Rheinische Friedrich-Wilhelms-Universität Bonn
Mathematisch-Naturwissenschaftliche Fakultät

Theoretical Aspects of Intruder Search

MA-INF 1318 Manuscript Wintersemsester 2015/2016

Elmar Langetepe

Bonn, 19. October 2015

The manuscript will be successively extended during the lecture in the Wintersemester. Hints and comments for improvements can be given to Elmar Langetepe by E-Mail `elmar.langetepe@informatik.uni-bonn.de`. Thanks in advance!

Chapter 4

Randomized variants

In this chapter and before turning over to some geometric variants of the intruder search problem, we would like to resume the graph decontamination problem for stationary guards in order to show that there are also randomized strategies and problem variants that can be discussed.

We show a slightly better approximation as the greedy algorithm for trees by a randomized strategy. Additionally, we interpret the search number of a graph in the configuration that the fire spreads from any vertex with the same probability. We concentrate on positive results.

4.1 Better approximations for trees by randomization

We pick up the firefighter problem for trees again. As already asked for in Exercise 10 we can formulate the problem as an integer LP by the following rules. Let $v \leq u$ denote that v equals u or is a predecessor of u w.r.t. the root r of tree T.

$$\begin{array}{llcll} \text{Minimize} & \sum_{v \in V} x_v w_v & & & \\ \text{so that} & x_r = 0 & = & 0 & \\ & \sum_{v \leq u} x_v & \leq & 1 & : \text{ for every leaf } u \\ & \sum_{v \in L_i} x_v & \leq & 1 & : \text{ for every level } L_i, i \geq 1 \\ & x_v & \in & \{0,1\} & : \forall\, v \in V \end{array}$$

In the above integer LP the weights w_v denote the number of vertices in the subtree T_v of vertex v w.r.t. the root r.

Let opt_{ILP} denote the optimal solution of the above integer LP. For the approximation we solve the problem in polynomial time for $x_v \in \mathbb{R}^{\geq 0}$. The optimal solution, opt_{RLP}, is a fractional solution so that a subtree T_v with $x_v = a \leq 1$ is called *a-saved*, a portion $a \cdot w_v$ of the subtree is saved. For two vertices v_1 and v_2 on the same path from the root r to a leaf u and v_1 is ancestor of v_2 and $x_{v_1} = a_1$ and $x_{v_2} = a_2$ the vertices of T_{v_2} are $(a_1 + a_2)$-saved. The remaining vertices of T_{v_1} are only a_1-saved.

The simple idea is that we would like to use a rounding scheme. But we do this in a randomized fashion. For each level we interpret the a-values as a probability dsitribution for choosing a vertex to be safe. This is a rounded strategy w.r.t. the distribution. On each level we simply choose a vertex at random according to its distribution. Note that the sum of the a-values for level i could be smaller than 1. We interpret the remaining portion as the probability of choosing none of the vertices in this level. The main problem is that we might choose vertices that are on the same path from the root to a leaf. If no such *double-protections* occur the expected value of the rounded strategy would be at least opt_{ILP} and the expected approximation value would be indeed 1.

If also a successor of a vertex is choosen by our procedure, we simply delete it in the final solution and do not choose another vertex at this level. This makes the choosing procedure independent for every level. Altogether, the only loss we have is for the double-protections. Let us assume that they can occur. What happens if the a tree T_{v_i} at level i is *fully* saved by the fractional strategy? We would like to argue that in the worst-case the fractional strategy has assigned a $1/i$ fraction to all vertices on the path from r to v_i and the subtree is fully saved by the rounding scheme with probability

$$1-(1-1/i)^i \geq 1-\frac{1}{e}\,.$$

We put this intuition into a formal argument.

Theorem 44 *Consider an algorithm that protects the vertices w.r.t. the probability distribution given by opt_{RLP}. The expected approximation ratio of the above strategy for the number of vertices protected is $\left(1-\frac{1}{e}\right)$.*

Proof. Let S_F denote the fractional solution for opt_{RLP}. For an integer solution, we choose a vertex on each level w.r.t. the probability distribution from opt_{RLP}. Let S_I denote the outcome of this assignment. We would like to show, that the expected value of S_I is larger than $\left(1-\frac{1}{e}\right)$ times the value of S_F which in turn outperforms opt_{ILP}.

Let x_v^F denote the value of x_v for the fractional strategy and let x_v^I denote the value $\{0,1\}$ of the integer strategy. For convenience we denote $y_v = \sum_{u\leq v} x_u \in \{0,1\}$, which indicates whether v is finally saved or not. Let $y_v^F = \sum_{u\leq v} x_u^F \leq 1$ denote the fraction of v saved by the fractional strategy. For $y_v = 1$ it suffices that one of the predecessor of v was chosen. Let $r = v_0, v_1, v_2, \ldots, v_k = v$ be the path from r to v, so we have

$$\mathbf{Pr}[y_v = 1] = 1-\prod_{i=1}^{k}(1-x_{v_i}^F)\,.$$

For example, the probability that v_2 is safe is $x_1 + (1-x_1)x_2 = 1-(1-x_1)(1-x_2)$ and the probability that v_3 is safe is $1-(1-x_1)(1-x_2)+(1-x_1)(1-x_2)x_3 = 1-(1-x_1)(1-x_2)(1-x_3)$ and so on.

Thus we compute

$$\begin{aligned}
\mathbf{Pr}[y_v = 1] &= 1-\prod_{i=1}^{k}(1-x_{v_i}^F)\\
&\geq 1-\left(\frac{\sum_{i=1}^{k}(1-x_{v_i}^F)}{k}\right)^k = 1-\left(\frac{k-\sum_{i=1}^{k}x_{v_i}^F}{k}\right)^k\\
&= 1-\left(\frac{k-y_v^F}{k}\right)^k\\
&= 1-\left(1-\frac{y_v^F}{k}\right)^k \geq 1-e^{-y_v^F} \geq \left(1-\frac{1}{e}\right)y_v^F\,.
\end{aligned}\tag{4.1}$$

The first inequality is a standard inequality for means of positive real values $\frac{x_1+x_2+\cdots+x_n}{n} \geq \sqrt[n]{x_1\cdot x_2\cdots x_n}$. The second and third inequalities stem from classical analysis where we use the fact that $0 \leq y_v^F \leq 1$ holds.

The value of S_F is simply the sum of all y_v^F. Thus, we conclude

$$\mathbf{E}(|S_I| = \sum_{v \in V} \mathbf{Pr}[y_v = 1] \geq \left(1 - \frac{1}{e}\right) \sum_{v \in V} y_v^F = \left(1 - \frac{1}{e}\right) |S_F| .$$

□

Altogether, we have a randomized polynomial time algorithm for trees with an expected approximation ratio better than the $\frac{1}{2}$-approximation of greedy.

Exercise 18 *Prove* $\frac{x_1+x_2+\cdots+x_n}{n} \geq \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$ *for positive real values* x_i. *Also prove the last two inequalities of Equation 4.1 in the proof of Theorem 44.*

4.2 Search numbers for random fire sources

The second part on randomization is that we might use it to consider the situation that the starting vertex of the fire has some influence on the number of agents required. Therefore, in this section we again consider the firefigther problem on graphs but the start of the fire is choosen uniformly at random among all vertices. The question is, what is the number of agents required so that for a given class C of graphs it can be expected that at least linear number of vertices can be saved.

This subsumes many questions handled before. We would like to have a classification by the properties of C, we would like to find a minimum number k of agents required and we use an expected value for assuming that the fire can start in any vertex with the same probability.

For a graph $G = (V, E)$ and a fixed number k of agents, the k-surviving rate, $s_k(G)$, is the expectation of the *proportion* of vertices that can be saved if the fire can start from any vertex with the same probability. We are looking for classes, C, of graphs G so that for a fixed constant ϵ, $s_k(G) \geq \epsilon$ holds for any $G \in C$. This means that at least $\epsilon \cdot |V|$ vertices will be saved. For a given graph G, a given k and a vertex $v \in V$ let $\mathrm{sn}_k(G, v)$, denote the number of vertices that can be protected by k agents, if the fire starts at v.

We are also searching for the minimal number k that guarantees $s_k(G) \geq \epsilon$. This means that

$$\frac{1}{|V|} \sum_{v \in V} \mathrm{sn}_k(G, v) \geq \epsilon |V|$$

has to be shown. For a class C let the minimum number k that guarantees $s_k(G) > \epsilon$ for any $G \in C$ be denoted as the firefighter-number, $\mathrm{ffn}(C)$, of C.

Firefighter-Number for a class C of graphs:
Instance: A class C of graphs $G = (V, E)$.
Question: Assume that the fire breaks out at any vertex of a graph $G \in C$ with the same probability. Compute $\mathrm{ffn}(C)$.

Theorem 45 *For planar graphs we have* $2 \leq \mathit{ffn}(C) \leq 4$.

There is a simple argument for the lower bound $2 \leq \mathrm{ffn}(C)$. Consider a planar bipartite complete graph with 2 and $n-2$ vertices on the corresponding sides. For any starting vertex at most one vertex can be saved and $\frac{1}{n}$ will become arbitrarily small.

For the upper bound we first show a somewhat easier result that shows the main idea. The vertices are subdivided into classes X and Y, where a root vertex from set X allows to save many (a linear number of) vertices and a root vertex from the set Y allows to save only few (almost zero) vertices. Finally, $|Y| \leq c|X|$ gives the bound.

Theorem 46 *For planar graphs G with no 3- and 4-cycle, we have $s_2(G) \geq 1/22$.*

Proof. We make use of the Euler formula, $c + 1 = v - e + f$, for planar graphs with e edges, v vertices, f faces and c components. We assume that the graph is connected, that is $c = 1$. A planar graph with no 3- and 4-cycle has average degree less than $\frac{10}{3}$. If not, we assume $\frac{10}{3}v \geq 2e$, summing up the degrees of al vertices gives twice the number of edges. We can also conclude $5f \leq 2e$, since any face has at least 5 edges that can neighbor two faces. This means $f \leq \frac{2}{5}$. Inserting $v \geq \frac{3}{5}e$ into the formula gives $f \geq 2 + \frac{2}{5}$, a contradiction. With similar arguments we can show that a graph with no 3-, 4 and 5-cylces has average degree less than three, which is the question of Exercise 19.

We subdivide the vertices V of G into groups w.r.t. the degree and the neighborship.

- Let X_2 denote the vertices of degree ≤ 2.
- Let Y_4 denote the vertices of degree ≥ 4.
- Let X_3 denote the vertices of degree exactly 3 but with at least one neighbor of degree ≤ 3.
- Let Y_3 denote the vertices of degree exacly 3 but with all neighbors having degree > 3 (degree 3 vertices not in X_3).

Let x_2,x_3,y_3 and y_4 denote the cardinality of the sets, respectively.

Let $|V| = n$. For a vertex starting in X_2, by two agents we protect the neighbors and safe $n - 2$ vertices. For a vertex in X_3, we save two neighbors so that the fire spreads to the neighbor u of degree ≤ 3 and in the next step we protect the remaining neighbors of u, thus protecting $n - 2$ vertices in total. For starting vertices in Y_3 and Y_4, we assume that we can save no vertex.

We have to show that $\frac{1}{n}\sum_{v\in V} \mathrm{sn}_k(G, v) \geq \epsilon \cdot n$ holds and we consider

$$s_2(G) = \frac{1}{n^2}\sum_{v\in V} \mathrm{sn}_k(G, v) \geq \frac{1}{n^2}(x_2 + x_3)(n - 2) = \frac{n-2}{n} \cdot \frac{x_2 + x_3}{x_2 + x_3 + y_3 + y_4} \tag{4.2}$$

since $x_2 + x_3 + y_3 + y_4 = n$ holds.

We first would like to compute a correspondance between Y_3 and Y_4 and consider the graph $G_Y = (V_Y, E_Y)$ that consists of the edges of G with precisely one vertex in Y_3 and one vertex in Y_4. The graph G_Y has precisely $3y_3$ edges and at most $y_3 + y_4$ vertices. Note that some of the vertices of Y_4 might be neighbors for more than one vertex of Y_3. The graph G_Y is bipartite and a subgraph of G. A cycle of size 5 has to go forth and back from Y_3 to Y_4 vertices and has to end at the same class Y_4 or Y_3. Therefore in G_Y we only have cycles of size at least 6 and by Exercise 19 the average degree of vertices of G_Y is at most 3. This means by counting $3(y_3 + y_4)$, we have counted at least any edge twice, which gives $3(y_3 + y_4) \geq 6y_3$ and $y_3 \leq y_4$.

Now we would like to compute a fixed relation between $x_2 + x_3$ and $y_3 + y_4$. By the average degree, by counting $\frac{10}{3}(x_2 + x_3 + y_3 + y_4)$ edges we have at least counted $3x_3 + 3y_3 + 4y_4$ edges, which gives $9x_3 + 9y_3 + 12y_4 \leq 10(x_2 + x_3 + y_3 + y_4)$ and in turn $2y_4 - y_3 \leq 10x_2 + x_3$. By $y_3 \leq y_4$ we have $y_4 \leq 10x_2 + x_3$ and also $y_3 + y_4 \leq 20x_2 + 2x_3 \leq 20(x_2 + x_3)$.

Now insertion into Equation 4.2 gives

$$s_2(G) \geq \frac{n-2}{n} \cdot \frac{x_2 + x_3}{x_2 + x_3 + y_3 + y_4} \geq \frac{n-2}{n} \cdot \frac{x_2 + x_3}{21(x_2 + x_3)} = \frac{n-2}{21n}. \tag{4.3}$$

If G has only two vertices, in any case the vertex distinct from the root can be saved. If G has $3 \leq n \leq 44$ vertices, at least $\frac{2}{44}$ are saved in a single step. For $n \geq 44$ we have $s_2(G) \geq \frac{42}{21\cdot 44} = \frac{1}{22}$. So the expected value of saved vertices is always $\frac{1}{22}n$. □

Exercise 19 *Prove by the Euler formula that a graph with no* 3-, 4-*cycle and* 5-*cylces has average degree less than three.*

Finally, we would like to prove the statement $\mathrm{ffn}(C) \leq 4$ of Theorem 45. To this end we prove the following result with a precise value of $s_4(G)$ for planar graphs.

Theorem 47 *Using four firefighters in the first step and then always three firefighters in each step, for every planar graph G there is a strategy such that $s_4(G) \geq \frac{1}{2712}$ holds.*

Proof. We can assume that G is a maximal planar graph without multi-edges. Inserting more edges will only help the fire but multi-edges will not. This means that G is a triangulation and we can assume that any face has exactly three edges.

We provide the proof in several steps and the proof contains four Lemmate, namely Lemma 48, Lemma 49, Lemma 50, and Lemma 51.

Similarily, to the proof above we subdivide the vertices V of G into sets X and Y. Where X will be the set of vertices where a strategy saves at least $n-6$ vertices and for Y we do not expect to save any vertex, for $|V| = n$.

The final conclusion is that for some $\alpha = \frac{1}{872}$ we will conclude

$$|Y| \leq \left(93 + \frac{3}{\alpha}\right)|X| = 2709|X|\,. \tag{4.4}$$

Thus from $|X| + |Y| = n$ and Equation 4.4 we conclude

$$s_4(G) \geq \frac{n-6}{n} \cdot \frac{|X|}{|X|+|Y|} > \frac{n-2}{n} \cdot \frac{|X|}{2710|X|} = \frac{n-6}{2710n}\,. \tag{4.5}$$

For $n \geq 10846$ we have

$$s_4(G) \geq \frac{1}{2710} - \frac{6}{4 \cdot 2710^2} \geq \frac{2710 - 3/2}{2710^2} \geq \frac{1}{2712}$$

For $2 \leq n < 10846$ we save at least $\min(4, n-1)$ in the first step, which gives also $s_4(G) \geq \frac{1}{2712}$.

The remaining task is, to establish the above bounds. First, we subdivide the vertices accordingly. Note that for starting vertices of degree 3 or four we can save $n-1$ vertices in the first step.

- For degree $3 \leq d \leq 6$ let X_d denote the vertices that guarantee to save at least $|V| - 6$ vertices.
- All other vertices form the set Y_d for $d \geq 5$.

Also note that a starting vertex v of degree 5 with a neighbor u of degree at most 6 is in X_5. Because of the triangulation u and v have two common neighbors n_1 and n_2 . In the first step, we let the fire only spread to u by protecting 4 neigbors at u. Then the neighbors v, n_1 and n_2 of u are already protected. So we fully protect the graph in the next step by 3 agents.

We require some more structural properties for the relationsship between X and Y which stem from the triangulation. The length of a path in the graph is given by the number of edges.

Lemma 48 *For a vertex $v \in Y_6$ there is a path of length at most* 3 *from v to a vertex u that has degree distinct from v (i.e., $\neq 6$) and the inner vertices of the path have degree exactly* 6.

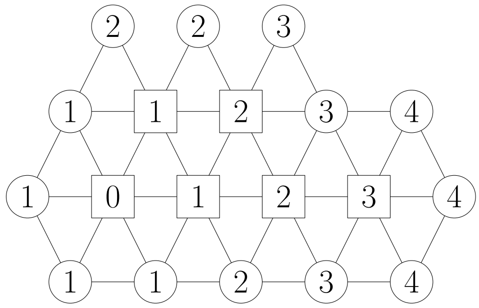

Figure 4.1: If a vertex v of degree 6 is in Y_6, we will find a vertex u as given in Lemma 48 in the neighborhood, or we finally end in a situation where v is blocked locally in a hexagon and thus belongs to X_6.

Proof. Let us assume that this is not the case. In the first step we can always protect 4 subsequent neighbors of v as depicted in Figure 4.1. If one of the remaining two neighbors (step 1) does not have degree 6, we are done. So assume that also these two neigbors have degree 6. Because of the triangulation, they will span a hexagon and we can protect 3 neighbors of these two as depicted in Figure 4.1. The fire spread to only two remaining neighbors (step 2). If one of them does not have degree 6 we are done again. So assume that both also have degree 6 and we extend the hexagonal grid. We can protect the neighbors by three agents as depicted in Figure 4.1 and only one vertex remains on fire after the fire spreads (step 3). If this vertex does not have degree 6 we are done again. But if this vertex also have degree 6 we will finally enclose the fire in the next step and only 6 vertices (v, 2 in (step 1), 2 in (step 2), 1 in (step3)) gets burned in total, a contradiction to $v \in Y_6$. Without the above property, v will be in X_5! □

The next lemma tells us something about vertices from Y_d with $d \geq 7$ related to y_5. Let $d(v)$ denote the degree of vertex v.

Lemma 49 *A vertex with $d(v) \geq 7$ has at most $\lfloor \frac{1}{2} d(v) \rfloor$ neighbors in Y_5.*

Proof. A neighbor u of v from Y_5 has two neighbors n_1 and n_2 in common with v. If one of them has degree strictly less than 7, the vertex u belongs to X_5. So the vertices from Y_5 around v are *seperated* by vertices of degree ≥ 7, which gives the bound. □

Finally, we make use of the following structural lemma that stems from the Euler formula and the simple, maximal planar triangulation.

Lemma 50 *Foir a simple, maximal planar graph we have*

$$\sum_{v \in V} (d(v) - 6) = -12\,. \tag{4.6}$$

Proof. For a maximal, simple planar graph we have $3f = 2e$, by counting the edges of every trianglular face, we count any edge exactly twice. Additionally, we have $\sum_{v \in V} d(v) = 2e$ because summing up the degree of the vertices counts any edge twice as well. The Euler formula says $v - e + f = 2$ and we conclude $v - e + \frac{2}{3} e = 2 \Longleftrightarrow 2e - 6v = -12$ which gives the conclusion. □

Now we present the main idea for obtaining Equation 4.4. The idea is that we distribute the *intitial potential* $p_1(v) := (d(v) - 6)$ of every vertex among the others so that finally any vertex has potential $p_2(v)$ and also $\sum_{v \in V} p_1(v) = \sum_{v \in V} p_2(v) = -12$ holds.

The rules for the distribution are as follows:

Rule A: A vertex v of degree at least 7 gives a value of $\frac{1}{4}$ to each neighbor vertex from Y_5.

Rule B: For a vertex $v \in Y_6$ we choose exactly one vertex u with $d(u) \neq 6$ and distance $d(v, u) \leq 6$ as in Lemma 48. The vertex u gives a value of $\alpha > 0$ to v.

We would like to choose α accordingly, the property $\sum_{v \in V} p_1(v) = \sum_{v \in V} p_2(v) = -12$ will hold since the distribution is cost neutral by construction. Such a distribution with desired properties exists.

Lemma 51 *There is a constant $\alpha > 0$ such that a distribution by Rule A and B gives $\sum_{v \in V} p_1(v) = \sum_{v \in V} p_2(v) = -12$ and for every $v \in X$ we have $p_2(v) > -3 - 93\alpha$ and for every $v \in Y$ we have $p_2(v) \geq \alpha$.*

Before we prove this final lemma, we use its conclusion. An $\alpha = \frac{1}{872}$ will do the job. We then conclude

$$-12 = \sum_{v \in V} p_2(v) \geq (-3 - 93\alpha)|X| + \alpha|Y|$$

and this gives

$$|Y| \leq \left(93 + \frac{3}{\alpha}\right)|X| < 2790|X|$$

which is Equation 4.4. It remains to prove Lemma 51.

Proof of Lemma 51.

For the claim we estimate the distribution of the potential as given by the rules A and B.

Considering Rule B, how often can a vertex u with $d(u) \neq 6$ give a potential of α to some vertex v. We give a rough upper bound with respect to the maximal distance ≤ 3 from u:

- Distance 1: $d(v)$ times to a direct neighbor, if all of them are in Y_6. This gives $1 \cdot d(u)$.
- Distance 2: For all $d(v)$ neighbors of the first case, at most 5 times, if the $d(v)$ neighbors of the above case have degree 6 and all 5 remaining neigbors are from Y_6. This gives $5 \cdot d(u)$.
- Distance 3: For all vertices of the second case, at most 5 times, if the vertices of the second case all have degree 6 and the remaining neighbors are from Y_6. This gives $25 \cdot d(u)$.

Altogether, any vertex u with $d(u) \neq 6$ can give a potential α at most $(1+5+25)d(u) = 31d(u)$ times. This gives upper bounds for the potential $p_2(v)$:

- $v \in X_3$: We have $p_2(v) \geq -3 - 93\alpha$ because $d(v) = 3$ and $p_1(v) = -3$.
- $v \in X_4$: We have $p_2(v) \geq -2 - 124\alpha$ because $d(v) = 4$ and $p_1(v) = -2$.
- $v \in X_5$: We have $p_2(v) \geq -1 - 155\alpha$ because $d(v) = 5$ and $p_1(v) = -1$.

For vertices of degree 6 we have the following:

- $v \in X_6$: $p_2(v) = 0$ because $d(v) = 6$ and $p_1(v) = 0$.
- $v \in Y_6$: $p_2(v) = p_1(v) + \alpha = \alpha$ because Rule B gives a single value α from some u to v, and by Lemma 48 such a vertex exists, if v exists.

Note that the distributions of these α potentials are cost-neutral in total for p_1 and p_2.

Considering Rule A, for vertex v and $d(v) \geq 7$, we can apply Lemma 49 and the above estimate for an upper bound

$$p_2(v) \geq (d(v) - 6) - \left\lfloor \frac{1}{2} d(v) \right\rfloor \cdot \frac{1}{4} - 31 d(v) \alpha \,. \tag{4.7}$$

So the remaining cases can be estimated by

- $v \in X \cup Y$ with $d(v) = 7$: $p_2(v) \geq \frac{1}{4} - 217\alpha$.
- $v \in X \cup Y$ with $d(v) \geq 8$: $p_2(v) \geq d(v) \left(\frac{7}{8} - 31\alpha\right) - 6$ by $\lfloor \frac{1}{2} d(v) \rfloor \cdot \frac{1}{4} \leq \frac{1}{8} d(v)$.

We choose $\alpha > 0$ so that $\frac{1}{4} - 217\alpha \geq \alpha$ and $d\left(\frac{7}{8} - 31\alpha\right) - 6 \geq \alpha$ holds for all $d \geq 8$. The first inequality is fulfilled for $\alpha = \frac{1}{218 \cdot 4} = \frac{1}{872}$ and this also fulfills the second inequality. Thus, we have shown that for $v \in Y$ $p_2(v) \geq \alpha$ holds. The same value for α also guarantess $p_2(v) \geq -3 - 93\alpha$ for all $v \in X$.

It remains to note that the distribution is cost-neutral also for Rule A. But this is clear because the sum of the potential added and retracted is the same. Altogether, Lemma 51 holds and the full conclusion can be drawn, which is $s_4(G) \geq \frac{1}{2712}$. $\square$

Exercise 20 *Present the precise strategies that stem from Theorem 46 and Theorem 47. Analyse the corresponding running time.*