Distributed Algorithms

Election Algorithms

Univ.-Prof. Dr.-Ing. habil. Gero Mühl

Architecture of Application Systems (AVA)
Faculty for Informatics and Electrical Engineering
University of Rostock

Overview

> The election problem

> Election algorithms for
  > arbitrary connected topologies
  > unidirectional and bidirectional rings
  > trees

> Randomized election algorithms for
  > bidirectional rings
  > anonymous rings

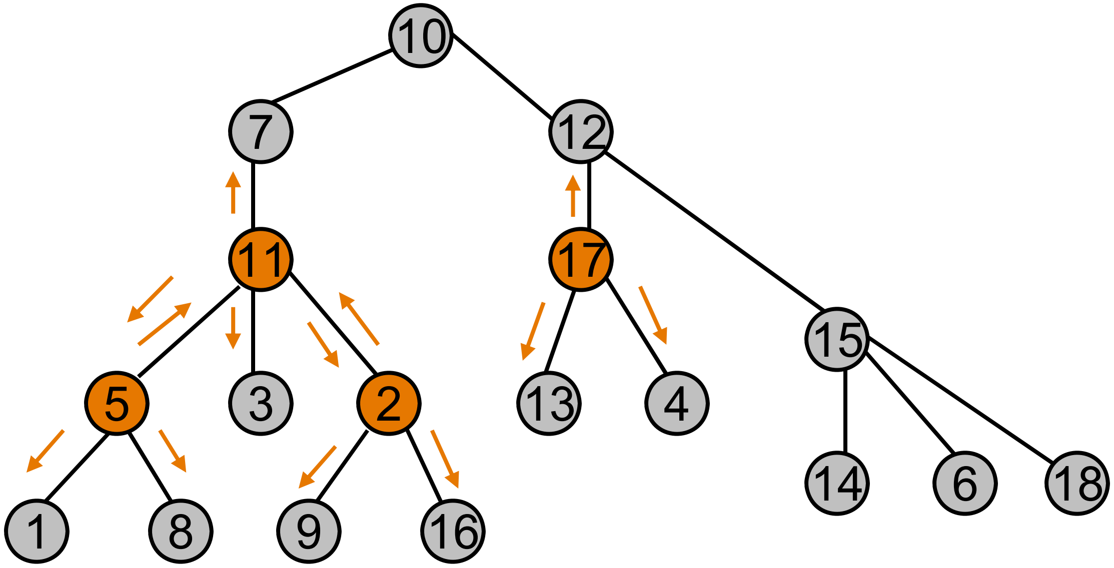
Election Algorithms for Trees

Election Algorithms on Trees

- Algorithm proceeds in three phases
  1. Explosion phase
     - Election request is propagated from the initiators to the leafs of the tree
  2. Contraction phase
     - From the leafs, the maximum of the already collected identities is propagated towards the center
  3. Information phase
     - Actual maximum is distributed from the center to all nodes in the network

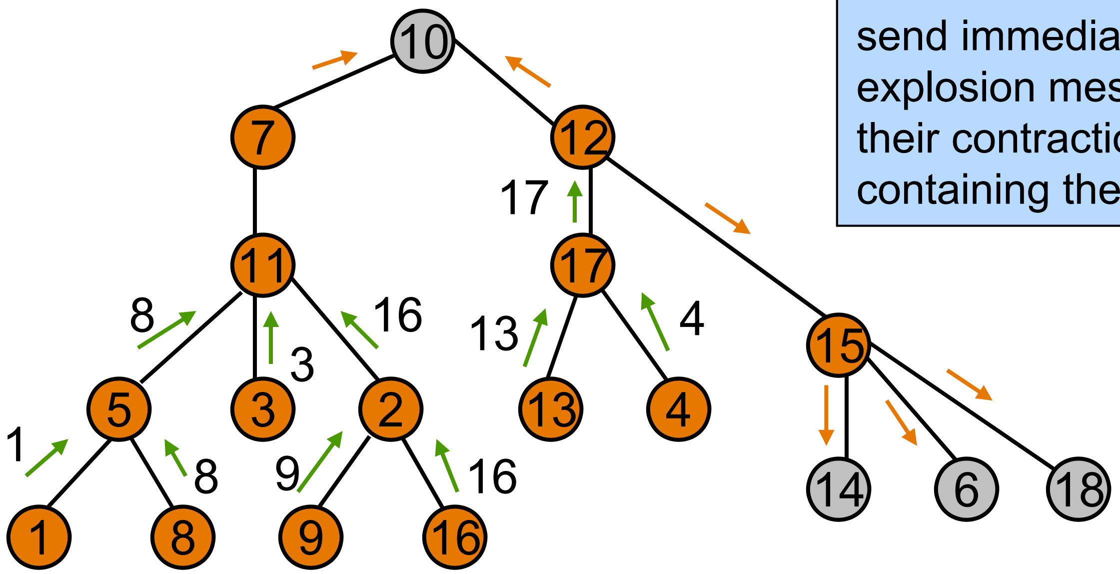
Explosion Phase

- Explosion starts at the initiators
- When a node receives an explosion message for the first time, the message is passed on to all other neighbors
- The explosion waves unite, where explosion messages meet on an edge



Contraction Phase

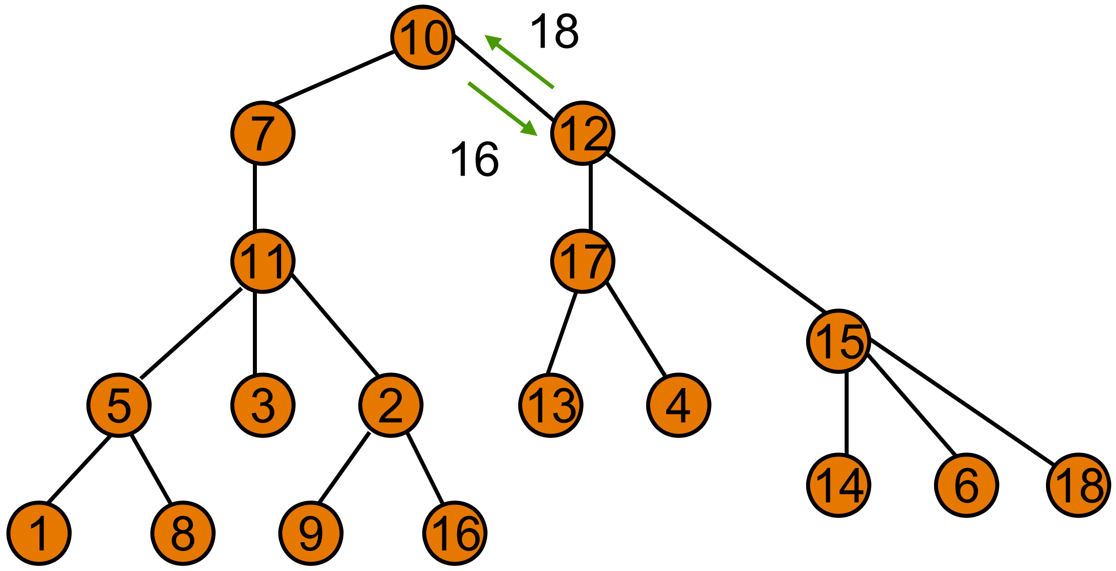
- Leafs answer an explosion message immediately with their own identity
- All other nodes send the maximum of the received identities and their own identity from the other edges over the last remaining edge



Leafs that are initiators send immediately after the explosion message also their contraction message containing their ID.

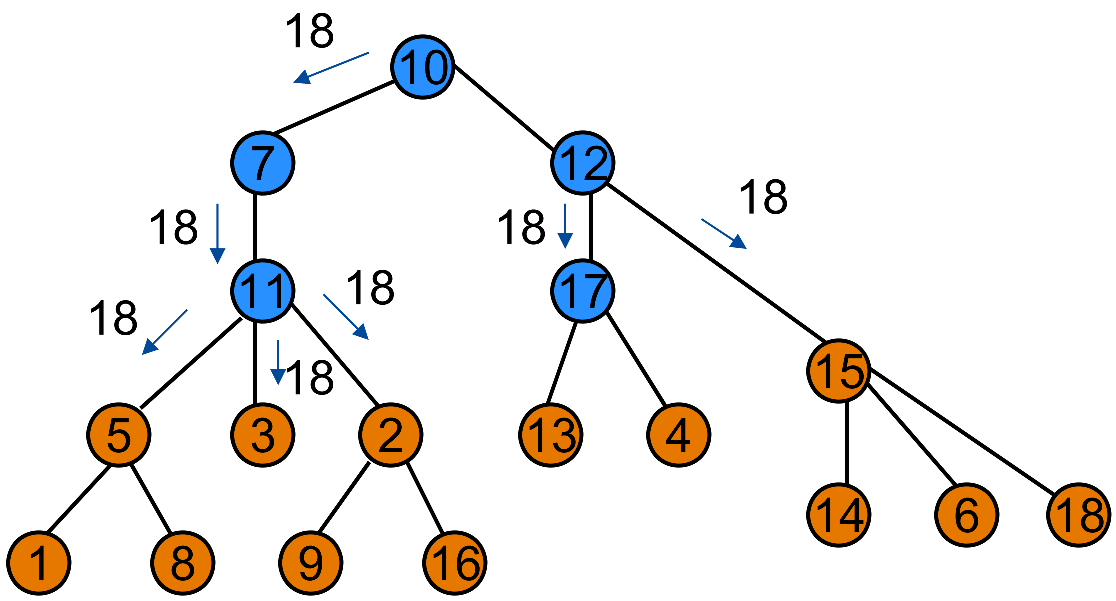
End of the Contraction Phase

> On *exactly* one edge two contraction messages meet with *different* maxima
> Both receiving nodes know the real maximum afterwards



Information Phase

> From both nodes the maximum is flooded into the network; hereby the edge between them is omitted



Message Complexity with *k* Initiators

- Explosion Phase: $(n-1)+(k-1)=n-2+k$
  - One message over each edge
  - Exception: 2 messages over the k - 1 meeting edges
- Contraction Phase: $(n-1)+1=n$
  - One message over each edge
  - Exception: Two messages over the central edge
- Information Phase: $(n-1)-1=n-2$
  - One message over every edge
  - Exception: No message over the central edge
- Altogether $3n+k-4$ messages → $O(n)$

⇒ Election on trees is much more efficient than on rings!

Randomized Election Algorithms

Randomized Algorithms

> Random algorithms are non-deterministic algorithms that influence their execution with random numbers

> Can be both determined or not determined

> Are often simpler than deterministic algorithms solving the same problem

> Through randomized algorithms some problems can be solved more efficiently (or at all)

Two Categories of Randomized Algorithms

- Las Vegas-Algorithms: weakened termination
  - Always provide a correct result,
    but the worst-case time complexity is unlimited
  - The limit of the termination probability approaches 1
    when the run time approaches infinity

- Monte Carlo-Algorithms: weakened partial correctness
  - Worst-case time complexity is limited,
    but they sometimes provide a wrong result

Randomized Election in Bidirectional Rings

```
I_p: {M_p == 0}
   M_p := p;
   <choose randomly one of two directions>;
   SEND <M_p> TO <next node in chosen direction>;


R_p: {A message <j> has arrived from either direction}
   IF M_p < j THEN
      M_p := j;
      SEND <M_p> TO <next node in other direction>;
   FI
   IF j == p THEN
      <node has won the election>
      <inform all by additional ring circuit>;
   FI
```

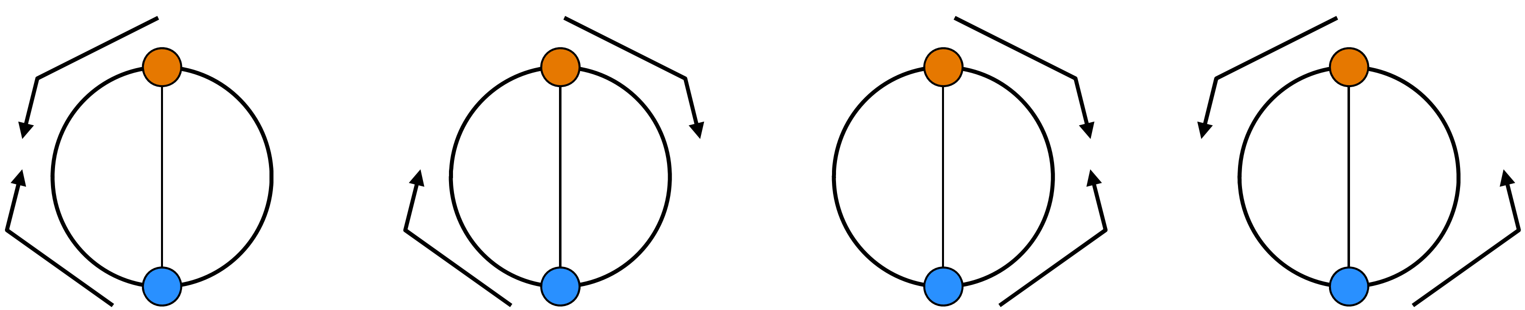
Slide shows randomized variant of the Chang/Roberts-Algorithm for bidirectional rings.

Randomized Election in Bidirectional Rings

- Average-case message complexity for $k = n$ is
  $0.5\sqrt{2}\, n \ln n \approx 0.71\, n \ln n = O(n \log n)$ (Lavault, 1990)
- That is about 30% better than the deterministic algorithm for unidirectional rings by Chang and Roberts!
- Why?

Randomized Election in Bidirectional Rings

- We first consider the simple case of only two initiators assuming equal message delays
- In half of the four cases (first and third figure below), the higher message and the lower message meet in the middle
- In this case, only half as much messages are needed on average for the lower message (i.e., n / 4 instead of n / 2)
- Therefore, on average 3/8 n messages instead of 1/2 n
  → 25% savings for the lower message

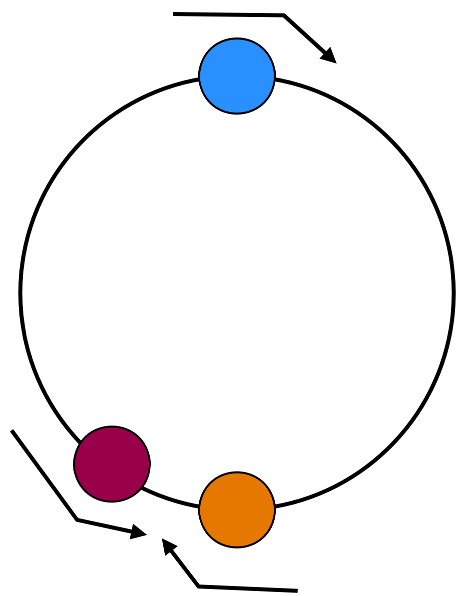


Randomized Election in Bidirectional Rings

- This argumentation can be generalized
- For the i-highest initiator ($i > 1$), the next higher initiator (called eliminator node) is in either direction on average n / i hops away
- In half of the cases, the messages from the initiator and from the eliminator node that lies in the chosen direction meet in the middle
- In these cases, on average $n / 2i$ messages are needed instead of n / i messages
- Therefore, on average $3 / 4 \cdot n / i$ instead of n / i messages are needed → again 25% savings for this message
- Asymptotically (i.e., for very large n), the fact that the highest message has to make a whole round can be neglected
- Thus, asymptotically 25% of the messages are saved on average

Randomized Election in Bidirectional Rings

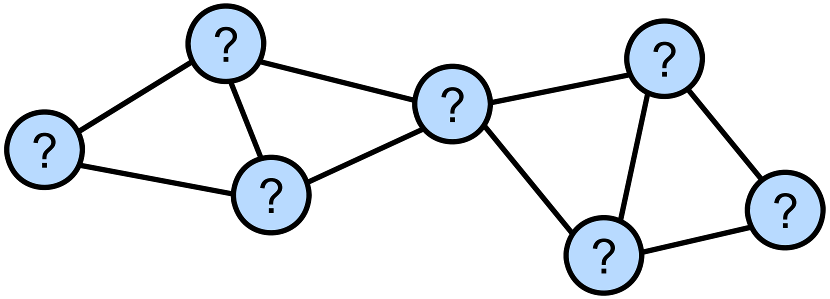
- Even if the respective eliminator node sends its message in the wrong direction, there can be a higher order eliminator that sends its message in the right direction
- This accounts for the remaining 4% savings



Considered initiator
Next-higher initiator
Higher-order initiator

Election in Anonymous Networks

- In anonymous networks, nodes do not have permanent unique identities
- Is it possible to determine a unique winner of the election, then?
- If so, under which conditions?



Las Vegas Election for Anonymous Rings

- Algorithm of Itai and Rodeh is based on the algorithm of Chang and Roberts
- Assumption: Ring size *n* is known
- Each node is an initiator and randomly chooses a temporary identity from the set of numbers $\{1, 2, \ldots, c\}$ with $c \geq 2$
- Thus, several initiators may choose the same identity
- Message extinction as usual, but messages contain
  - hop count *h* that is 1 initially and that is incremented by each relaying node
  - flag *f* that has the value 1 initially when sending
  - variable with the number of the corresponding election round

Las Vegas Election for Anonymous Rings

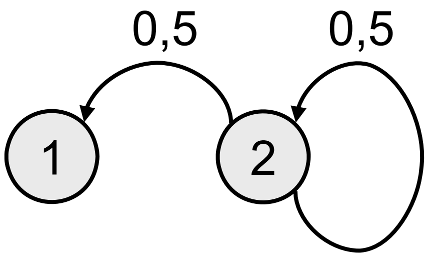
- If a node receives a message with its own identity, two cases are distinguished
  1. If h != n, there is at least one other node with the same identity. Thus, f is set to 0 and the message is relayed
  2. If $h = n$, the node has won the election
     - If $f = 1$, it is the only winner and it can send the win message
     - If $f = 0$, there are several winners and more election rounds are needed to determine a unique winner

Las Vegas Election for Anonymous Rings

- Elections are carried out, until there is a unique winner
- Only the winners of the most recent election round participate actively in the next election round under a new random ID
- Eliminated nodes are passive and only relay messages (with incremented hop count)
- Messages from earlier election rounds are simply swallowed
- Expected value E for the number of elections for $c = n$ is bounded by $\mathcal{E}(n / (n - 1))$ ($\mathcal{E}$ is Euler's number)
- A generalization of the algorithm to general networks is possible (with application of the echo election algorithm) provided that the number of nodes is known

Expectation Value for $n = c = 2$

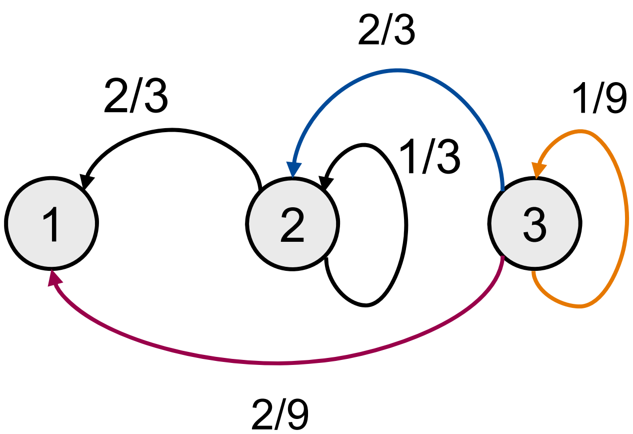
> Application of a Markov-Chain

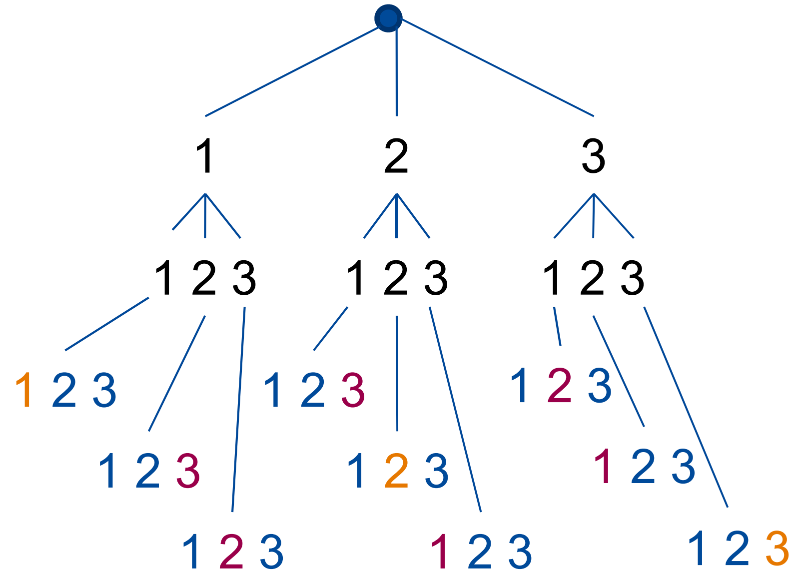


> $E = \sum_{i=1}^{\infty} i \cdot 2^{-i}$

$= 1 \cdot 2^{-1} + 2 \cdot 2^{-2} + 3 \cdot 2^{-3} + \ldots = 2$

Expectation Value for *n* = *c* = 3





$$\sum_{i=1}^{\infty} \left(\frac{1}{9}\right)^{i-1} \cdot \left(\frac{2}{9}\cdot i + \frac{2}{3}\cdot(i + 1.5)\right) = 2.251$$

$$\sum_{i=1}^{\infty} i \cdot \left(\frac{1}{3}\right)^{i-1} \cdot \frac{2}{3} = 1.5$$

> Derivation is also possible for a general *n*

Election in Rings with Unknown Size

- Through observing patterns of the identities of the nodes, the ring size can be estimated by several ring circulations
  - For large rings and large *c* it is probable that one estimates correctly; but what happens if one estimates incorrectly?
  - Then, there are several winners and this is not detected
- What amount of minimal (structural) information is necessary to break the symmetry either deterministically or randomized?
  - The ring size lies between *N* and *2N-1*
  - The unknown ring size is a prime number (only with further assumptions)
  - ...
- There is neither a deterministic nor a Las Vegas algorithm for the election in anonymous rings with unknown size!

Exemplary Exam Questions

1. Explain how an election algorithm can be implemented based on the Echo algorithm!
2. Describe the algorithm of Le Lann and the algorithm of Chang and Roberts for unidirectional rings!
3. Describe the Hirschberg-Sinclair election algorithm as well as the Peterson election algorithm for bidirectional rings!
4. How do the algorithms of questions 2 and 3 differ in terms of their message complexity?
5. How can the election on trees be performed and which messages complexity results from this election procedure?
6. What is a Las Vegas algorithm?
7. Explain the election algorithm of Itai and Rodeh for anonymous rings!

Literature

1. E. Chang and R. Roberts. An improved algorithm for decentralized extrema-finding in circular configurations of processes. Communications of the ACM (CACM), 22(5):281--283, 1979.
2. D. S. Hirschberg and J. B. Sinclair. Decentralized extrema-finding in circular configurations of processors. Communications of the ACM (CACM), 23(11):627--628, 1980.
3. Gary L. Peterson. An O(n log n) unidirectional algorithm for the circular extrema problem. ACM Transactions on Programming Languages and Systems (TOPLAS), 4(4):758--762, 1982.
4. C. Lavault. Average number of messages for distributed leader finding in rings of processors. Information Processing Letters, 30:167--176, 1989.
5. A. Itai and M. Rodeh. Symmetry breaking in distributed networks,. In Proceedings of the 22nd IEEE Symposium on Foundations of Computer Science, pages 150--158. IEEE Press, 1981.
6. Friedemann Mattern. Verteilte Basisalgorithmen. Springer-Verlag, 1989. Kapitel 2: Untersuchung von Election-Algorithmen

Thank you for your kind attention!

Univ.-Prof. Dr.-Ing. habil. Gero Mühl

gero.muehl@uni-rostock.de
http://wwwava.informatik.uni-rostock.de