Towards malware inspired management frameworks

Jérôme François, Radu State and Olivier Festor

MADYNES

Loria

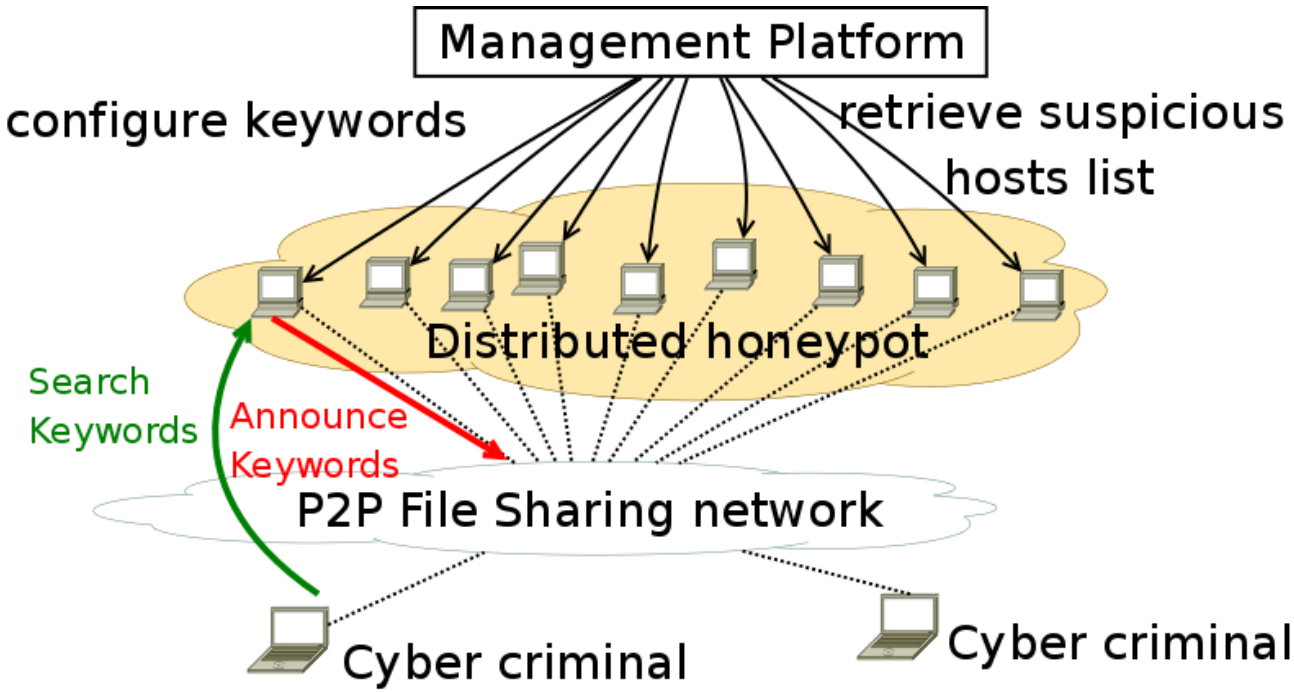
Nancy-Université

Outline

Outline

Motivation

- scalable management
- mass configuration
- distributed honeypots for tracking cyber-predators
- announce specific-keywords on P2P file sharing system



Research challenges

- scalability: open participation to honeypot
- efficiency: keywords changes $\longrightarrow$ fast keywords updates
- tracking prevention: controller and honeypots anonymity
- security: false keywords list updates
- reachability guarentees: knowing the impact of a request is needed provide additional operations

Outline

Malware communication paradigms

- attackers faced the same problems
  - control multiple machines through the Internet
  - goals: distributed denial of service attacks, mass collecting of sensitive data
- construction of a botnet
  - control mechanism to send orders to the bots and get the responses
  - decentralized and scalable: example of 400 000 zombies in one botnet

Botnet based network management

- use a botnet to perform management operations
- different types of botnet
  - IRC model[1]
  - P2P models : unstructered (Slapper) and structured (Chord)

$\longrightarrow$ study of performances of these types of botnets once they are deployed

[1] J. Francois, R. State, and O. Festor, 'Botnet based scalable network management', DSOM 2007

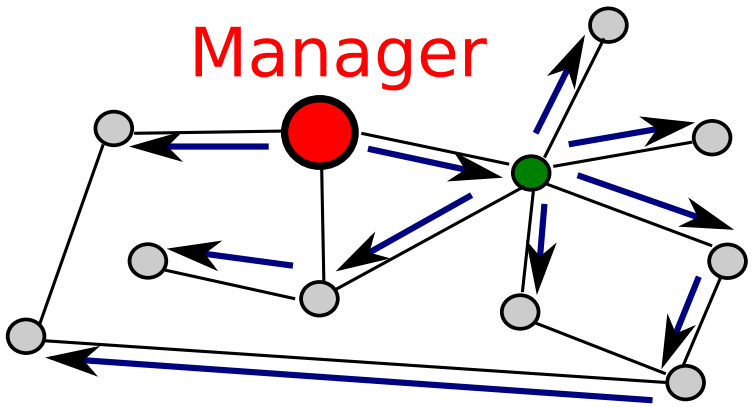
Outline

Parameters

- N: total number of devices/peers
- m is the maximal branching factor = the maximal number message sent by a peer at the same time (message forwarding)

N = 10 devices + 1 manager = 11

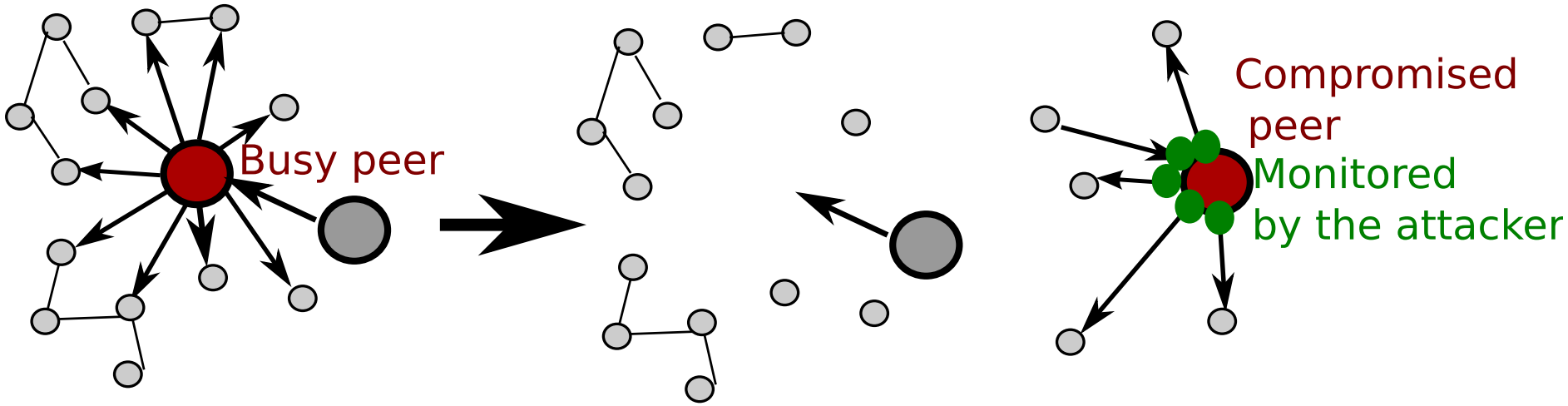
m = 5



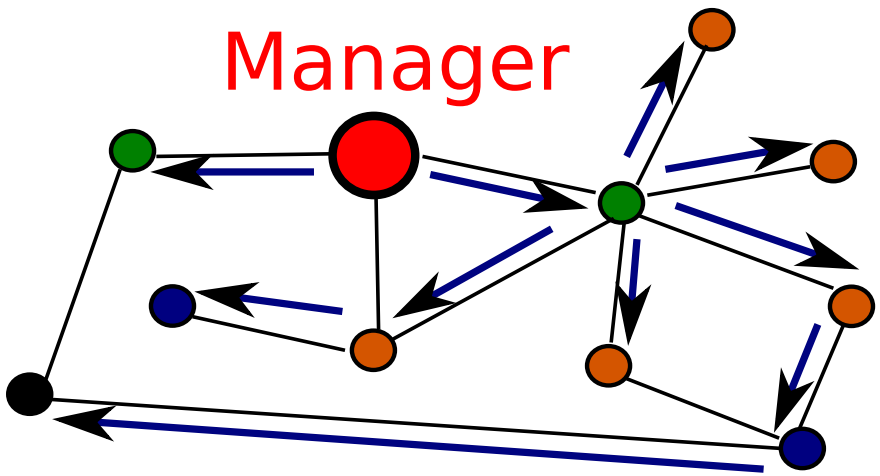
Broadcast message Peers

Parameters

- a peer can crash if it has to maintain too many connections $\longrightarrow \alpha(m)$ is the probability for a peer to be able to forward the messages, decreasing function
- the risk to be compromised by an attacker and to be attacked (network communication monitoring): β



Goal: determine the reachability = the number of peers reached at a certain distance



Distance 1: 2 peers

Distance 2: 5 peers

Distance 3: 4 peers

Distance 4: 1 peer

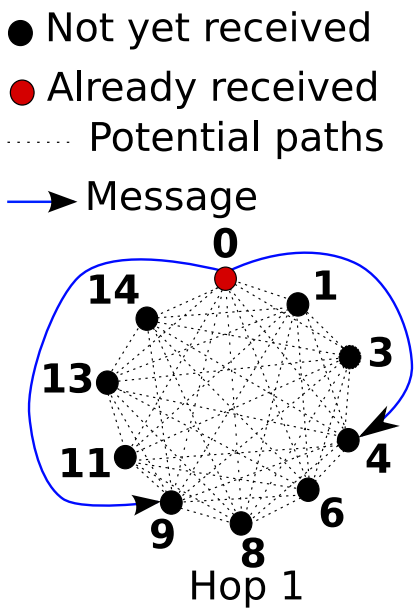
Broadcast message

Peers

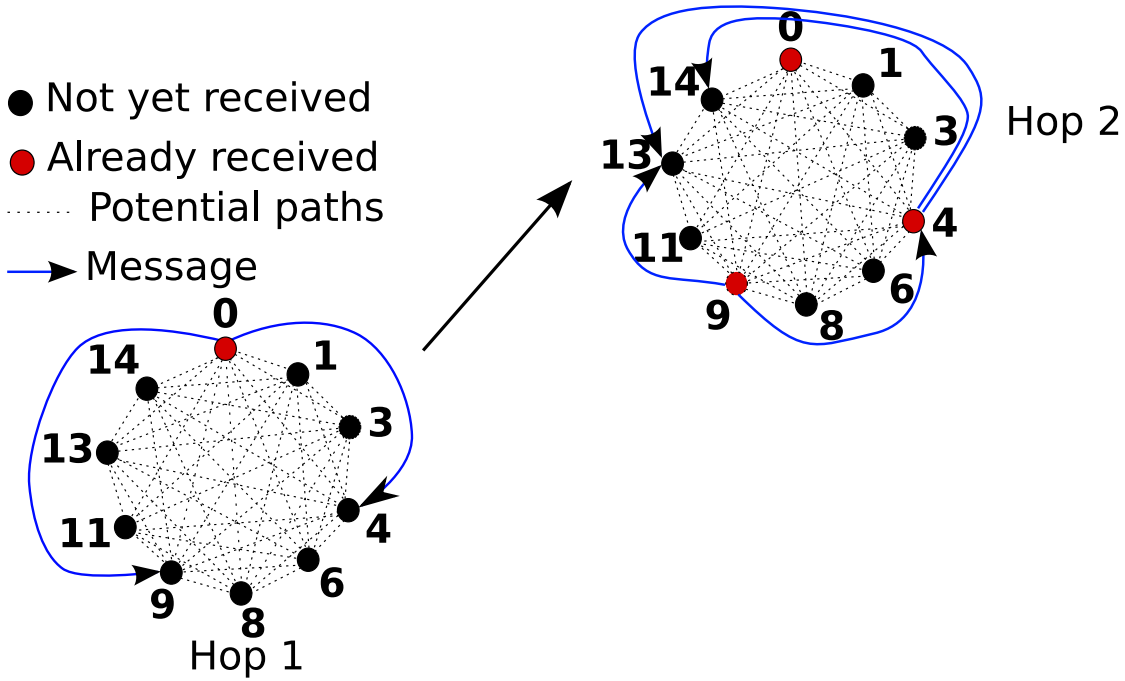
Slapper model

- a sophisticated worm
- infected computers form a botnet
  - full-meshed network
  - controller tracking prevention: the message is transmitted through several peers
- broadcast segmentation
  - the initiator (the controller) sends the messages to m random peers
  - when a peer receives a message, it sends the messages to m random peers
  - a maximal number of hops is fixed
  - original $m = 2$

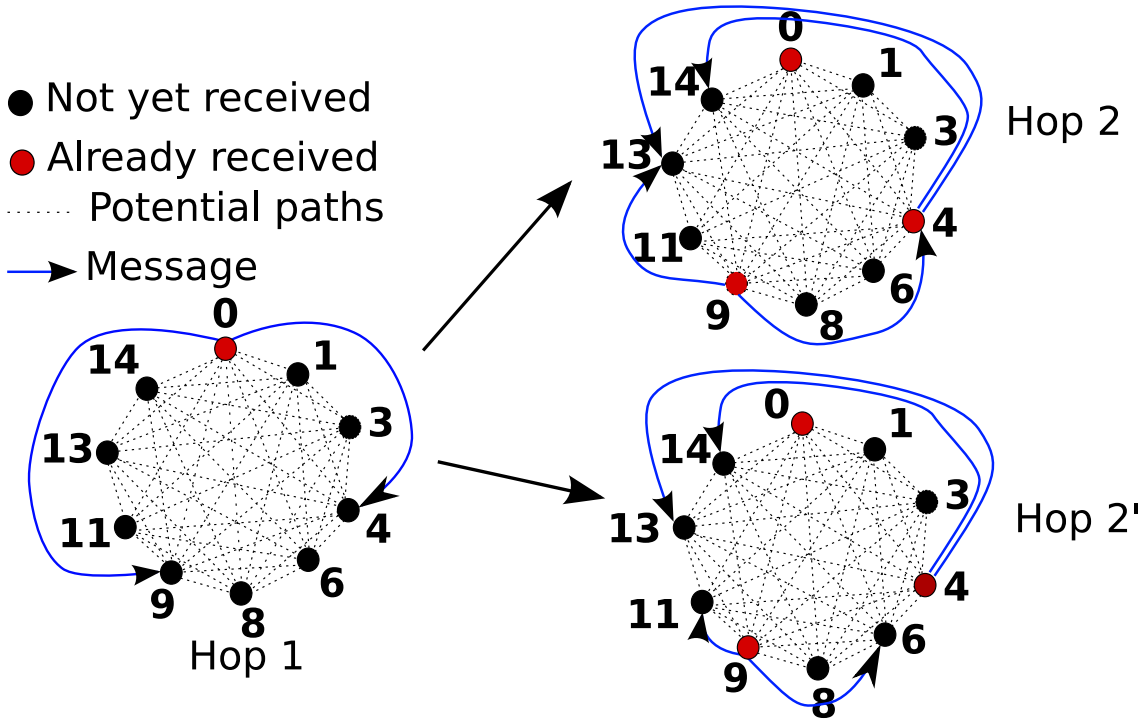
Slapper model



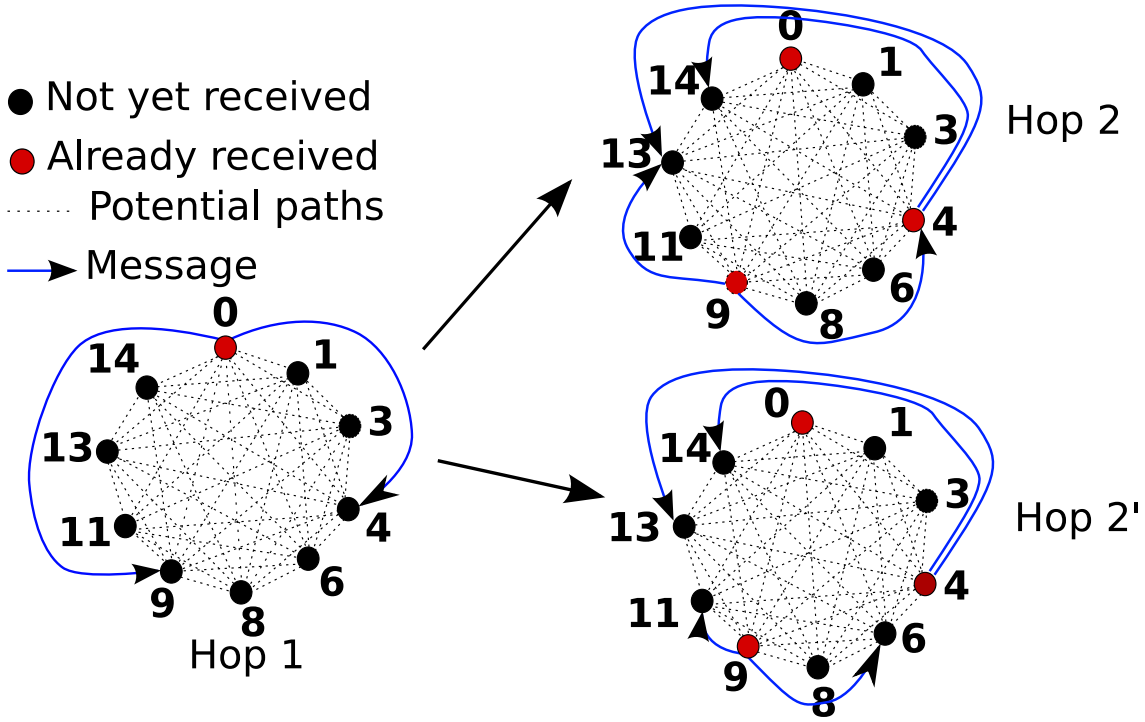
Slapper model



Slapper model



Slapper model



- the same message can be sent to the same peers two times
- no guarentee to reach all peers

Chord model

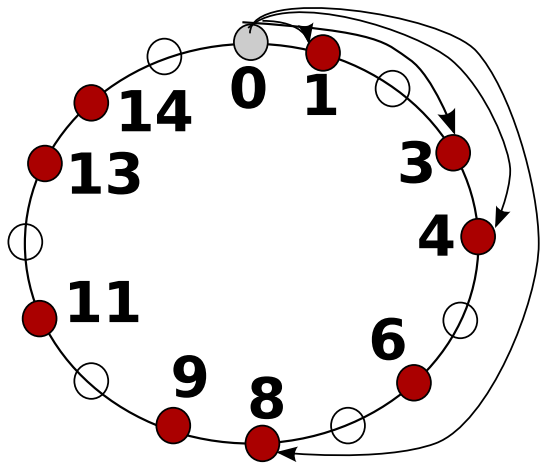
- each peer has an id: $0 \leq id < N_{MAX}$
- routing table of each node p:
  - $log(N_{MAX})$ entries
  - ith entry: first id at a distance from p at least 2^{i-1}

○ Non used identifier

● Peers

Routing table
of node 0

i	1	2	3	4
node	1	3	4	8



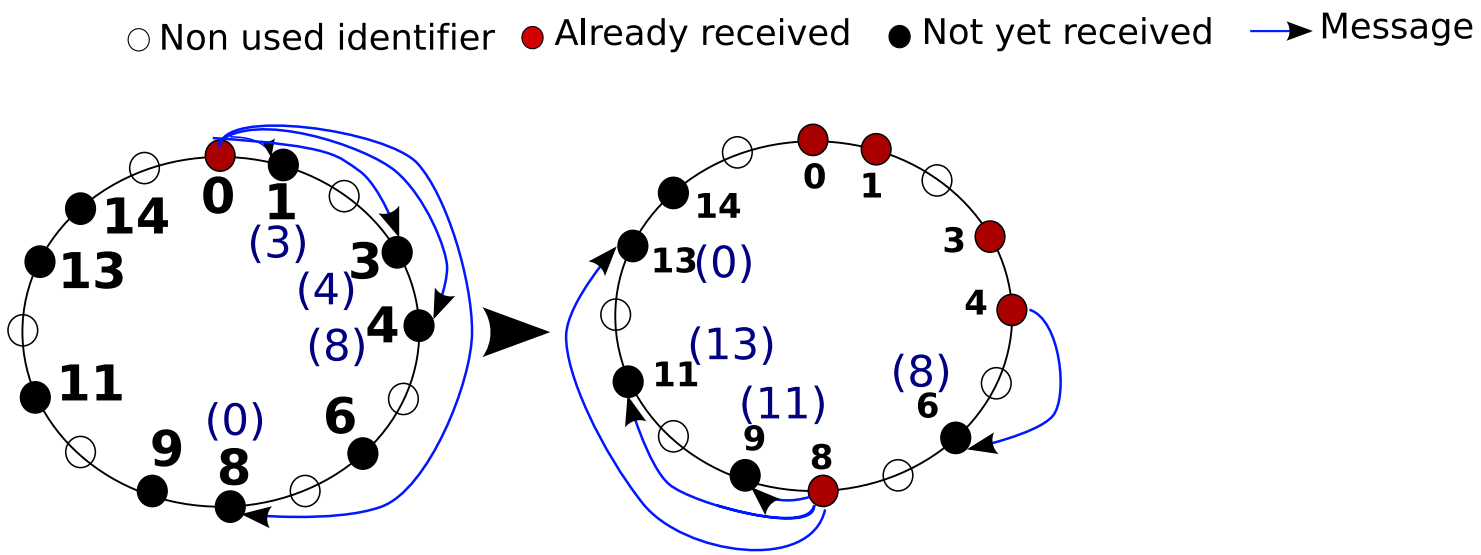
Chord model

- broadcast[2]:
  - forward the messages to each peers of the routing table
  - each peer has an exploration limit = min(the next peers in the routing table of the message sender, sender exploration limit)

[2]S. El-Ansary et-al, 'Efficient broadcast in structured p2p networks' IPTPS 03

Chord model

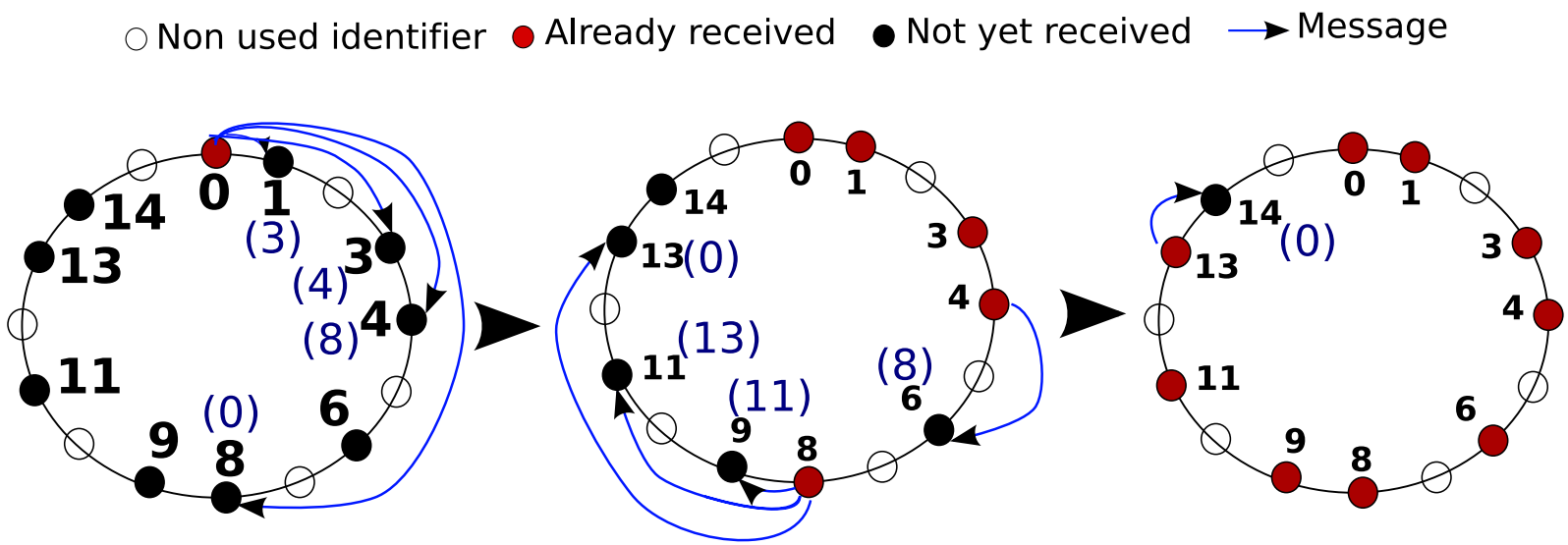
- broadcast[2]:
  - forward the messages to each peers of the routing table
  - each peer has an exploration limit = min(the next peers in the routing table of the message sender, sender exploration limit)



[2]S. El-Ansary et-al, 'Efficient broadcast in structured p2p networks' IPTPS 03

Chord model

- broadcast[2]:
  - forward the messages to each peers of the routing table
  - each peer has an exploration limit = min(the next peers in the routing table of the message sender, sender exploration limit)

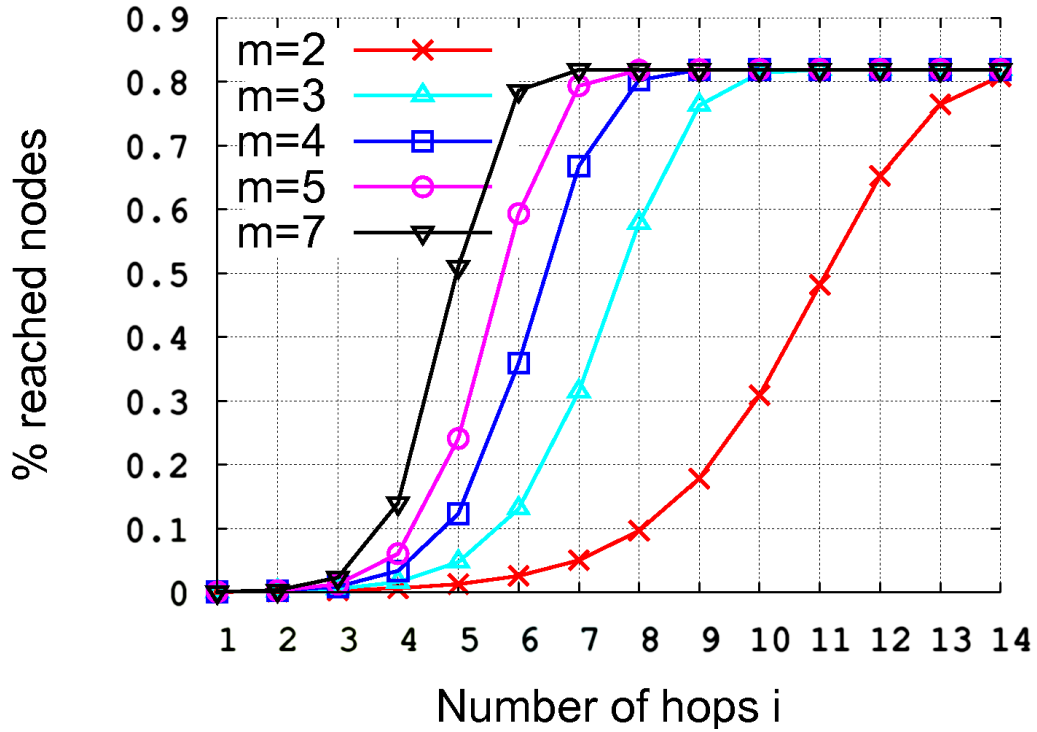


[2]S. El-Ansary et-al, 'Efficient broadcast in structured p2p networks' IPTPS 03

Outline

Slapper

- $N = 2000$ peers
- i varies from 1 to 14 hops

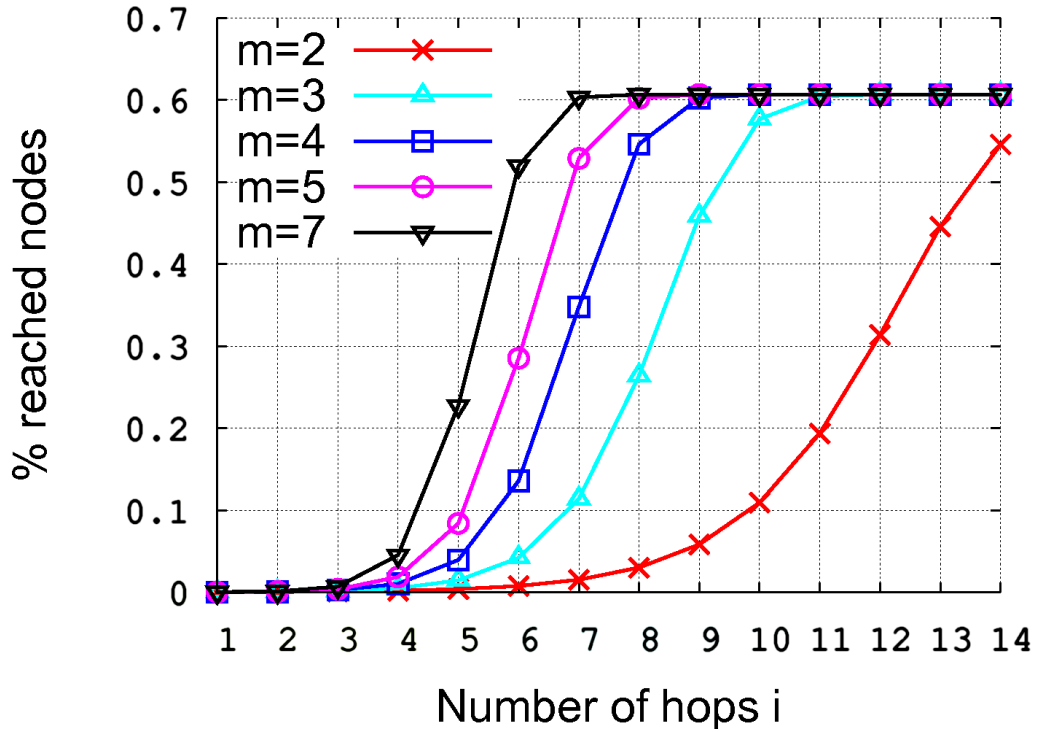


- maximal value = reach all peers except discovered peers
- $\longrightarrow$ limited by β (probability for each node to be compromised)

- higher branching factor $\longrightarrow$ higher reachability

Slapper

- $N = 5000$ peers
- i varies from 1 to 14 hops

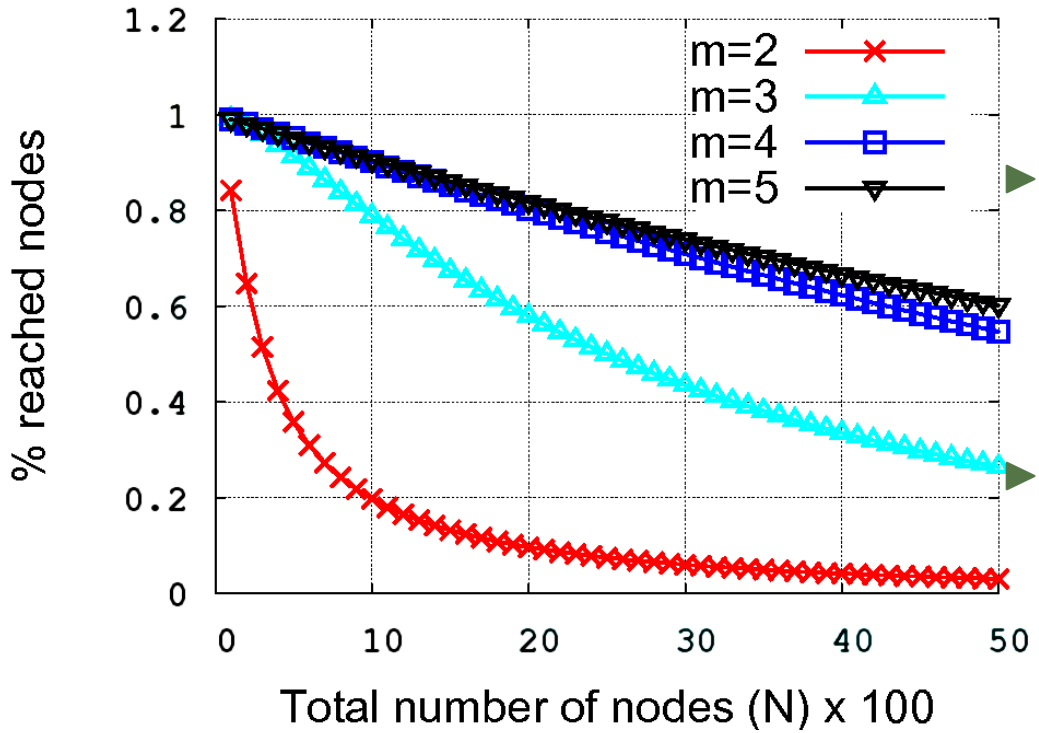


- compromised probability β has a higher impact when the number of peers increases
- N increases $\longrightarrow$ curves increase less at the begin and more at the end

- same number of hops to reach the maximal value

Slapper

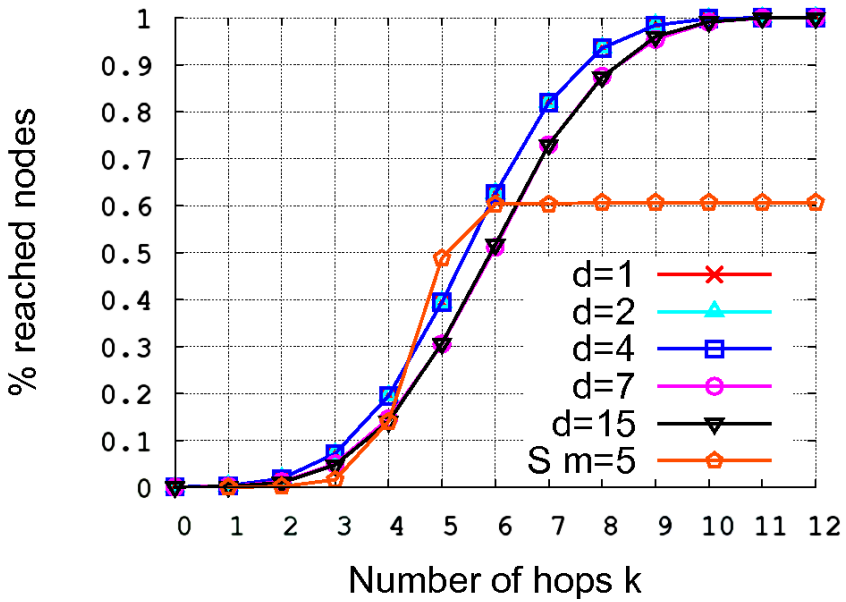
- number of hops $= 8$
- N varies from 100 to 5000



- curves converge to a fixed limit depending on β and N
- very bad performances for $m = 2$ (not suitable)
- high distance $\rightarrow$ no impact of the branching factor

Chord

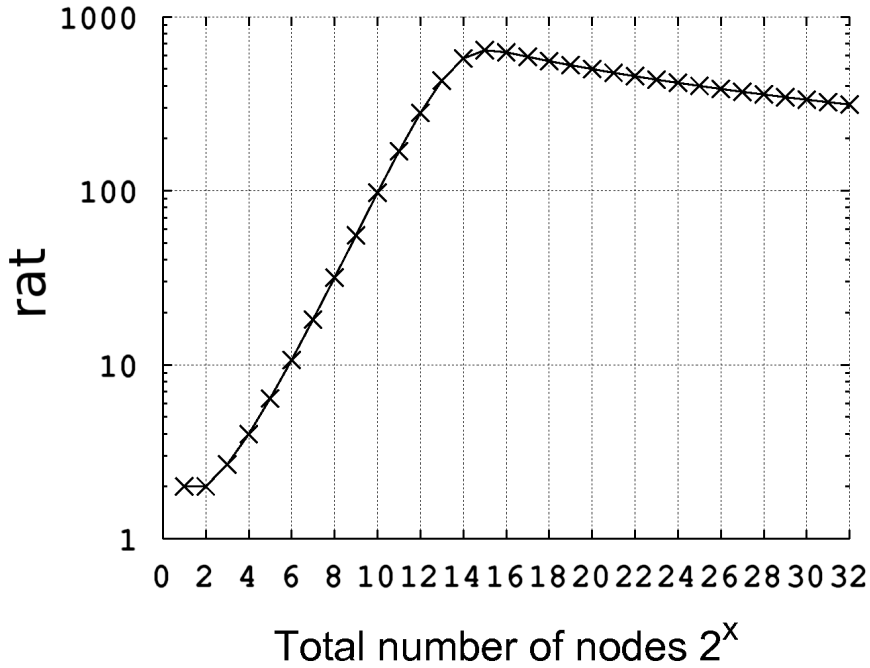
- number of hops varies from 1 to 13



- $N = 5000$ peers
- very close curve $\longrightarrow$ limited impact of the average distance between two node
- Slapper is about equivalent until a certain distance
- Chord $\longrightarrow$ all the peers can be reached
- Chord has a better reachability

Impact of attacks

- $rat(n) = \frac{\#discovered\_peers_{Slapper}}{\#discovered\_peers_{Chord}}$

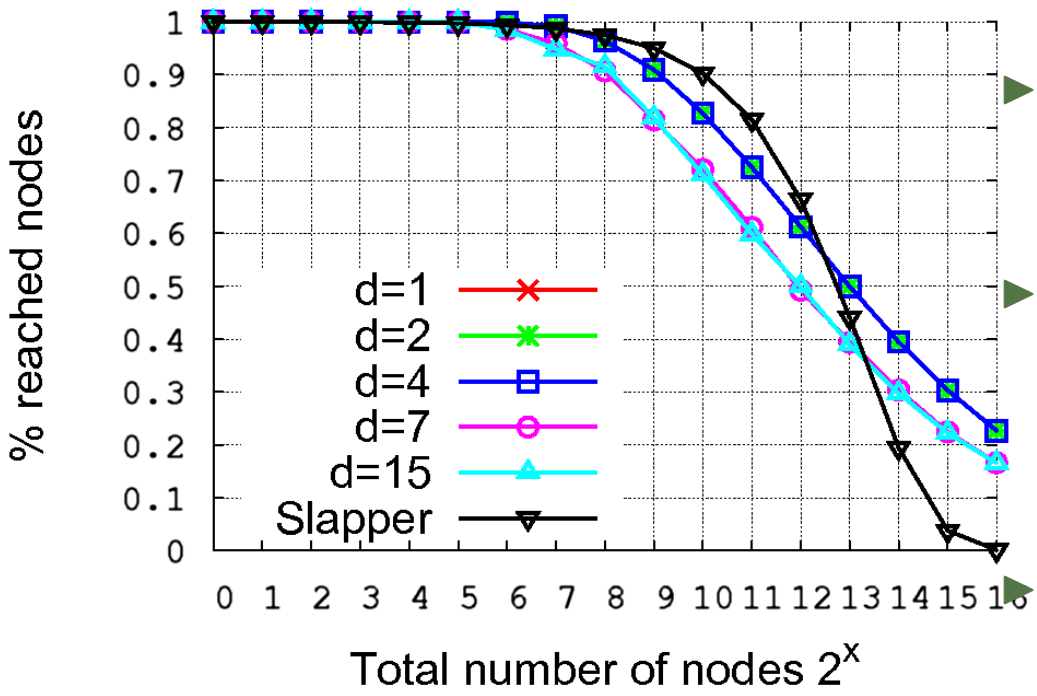


- independant from the distance d
- important benefit of Chord
- ratio decreases at the end

- ratio is still 20 for 2^{512} peers

Chord

- number of hops = 6
- N varies from 1 to 2^{16}



- Slapper: limitation by *beta* (best case)
- 6 hops = number of hops to have a reachability equivalent to Slapper
- increasing distance → better results for Chord

- Slapper is better between 2^{10} and 2^{12} peers
- Chord can be better from 2^{12} peers

Outline

What to choose ?

	IRC	Slapper	Chord
Efficiency	The lowest number of hops	The lowest delays	
Resiliency	very constrained (unavaibility, attacks)	very constrained by attacks, few connections	high resiliency, few connections, partial view
Scalability	$\#\text{devices} < 2^{12}$		$\#\text{devices} \geq 2^{12}$
Security	The manager can be tracked	Tracking the manager is very difficult (the intermediary nodes)	
Interest	Large and closed networks + central authority	Large networks of checked partners (research distributed honeypot)	Huge and public networks (honeypot where everyone can participate)

Questions ?

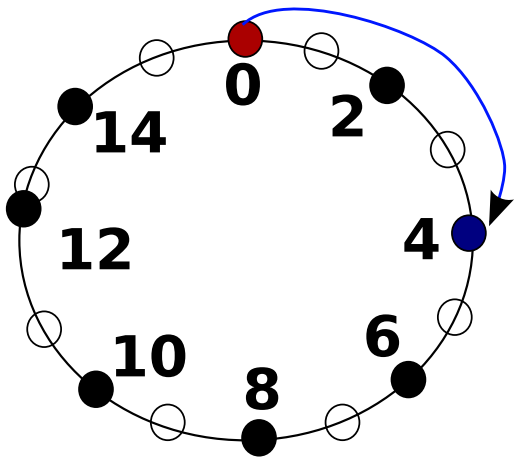
Slapper model

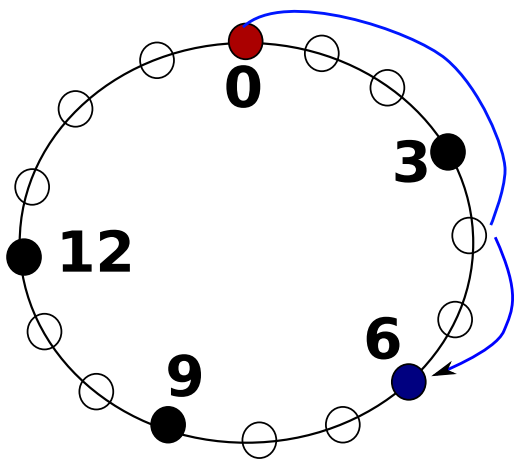
- assumptions:
  - $reach_{i-1}$ total number of reached peers at a maximal distance $i-1$
  - $p(t, c, j)$: probability to contact j not yet reached peers from already contacted c peers and with c messages to sent
- maximal number of messages sent at the *ith* hop :
  - 1st hop: m, 2nd hop: $m \times m \longrightarrow m^i$
  - limited by avability factor: $msg = (m \times \alpha(m))^i$
- maximum number of new reached peers at the *ith* hop: $max = min((m \times \alpha(m))^i, N - reach_{i-1})$
- average number of reached peers at an exact distance of $i = \sum_{k=0}^{max} p(reach_{i-1}, msg, k) \times k$

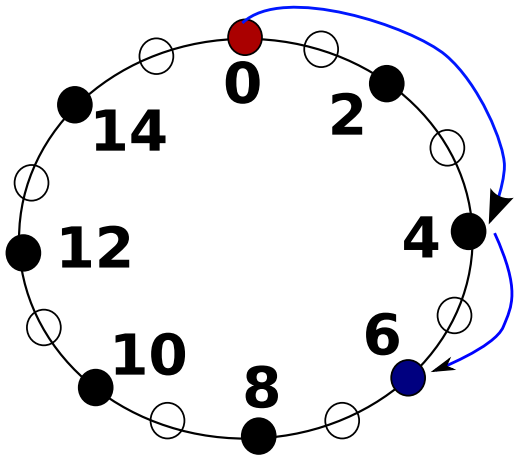
Chord model

- compute the number of hops to reach a peer p from the peer 0
  - $p = 2^k \rightarrow$ single hop
  - $p - d < 2^k \rightarrow$ no peers between p and $2^k \rightarrow$ single hop
  - else there is an intermediary peer $\rightarrow$, do the same process from this peer

○ Non used identifier







Deployment

Evaluation $\longrightarrow$ help an administrator to choose the right topology and to know the attended performances

	IRC	Slapper	Chord
Number of hops to have the best reachability	a fixed knowed value whatever the number of devices		A maximal value depending on the identifiers space size
Impact of an high branching factor	negative impact (m=5)	Positive impact	