



HAL
open science

Is it possible to find the maximum clique in general graphs?

José Ignacio Alvarez-Hamelin

► **To cite this version:**

José Ignacio Alvarez-Hamelin. Is it possible to find the maximum clique in general graphs?. 2011. hal-00625917v3

HAL Id: hal-00625917

<https://hal.science/hal-00625917v3>

Preprint submitted on 17 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IS IT POSSIBLE TO FIND THE MAXIMUM CLIQUE IN GENERAL GRAPHS?

JOSÉ IGNACIO ALVAREZ-HAMELIN

ABSTRACT. Finding the maximum clique is a known NP-Complete problem and it is also hard to approximate. This work proposes two efficient algorithms to obtain it. Nevertheless, the first one is able to find the maximum for some special cases, while the second one has its execution time bounded by the number of cliques that each vertex belongs to.

1. INTRODUCTION

Finding cliques in graphs is a well known problem, mainly the maximum clique was found to be a NP-Complete problem [Karp, 1972]. Indeed, from any vertex, to discover the maximum clique to which it belongs, we should take all combinations of k neighbors and verify whenever they are mutually adjacent, which yields an exponential time as a function of the vertex degree. Moreover, [Johnson, 1973] shows that there exists no sublinear approximation algorithm. This problem was largely treated and an extensive survey can be found in [Bomze et al., 1999]. In this work, we present two algorithms to find maximum cliques. We decompose the graph vertex by vertex until no vertices remain; then, we re-build the graph restoring each of the vertices, one by one, in an inverted order computing maximal cliques at each step. Next section is devoted to present the algorithms and the theorems showing their correctness. A section showing the algorithms applied to real graphs is presented to illustrate how they work. The paper is concluded with a discussion about the complexity of the problem.

2. ALGORITHMS

Let $G = (V, E)$ be a simple undirected graph with $n = |V|$ vertices and $m = |E| \leq |V| \times |V|$ edges. The neighborhood of a vertex is the set composed by vertices directly connected to it, i.e., $w \in N(v)$ such as $\{v, w\} \in E$. Then, the degree of vertex v is denoted as $d(v) = |N(v)|$. Let us call $H = (V, E, A)$ the annotated graph G , where $|A| = |V|$ such that if $v \in V$ then a_v is a list of attributes of vertex v .

Attributes a_v are list of sets, each of one a clique, and they are computed by the proposed algorithms. For Algorithm 2, we denote an element of the list as a set $L \in a_v^x$, its initialization with set L as $a_v^x \leftarrow \{L\}$, and the append of set L as $a_v \leftarrow \{a_v, L\}$ and the elimination of a set L as $a_v \leftarrow \{a_v \setminus L\}$. The elements of this

Date: February 17, 2012.

Key words and phrases. algorithms, graphs, complexity.

Acknowledgements: Juan Ignacio Giribet, Pablo Jacovkis, Mario Valencia-Pabon, Ariel Futoransky.

list can also interpreted as a_v^i for $i \in N(v)$ (see Algorithm 1). A list of maximal cliques is stated by Definition 1.

Definition 1. *Given a vertex v and the list l_v where its elements are sets S_i having v , i.e., $S_i \supset v$. These sets $S_i \in l_v$ are maximal if for all $i \neq j$ they verify the following properties:*

- (i) $v \subseteq S_i \cap S_j$
- (ii) $S_i \not\subseteq S_j$
- (iii) $S_i \not\supseteq S_j$

Therefore, lists a_v have the properties presented in the following Definition 2.

Definition 2. *Let a vertex $v \in V$ the a_v be a list composed by sets of vertices A_i , such that the following properties hold:*

- (i) each set $A_i \in a_v$ denotes a maximal clique having v (see Definition 1);
- (ii) the maximum clique $K_{\max}(G)$ of the graph G is found as

$$(1) \quad K_{\max}(G) = A \text{ if } A = \max(|A_i| : A_i \in a_v, i \in [1, |a_v|]) ,$$

where $|a_v|$ refers to the length of a_v list.

Notice, firstly, each set A_i is maximal in the sense that there not exist other set $B \supseteq A_i$; secondly sets $A_i \in a_v$ are maximal cliques at the time that they are computed (see Algorithm 2), but at any later time can exist other clique maximal containing the A_i . As we demonstrate, computing Equation 1 when Algorithm 2 finished leads to obtain the maximum clique. We analyze it cost later.

We introduce first an algorithm of low complexity that could find the maximum clique in certain cases, this is the Algorithm 1.

Theorem 1. *Algorithm 1 ends if $G = (V, E)$ has finite size, computing the attributes a_v^x for all vertices $v \in V$ and their corresponding neighbors, according to the property (i) in Definition 2.*

Proof. Let us start with the winding phase of the recursion, that is, steps 1.2, 1.4 and 1.8 are executed until we reach an empty graph G' (see 1.7). At this step, the end of recursion (step 1.7) is found because each call to `find_cliques(G')` function is done with a reduced set of vertices $|V'| = |V - 1|$ (and its induced graph), and G has finite size; i.e., the recursion is done n times.

From there, we analyze the unwinding phase. Let's start when steps 1.19 and 1.20 are executed for the first time: the return of the function will carry a graph with just the vertex v , no edges and $a_v = \emptyset$ (step 1.5), that is $H = (\{v\}, \emptyset, a_v)$. Then, the following instance(s) can add vertices of degree zero until one instance begins to add the first edges (edge), getting a star with leaves (one leaf), because the degree is an increasing function (the winding phase was carried out taking the maximum degree at step 1.2, so the unwinding one reconnects vertices with the same degree or greater one). At this instance, steps 1.9 and 1.10 are executed and step 1.10 yields $L = \{v, x\}$ because a_x is empty (the vertex x has no registered neighbors until now). The following conditional sentence is true (step 1.5 assures $a_v^x = \emptyset$) setting each neighbor as a clique, on both sides, the neighbor vertex $a_x^v = \{v, x\}$ and the local one $a_v^x = \{v, x\}$ (see 1.12, 1.13 and 1.14, we consider objects a_w as mutable).

From this point of the algorithm execution any one of the next vertices could build a $K_s/s \in \mathbb{N}$ (e.g., $s = 3$) because a new vertex joining former vertices constituting a K_{s-1} could appear. It is worth remarking that it is not possible to build

Algorithm 1: Function $H \leftarrow \text{find_cliques}(G)$

Input: a graph $G = (V, E)$
Output: a graph $H = (V, E, A)$

```

1.1 begin
1.2   find a vertex  $v$  such that  $d(v)$  is maximum ;
1.3   set  $M = N(v)$ ;
1.4    $G' \leftarrow (V', E') : V' = V \setminus v, E' = E \setminus \{v \times N(v)\}$  ;
1.5    $a_v^x \leftarrow \emptyset$  for all  $x \in M$ , or  $a_v \leftarrow \emptyset$  if  $M = \emptyset$  ;
1.6   set  $H' \leftarrow \{\emptyset, \emptyset, \emptyset\}$  ;
1.7   if  $G' \neq \emptyset$  then
1.8     set  $H' \leftarrow \text{find\_cliques}(G')$  ;
1.9     for each  $x \in M$  do
1.10      for each set  $a_x^i$ , or  $a_x = \emptyset$  do
1.11       if  $(a_x \neq \emptyset)$ , then  $L \leftarrow (N(v) \cap a_x^i) \cup v \cup x$ , else  $L \leftarrow v \cup x$  ;
1.12       if  $(|L| \geq |a_v^x|$  or  $a_x = \emptyset)$  then
1.13          $a_x^v \leftarrow L$  ;
1.14          $a_v^x \leftarrow L$  ;
1.15       end
1.16     end
1.17   end
1.18 end
1.19 set  $H \leftarrow H' \cup (v, v \times M, a_v)$  ;
1.20 return  $H$  ;
1.21 end

```

a K_{s+1} at this stage (e.g., $s + 1 = 4$). The reason why it is not possible is that there are only K_{s-1} and a new vertex just adds edges between this new vertex and the present vertices, although this new vertex will never add an edge between the present vertices. In this way, the size of new cliques is an increasing function (either the maximum clique remains at same size or it is increased by one vertex).

Now, we will show the a_v^x is always a clique. We have also shown that the first elements in a_v^x constitute a clique of two vertices: v and its neighbor x . Considering, at any instance in the unwinding phase of the Algorithm 1, a vertex v has a clique stored for each one of its neighbors x in a_v^x . Let's consider, without loss of generality, a new neighbor of v , called w , having as neighbors $B \subseteq a_v^x = K_t$, that is $N(w) \supseteq B \cup v$. When step 1.12 is executed, either $a_w^v = \emptyset$ or $a_w^v = \{w, v\}$ (because it is possible that $a_v^x \cap N(w)$ is empty in 1.10), and then $a_w^v \leftarrow B \cup w$, which is also a clique because $B \subseteq a_v^x$ is a clique and w is a neighbor of all vertices in B by hypothesis.

To conclude the proof is enough to determine if the property (i) in Definition 2 is obtained by Algorithm 1. The initial case was already shown in the second paragraph of this proof. Then, considering a case where a maximal clique of vertices w and y is K_s , and there exists another clique K_t such that $w, y \in K_t$ and $t \leq s$. As seen in a previous paragraph, vertices can only build cliques that increase the previous one by just one vertex. Let's consider that the next vertex z is connected with all vertices in K_t and z is at most connected with $t - 1$ vertices in K_s . The

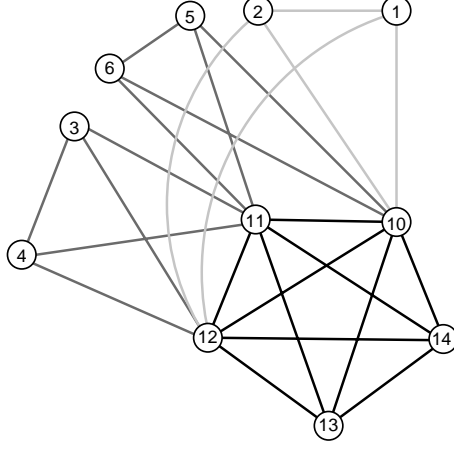


FIGURE 1. Counterexample for Algorithm 1, which gives one of size 4 instead of 5 when vertices are treated an increasing order.

minimum difference that is needed to distinguish between two cliques is one vertex. Taking into account, without loss of generality, that the maximum clique for z is $K_s \cup z$, steps 1.10 and 1.12 will select the clique $K_s \cup z$ because at least one neighbor of z , let's call it u , has $a_u^i = K_s/i \in N(u)$. At this point, the values a_z^w , a_z^y , a_w^z , and a_y^z will be updated because their size is greater or equal to the size of a clique found before (see 1.12, 1.13 and 1.14). Thus, if $t + 1 < s$ then K_s remains a maximal clique for w and y ; or else $t + 1 \geq s$ and the set of t vertices in K_t plus z is a maximal clique for w and y . It is worth remarking that it is possible that a vertex v has several cliques with a neighbor x , and the condition in 1.12 assures that the maximum clique, among the known cliques, is taken. Notice that a_w^y and a_y^w have the other clique K_s stored, but still $K_{t+1} = a_w^x \max(|a_w^x|, \forall x \in N(v))$ is the maximum clique, among the known cliques, for w because $t + 1 \geq s$ and the previous maximum clique was K_s ; the same occurs to y . \square

Before to analyze its complexity, we remark that Algorithm 1 not guarantee that the operation $A = \max(|A_i| : A_i \in a_v^x, v \in V, x \in N(v))$ gives the maximum clique $K_{\max}(G)$. In fact, Figure 1 presents a counterexample. Imagine that the unwinding phase takes vertices in the order 1, 2, 3, 4, 5, 6, 10, 11, 12, 13, 14. Regarding that list a_{12} has not the clique $\{10, 11, 12\}$ because $a_{11}^{10} = \{5, 6, 10, 11\}$, $a_{12}^{11} = \{3, 4, 11, 12\}$ and $a_{12}^{10} = \{1, 2, 10, 12\}$; therefore, when vertex 13 is added it never find the clique $\{10, 11, 12, 13\}$, so the maximum clique is not found. Nevertheless, the maximum clique is often found when very few vertices have degree close to d_{\max} and the graph is sparse, as it is the case for the so called scale free networks ¹.

Time complexity of Algorithm 1. Step 1.10 of Algorithm 1 is an intersection of two sets of size d_{\max} (maximum degree), if both are ordered² taking $O(d_{\max})$. Next, we consider loop 1.10, taking d_{\max} times, what determines $O(d_{\max}^2)$. Loop 1.9 on vertex v neighbors 1.9 takes an extra d_{\max} , giving $O(d_{\max}^3)$. Step 1.2 has a complexity of $O(n)$ if vertices are not ordered, but there is an additive cost expected

¹Graphs having a heavy tailed degree distribution which can be bound by a power law.

²This can be done at the beginning for all vertices in graph G , taking $O(n \cdot d_{\max} \cdot \log(d_{\max}))$.

smaller than $O(d_{\max}^3)$. Finally, the recursion is done for each vertex of the graph, producing a total time complexity of $O(n \cdot d_{\max}^3)$. In the case when neighbors are not ordered, we get $O(n \cdot d_{\max}^4)$. Considering connected graph, we can express the n recursions and the visit to all neighbors in loop 1.9 as visiting all the edges, yielding a time complexity of $O(m \cdot d_{\max}^2)$.

For graphs in general, where d_{\max} could be bound by n , time complexity is $O(n^4)$. However, for graphs having a heavy tailed degree distribution which can be bound by a power law³: $P(d) \propto d^{-\beta}$, it is possible to find a lower bound. Indeed, these graphs have $n^{\frac{1}{\beta}}$ as a bound of d_{\max} , therefore the complexity yields $O(n^{1+\frac{3}{\beta}})$ for $\beta \leq 3$; and for $\beta > 3$ either the search in 1.2 or the elimination in 1.4 dominates, reaching the bound $O(n^2)$ or $O(n \cdot d_{\max} \cdot \log d_{\max})$ respectively.

Storage complexity. It can be computed as the space to storage graph G , which is $O(n \cdot d_{\max})$, and the space occupied by all the a_v^x sets. This last quantity can be computed as the length of each set a_v^x , which is bound by d_{\max} and the number of them per vertex, which is also bound by d_{\max} , yielding $O(d_{\max}^2)$ per vertex. Thus, the total storage complexity is $O(n \cdot d_{\max}^2)$.

Now, we present another Algorithm 2 capable to find the maximum clique in any graph.

The following theorem proof the correctness of Algorithm 2, verifying how this algorithm accords with Definition 2.

Theorem 2. *Algorithm 1 ends if $G = (V, E)$ has finite size, computing correctly attributes in the list a_v for all vertices $v \in V$, according to Definition 2.*

Proof. Let us start with the winding phase of the recursion, that is, steps 2.2, 2.4 and 2.8 are executed until we reach an empty graph G' (see 2.7). At this step, the end of recursion (step 2.7) is found because each call to `findCliques(G')` function is done with a reduced set of vertices $|V'| = |V - 1|$ (and its induced graph), and G has finite size; i.e., the recursion is done n times.

Next, we analyze if steps 2.13 to 2.26 maintains the list a_v according to Definition 1. Firstly, suppose that a certain set L in a subset of one or more sets in a_v list, that is $L \subset S_j \in a_v$. In this case, for every $S_j \in a_v$, the conditional step 2.14 is false because the size of L is necessarily smaller than S_j because $L \subset S_j$; then the ‘else’ clause is executed. The condition in step 2.20 is also ‘false’ by hypothesis $L \subset S_j$, and again the ‘else’ clause is selected, setting the *old* variable as *TRUE* (notice this variable was initialized as *FALSE* in step 2.11). Therefore, the conditional sentence in step 2.27 will be ‘false’ because the value of *old* is not changed until step 2.11 is executed again, and L will not included in a_v . Secondly, suppose that a certain L in not a subset of any $A_i \in a_v$, and certain A_j are subsets of L . Considering when condition in 2.14 is true, then *new* will be set ‘true’ only if sets A_j are been considered, and also the set A_j is eliminated form list a_v in step 2.16 (notice that the case $a_v = \emptyset$ conditions 2.14 and 2.15 are true because B size is zero and $B \cap L = B$ when $B = \emptyset$). Finally, for the case in which L is different to all sets $A_i \in a_v$, *new* is set to *TRUE* in step 2.18 or step 2.21 because, either $|A_i| < |L|$ and $A_i \not\subseteq L$, or $A_i \not\supseteq L$, respectively. Consequently, it is shown that the algorithm maintains list a_v verifying Definition 1.

³Most of the real problems in Complex Systems field has $2 \leq \beta \leq 3$.

Algorithm 2: Function $H \leftarrow \text{find_cliques}(G)$

Input: a graph $G = (V, E)$
Output: a graph $H = (V, E, A)$

```

2.1 begin
2.2   find a vertex  $v \in V$  ;
2.3   set  $M = N(v)$ ;
2.4    $G' \leftarrow (V', E') : V' = V \setminus v, E' = E \setminus \{v \times N(v)\}$  ;
2.5    $a_v \leftarrow \emptyset$  ;
2.6   set  $H' \leftarrow \{\emptyset, \emptyset, \emptyset\}$  ;
2.7   if  $G' \neq \emptyset$  then
2.8     set  $H' \leftarrow \text{find\_cliques}(G')$  ;
2.9     for each  $x \in M$  do
2.10      for each set  $A \in a_x$ , or  $a_x = \emptyset$  do
2.11        set  $new \leftarrow FALSE$ , and  $old \leftarrow FALSE$  ;
2.12        if  $a_x \neq \emptyset$  then  $L \leftarrow (N(v) \cap A) \cup v \cup x$  else  $L \leftarrow v \cup x$  ;
2.13        for each set  $B \in a_v$ , or  $a_v = \emptyset$  do
2.14          if  $|B| < |L|$  then
2.15            if  $B \neq \emptyset$  and  $B \subset L$  then
2.16               $a_v \leftarrow \{a_v \setminus B\}$  ;
2.17            end
2.18            set  $new = TRUE$  ;
2.19          else
2.20            if  $B \not\subseteq L$  then
2.21              set  $new \leftarrow TRUE$  ;
2.22            else
2.23              set  $old \leftarrow TRUE$  ;
2.24            end
2.25          end
2.26        end
2.27        if  $new == TRUE \ \& \ old == FALSE$  then  $a_v \leftarrow \{a_v, L\}$ 
2.28      end
2.29    end
2.30  end
2.31  set  $H \leftarrow H' \cup (v, v \times M, a_v)$  ;
2.32  return  $H$  ;
2.33 end

```

We continue with the algorithm from the end of the winding phase. Let's start with the unwinding phase when steps 2.31 and 2.32 are executed for the first time: the return of the function will carry a graph with just the vertex v , no edges and $a_v = \emptyset$ (step 2.5), that is $H = (\{v\}, \emptyset, a_v)$. Then, the following instance(s) can add vertices of degree zero until one instance begins to add the first edges (edge), getting a star with leaves (one leaf). At this instance, steps 2.9 and 2.10 are executed and step 2.12 yields $L = \{v, x\}$ because a_x is empty (the vertex x has no registered neighbors until now). Steps 2.13 to 2.26 are executed yielding as result

$new = TRUE$ and $old = FALSE$ because $a_v = \emptyset$, and getting finally $a_v = \{\{v, x\}\}$ in step 2.27. Until now, we shown the initial phase, where the first cliques of size 2 are stored.

We continue to treat the case at any instance in the unwinding phase of the Algorithm 1.

To verify property (i) of Definition 2, we need to show that each set $A_i \in a_v$ is a clique. As we shown, the initialization phase let always sets composed by a vertex w and its neighbors $z \in N(w)$, which are cliques of size 2. A new joining vertex v neighbor of w will compute the step 2.12 with $a_w \neq \emptyset$ as $L \leftarrow (N(v) \cap A) \cup v \cup w$, where can be $A = \{w, z\}$. Thus, if z is also a neighbor of v , the result will be $L = \{v, w, z\}$, which is also a clique. In general, if the set A is a clique, then the result of step 2.12 is also a clique because we find the intersection of A with the neighbors of the actual connected vertex v . Then, all the sets $A_i \in a_v$ are always cliques; and as we already shown that they verify the Definition 1, they are also maximal.

To conclude the proof is enough to determine if the property (ii) of Definition 2 is obtained by Algorithm 1. The initial phase was already shown some paragraphs before. steps 2.13 to 2.26 assure that a new clique L that the joining vertex v is stored when it is different to all the previous ones, or any of the previous $A_i \in a_v$ are a subset of L . Then, when the vertex v is reconnected to the graph, all the maximal cliques are computed and stored, and among them the maximum clique that this vertex v belongs to at this time of the algorithm; we call it $K_{max,j}(v)$. Moreover, it is possible that in a later time of the execution of the algorithm, other vertex w found a greater clique containing the last seen $K_{max,j}(v)$, but this clique will be always found because it will stored in the last reconnected vertex belonging to this clique. Therefore, reading all the $A_i \in a_v$ for all vertices the maximum clique is obtained. \square

Time complexity of Algorithm 2. We firstly analyze the cost of the central loop in steps 2.13 to 2.26. Considering that the size of list a_v can be bound by the function $S(d_{max})$, and the set operations in steps 2.15 and 2.20 are bound by the maximum clique size in v , that is $O(d_{max})$, therefore this loop is done in $O(d_{max} \cdot S(d_{max}))$.

Then, step 2.12 of Algorithm 1 is an intersection of two sets of size d_{max} (maximum degree), if both are ordered⁴ taking $O(d_{max})$. Next, we consider loop 2.10, taking $S(d_{max})$ times because $|a_x| \leq S(d_{max})$, giving $O(d_{max} \cdot S^2(d_{max}))$. Loop 2.9 on vertex v neighbors takes an extra d_{max} times, giving $O(d_{max}^2 \cdot S^2(d_{max}))$. Step 2.2 has a complexity of $O(1)$. Finally, the recursion is done for each vertex of the graph, producing a total time complexity of $O(n \cdot d_{max}^2 \cdot S^2(d_{max}))$. Considering graphs with $m \gg n$, we can express the n recursions and the visit to all neighbors in loop 2.9 as visiting all the edges, yielding a time complexity of $O(m \cdot d_{max} \cdot S^2(d_{max}))$.

For graphs in general, where d_{max} could be bound by n , time complexity is $O(n^3 \cdot S^2(n))$. Now, the main problem is to bound $S(d_{max})$ function. In general, this function can be exponential but for some family of graphs it can be polynomial. This last case is observed in most of the graphs issues form the Complex System field, whose have a heavy tailed degree distribution which can be bound by a power law. These last have also the property that $m \ll n^2$, that is they are sparse. As

⁴This can be done at the beginning for all vertices in graph G , taking $O(n \cdot d_{max} \cdot \log(d_{max}))$.

the number of cliques is highly related to the number of triangles in the graph, and this is low because the graph is sparse, the number of cliques per vertex is also low. For these cases, we can model this function as $S(d_{\max}) = d_{\max}^\alpha$ with $\alpha \in \mathbb{N}$. Then, knowing that $n^{\frac{1}{\beta}}$ is a bound of d_{\max} , therefore the complexity yields $O(n^{1+\frac{2+\alpha}{\beta}})$ for $\beta \leq 2 + \alpha$; and for $\beta > 2 + \alpha$ the elimination in 1.4 can dominate, reaching the bound $O(n^2)$ or $O(n \cdot d_{\max} \cdot \log d_{\max})$.

Storage complexity. It can be computed as the space to storage graph G , which is $O(n \cdot d_{\max})$, and the space occupied by all the a_v list of sets. This last quantity can be computed as the length of each set in a_v , which is bound by $S(d_{\max})$. Thus, counting d_{\max} neighbors, the total storage complexity is $O(n \cdot d_{\max} \cdot \max(d_{\max}, S(d_{\max})))$.

3. APPLICATIONS

In this section we illustrate Algorithm 1 and 2 through an implementation developed in `python` programming language [Alvarez-Hamelin, 2011], and its application to some graphs showing their maximum clique.

Firstly, we apply our algorithm to find cliques to some random graphs defined by [Erdős and Rényi, 1959]. It is shown in [Bollobás, 2001, Bollobás and Erdős, 1976] that an ER random graph has a high probability to contain a clique of size,

$$(2) \quad r = \frac{2 \cdot \log n}{\log 1/p} ,$$

where n is the number of vertices and p is the probability that an edge exist between any pair of vertices.

n	p	\bar{d}	r	$ K_{\max} $	$ \mathcal{K} : \mathcal{K} = \{K \in K_{\max}\}$	induced $r + 1$
100	0.01	1	2	2	60	yes
1000	0.01	10	3	3	159	yes
10000	0.01	100	4	4	372	yes
10000	0.04642	464.2	6	6	5	yes

TABLE 1. Maximum cliques in Erdős Renyi graphs.

Table 1 show the results, where the columns are: the size of the graph, the probability p , the average degree \bar{d} , the computed r according to Equation 2, the size found by Equation 1, the number of different cliques (i.e., at least one vertex is different), and if an induced clique of size $r + 1$ were found. The last column is obtained adding new edges to build a greater clique than the maximum, it shows 'yes' when this clique is found and 'not' if this is not found. Moreover, the 'yes' answer also means that we find just one clique of that size (see the number of clique r find in the original graph). We tested the algorithm on several graphs of each kind, obtaining the same results (excluding $|\mathcal{K}|$ which changed some times). We display just one result of each kind.

The result is evident, we always find the predicted maximum clique, even when an artificial one is introduced.

Secondly, we applied our algorithm to a AS Internet graph. This graph has, as main properties, a power law degree distribution and most vertices of low degree are connected to the high degree ones. We used an exploration of [CAIDA, 1998] performed in September 2011. Figure 2 shows a visualization of this map obtained

by LaNet-vi [Beiró et al., 2008]. This visualization is based on k -core decomposition. A k -core is a the maximum induced subgraph such that all vertices have at least k degree [Seidman, 1983, Bollobás, 1984]. LaNet-vi paints each vertex with the rain-bow colors according to its *shell index*, i.e., the maximum core that a vertex belongs to. It also makes a greedy clique decomposition of the top core, i.e., the core with maximum k , placing each clique in circular sector according to its size.

In this graph Algorithm 2 found a K_{29} while LaNet-vi found a K_{24} ; this is displayed as the largest circular sector of red vertices in Figure 2. Moreover, this figure shows vertices of the K_{29} as those enumerated from 01 to 29. It is possible to appreciate that, even if all vertices are in the top core, the heuristic of LaNet-vi do not find this clique. For cases where some vertex is not at the top core LaNet-vi never find the maximum clique. Algorithm 1 runs several times faster than Algorithm 2, but it not always find the maximum clique. For instance, for some starting vertices, Algorithm 1 found a K_{27} instead of K_{29} .

Comparing the execution time of this graph and a ER graph of the same size, e.g., the same number of edges, we find that AS graph ends quicker than ER graph, since that its degree distribution follows a power law with $\beta \simeq 2.2$.

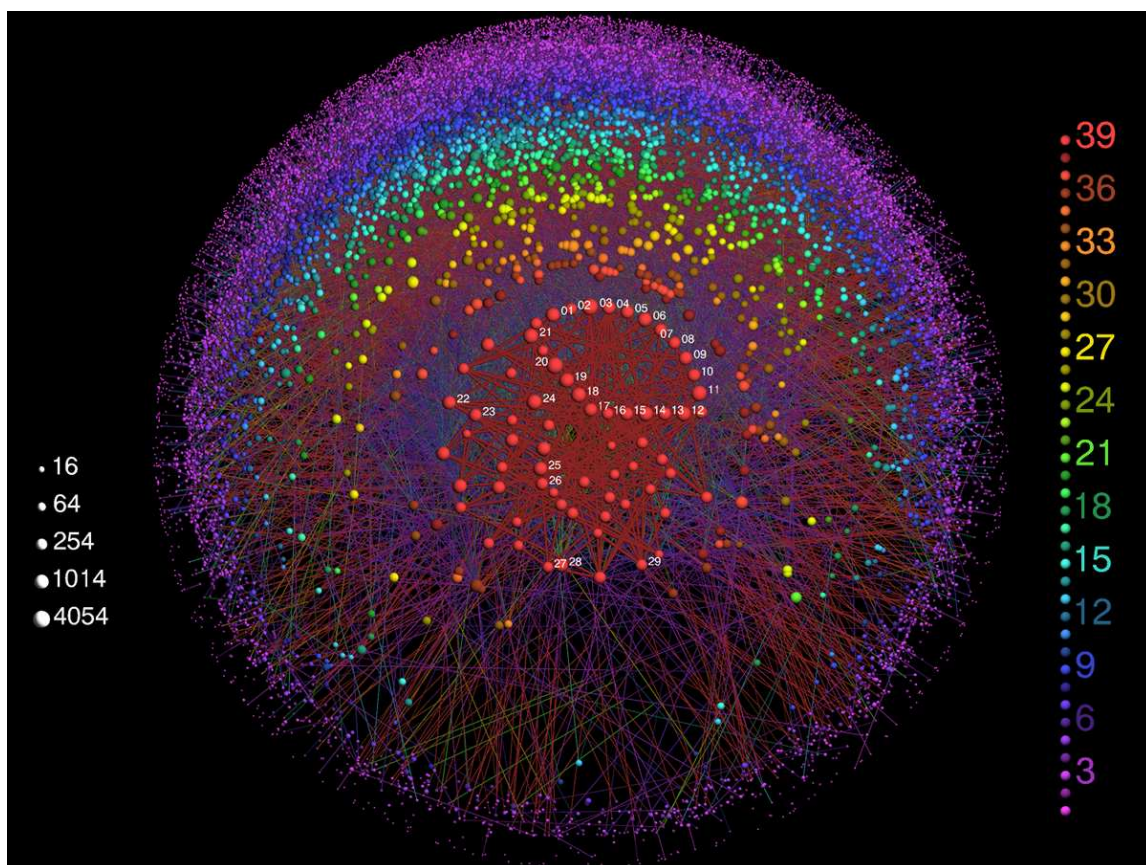


FIGURE 2. Visualization of AS Internet map by LaNet-vi. Vertices in the maximum clique are labeled with numbers.

4. DISCUSSION

As we have already remarked this problem is NP-Complete, and this can be seen from the complexity of Algorithm 2, which depends on the number of cliques that a vertex belongs to, or the Algorithm 1 do not find always the maximum clique. However, this paper aims to introduce a new approach to find cliques, that seems to be faster than the classical algorithms.

REFERENCES

- [Alvarez-Hamelin, 2011] Alvarez-Hamelin, J. I. (2011). `find_cliques`: a library for graphs in python: <http://findcliques.sourceforge.net/>.
- [Beiró et al., 2008] Beiró, M. G., Alvarez-Hamelin, J. I., and Busch, J. R. (2008). A low complexity visualization tool that helps to perform complex systems analysis. *New J. Phys*, 10(12):125003.
- [Bollobás, 1984] Bollobás, B. (1984). The evolution of sparse graphs. *Graph Theory and Combinatorics*, pages 35–57.
- [Bollobás, 2001] Bollobás, B. (2001). *Random Graphs*. Cambridge University Press.
- [Bollobás and Erdős, 1976] Bollobás, B. and Erdős, P. (1976). Cliques in random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80:419–427.
- [Bomze et al., 1999] Bomze, I., Budinich, M., Pardalos, P., and Pelillo, M. (1999). The maximum clique problem. In Du, D.-Z. and Pardalos, P. M., editors, *Handbook of Combinatorial Optimization, volume 4*. Kluwer Academic Publishers.
- [CAIDA, 1998] CAIDA (1998). Cooperative Association for Internet Data Analysis , Router-Level Topology Measurements. <http://www.caida.org/tools/measurement/skitter/>.
- [Erdős and Rényi, 1959] Erdős, P. and Rényi, A. (1959). On random graphs I. *Publ. Math. (Debrecen)*, 6:290–297.
- [Johnson, 1973] Johnson, D. S. (1973). Approximation algorithms for combinatorial problems. In *Proceedings of the fifth annual ACM symposium on Theory of computing*, STOC '73, pages 38–49, New York, NY, USA. ACM.
- [Karp, 1972] Karp, R. (1972). Reducibility among combinatorial problems. In Miller, R. and Thatcher, J., editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press.
- [Seidman, 1983] Seidman, S. B. (1983). Network structure and minimum degree. *Social Networks*, 5:269–287.

INTECIN (UBA-CONICET), FACULTAD DE INGENIERÍA, PASEO COLÓN 850, C1063ACV
BUENOS AIRES – ARGENTINA

E-mail address: ignacio.alvarez-hamelin@cnet.fi.uba.ar