pdf

[7] J. Moncel, *Codes identifiants dans les graphes*, Thèse de Doctorat, Université Joseph Fourier - Grenoble I, France, 2005.