Epidemic Information Dissemination

Lightweight Probabilistic Broadcast

Éireann Leverett

Advanced Computer Science MPhil
University of Cambridge

DataCentric Networking, 2011

Outline

Outline

Variables and Equations

Variables

- r is generation or 'round' of infection
- X is the number of infected individuals in a given round
- p_k is the probability of infecting$k$ (previously uninfected) individuals in a given round
- f is the mean number of infections per individual per round(Fan out)

Equations

- $P_{ext} = \sum_{k \geq 1} p_k (P_{ext})^k$is the probability of extinction of the epidemic (note recursive definition)
- $f = \sum_{k \geq 1} k p_k$ Note: $P_{ext} = 1$if $f \leq 1$ and $P_{ext} < 1$if $f > 1$

Variables and Equations

Variables

- r is generation or 'round' of infection
- X is the number of infected individuals in a given round
- p_k is the probability of infecting$k$ (previously uninfected) individuals in a given round
- f is the mean number of infections per individual per round(Fan out)

Equations

- $P_{ext} = \sum_{k \geq 1} p_k (P_{ext})^k$is the probability of extinction of the epidemic (note recursive definition)
- $f = \sum_{k \geq 1} k p_k$ Note: $P_{ext} = 1$if $f \leq 1$ and $P_{ext} < 1$if $f > 1$

Finite Population Models

Variables Redefined

- n is the population size, so
- X is now the number of infected individuals in the r^{th} round
- p_k is the the same
- t is now the number of rounds infection continues

Models

- Infect and Die ($t = 1$)
- Infect Forever ($t = \infty$)

Finite Population Models

Variables Redefined

- n is the population size, so
- X is now the number of infected individuals in the r^{th} round
- p_k is the the same
- t is now the number of rounds infection continues

Models

- Infect and Die $(t = 1)$
- Infect Forever $(t = \infty)$

Outline

How fast is the virus?

Values of interest

- Z_r is the number of individuals infected before a given round r
- $Y_r = Z_r/n$ is proportion of the population infected after a given round r

Rate of infection grows exponentially

- Infect Forever $(t = \infty)$
- $Y_r \approx \frac{1}{1+ne^{-fr}}$
- Growth is e^f on average!

How fast is the virus?

Values of interest

- Z_r is the number of individuals infected before a given round r
- $Y_r = Z_r/n$ is proportion of the population infected after a given round r

Rate of infection grows exponentially

- Infect Forever $(t = \infty)$
- $Y_r \approx \frac{1}{1+ne^{-fr}}$
- Growth is e^f on average!

Latency of Infection

This is why we're interested

- In the infect forever model we need to wait some number of rounds before our subscription event

Logarithmic wait!

- $R = \frac{log(n)}{log(log(n))} + O(1)$

Latency of Infection

This is why we're interested

- In the infect forever model we need to wait some number of rounds before our subscription event

Logarithmic wait!

- $R = \frac{log(n)}{log(log(n))} + O(1)$

What Information is Neccesary

- Membership-How processes get to know each other
- Network Awareness-How to make connections
- Buffer Management-Which information to drop at what time (dynamic t)
- Message Filtering-Pub/Sub essentially

Varying t

So if infect forever is overkill, and infect and die is too slow, then how do we purge our buffers?

- Age based prioritisation (number of transmissions, implemented with transmission counter)
- Application defined obsolescence (it's up to you!)

Outline

Scalability over Reliability

- Give up certainty that a message will arrive.
- In return you get high probability that it will, with only a few 'rounds' of latency.
- Since this requires no centralisation, it is highly scalable.
- Ok, but what about barriers in networks?

Outline

Small World

- How far is 192.0.0.1/31?
- Barriers and bifurcations?
- What does a firewall do in a gossiping architecture?
- Infect nearby, but occasionally across the world (smallworld).
- 6 degrees of seperation you say?
- 6-7 rounds to infect 1000 processes!

Subscription and Unsubscription

Example

Subscription

- Piggyback on gossip messages
- Different subscription buffer called 'subs'
- Used to update another buffer called 'view'
- Note you gossip about yourself to subscribe

Example

Unsubscriptions

- Also piggybacks on gossip messages
- Allows removal from buffer 'subs'
- Unsub, then sub, then gossip

Subscription and Unsubscription

Example

Subscription

- Piggyback on gossip messages
- Different subscription buffer called 'subs'
- Used to update another buffer called 'view'
- Note you gossip about yourself to subscribe

Example

Unsubscriptions

- Also piggybacks on gossip messages
- Allows removal from buffer 'subs'
- Unsub, then sub, then gossip

Summary

- Epidemic Routing is fast and decentralised, assuming some delay tolerance.
- Useful for database replication, or crowdstore applications.
- Imagine a wildernet, where phones in Antarctica carry news to people who are offline.

- Future research
  - I'd like to see more research on 'message size' with respect to other parameters.
  - Epidemic routing with mobility needs to be studied in context of the rate of mobility/contact.
  - Try brainstorming yourself, there's plenty of applications.

Questions

???