A Next Generation Network for the Networked Generation?

Roch Guerin
University of Pennsylvania

NGI 2009 – Aveiro, Portugal

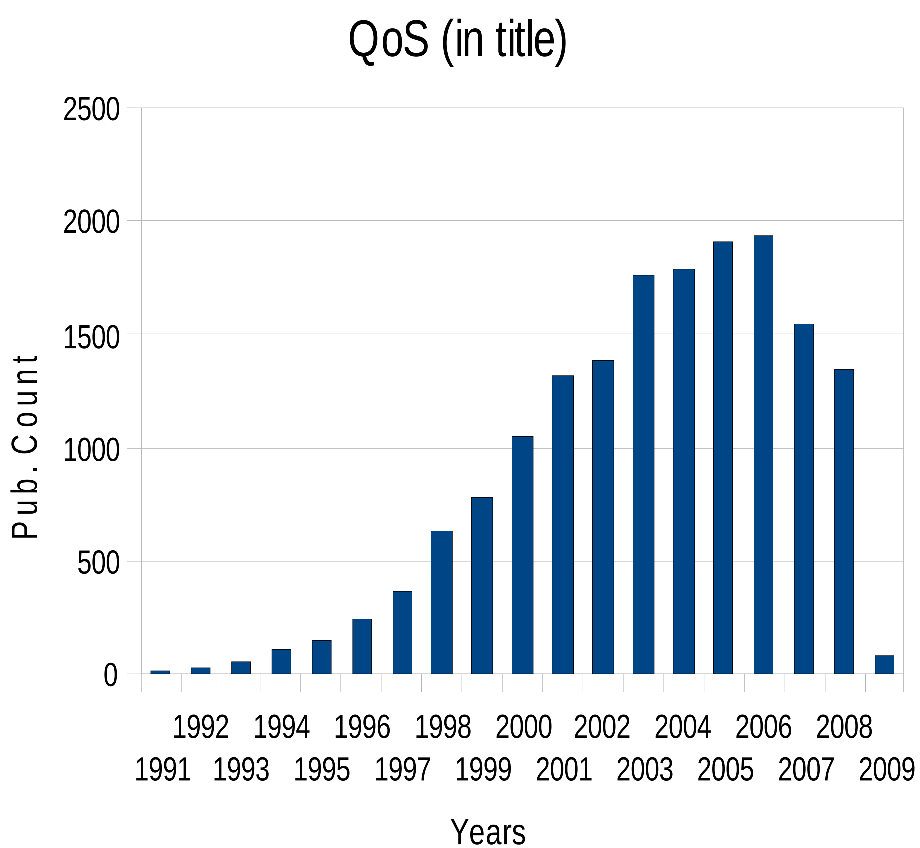
What I'm Going to Try to Convey

The challenges of network innovation

1. You need to make sure you need it
   - The Internet is on its last leg
     - This is not the first time and probably not the last
   - What is today's Internet preventing us to do?
     - A still healthy growth curve by all accounts
2. Once you have it, you need to make sure users adopt it
   - We've had a new architecture for 15 years and it's barely starting to take-off
   - Do we really understand what drives network migration?
3. Once the network has been built, many of the important problems are in using it, not building a new one
   - This may be where the real (and interesting) problems are

Internet Failure Predictions – (1)

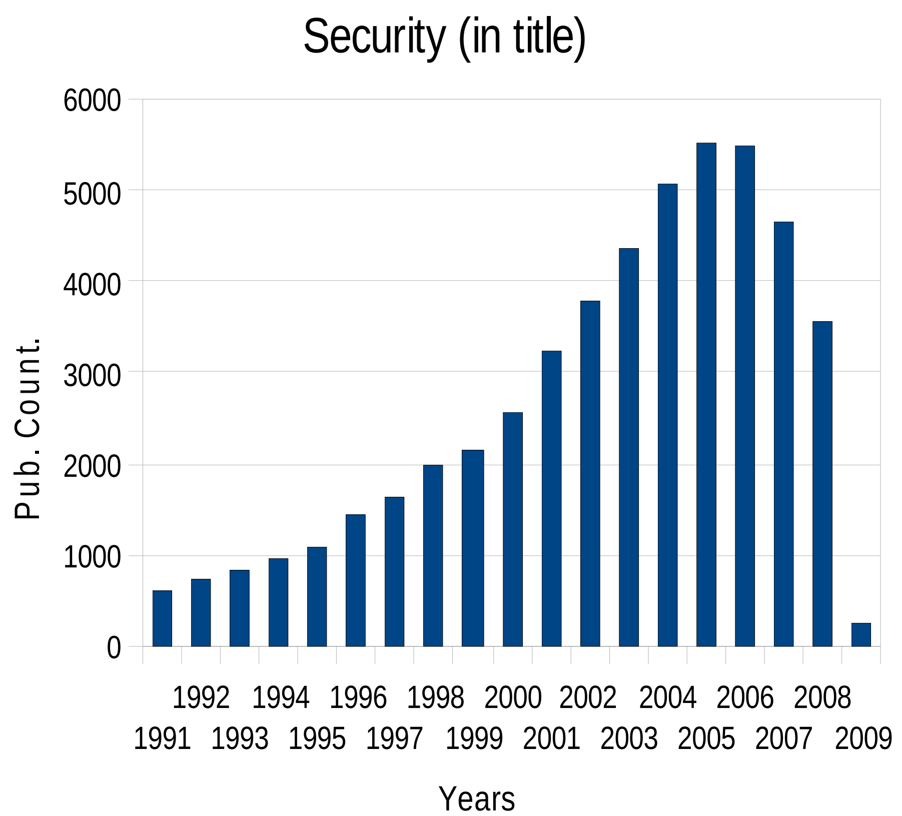
- Best effort can surely not be good enough!
  - The "race" for QoS
    - From Int-Serv, to Diff-Serv, to …
  - A phenomenal expenditure of intellectual resources
    - We've solved pretty much every QoS problem there was to solve
  - And… no one is really using the answers



From http://scholar.google.com/

Internet Failure Predictions – (2)

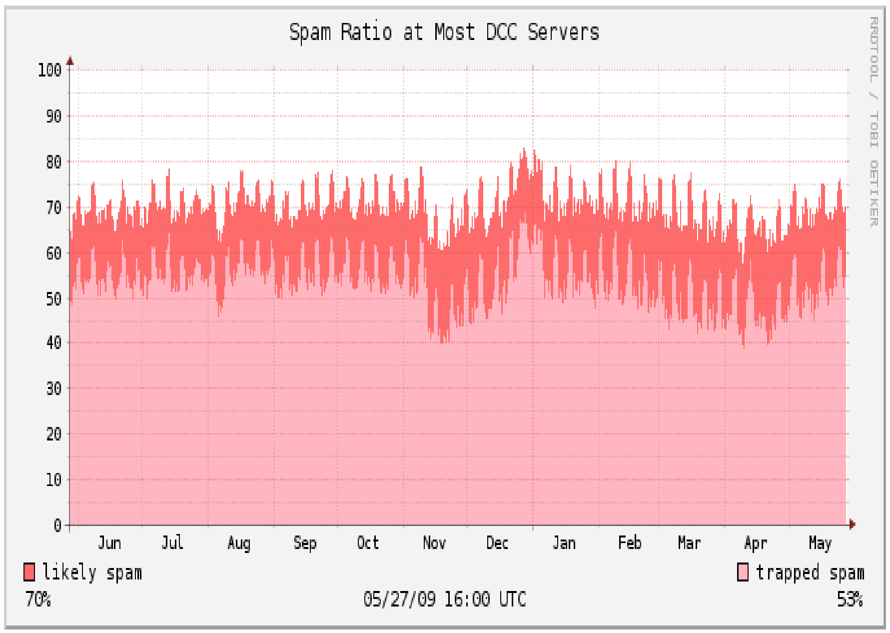
- An open network with distributed control can surely not be secure enough!
  - From BGP to S-BGP?
  - And lets not forget IPSEC, SPAM filters, honeypots, DDoS prevention, etc.
- It's not a perfect world, but things seem headed the "right" way

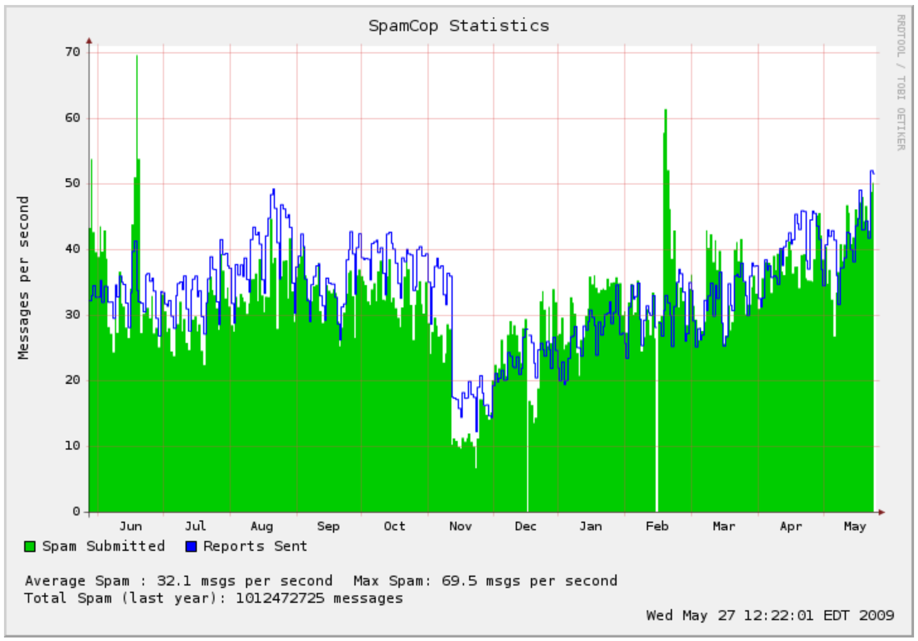


From http://scholar.google.com/

Spam – From Crisis to Boring Pain

This does not really look like exponential growth…



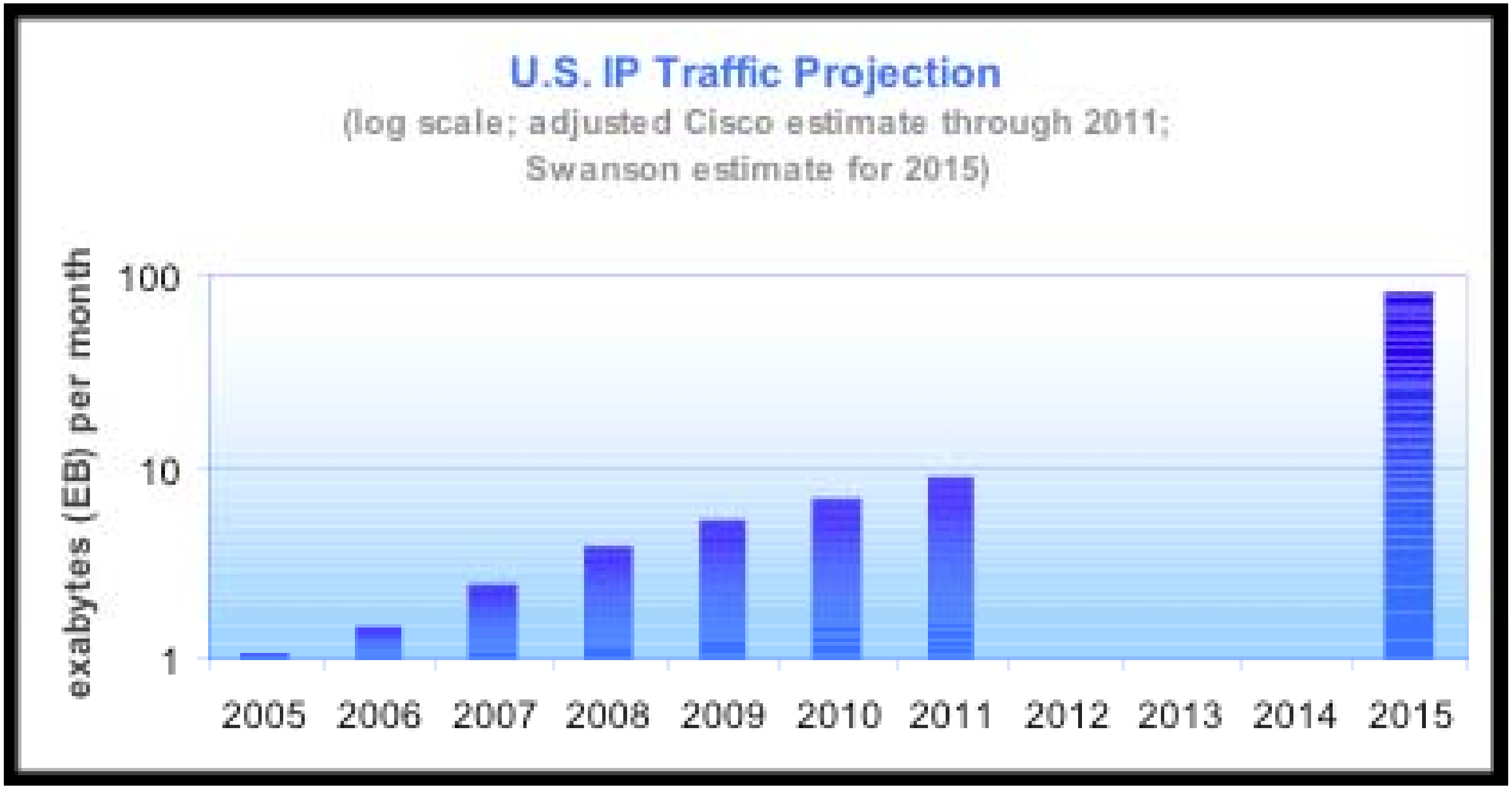


From http://www.dcc-servers.net/dcc/graphs/

From http://www.spamcop.net/spamstats.shtml

On the Flip Side

Roughly a steady 50% annual growth rate for Internet traffic



From http://www.discovery.org/a/4428

See also http://www.dtc.umn.edu/mints/home.php for additional growth info

More on Internet Traffic Growth

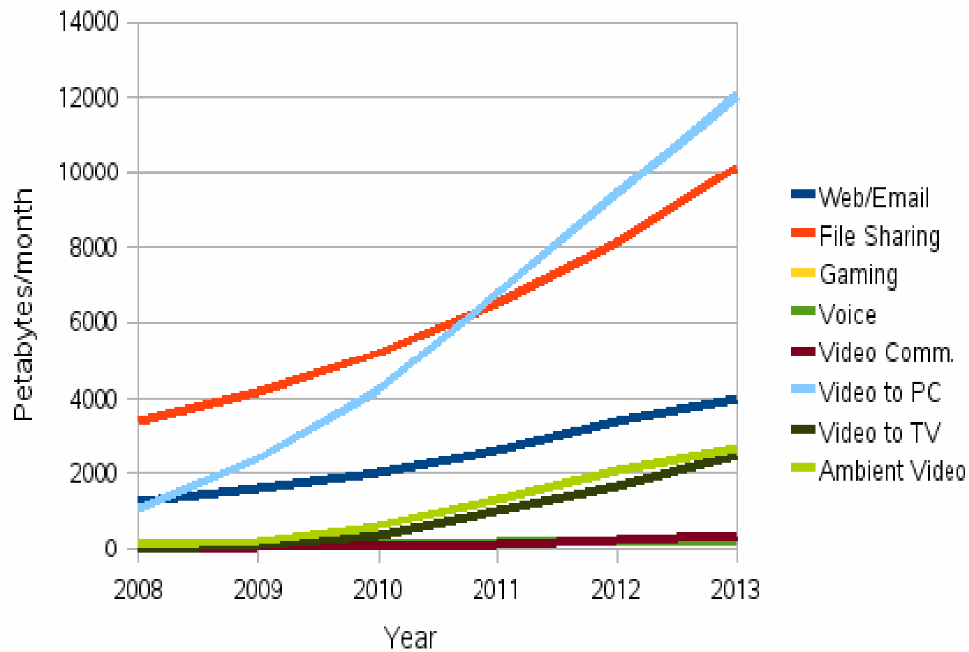
The Past ~ 75-100+% AGR

From http://www.dtc.umn.edu/mints

Year	AGR
2003	75%
2004	75%
2005	115%
2006	149%
2007	58%
2008	61%

The Future ~ 50% AGR

From [Cisco Visual Networking Index]()



But, video is growing at ~100% AGR and expected to represent 90% of Internet traffic in 2013!

Tracking Hulu

(http://www.yourbrandplan.com/forum/technology-innovation/11393-techcrunch-hulu-still-going-strong-but-growth-dropping-off-sharply.html)

Month	comScore Video Metrix (000s)				Nielsen VideoCensus (000s)			
	Ranking*	Overall Streams	Unique Viewers	Minutes Per Viewer	Ranking*	Overall Streams	Unique Viewers	Time Spent Viewing
Apr '08	N/A	N/A	N/A	N/A	10	63,228	2,428	N/A
May '08	10	88,284	6,765	50.7	9	80,045	2,724	N/A
Jun '08	9	95,093	8,277	47.7	10	83,838	2,611	N/A
Jul '08	8	119,357	9,779	50.0	8	105,830	3,293	N/A
Aug '08	9	122,124	10,201	52.6	8	107,622	2,632	N/A
Sep '08	7	145,815	12,535	54.8	6	142,261	6,323	114.7
Oct '08	6	235,096	23,993	113.8	3	206,068	9,069	119.2
Nov '08	6	226,540	22,456	119.7	3	220,536	7,509	147.4
Dec '08	6	240,585	24,572	99.0	4	216,344	6,679	177.8
Jan '09	6	250,473	24,448	79.0	3	232,444	7,238	187.6
Feb '09	4	332,504	34,731	64.5	2	308,806	9,473	176.9
Mar '09	3	380,102	41,564	57.9	2	348,520	8,865	260.0
Apr '09	**3**	**396,953**	**40,110**	**61.0**	**2**	**373,290**	**7,458**	**324.9**

The Health of Internet Innovation (It's video but not just video)

	# Videos (000)	# Viewers/mo (000)
Google sites	6,367,638	101,870
Fox Interactive	551,991	62,109
Yahoo! Sites	374,161	41,859
Viacom digital	287,615	24,126
Microsoft sites	267,475	30,042
Hulu	250,473	24,448
Turner Netw.	195,983	22,979
AOL	184,808	27,198
Disney Online	141,452	13,435

	Monthly users	Growth rates	Statistics
Facebook	54.5M	85%	150M users
Twitter	4.4M	752%	>1B tweets
Linked-in	11.9M	153%	30.1M users
YouTube	71.3M	21%	258M users 100M videos/day
Flickr	27.5M	29%	>3B images
Digg	33.4M	91%	
Wikepedia	59.6M	15%	>10M articles
Blogs	30M	68%	364M readers

http://www.comscore.com/Press_Events/Press_Releases/2009/3/YouTube_Surpasses_100_Million_US_Viewers

http://www.quantcast.com/youtube.com
http://nmlab.com/download/1/

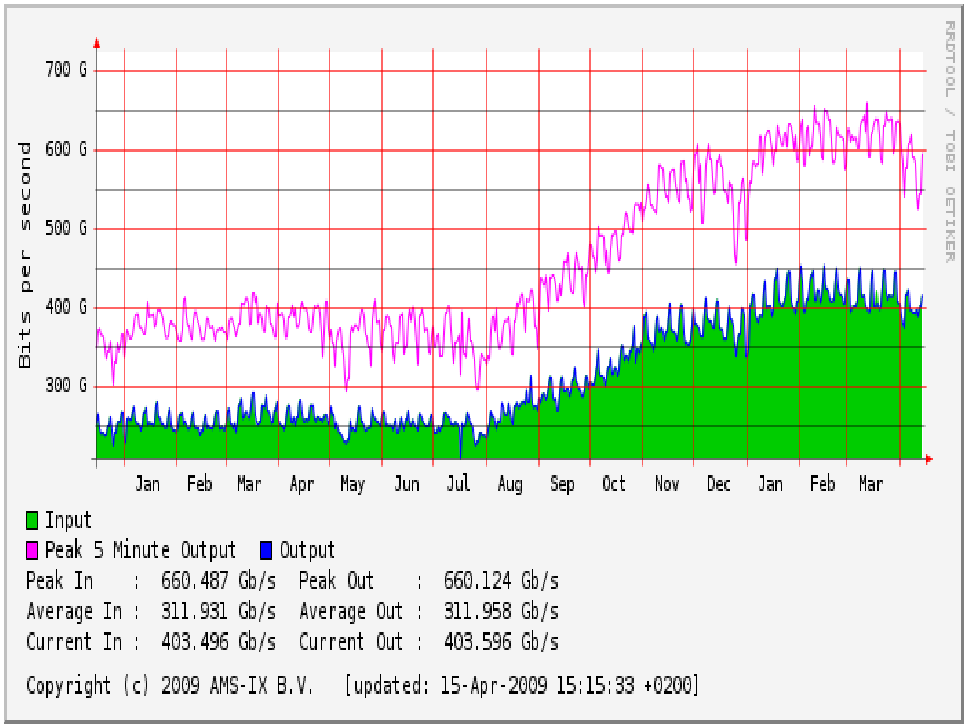
Taking Stock

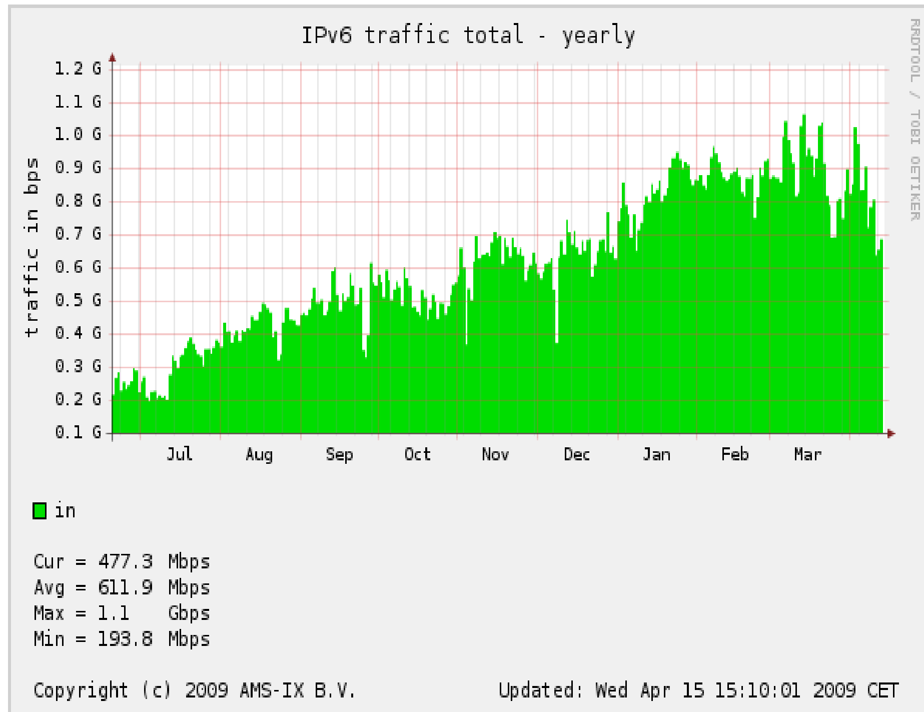
- There have been many past predictions of the Internet's demise
  - So far, they have been just that
- Today's Internet is still growing strong
  - About 50% AGR after over 20 years of ~100% AGR!
  - And some argue that video will give it a new boost
- It does not appear to be stifling innovation
  - A steady stream of new applications and uses
  - Solid growth across the board for existing apps and uses

We May Still Need a New Network
(some day)

- But, we've have had a new network for 15 years
  - It's called IPv6
  - It fixes a number of things with IPv4, though not everything
- But being better is not enough
  - Especially when dealing with a large incumbent
- We are starting to see some changes
  - Motivated by the emergence of a ***real*** problem and limitation of IPv4
  - But even now it's not obvious if/when IPv6 will really emerge

IPv4 & IPv6 Yearly AMS-IX Traffic



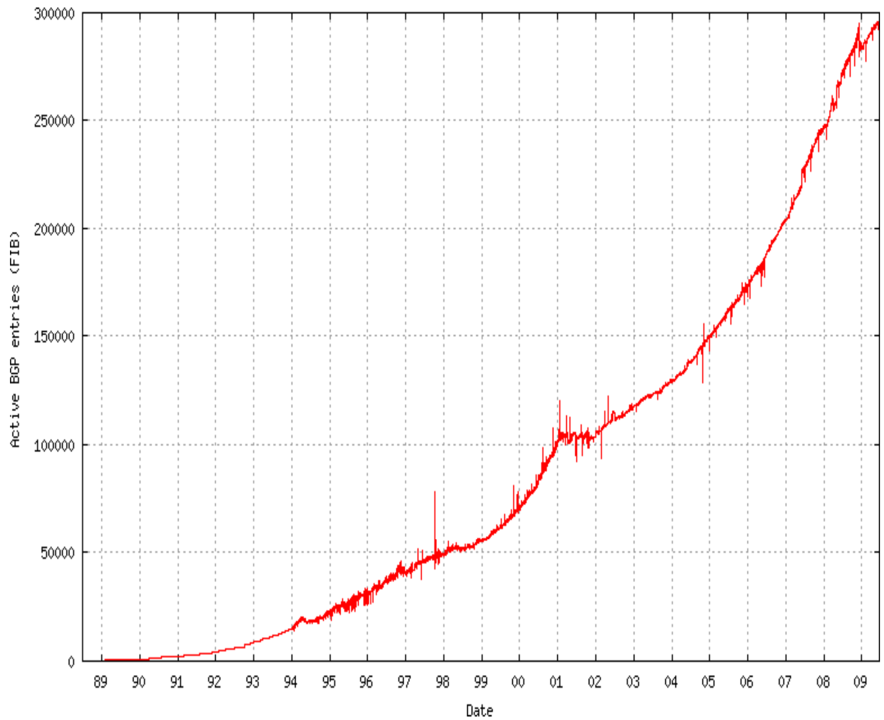


- After ~15 years since being standardized, IPv6 traffic amounts to about 0.2% of IPv4 traffic…

Source: AMS-IX web site - http://www.ams-ix.net/

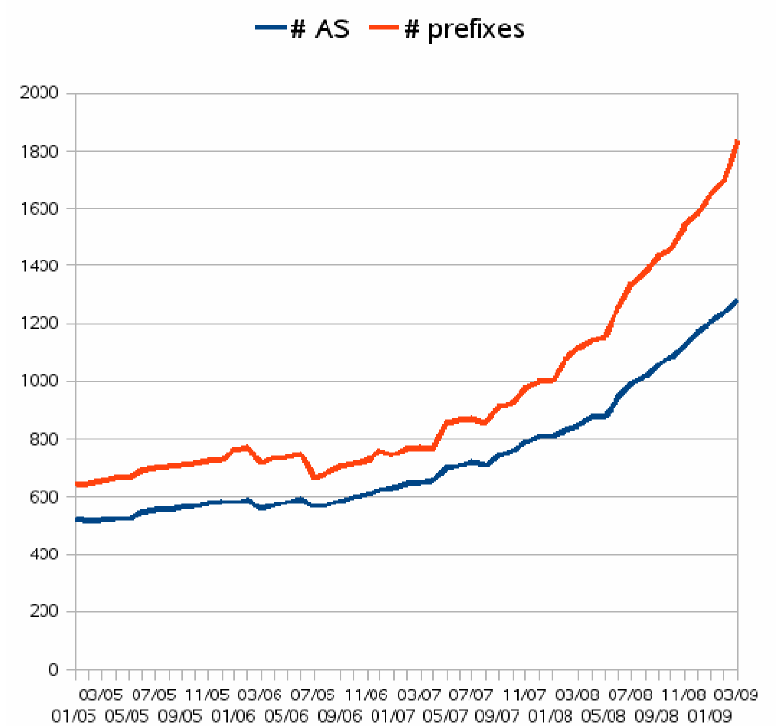
Another Look at IPv4 & IPv6 Growth
(routing)

IPv4



http://bgp.potaroo.net

IPv6



http://www.ipv6actnow.org/info/statistics/

The Challenges of Network Migration

- Lets assume that some time in the distant future
  - We have created a much better network architecture that allows us to do things we simply cannot do on today's Internet
- The Internet will be pretty big by then
- What will it take for the new network to successfully displace the current Internet technology?

An Attempt at a Simple Model*

- Two competing and incompatible networks, e.g., IPv4 and IPv6
  - Different qualities and price
  - Different installed base, e.g., one is starting from scratch
- Users individually (dis)adopt whichever technology gives them the highest positive ***utility***
  - Depends on technology intrinsic ***value*** and ***price***
  - Depends also on the number of users of each technology (***externality***)
- Gateways can offer a migration path
  - Overcome chicken-and-egg problem of first users
  - Effectiveness depends on gateways characteristics/performance
    - Duplex vs. simplex (independent in each direction or coupled)
    - Asymmetric vs. symmetric (performance/functionality wise)
    - Constrained vs. unconstrained (performance/functionality wise)

* http://repository.upenn.edu/ese_papers/496

A Basic User Model

Technology 1: $U_1(\theta, x_1, x_2) = \theta q_1 + (x_1 + \alpha_1 \beta x_2) - p_1$

Technology 2: $U_2(\theta, x_1, x_2) = \theta q_2 + (\beta x_2 + \alpha_2 x_1) - p_2$

- Users evaluate the relative benefits of each technology
  - Intrinsic value of the technology (θq_i)
    - Tech. 2 better than tech. 1 ($q_2 > q_1$)
    - θ denotes user valuation of technology (captures heterogeneity)
  - Externalities: linear in # users ($0 \leq x_1 + x_2 \leq 1$) – Metcalfe's law
    - Possibly different across technologies ($\beta \neq 1$)
    - α_i, $0 \leq \alpha_i \leq 1$, $i = 1,2$, captures gateways' performance
  - Cost (recurrent) for each technology (p_i)

How Do Users Decide?

- Decision based on *indifference points/thresholds* for each technology: $\theta_1^0(\underline{x})$, $\theta_2^0(\underline{x})$, $\theta_2^1(\underline{x})$
  - $U_1(\theta, \underline{x}) > 0$ if $\theta \geq \theta_1^0(\underline{x})$ - Tech. 1 becomes attractive
  - $U_2(\theta, \underline{x}) > 0$ if $\theta \geq \theta_2^0(\underline{x})$ - Tech. 2 becomes attractive
  - $U_2(\theta, \underline{x}) > U_1(\theta, \underline{x})$ if $\theta \geq \theta_2^1(x)$ - Tech. 2 over Tech. 1
- Users are "rational" and choose:
  - Neither technology if $U_1 < 0$, $U_2 < 0$
  - Technology 1 if $U_1 > 0$, $U_1 > U_2$
  - Technology 2 if $U_2 > 0$, $U_1 < U_2$
- Decisions change as $\underline{x}$ evolves
  - Can formulate a diffusion model to capture evolution of decisions
  - Solving the model identifies possible equilibria and trajectories

Two Possible Examples

1. IPv4 $\leftrightarrow$ IPv6
   – Duplex, asymmetric, constrained gateways

2. Low def. video conf. $\leftrightarrow$ High def. video conf.
   – Simplex, asymmetric, unconstrained converters

IPv4 (Tech. 1) $\leftrightarrow$ IPv6 (Tech. 2)

IPv4: $U_1(\theta, x_1, x_2) = \theta q_1 + (x_1 + \alpha_1 \beta x_2) - p_1$

IPv6: $U_2(\theta, x_1, x_2) = \theta q_2 + (\beta x_2 + \alpha_2 x_1) - p_2$

- IPv4 and IPv6 are similar as "technologies" ($q_1 \approx q_2$ and $\beta = 1$)
- As IPv4 addresses become scarce
  - Providers start assigning IPv6 addresses to new subscribers ($p_{\text{IPv4}} = p_1 > p_2 = p_{\text{IPv6}}$)
- IPv6<->IPv4 gateways for transition to happen
  - Most content is ***not*** yet available on IPv6
    - Little in way of incentives for content providers to do it
  - Duplex, asymmetric, constrained converters
- Users choose technology primarily as a function of
  - Price (p_{IPv4} vs. p_{IPv6}) and accessible content (x_1 vs. x_2)

Low-def. video ↔ High-def. video

Low-def: $U_1(\theta, x_1, x_2) = \theta q_1 + (x_1 + \alpha_1 \beta x_2) - p_1$

High-def: $U_2(\theta, x_1, x_2) = \theta q_2 + (\beta x_2 + \alpha_2 x_1) - p_2$

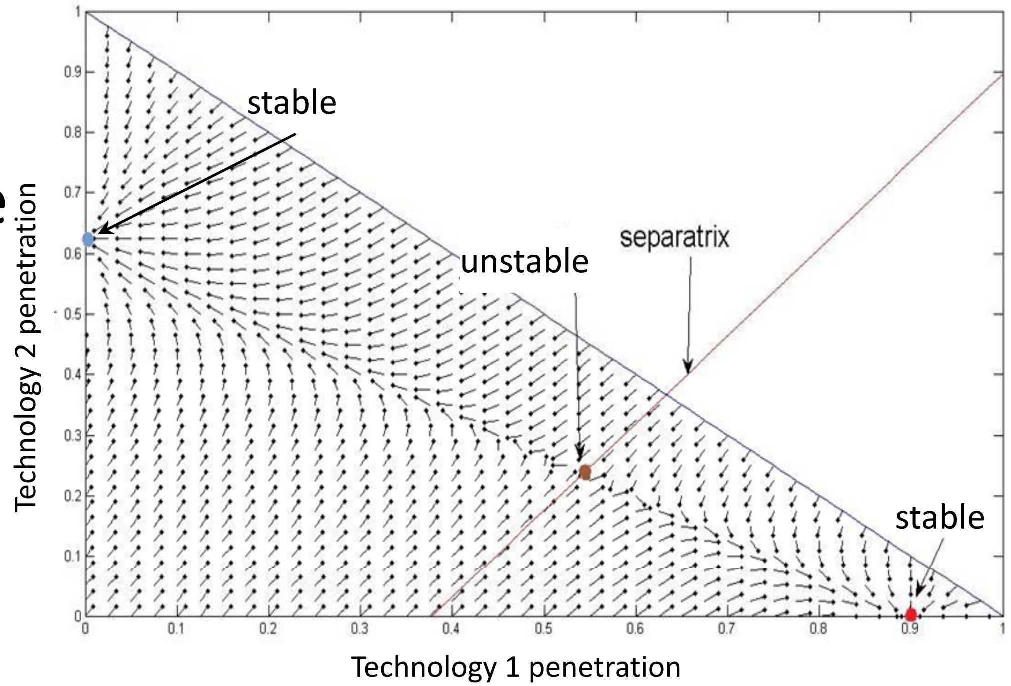
- Two video-conf service offerings: Low-def & High-def
  - Low-def has lower price ($p_1 < p_2$), but lower quality ($q_1 < q_2$)
  - Video is an *asymmetric* technology
    - Encoding is hard, decoding is easy
      - Low-def subscribers could ***display*** high-def signals but not generate them
    - Externality benefits of High-def are higher than those of Low-def ($\beta > 1$)
- Converters characteristics
  - High/Low-def user can decode Low/High-def video signal
  - Simplex, asymmetric, unconstrained
- Users choose technology as a function of
  - Price vs. quality trade-off
  - The level of externality benefits they can enjoy

What Do We Learn from the Model?

- What are possible outcomes?
  - Combinations of equilibria
- What trajectories to equilibria?
  - Monotonic vs. chaotic
- What roles for gateways?
  - Do they help and how much?

A “Typical” Outcome

- At most two stable equilibria
- Coexistence is possible
- Final outcome is hard to predict simply from observing the evolution of adoption decisions



Applying the Model to IPv4→IPv6

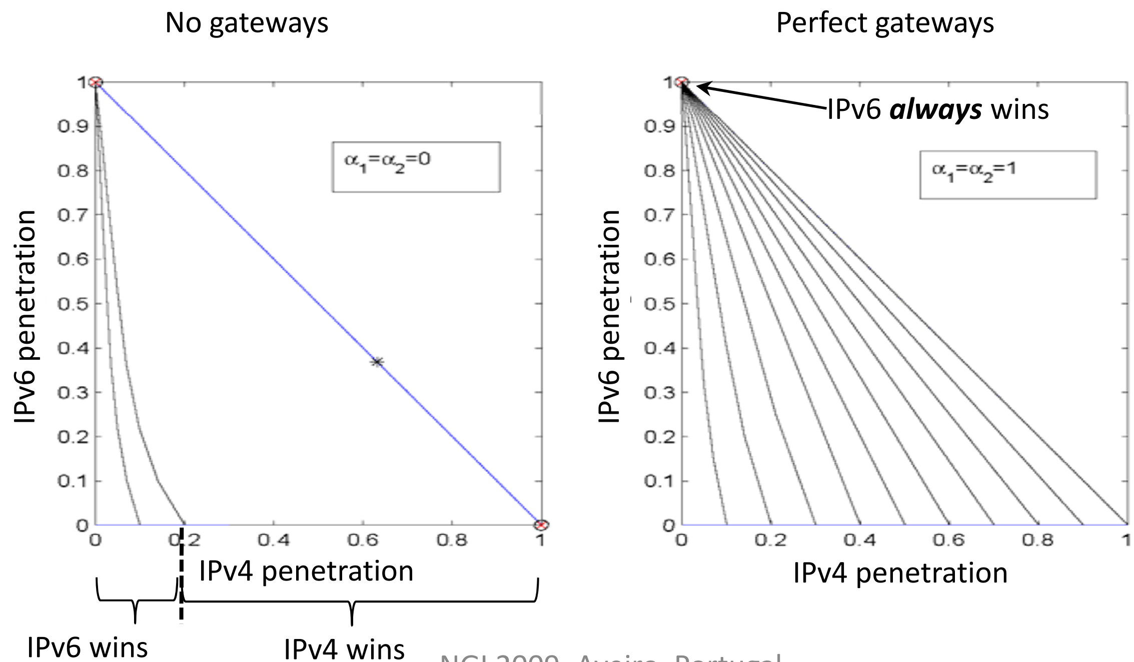
- Two possible scenarios (nothing surprising in either)

  1. IPv4 slightly "better" than IPv6
     - Greater user familiarity with technology
  2. IPv6 slightly "better" than IPv4
     - More addresses, better security and/or mobility

- Both yield similar behaviors and highlight the role of gateways

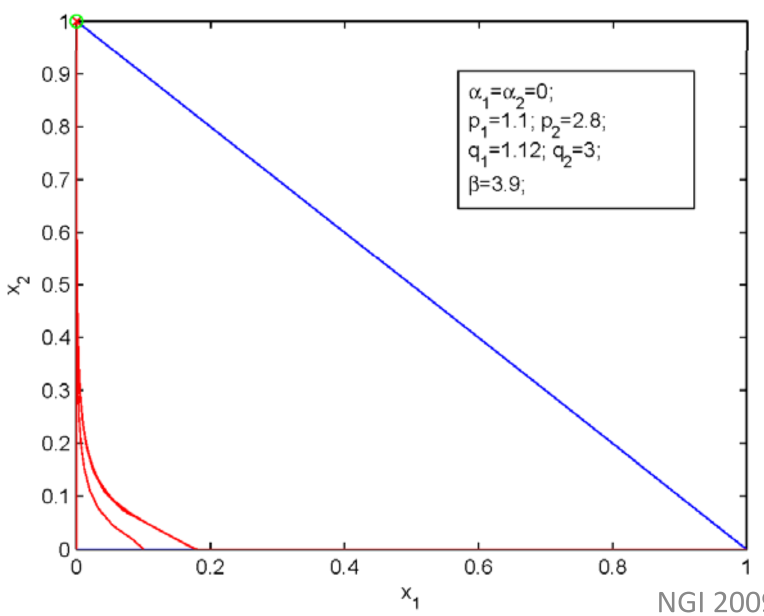
IPv6 “Better” than IPv4

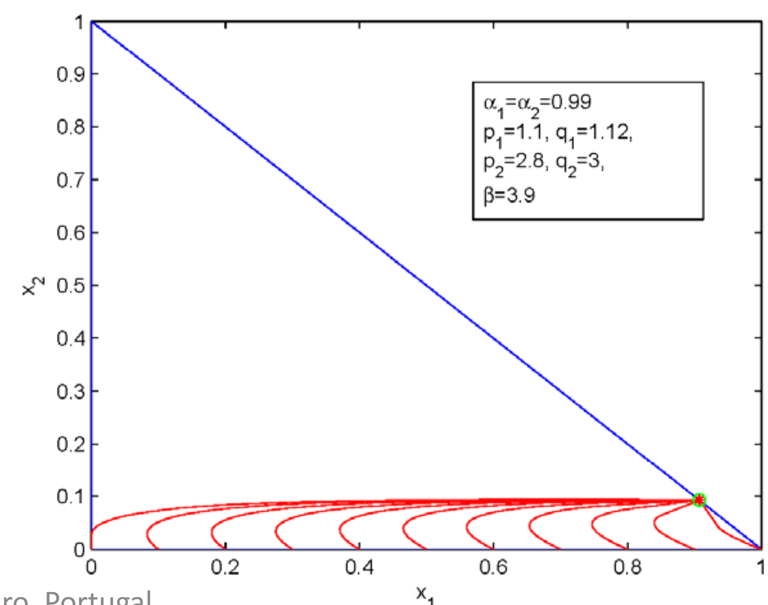
- Without gateways, IPv6 never takes off if it starts late
- With “perfect” gateways, IPv6 always eventually win
  - But gateways must be better than a minimum threshold
- This is an instance where gateways help defeat the incumbent



Gateways Can Also Help the Incumbent

- No gateways: Tech. 2 wipes out Tech. 1
- Perfect gateways: Tech. 1 nearly wipes out Tech. 2



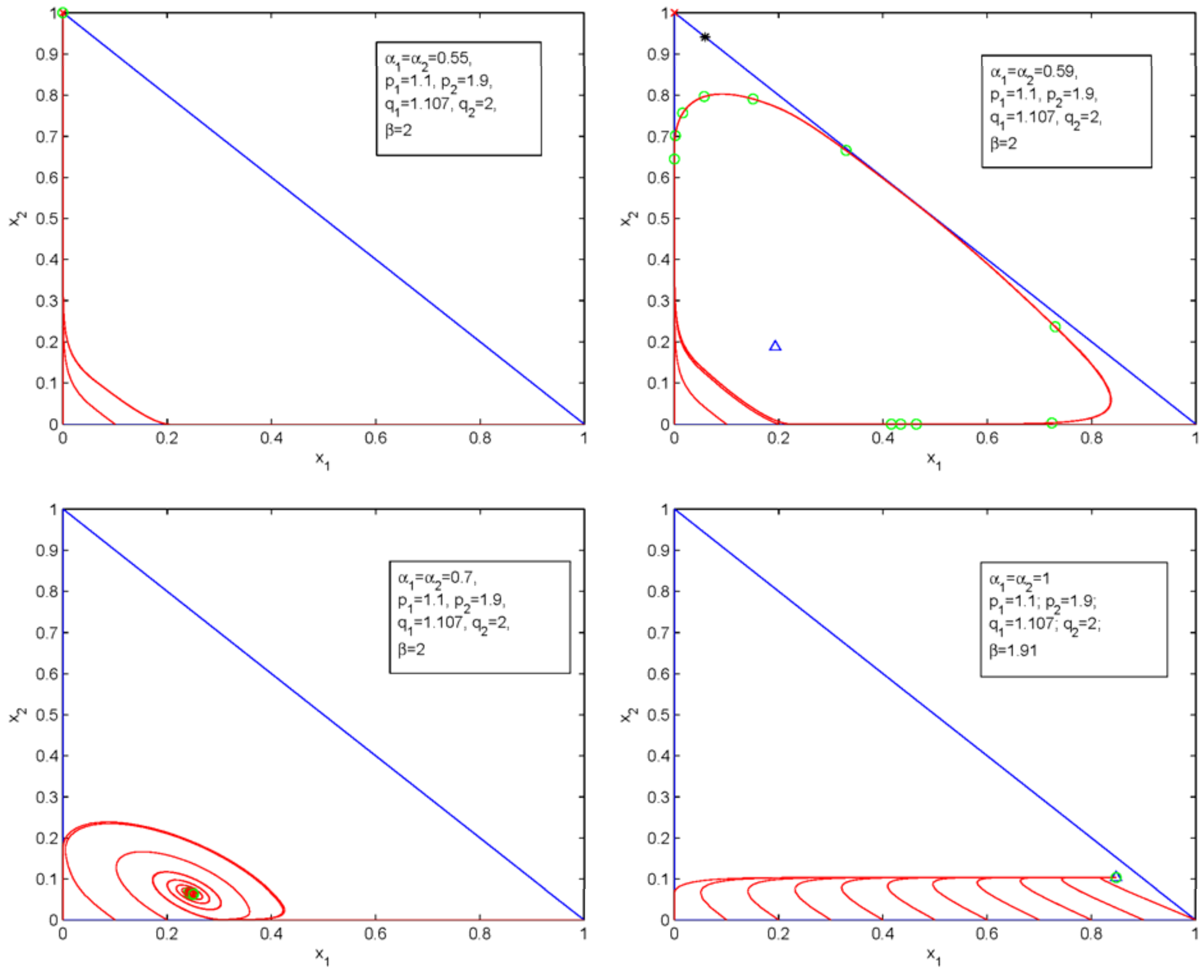


More Bad Gateway Behaviors

- Better gateways can harm overall market penetration
- Gateways can also render the adoption process unstable
  - Perpetual cycles of adoption/disadoption
  - This only happens when the new technology is significantly better, and users of the incumbent can tap into those benefits through gateways (the video example)

When Things Go Really Wrong

- No gateways: Tech. 2 captures full market
- Low efficiency gateways: No stable outcome
- Medium efficiency gateways: Pitiful overall market penetration
- High efficiency gateways: Tech. 1 dominates at close to full market penetration



How Serious is This?

- Most/all results are actually robust to a wide range of modeling changes (not just a modeling artifact)

  - User preferences (θ)
    - Arbitrary distributions
    - Extended to externality benefits
  - Externality effect
    - Sub-linear: x^{α}, $0<\alpha<1$
    - Super-linear: x^{α}, $\alpha>1$
    - Logarithmic: $\log(x+1)$

The Net of It

Caution is in order when

- Deploying a new network technology with strong externality effects, an entrenched incumbent, and
- Deciding how good a gateway to build

If you build it, they may not come…

If Building New Networks Is Dicey, What Else Can We Do?

- There are lots of interesting problems that arise when everything is ***networked***
- Broadly speaking, this is what people have recently been calling *NETWORK SCIENCE*

- It's an abused term that nevertheless spans some really interesting areas

One Out Of Many Examples

- Consider a networked system, e.g., a social network à la Facebook
- We want to deploy a new application/feature
  - Its value to users depends on how many others are using it (another instance of externalities)

$$U_i(a_i, \mathcal{A}) = a_i \left[\sum_{j \in \mathcal{A} \setminus \{i\}} w_{ij} a_j + w_{ii} - p \right]$$

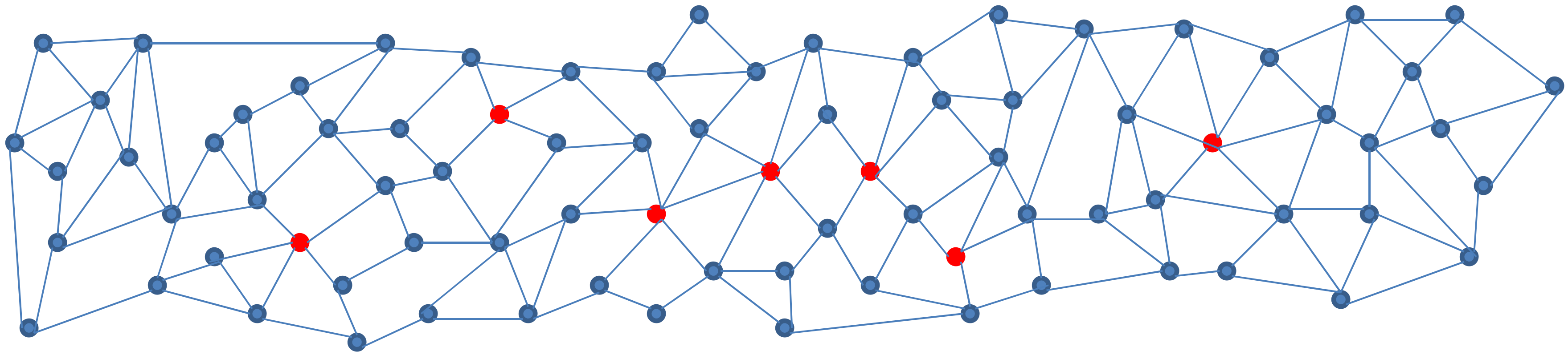
    - a_i: adoption decision, $\mathcal{A}$: set of adopters, w_{ij}: edge weight between i and j, w_{ii}: intrinsic value, p: price
- Chicken and egg adoption decision

Fostering Adoption

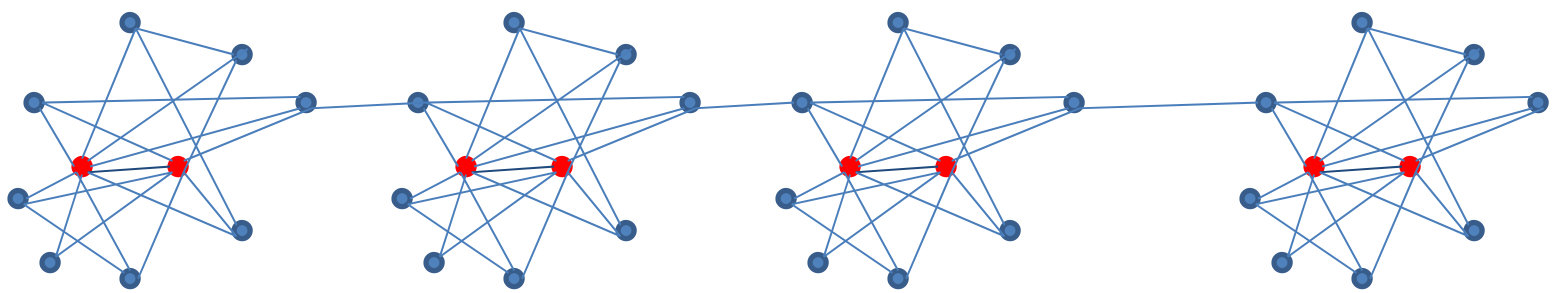
- A strategy based on *seeding*
  - Give technology away (free or cheap) to a small number of users to bootstrap the adoption process
- Basic question: Who should I give it to?
- Can be formulated as an optimization problem
  - Like many network optimization problems, it is NP-hard in most settings
  - Many "folks" heuristics have been used and proposed, e.g., the concept of influentials [1]

[1] D.J. Watts and P.S. Dodds, "Influentials, Networks, and Public Opinion Formation." Journal of Consumer Research, Dec. 2007.

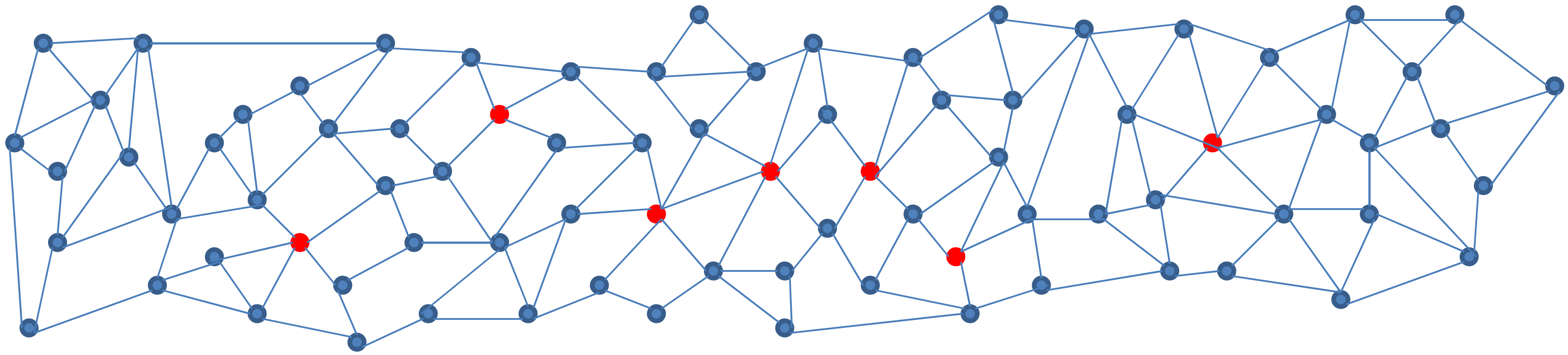
Network Structure and Seeding Strategy



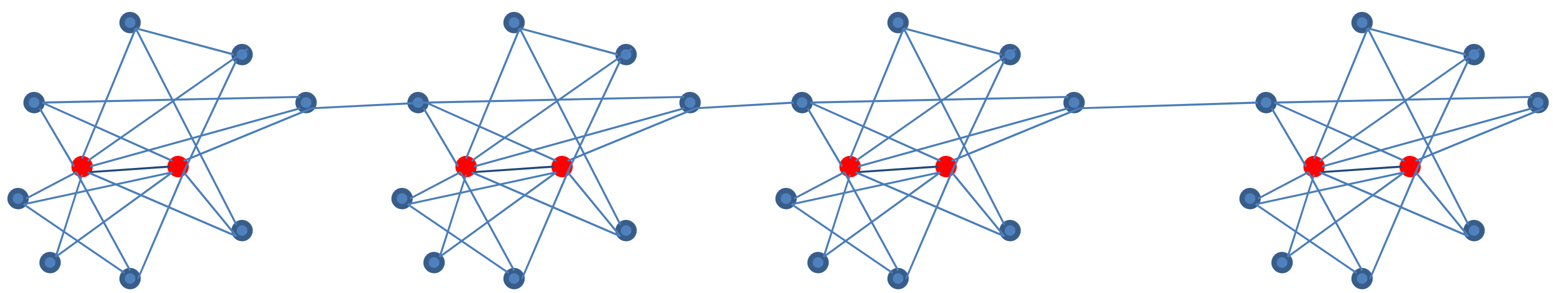
Seeding one city or several villages?



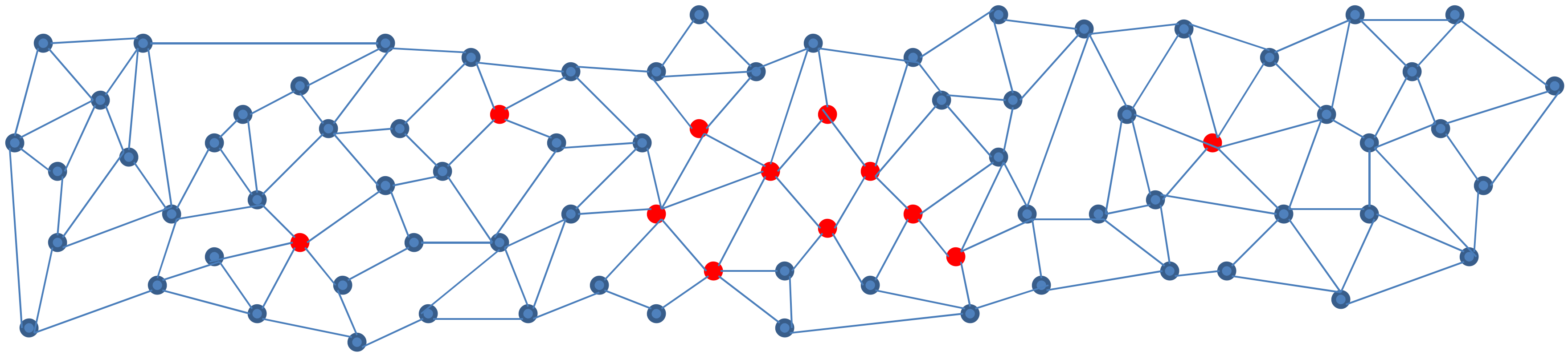
Network Structure and Seeding Strategy



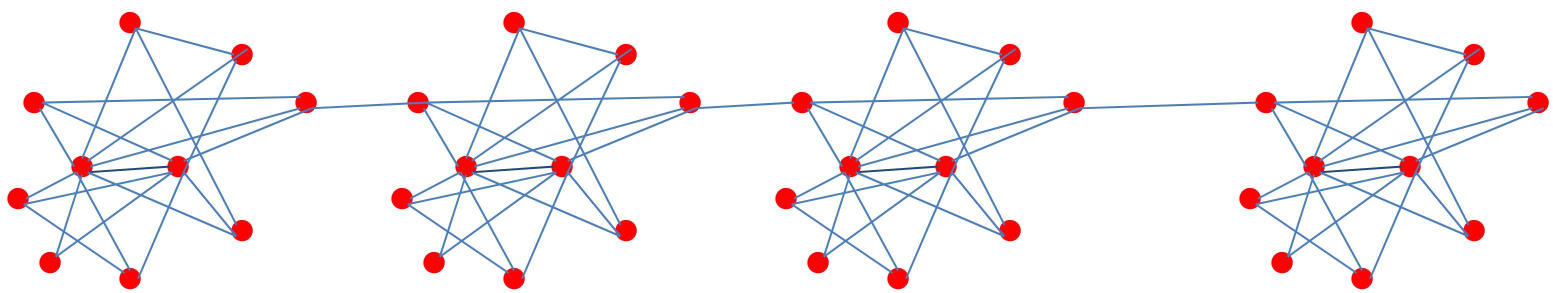
Seeding one city or several villages?



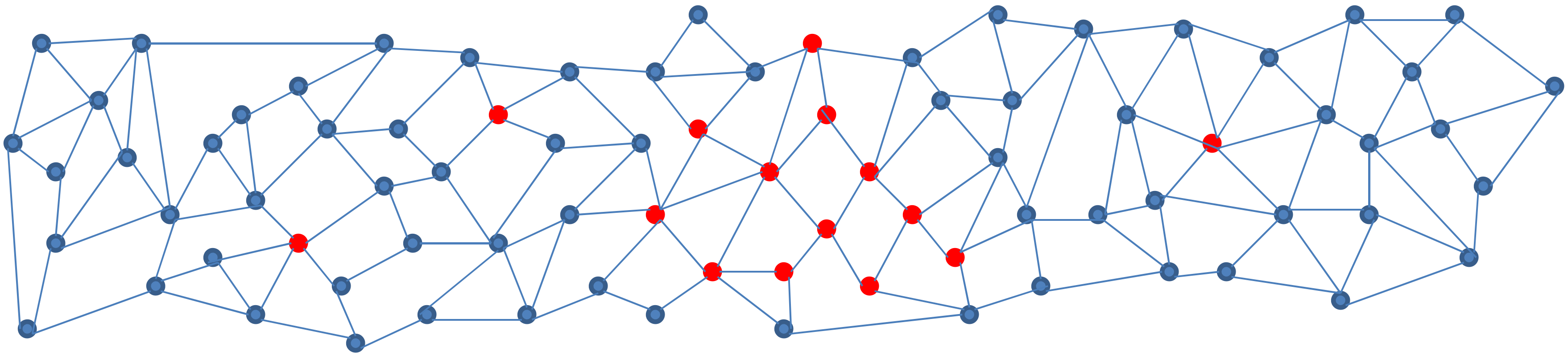
Network Structure and Seeding Strategy



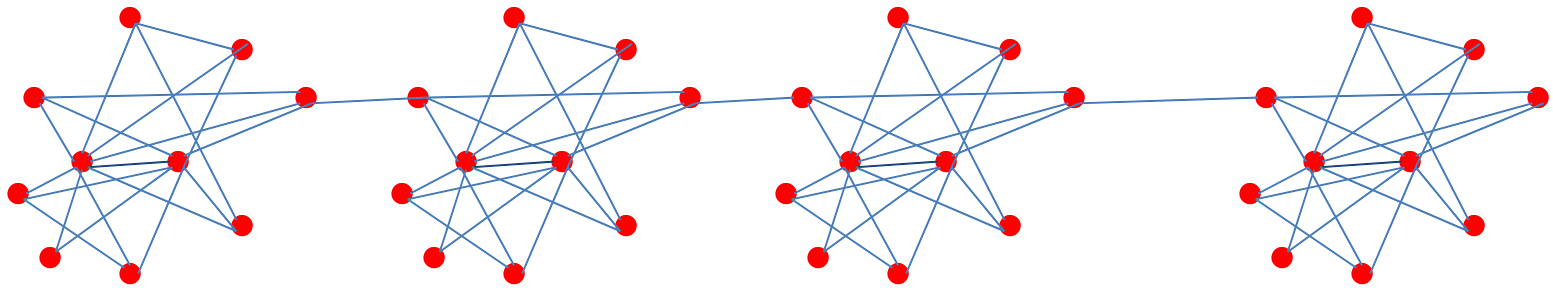
Seeding one city or several villages?



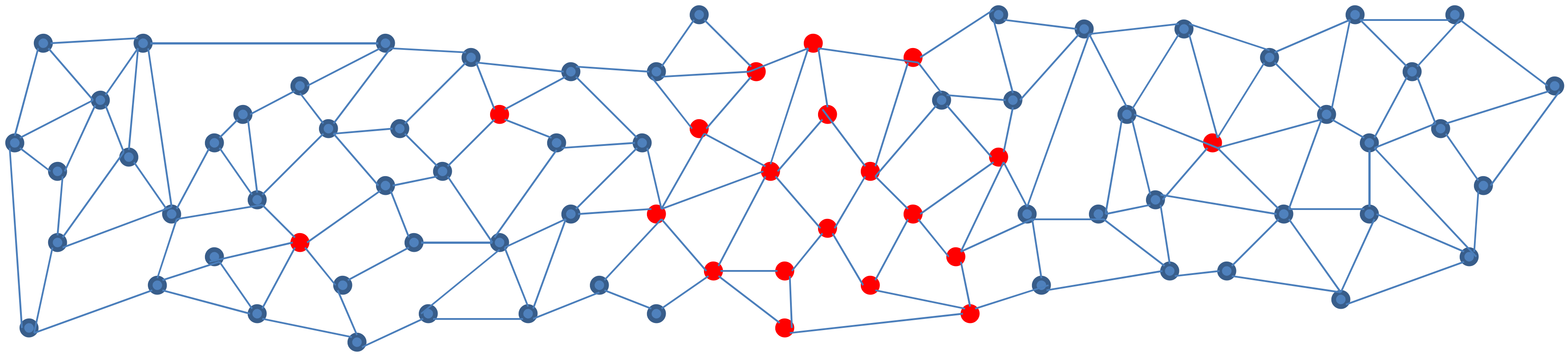
Network Structure and Seeding Strategy



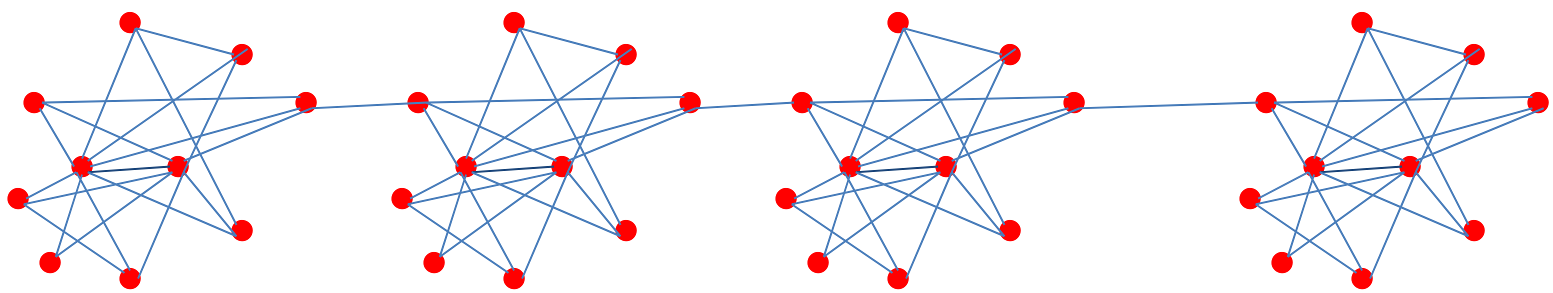
Seeding one city or several villages?



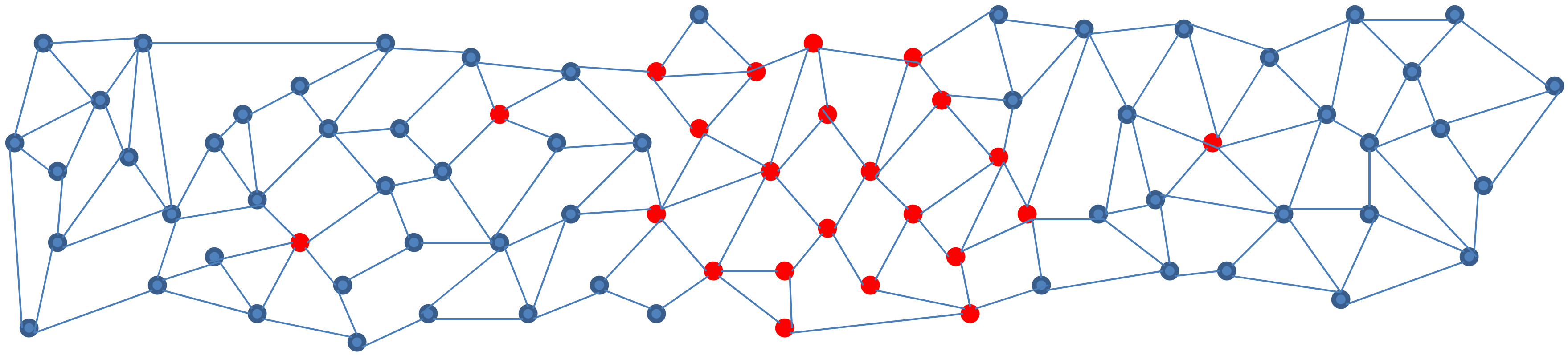
Network Structure and Seeding Strategy



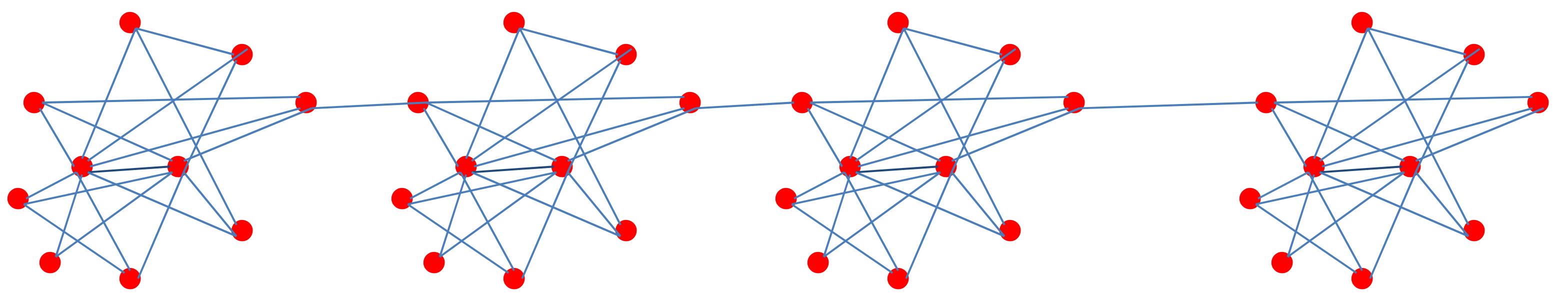
Seeding one city or several villages?



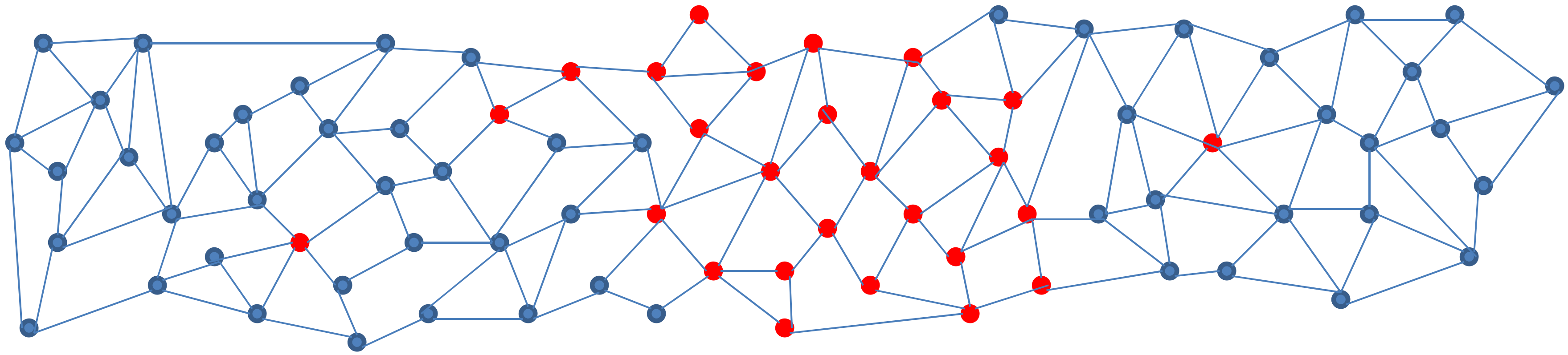
Network Structure and Seeding Strategy



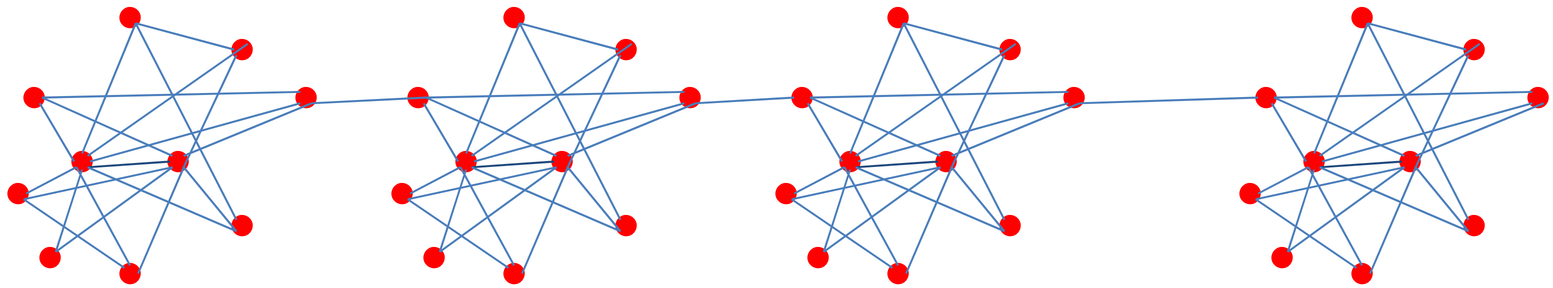
Seeding one city or several villages?



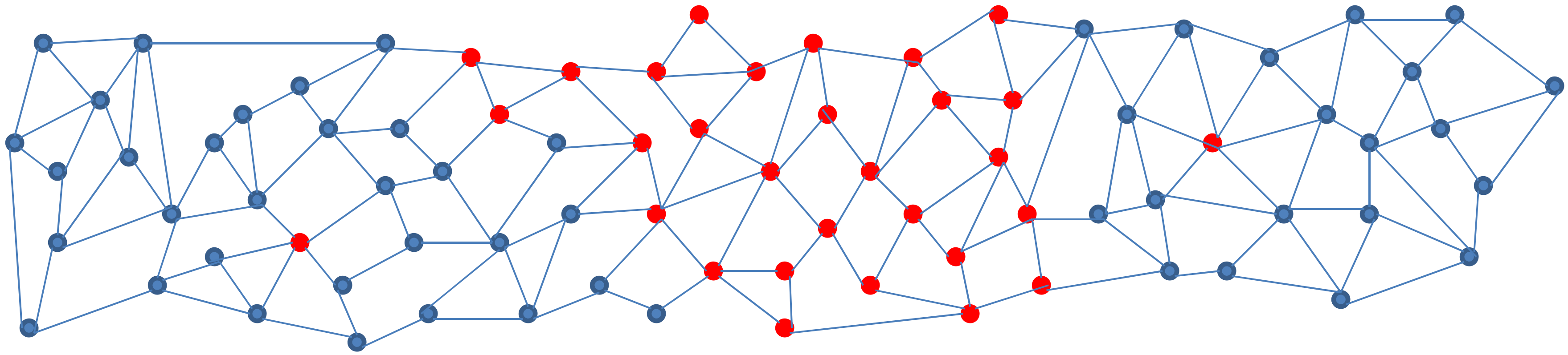
Network Structure and Seeding Strategy



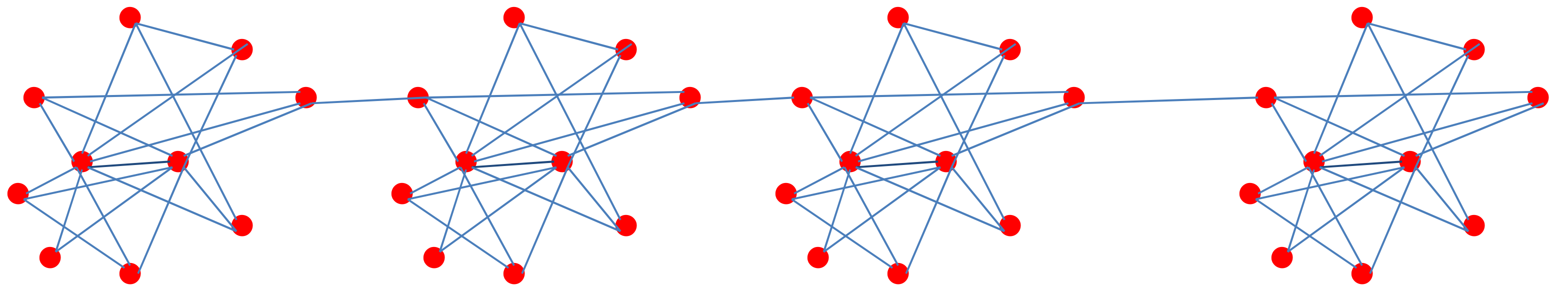
Seeding one city or several villages?



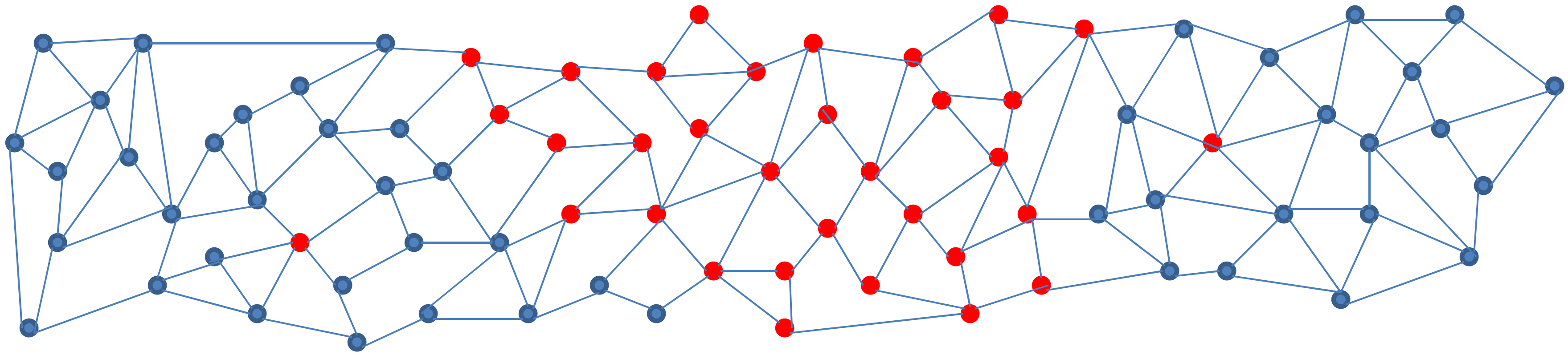
Network Structure and Seeding Strategy



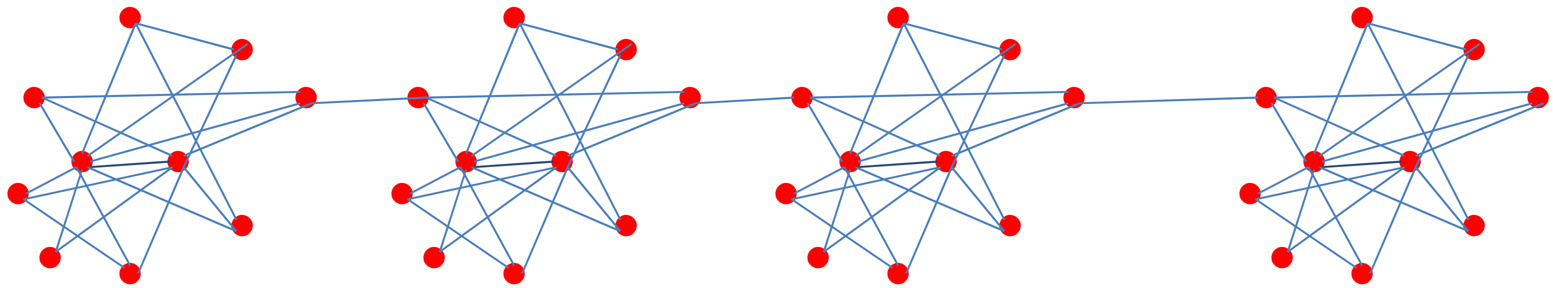
Seeding one city or several villages?



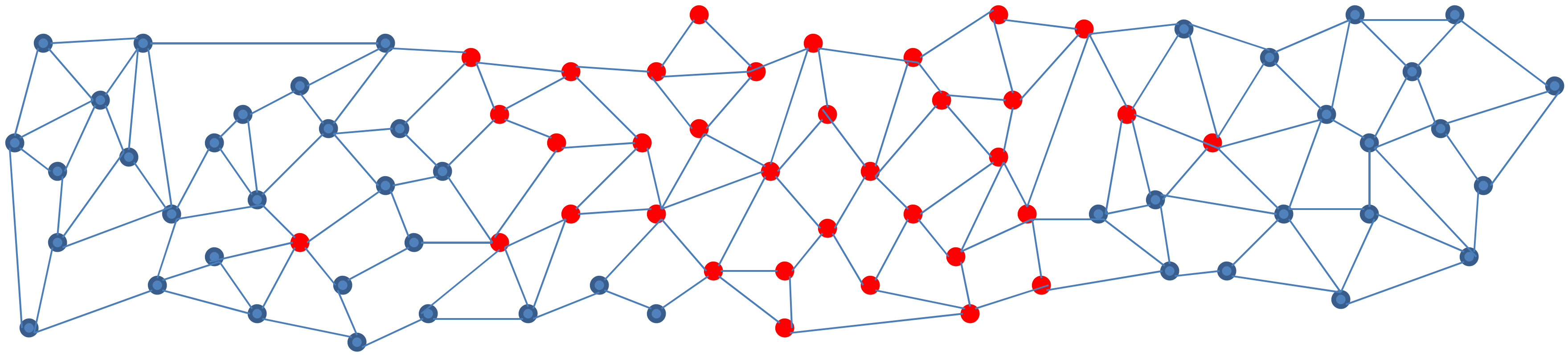
Network Structure and Seeding Strategy



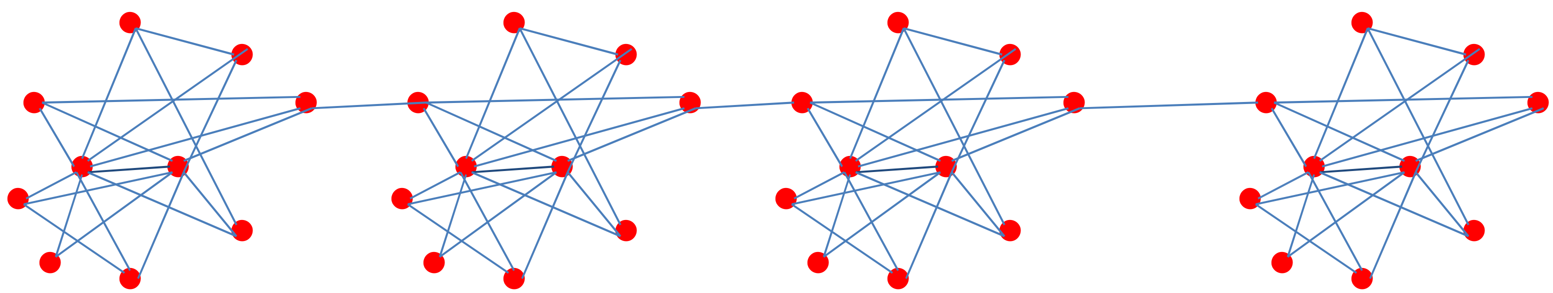
Seeding one city or several villages?



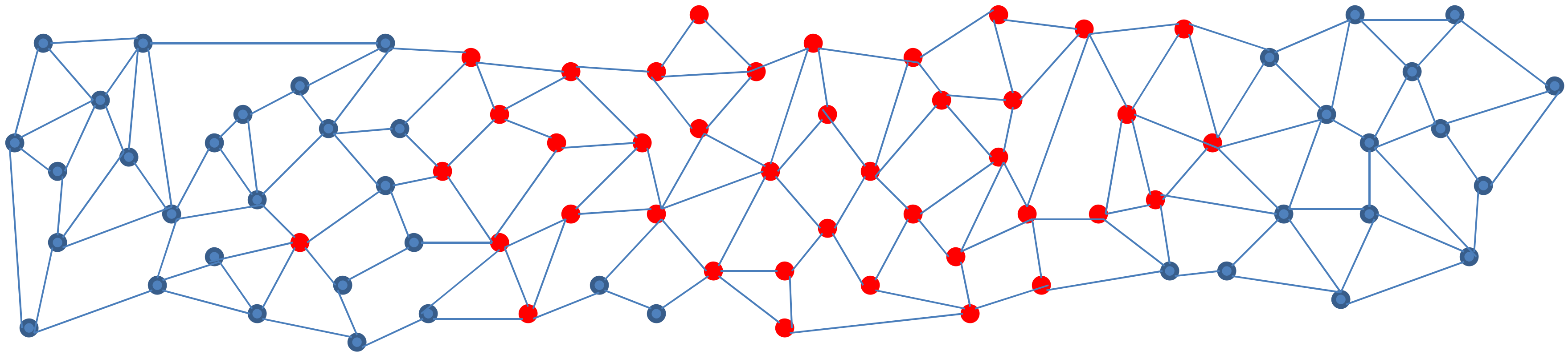
Network Structure and Seeding Strategy



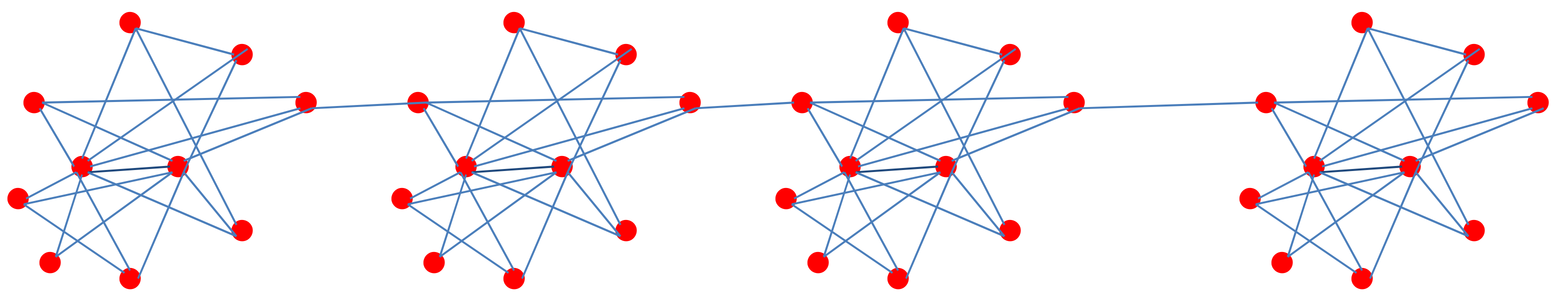
Seeding one city or several villages?



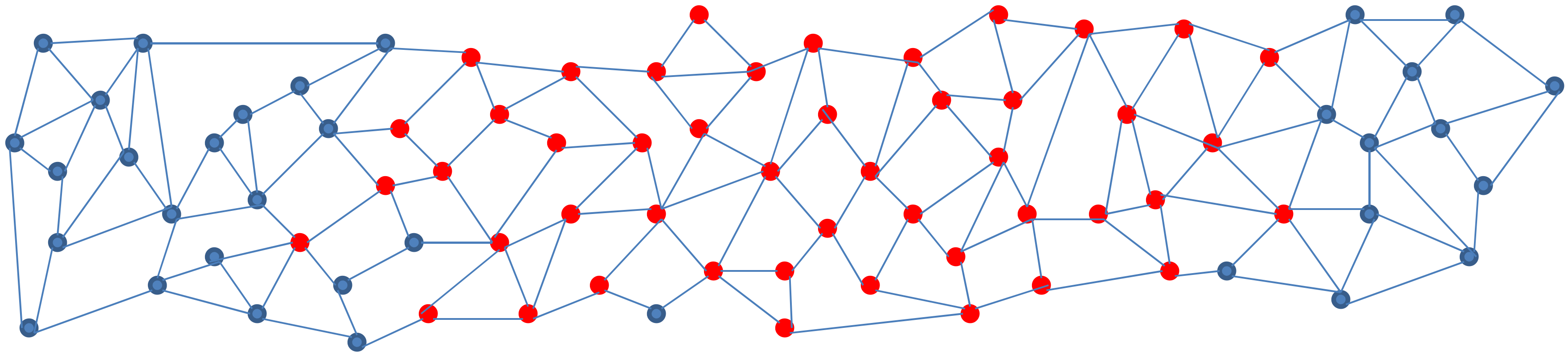
Network Structure and Seeding Strategy



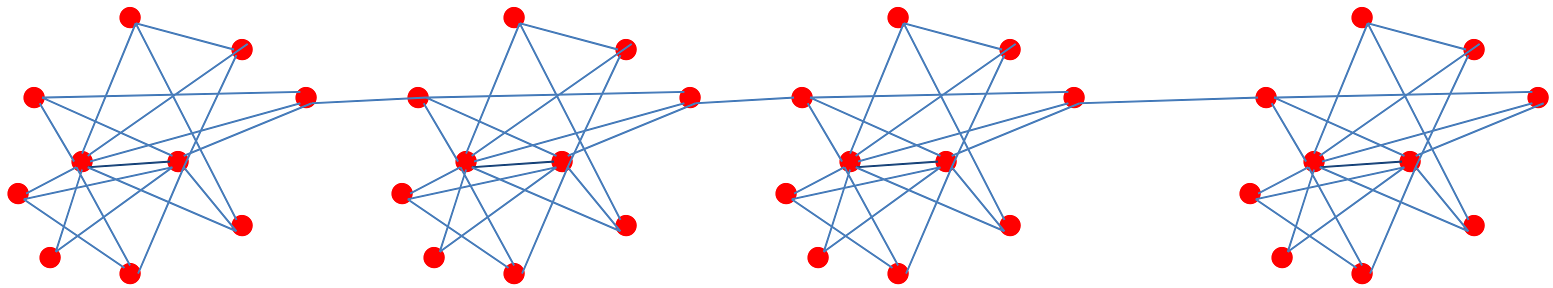
Seeding one city or several villages?



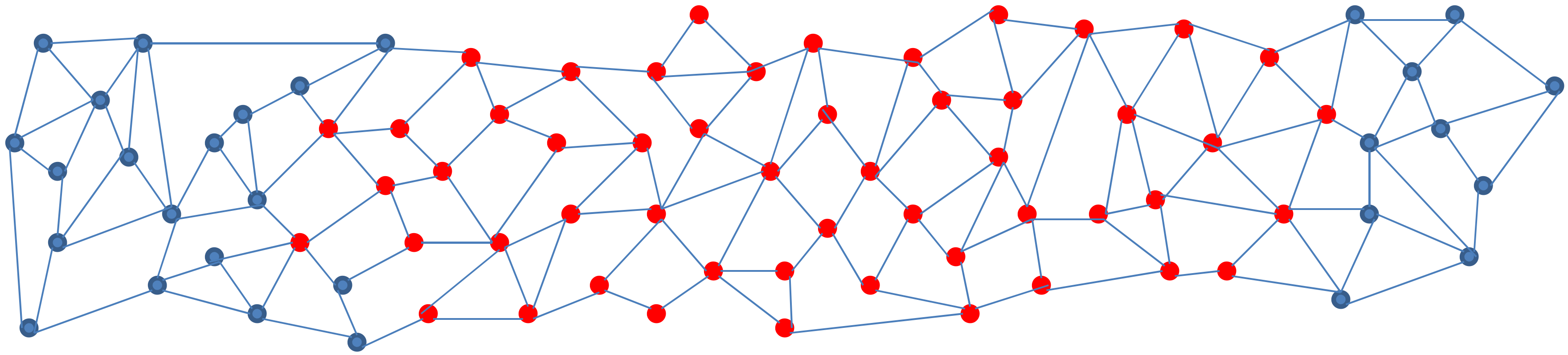
Network Structure and Seeding Strategy



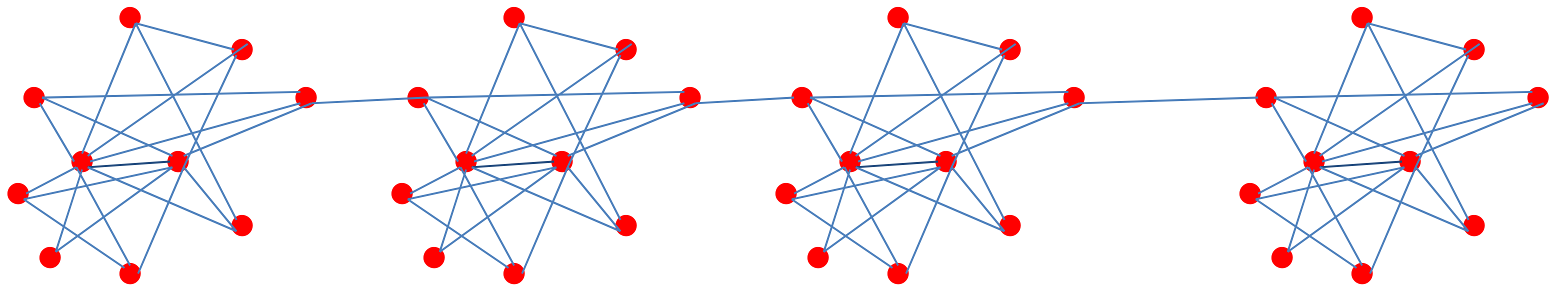
Seeding one city or several villages?



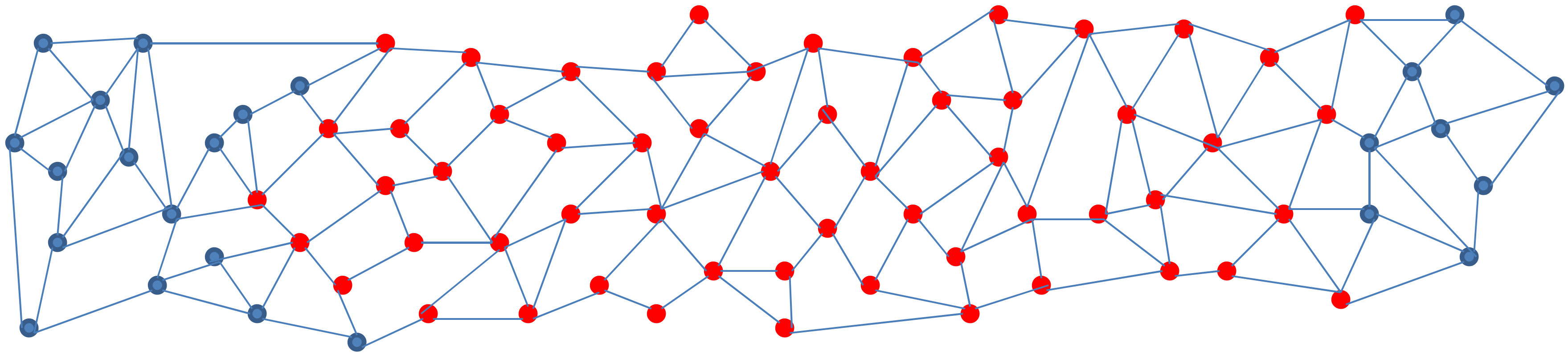
Network Structure and Seeding Strategy



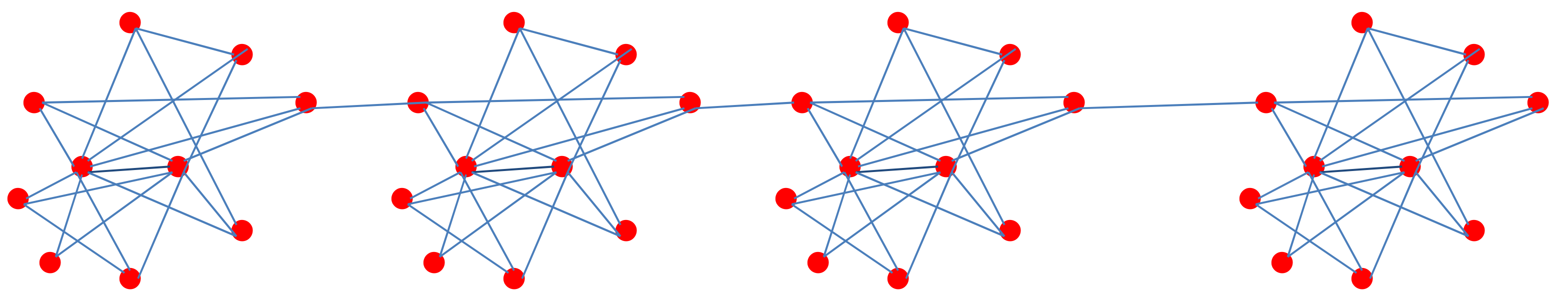
Seeding one city or several villages?



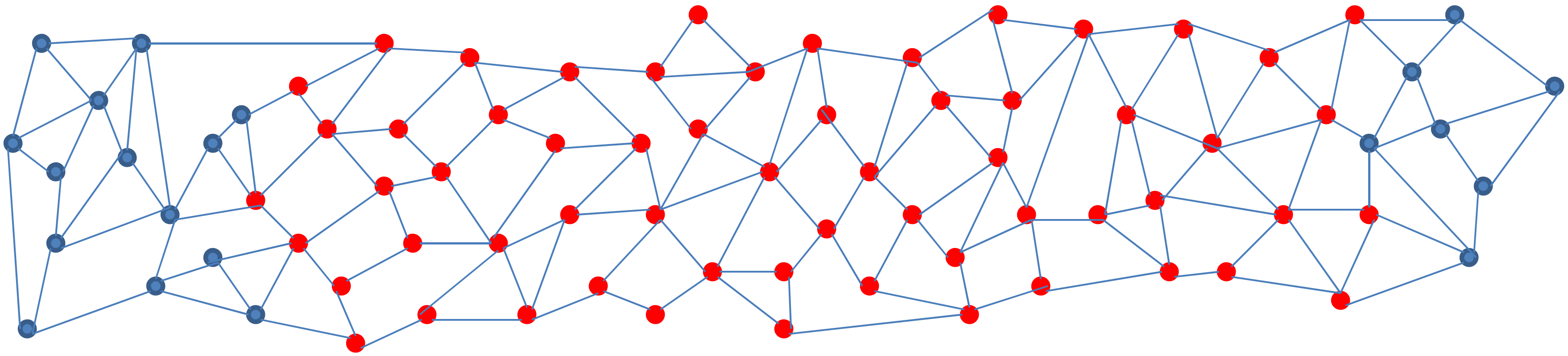
Network Structure and Seeding Strategy



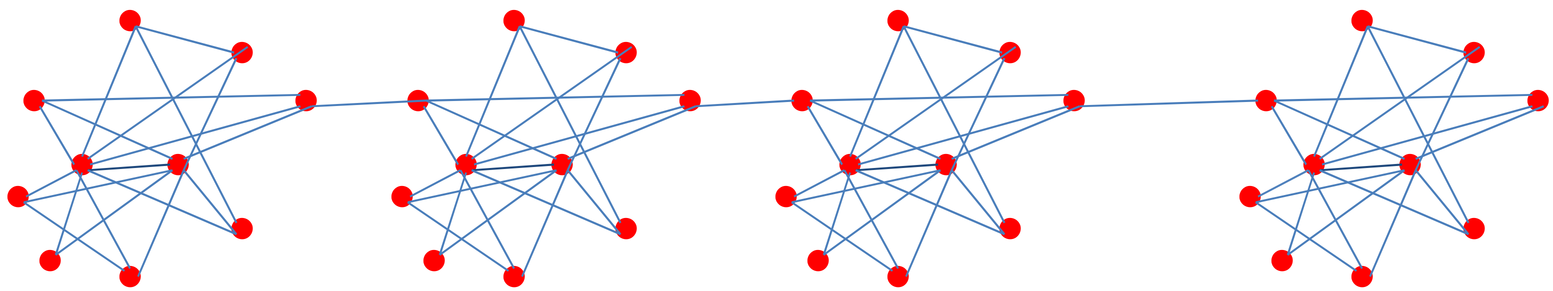
Seeding one city or several villages?



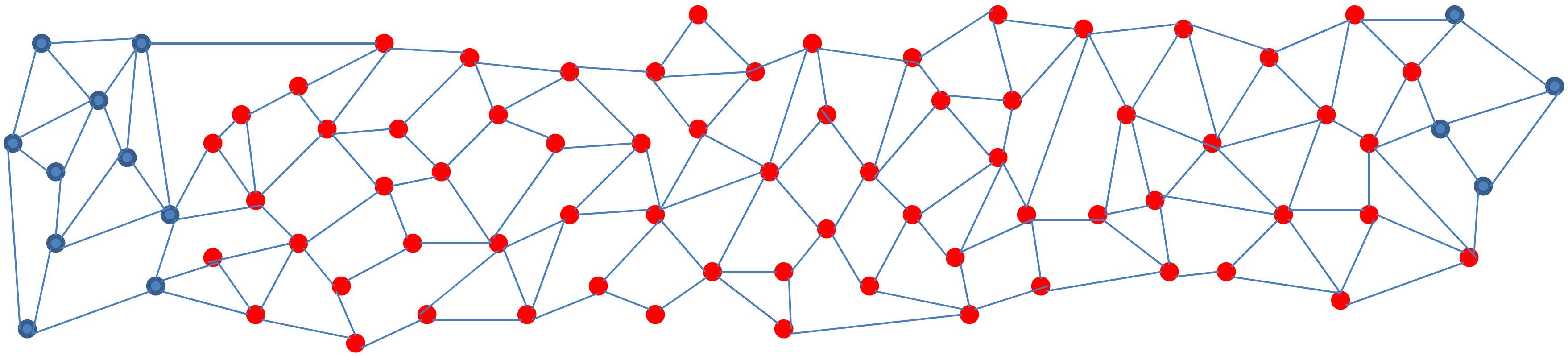
Network Structure and Seeding Strategy



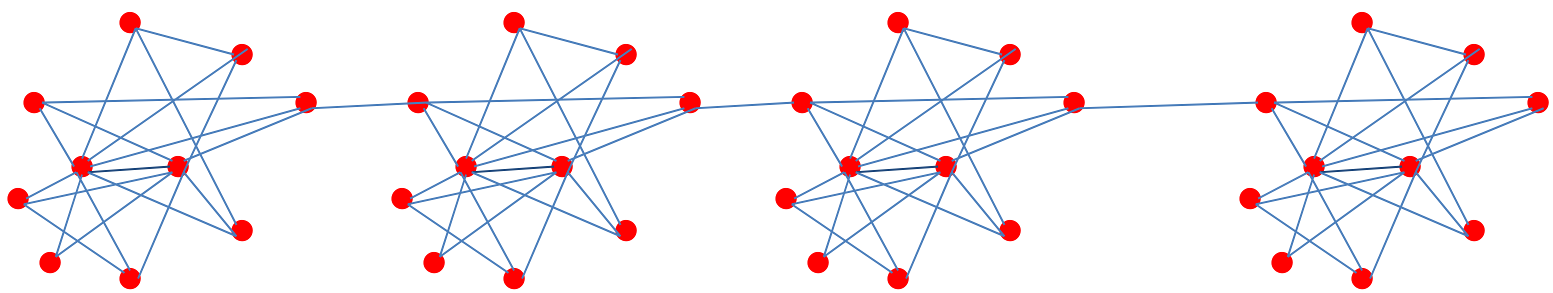
Seeding one city or several villages?



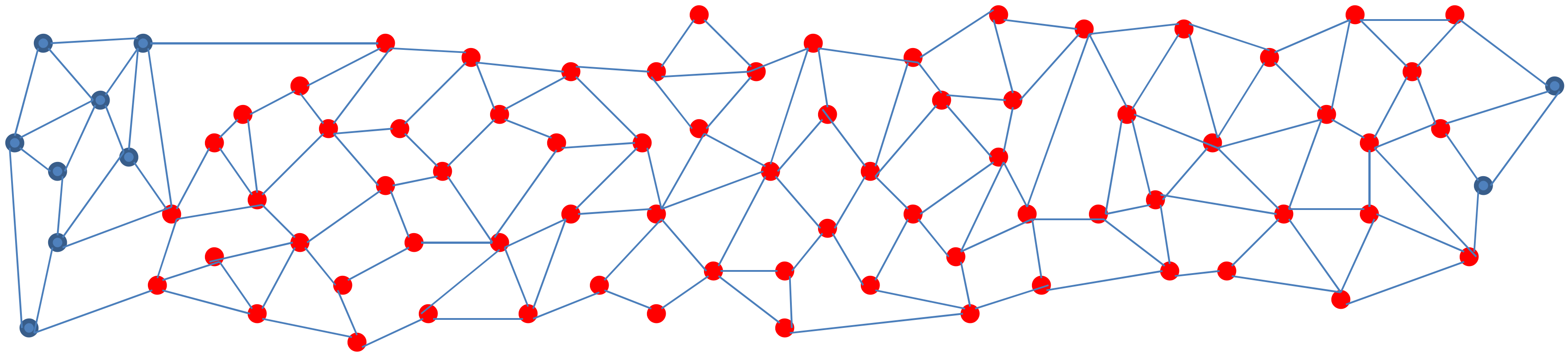
Network Structure and Seeding Strategy



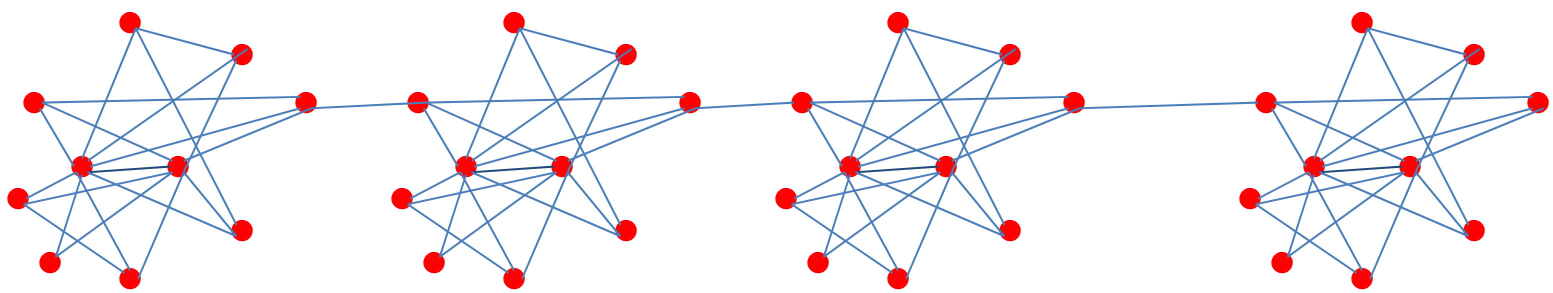
Seeding one city or several villages?



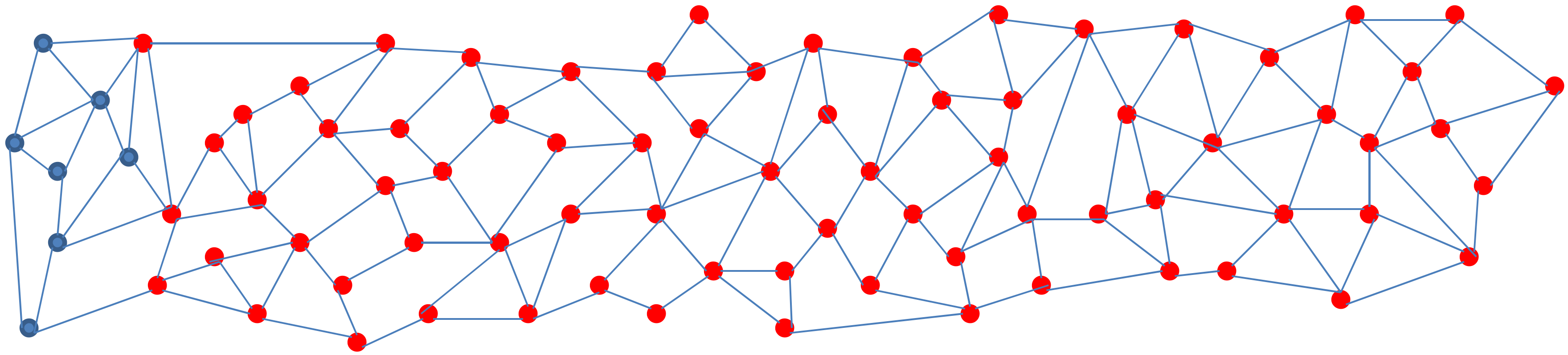
Network Structure and Seeding Strategy



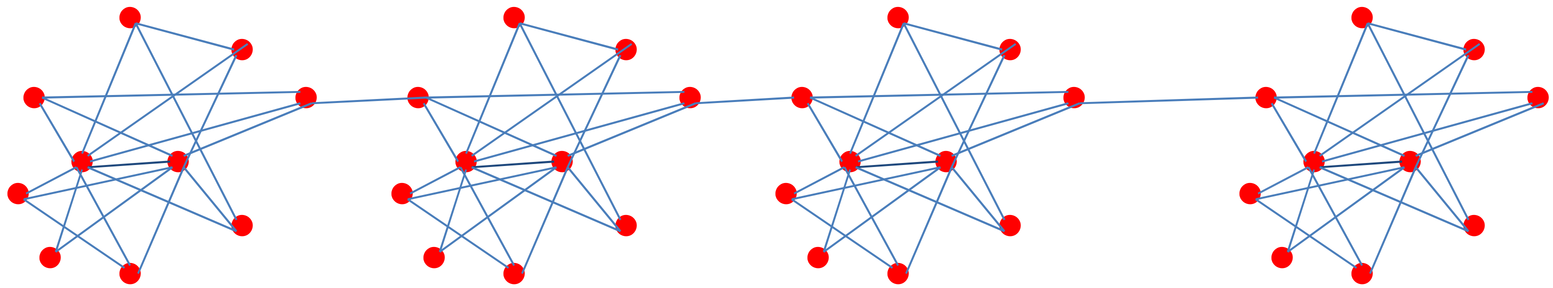
Seeding one city or several villages?



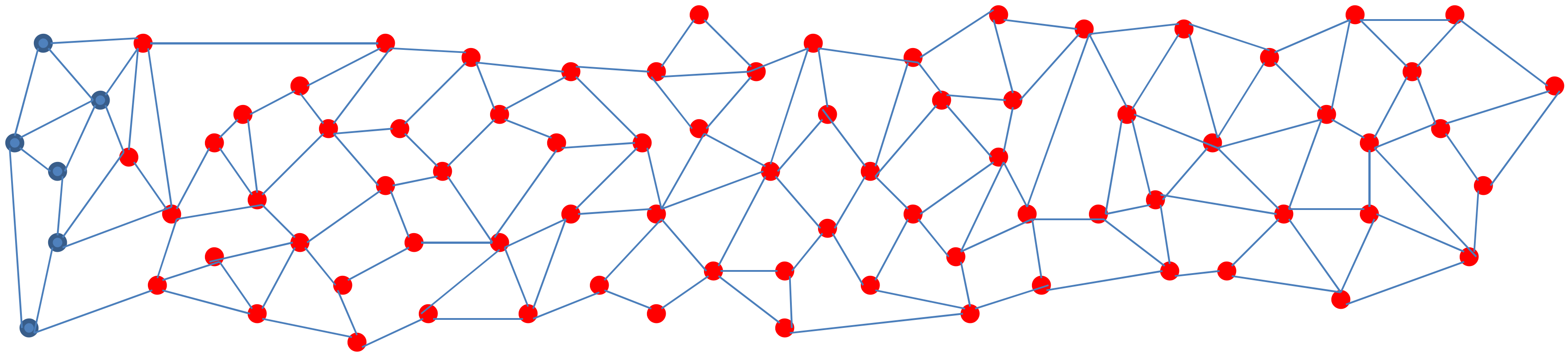
Network Structure and Seeding Strategy



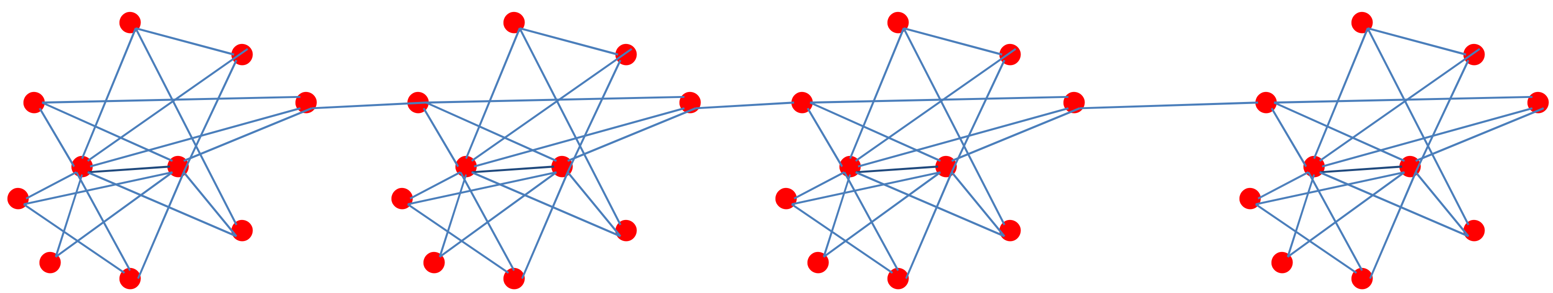
Seeding one city or several villages?



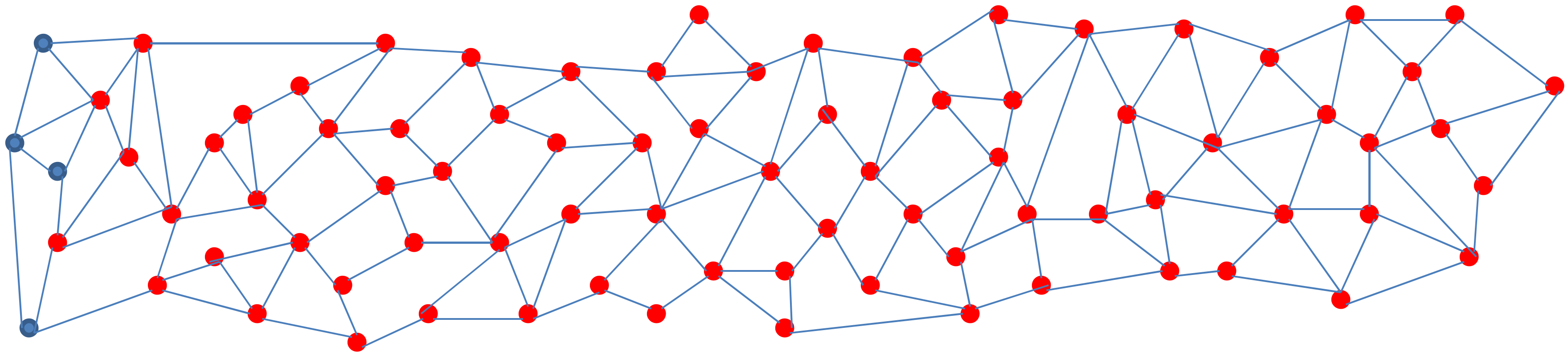
Network Structure and Seeding Strategy



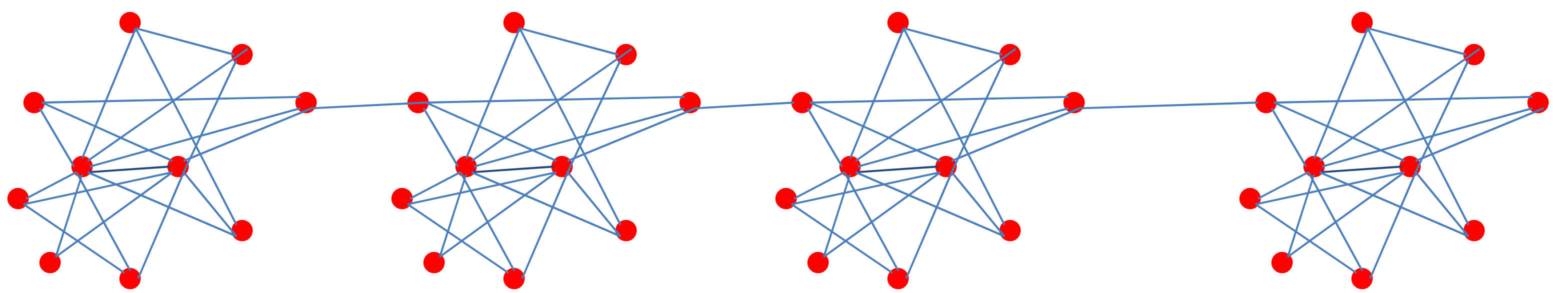
Seeding one city or several villages?



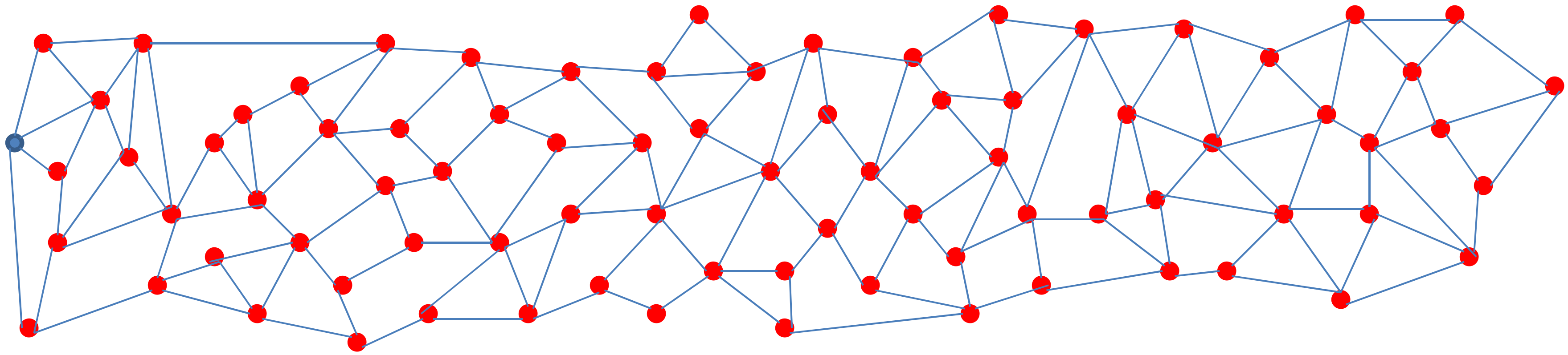
Network Structure and Seeding Strategy



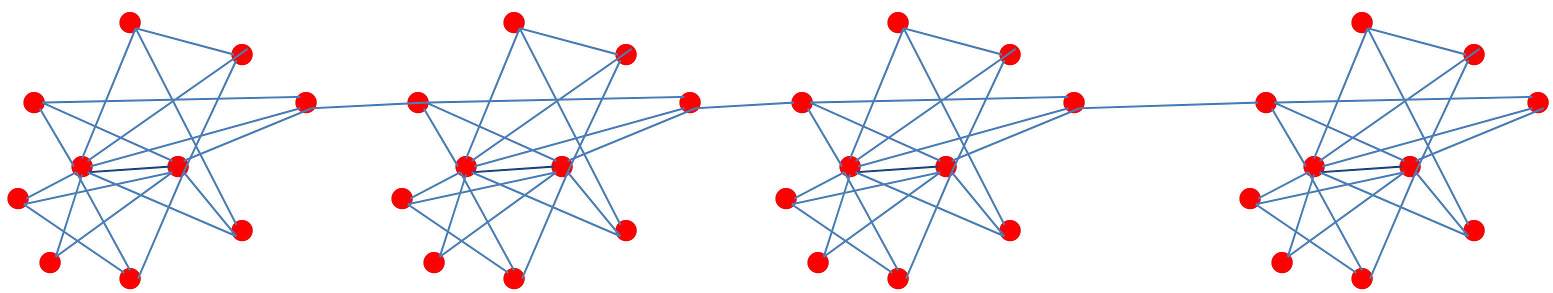
Seeding one city or several villages?



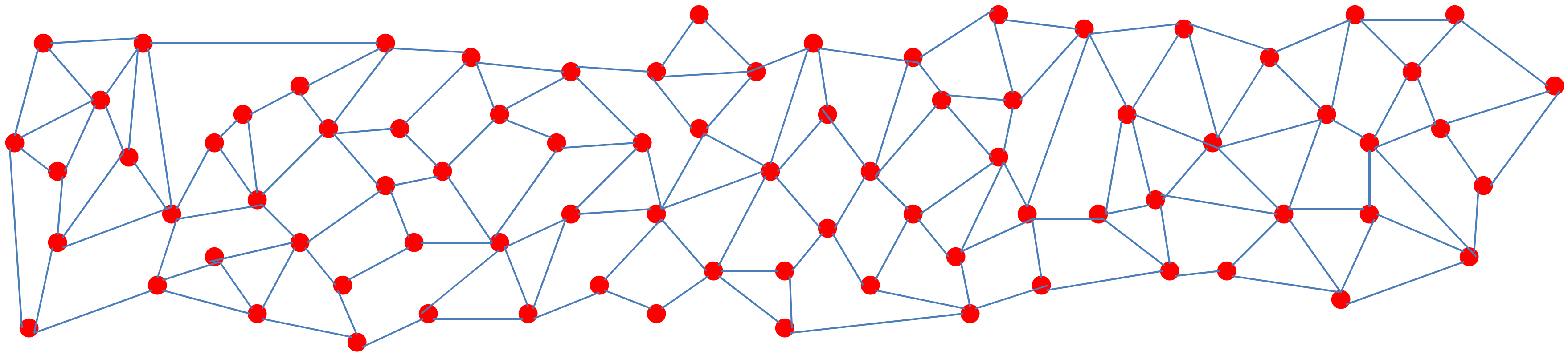
Network Structure and Seeding Strategy



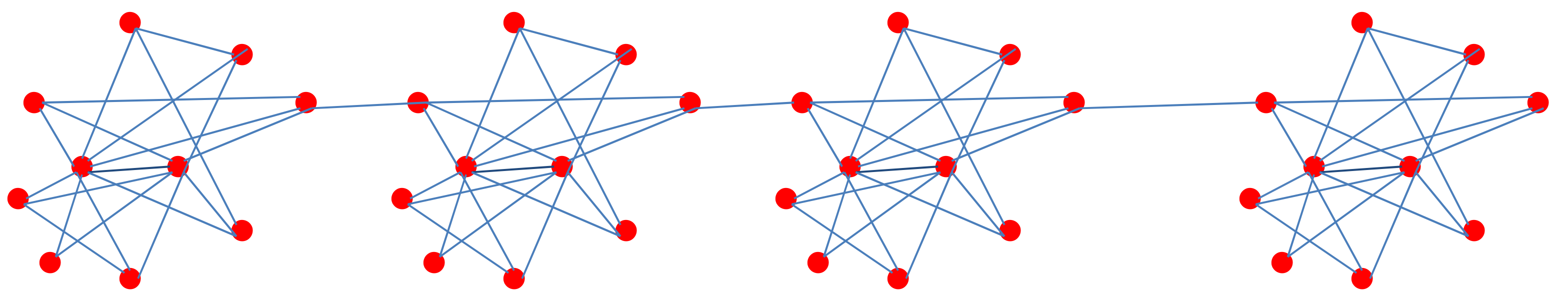
Seeding one city or several villages?



Network Structure and Seeding Strategy



Seeding one city or several villages?

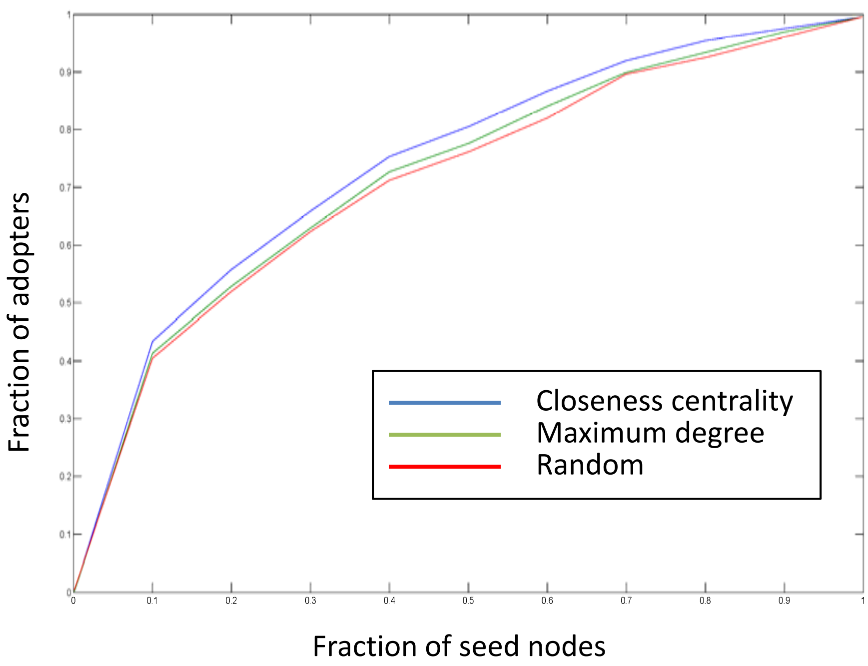


Seeding Strategies

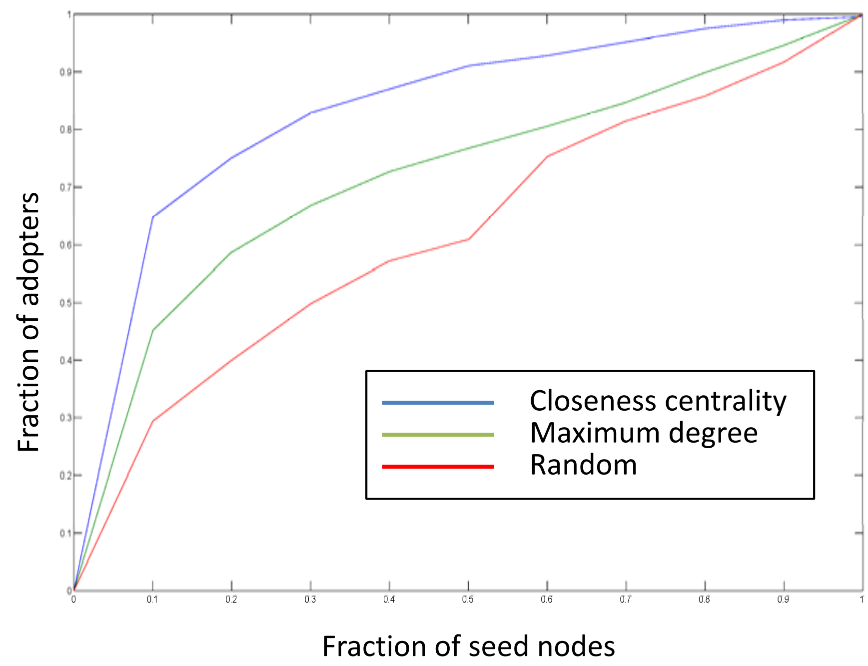
- Understanding the effect of *network structure* is important
- Three sample strategies
  - Random: No information on users or the network
  - Largest degree: Local user information only
  - Closeness centrality: Captures user and network information

Final Adoption Levels
10,000 Node Generalized Random Graphs

Low-variance graph



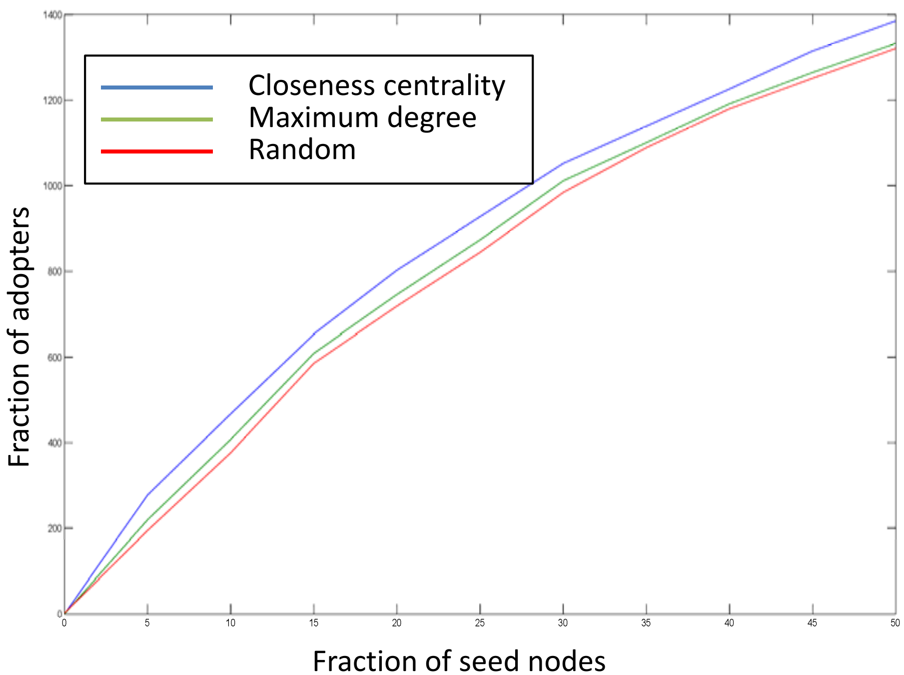
High-variance graph



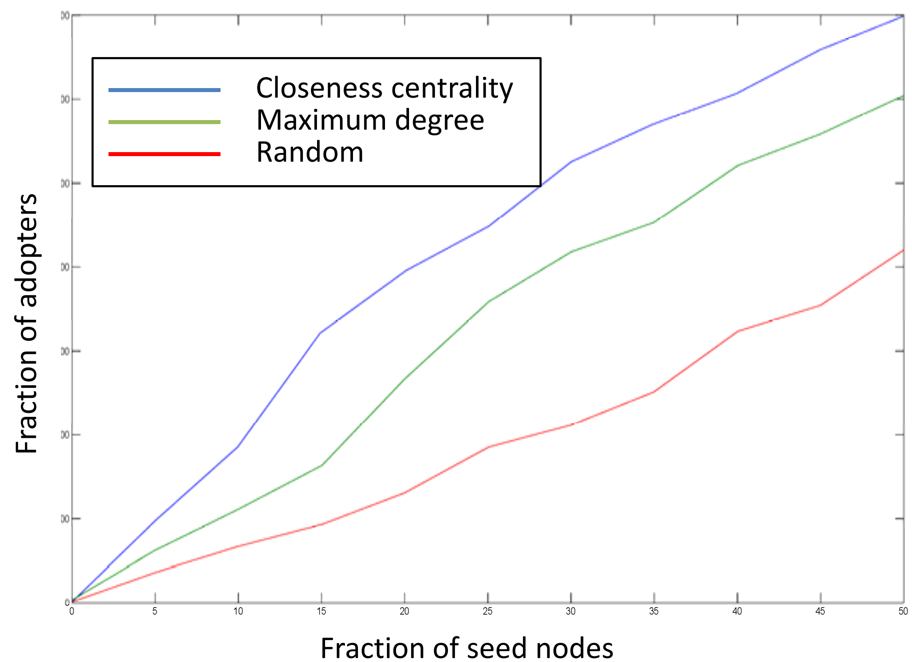
Cascade Sizes

10,000 Node Preferential Attachment Graphs

Low-variance graph



High-variance graph



Closing the Loop

- I did not say you don't/wont need a new network
  - And there is still quite a bit of fun stuff to do there
- I did say that
  - It better allow us do (or imagine doing) things we ***really*** want or need to do and cannot do with today's Internet
  - It better be ***much*** better to convince people to switch
- I also did say that
  - As the field of networking matures, many of the interesting problems arise in *using* networks not just *building* them
    - This is no different from what happens with most other technologies

ACKNOWLEDGMENTS

- This talk is based on joint work with many colleagues and students, and has benefited from their inputs. In particular, I would like to acknowledge (in alphabetical order)

J. Corbo, K. Hosanagar, Y. Jin, A. Odlyzko, S. Sen, and Z.L. Zhang

Thank You!