

## Chapter 4

# Maximal Independent Set

In this chapter we present a first highlight of this course, a fast maximal independent set (MIS) algorithm. The algorithm is the first randomized algorithm that we study in this class. In distributed computing, randomization is a powerful and therefore omnipresent concept, as it allows for relatively simple yet efficient algorithms. As such the studied algorithm is archetypal.

A MIS is a basic building block in distributed computing, some other problems pretty much follow directly from the MIS problem. At the end of this chapter, we will give two examples, matching and vertex coloring (see Chapter 1.1).

### 4.1 MIS

**Definition 4.1** (Independent Set). *Given an undirected Graph  $G = (V, E)$  an independent set is a subset of nodes  $U \subseteq V$ , such that no two nodes in  $U$  are adjacent. An independent set is maximal if no node can be added without violating independence. An independent set of maximum cardinality is called maximum.*

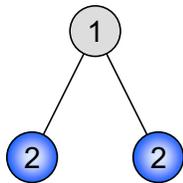


Figure 4.1: Example graph with 1) a maximal independent set (MIS) and 2) a maximum independent set (MaxIS).

**Remarks:**

- Computing a maximum independent set (MaxIS) is a notoriously difficult problem. It is equivalent to maximum clique on the complementary graph. Both problems are NP-hard, in fact not approximable within  $n^{\frac{1}{2}-\epsilon}$ .
- In this course we concentrate on the maximal independent set (MIS) problem. Please note that MIS and MaxIS can be quite different, indeed there are graphs where the MIS is  $\Theta(n)$  smaller than the MaxIS.
- Computing a MIS sequentially is trivial: Scan the nodes in arbitrary order. If a node  $u$  does not violate independence, add  $u$  to the MIS. If  $u$  violates independence, discard  $u$ . So the only question is how to compute a MIS in a distributed way.

---

**Algorithm 16** Slow MIS

---

**Require:** Node IDs**Every node**  $v$  executes the following code:

- 1: **if** all neighbors of  $v$  with larger identifiers have decided not to join the MIS **then**
  - 2:    $v$  decides to join the MIS
  - 3: **end if**
- 

**Remarks:**

- Not surprisingly the slow algorithm is not better than the sequential algorithm in the worst case, because there might be one single point of activity at any time. Formally:

**Theorem 4.2** (Analysis of Algorithm 16). *Algorithm 16 features a time complexity of  $O(n)$  and a message complexity of  $O(m)$ .*

**Remarks:**

- This is not very exciting.
- There is a relation between independent sets and node coloring (Chapter 1), since each color class is an independent set, however, not necessarily a MIS. Still, starting with a coloring, one can easily derive a MIS algorithm: We first choose all nodes of the first color. Then, for each additional color we add “in parallel” (without conflict) as many nodes as possible. Thus the following corollary holds:

**Corollary 4.3.** *Given a coloring algorithm that needs  $C$  colors and runs in time  $T$ , we can construct a MIS in time  $C + T$ .*

**Remarks:**

- Using Theorem 1.17 and Corollary 4.3 we get a distributed deterministic MIS algorithm for trees (and for bounded degree graphs) with time complexity  $O(\log^* n)$ .

- With a lower bound argument one can show that this deterministic MIS algorithm for rings is asymptotically optimal.
- There have been attempts to extend Algorithm 5 to more general graphs, however, so far without much success. Below we present a radically different approach that uses randomization. Please note that the algorithm and the analysis below is not identical with the algorithm in Peleg's book.

## 4.2 Fast MIS from 1986

---

### Algorithm 17 Fast MIS

---

The algorithm operates in synchronous rounds, grouped into phases.

**A single phase** is as follows:

- 1) Each node  $v$  marks itself with probability  $\frac{1}{2d(v)}$ , where  $d(v)$  is the current degree of  $v$ .
  - 2) If no higher degree neighbor of  $v$  is also marked, node  $v$  joins the MIS. If a higher degree neighbor of  $v$  is marked, node  $v$  unmarks itself again. (If the neighbors have the same degree, ties are broken arbitrarily, e.g., by identifier).
  - 3) Delete all nodes that joined the MIS and their neighbors, as they cannot join the MIS anymore.
- 

#### Remarks:

- Correctness in the sense that the algorithm produces an independent set is relatively simple: Steps 1 and 2 make sure that if a node  $v$  joins the MIS, then  $v$ 's neighbors do not join the MIS at the same time. Step 3 makes sure that  $v$ 's neighbors will never join the MIS.
- Likewise the algorithm eventually produces a MIS, because the node with the highest degree will mark itself at some point in step 1.
- So the only remaining question is how fast the algorithm terminates. To understand this, we need to dig a bit deeper.

**Lemma 4.4** (Joining MIS). *A node  $v$  joins the MIS in step 2 with probability  $p \geq \frac{1}{4d(v)}$ .*

Proof: Let  $M$  be the set of marked nodes in step 1. Let  $H(v)$  be the set of neighbors of  $v$  with higher degree, or same degree and higher identifier. Using independence of  $v$  and  $H(v)$  in step 1 we get

$$\begin{aligned}
 \Pr[v \notin \text{MIS} | v \in M] &= \Pr[\exists w \in H(v), w \in M | v \in M] \\
 &= \Pr[\exists w \in H(v), w \in M] \\
 &\leq \sum_{w \in H(v)} \Pr[w \in M] = \sum_{w \in H(v)} \frac{1}{2d(w)} \\
 &\leq \sum_{w \in H(v)} \frac{1}{2d(v)} \leq \frac{d(v)}{2d(v)} = \frac{1}{2}.
 \end{aligned}$$

Then

$$Pr[v \in \text{MIS}] = Pr[v \in \text{MIS} | v \in M] \cdot Pr[v \in M] \geq \frac{1}{2} \cdot \frac{1}{2d(v)}$$

□

**Lemma 4.5** (Good Nodes). *A node  $v$  is called good if*

$$\sum_{w \in N(v)} \frac{1}{2d(w)} \geq \frac{1}{6}.$$

*Otherwise we call  $v$  a bad node. A good node will be removed in step 3 with probability  $p \geq \frac{1}{36}$ .*

Proof: Let node  $v$  be good. Intuitively, good nodes have lots of low-degree neighbors, thus chances are high that one of them goes into the independent set, in which case  $v$  will be removed in step 3 of the algorithm.

If there is a neighbor  $w \in N(v)$  with degree at most 2 we are done: With Lemma 4.4 the probability that node  $w$  joins the MIS is at least  $\frac{1}{8}$ , and our good node will be removed in step 3.

So all we need to worry about is that all neighbors have at least degree 3: For any neighbor  $w$  of  $v$  we have  $\frac{1}{2d(w)} \leq \frac{1}{6}$ . Since  $\sum_{w \in N(v)} \frac{1}{2d(w)} \geq \frac{1}{6}$  there is a subset of neighbors  $S \subseteq N(v)$  such that  $\frac{1}{6} \leq \sum_{w \in S} \frac{1}{2d(w)} \leq \frac{1}{3}$

We can now bound the probability that node  $v$  will be removed. Let therefore  $R$  be the event of  $v$  being removed. Again, if a neighbor of  $v$  joins the MIS in step 2, node  $v$  will be removed in step 3. We have

$$\begin{aligned} Pr[R] &\geq Pr[\exists u \in S, u \in \text{MIS}] \\ &\geq \sum_{u \in S} Pr[u \in \text{MIS}] - \sum_{u, w \in S; u \neq w} Pr[u \in \text{MIS} \text{ and } w \in \text{MIS}]. \end{aligned}$$

For the last inequality we used the inclusion-exclusion principle truncated after the second order terms. Let  $M$  again be the set of marked nodes after step 1. Using  $Pr[u \in M] \geq Pr[u \in \text{MIS}]$  we get

$$\begin{aligned} Pr[R] &\geq \sum_{u \in S} Pr[u \in \text{MIS}] - \sum_{u, w \in S; u \neq w} Pr[u \in M \text{ and } w \in M] \\ &\geq \sum_{u \in S} Pr[u \in \text{MIS}] - \sum_{u \in S} \sum_{w \in S} Pr[u \in M] \cdot Pr[w \in M] \\ &\geq \sum_{u \in S} \frac{1}{4d(u)} - \sum_{u \in S} \sum_{w \in S} \frac{1}{2d(u)} \frac{1}{2d(w)} \\ &\geq \sum_{u \in S} \frac{1}{2d(u)} \left( \frac{1}{2} - \sum_{w \in S} \frac{1}{2d(w)} \right) \geq \frac{1}{6} \left( \frac{1}{2} - \frac{1}{3} \right) = \frac{1}{36}. \end{aligned}$$

□

**Remarks:**

- We would be almost finished if we could prove that many nodes are good in each phase. Unfortunately this is not the case: In a star-graph, for instance, only a single node is good! We need to find a work-around.

**Lemma 4.6** (Good Edges). *An edge  $e = (u, v)$  is called bad if both  $u$  and  $v$  are bad; else the edge is called good. The following holds: At any time at least half of the edges are good.*

Proof: For the proof we construct a directed auxiliary graph: Direct each edge towards the higher degree node (if both nodes have the same degree direct it towards the higher identifier). Now we need a little helper lemma before we can continue with the proof.

**Lemma 4.7.** *A bad node has outdegree at least twice its indegree.*

Proof: For the sake of contradiction, assume that a bad node  $v$  does not have outdegree at least twice its indegree. In other words, at least one third of the neighbor nodes (let's call them  $S$ ) have degree at most  $d(v)$ . But then

$$\sum_{w \in N(v)} \frac{1}{2d(w)} \geq \sum_{w \in S} \frac{1}{2d(w)} \geq \sum_{w \in S} \frac{1}{2d(v)} \geq \frac{d(v)}{3} \frac{1}{2d(v)} = \frac{1}{6}$$

which means  $v$  is good, a contradiction.  $\square$

Continuing the proof of Lemma 4.6: According to Lemma 4.7 the number of edges directed into bad nodes is at most half the number of edges directed out of bad nodes. Thus, the number of edges directed into bad nodes is at most half the number of edges. Thus, at least half of the edges are directed into good nodes. Since these edges are not bad, they must be good.

**Theorem 4.8** (Analysis of Algorithm 17). *Algorithm 17 terminates in expected time  $O(\log n)$ .*

Proof: With Lemma 4.5 a good node (and therefore a good edge!) will be deleted with constant probability. Since at least half of the edges are good (Lemma 4.6) a constant number of edges will be deleted in each phase.

More formally: With Lemmas 4.5 and 4.6 we know that an edge will be removed with probability at least  $1/72$ . Let  $R$  be the number of edges to be removed. Using linearity of expectation we know that  $\mathbb{E}[R] \geq m/72$ ,  $m$  being the total number of edges at the start of a phase. Now let  $p := \Pr[R \leq \mathbb{E}[R]/2]$ . Bounding the expectation yields

$$\mathbb{E}[R] = \sum_r \Pr[R = r] \cdot r \leq p \cdot \mathbb{E}[R]/2 + (1 - p) \cdot m.$$

Solving for  $p$  we get

$$p \leq \frac{m - \mathbb{E}[R]}{m - \mathbb{E}[R]/2} < \frac{m - \mathbb{E}[R]/2}{m} \leq 1 - 1/144.$$

In other words, with probability at least  $1/144$  at least  $m/144$  edges are removed in a phase. After expected  $O(\log m)$  phases all edges are deleted. Since  $m \leq n^2$  the Theorem follows.  $\square$

**Remarks:**

- With a bit of more math one can even show that Algorithm 17 terminates in time  $O(\log n)$  “with high probability”.
- The presented algorithm is a simplified version of an algorithm by Michael Luby, published 1986 in the SIAM Journal of Computing. Around the same time there have been a number of other papers dealing with the same or related problems, for instance by Alon, Babai, and Itai, or by Israeli and Itai. The analysis presented here takes elements of all these papers, and from other papers on distributed weighted matching. The analysis in the book by David Peleg is different, and only achieves  $O(\log^2 n)$  time.
- Though not as incredibly fast as the  $\log^*$ -coloring algorithm for trees, this algorithm is very general. It works on any graph, needs no identifiers, and can easily be made asynchronous.
- Surprisingly, much later, there have been half a dozen more papers published, with much worse results!! In 2002, for instance, there was a paper with linear running time, improving on a 1994 paper with cubic running time, restricted to trees!
- In 2009, Métivier, Robson, Saheb-Djahromi and Zemmari found a slightly different (and simpler) way to compute a MIS in the same logarithmic time:

### 4.3 Fast MIS from 2009

---

**Algorithm 18** Fast MIS 2
 

---

The algorithm operates in synchronous rounds, grouped into phases.

A **single phase** is as follows:

- 1) Each node  $v$  chooses a random value  $r(v) \in [0, 1]$  and sends it to its neighbors.
  - 2) If  $r(v) < r(w)$  for all neighbors  $w \in N(v)$ , node  $v$  enters the MIS and informs its neighbors.
  - 3) If  $v$  or a neighbor of  $v$  entered the MIS,  $v$  terminates ( $v$  and all edges adjacent to  $v$  are removed from the graph), otherwise  $v$  enters the next phase.
- 

**Remarks:**

- Correctness in the sense that the algorithm produces an independent set is simple: Steps 1 and 2 make sure that if a node  $v$  joins the MIS, then  $v$ 's neighbors do not join the MIS at the same time. Step 3 makes sure that  $v$ 's neighbors will never join the MIS.
- Likewise the algorithm eventually produces a MIS, because the node with the globally smallest value will always join the MIS, hence there is progress.
- So the only remaining question is how fast the algorithm terminates. To understand this, we need to dig a bit deeper.

- Our proof will rest on a simple, yet powerful observation about expected values of random variables that *may not be independent*:

**Theorem 4.9** (Linearity of Expectation). *Let  $X_i$ ,  $i = 1, \dots, k$  denote random variables, then*

$$\mathbb{E} \left[ \sum_i X_i \right] = \sum_i \mathbb{E} [X_i].$$

*Proof.* It is sufficient to prove  $\mathbb{E} [X + Y] = \mathbb{E} [X] + \mathbb{E} [Y]$  for two random variables  $X$  and  $Y$ , because then the statement follows by induction. Since

$$\begin{aligned} \Pr [(X, Y) = (x, y)] &= \Pr [X = x] \cdot \Pr [Y = y | X = x] \\ &= \Pr [Y = y] \cdot \Pr [X = x | Y = y] \end{aligned}$$

we get that

$$\begin{aligned} \mathbb{E} [X + Y] &= \sum_{(X, Y) = (x, y)} \Pr [(X, Y) = (x, y)] \cdot (x + y) \\ &= \sum_{X=x} \sum_{Y=y} \Pr [X = x] \cdot \Pr [Y = y | X = x] \cdot x \\ &\quad + \sum_{Y=y} \sum_{X=x} \Pr [Y = y] \cdot \Pr [X = x | Y = y] \cdot y \\ &= \sum_{X=x} \Pr [X = x] \cdot x + \sum_{Y=y} \Pr [Y = y] \cdot y \\ &= \mathbb{E} [X] + \mathbb{E} [Y]. \end{aligned}$$

#### Remarks:

- How can we prove that the algorithm only needs  $O(\log n)$  phases in expectation? It would be great if this algorithm managed to remove a constant fraction of nodes in each phase. Unfortunately, it does not.
- Instead we will prove that the number of *edges* decreases quickly. Again, it would be great if any single edge was removed with constant probability in step 3. But again, unfortunately, this is not the case.
- Maybe we can argue about the expected number of edges to be removed in one single phase? Let's see: A node  $v$  enters the MIS with probability  $1/(d(v) + 1)$ , where  $d(v)$  is the degree of node  $v$ . By doing so, not only  $v$ 's edges are removed, but indeed all the edges of  $v$ 's neighbors as well – generally these are much more than  $d(v)$  edges. So there is hope, however, we need to be careful: If we do this the most naive way, we will count the same edge many times.
- How can we fix this? The nice observation is that it is enough to count just some of the removed edges. Given a new MIS node  $v$  and a neighbor  $w \in N(v)$ , we count the edges only if  $r(v) < r(x)$  for all  $x \in N(w)$ . This looks promising. In a star graph, for instance, only the smallest random value can be accounted for removing all the edges of the star.

**Lemma 4.10** (Edge Removal). *In a single phase, we remove at least half of the edges in expectation.*

Proof: To simplify the notation, at the start of our phase, the graph is simply  $G = (V, E)$ . Suppose that a node  $v$  joins the MIS in this phase, i.e.,  $r(v) < r(w)$  for all neighbors  $w \in N(v)$ . In addition,  $r(v) < r(x)$  for all neighbors  $x$  of a neighbor  $w$  of  $v$ . The probability of this event is at least  $1/(d(v) + d(w))$ , since  $d(v) + d(w)$  is the maximum number of nodes adjacent to  $v$  or  $w$  (or both). As  $v$  joins the MIS, all edges  $(w, x)$  will be removed; there are  $d(w)$  of these edges.

Whether we remove the edges adjacent to  $w$  is a random variable. If it occurs it has the value  $d(w)$ , if not it has the value 0. For each edge  $\{v, w\}$  we have two such variables, the event  $(v \rightarrow w)$  for  $v$  joining the MIS, and symmetrically the event  $(w \rightarrow v)$  for  $w$  joining the MIS. Due to Theorem 4.9, the expected value of the sum  $X$  of all these random variables is at least

$$\begin{aligned} \mathbb{E}[X] &= \sum_{(v,w) \in E} Pr[\text{Event } (v \rightarrow w)] \cdot d(w) + Pr[\text{Event } (w \rightarrow v)] \cdot d(v) \\ &\geq \sum_{(v,w) \in E} \frac{d(w)}{d(v) + d(w)} + \frac{d(v)}{d(w) + d(v)} \\ &= \sum_{(v,w) \in E} 1 = |E|. \end{aligned}$$

In other words, in expectation all edges are removed in a single phase!?!? Probably not. This means that we still counted some edges more than once. Indeed, for an edge  $\{v, w\} \in E$  our random variable  $X$  includes the edge if the event  $(u \rightarrow v)$  happens, but  $X$  also includes the edge if the event  $(x \rightarrow w)$  happens. So we may have counted the edge  $\{v, w\}$  twice. Fortunately however, not more than twice, because at most one event  $(\cdot \rightarrow v)$  and at most one event  $(\cdot \rightarrow w)$  can happen. If  $(u \rightarrow v)$  happens, we know that  $r(u) < r(w)$  for all  $w \in N(v)$ ; hence another  $(u' \rightarrow v)$  cannot happen because  $r(u') > r(u) \in N(v)$ . Therefore the random variable  $X$  must be divided by 2. In other words, in expectation at least half of the edges are removed.

**Remarks:**

- This enables us to follow a bound on the expected running time of Algorithm 18 quite easily.

**Theorem 4.11** (Expected running time of Algorithm 18). *The expected running time of Algorithm 18 is  $3 \log_{4/3} m \in O(\log n)$ .*

Proof: The probability that in a single phase at least a quarter of all edges are removed is at least  $1/3$ . For the sake of contradiction, assume not. Then with probability less than  $1/3$  we may be lucky and many (potentially all) edges are removed. With probability at least  $2/3$  less than  $1/4$  of edges are removed. Hence the expected fraction of removed edges is strictly less than  $1/3 \cdot 1 + 2/3 \cdot 1/4 = 1/2$ . This contradicts Lemma 4.10.

Hence, at least every third phase is “good” and removes at least a quarter of the edges. To get rid of all edges (but a small constant) we need  $\log_{4/3} m$  good phases, hence a total of  $3 \log_{4/3} m$  phases are enough in expectation.

**Remarks:**

- Sometimes one expects a bit more of an algorithm: Not only should the expected time to terminate be good, but the algorithm should *always* terminate quickly. As this is impossible in randomized algorithms (after all, the random choices may be “unlucky” all the time!), researchers often settle for a compromise, and just demand that the probability that the algorithm does not terminate in the specified time can be made absurdly small. For our algorithm, this can be deduced from Lemma 4.10 and another standard tool, namely Chernoff’s Bound.

**Definition 4.12** (W.h.p.). *We say that an algorithm terminates w.h.p. (with high probability) within  $O(t)$  time if it does so with probability at least  $1 - 1/n^c$  for any choice of  $c \geq 1$ . Here  $c$  may affect the constants in the Big- $O$  notation because it is considered a “tunable constant” and usually kept small.*

**Definition 4.13** (Chernoff’s Bound). *Let  $X = \sum_{i=1}^k X_i$  be the sum of  $k$  independent  $0-1$  random variables, taking the value 1 with probability  $p \leq \frac{1}{2}$ . Then Chernoff’s bound states for every  $\alpha > 0$*

$$\Pr \left[ |X - \mathbb{E}[X]| \geq \sqrt{\alpha k \mathbb{E}[X]} \right] \leq 2^{-\Omega(\alpha k)}.$$

**Corollary 4.14** (Running Time of Algorithm 18). *Algorithm 18 terminates w.h.p. in  $O(\log n)$  time.*

Proof: We have  $p = 1/3$  and choose  $\alpha = 1/12$ . Thus, the inequality becomes

$$\Pr [|X - k/3| \geq k/6] \leq 2^{-\Omega(k)}.$$

Hence, for some sufficiently large  $k \in O(c \log n)$  we have that

$$\Pr \left[ X < \log_{4/3} m \right] \leq \Pr [X < k/6] \leq 2^{-c \log n} = 1/n^c,$$

i.e., the probability that the algorithm has not terminated after  $O(\log n)$  steps is smaller than  $1/n^c$  as claimed.

**Remarks:**

- The algorithm can be improved a bit more even. Drawing random real numbers in each phase for instance is not necessary. One can achieve the same by sending only a total of  $O(\log n)$  random (and as many non-random) bits over each edge.
- One of the main open problems in distributed computing is whether one can beat this logarithmic time, or at least achieve it with a deterministic algorithm.
- Let’s turn our attention to applications of MIS next.

## 4.4 Applications

**Definition 4.15** (Matching). *Given a graph  $G = (V, E)$  a matching is a subset of edges  $M \subseteq E$ , such that no two edges in  $M$  are adjacent (i.e., where no node is adjacent to two edges in the matching). A matching is maximal if no edge can be added without violating the above constraint. A matching of maximum cardinality is called maximum. A matching is called perfect if each node is adjacent to an edge in the matching.*

**Remarks:**

- In contrast to MaxIS, a maximum matching can be found in polynomial time (Blossom algorithm by Jack Edmonds), and is also easy to approximate (in fact, already any maximal matching is a 2-approximation).
- An independent set algorithm is also a matching algorithm: Let  $G = (V, E)$  be the graph for which we want to construct the matching. The auxiliary graph  $G'$  is defined as follows: for every edge in  $G$  there is a node in  $G'$ ; two nodes in  $G'$  are connected by an edge if their respective edges in  $G$  are adjacent. A (maximal) independent set in  $G'$  is a (maximal) matching in  $G$ , and vice versa. Using Algorithm 18 directly produces a  $O(\log n)$  bound for maximal matching.
- More importantly, our MIS algorithm can also be used for vertex coloring (Problem 1.1):

---

**Algorithm 19** General Graph Coloring

---

- 1: Given a graph  $G = (V, E)$  we virtually build a graph  $G' = (V', E')$  as follows:
  - 2: Every node  $v \in V$  clones itself  $d(v)+1$  times ( $v_0, \dots, v_{d(v)} \in V'$ ),  $d(v)$  being the degree of  $v$  in  $G$ .
  - 3: The edge set  $E'$  of  $G'$  is as follows:
  - 4: First all clones are in a clique:  $(v_i, v_j) \in E'$ , for all  $v \in V$  and all  $0 \leq i < j \leq d(v)$
  - 5: Second all  $i^{\text{th}}$  clones of neighbors in the original graph  $G$  are connected:  $(u_i, v_i) \in E'$ , for all  $(u, v) \in E$  and all  $0 \leq i \leq \min(d(u), d(v))$ .
  - 6: Now we simply run (simulate) the fast MIS Algorithm 18 on  $G'$ .
  - 7: If node  $v_i$  is in the MIS in  $G'$ , then node  $v$  gets color  $i$ .
- 

**Theorem 4.16** (Analysis of Algorithm 19). *Algorithm 19  $(\Delta + 1)$ -colors an arbitrary graph in  $O(\log n)$  time, with high probability,  $\Delta$  being the largest degree in the graph.*

Proof: Thanks to the clique among the clones at most one clone is in the MIS. And because of the  $d(v)+1$  clones of node  $v$  every node will get a free color! The running time remains logarithmic since  $G'$  has  $O(n^2)$  nodes and the exponent becomes a constant factor when applying the logarithm.

**Remarks:**

- This solves our open problem from Chapter 1.1!
- Together with Corollary 4.3 we get quite close ties between  $(\Delta+1)$ -coloring and the MIS problem.
- However, in general Algorithm 19 is not the best distributed algorithm for  $O(\Delta)$ -coloring. For fast distributed vertex coloring please check Kothapalli, Onus, Scheideler, Schindelhauer, IPDPS 2006. This algorithm is based on a  $O(\log \log n)$  time *edge* coloring algorithm by Grable and Panconesi, 1997.