Growing up in the Internet Shadow and What Next?

Roch Guérin
Dept. Comp. Sci. & Eng.
Washington University in St. Louis

WASHINGTON·UNIVERSITY·IN·ST·LOUIS

Acknowledgments

- The technical part of this talk is based on joint work with several talented collaborators, who deserve much/most of the credit
  - They are, in alphabetical order, M. H. Afrasiabi (Penn), J. C. de Oliveira (Drexel), S. Venkatesh (Penn) and S. Weber (Drexel)

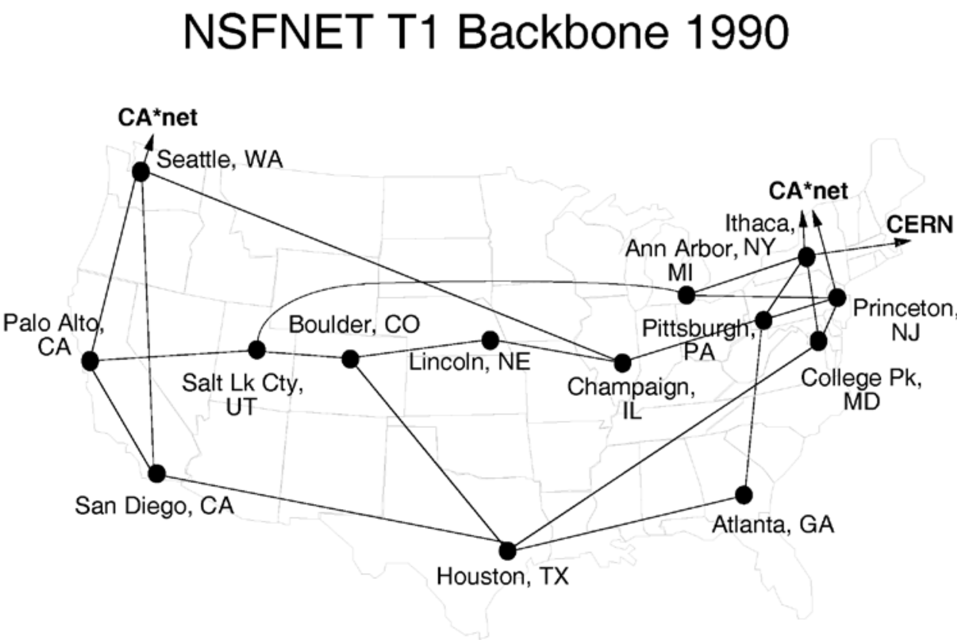
- Errors, inaccuracies, and omissions in the non-technical part are, however, all mine

The technical part of this work was made possible in part by generous support from the NSF through grants CNS-1116039, CCF-1137519, and CNS-1319684

The Internet (R)evolution

From less than 4 millions users in 1990

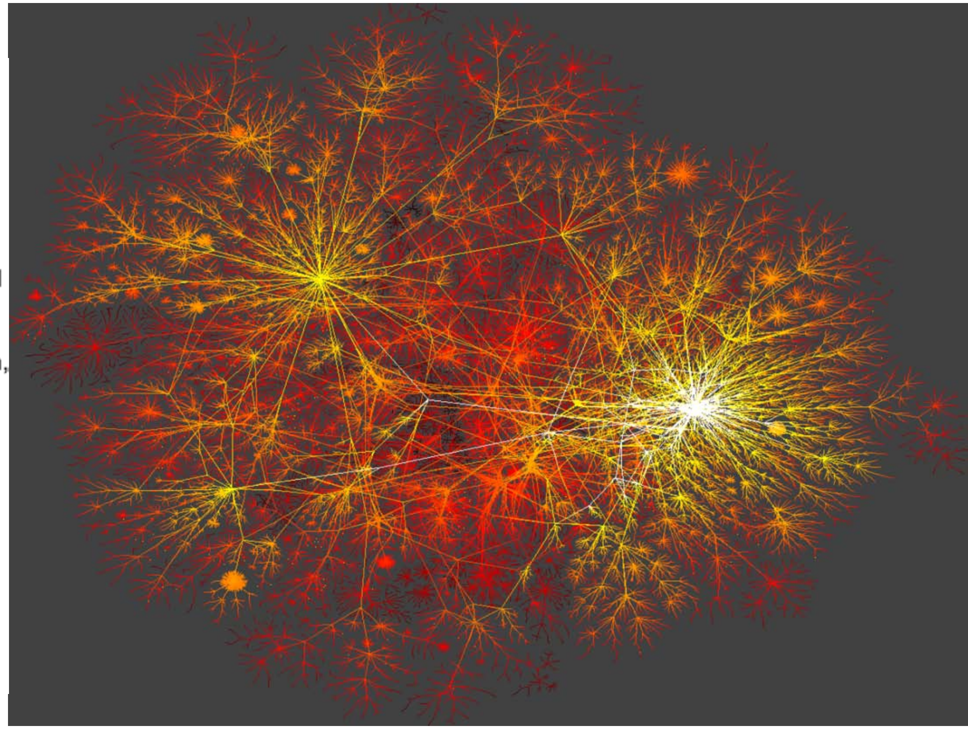
http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet-related-networks



From ftp://ftp.cs.toronto.edu/doc/maps/nsfnet.ps

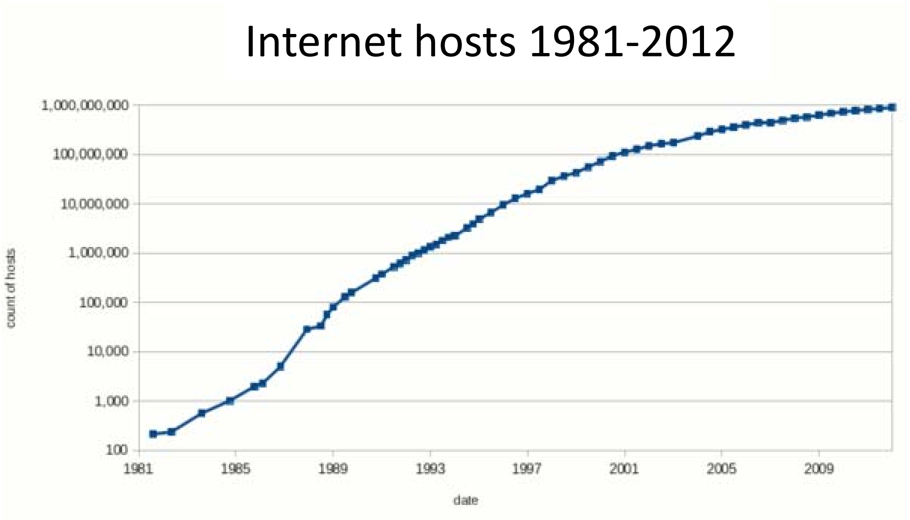
To about 2.5 billions users in 2013

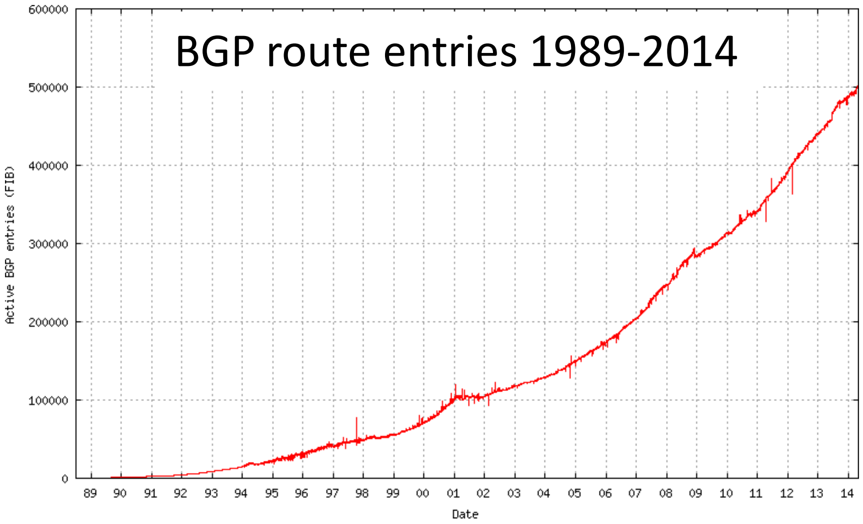
http://www.internetworldstats.com/stats.htm



From http://www3.nd.edu/~networks/Image%20Gallery/gallery.htm

Quantifying the Internet Phenomenon





- From a few hundred hosts to over a billion in a span of about 30 years (~**seven** orders of magnitude), and the growth is not over (only ~40% of the world's population is connected and IoT promises further growth)

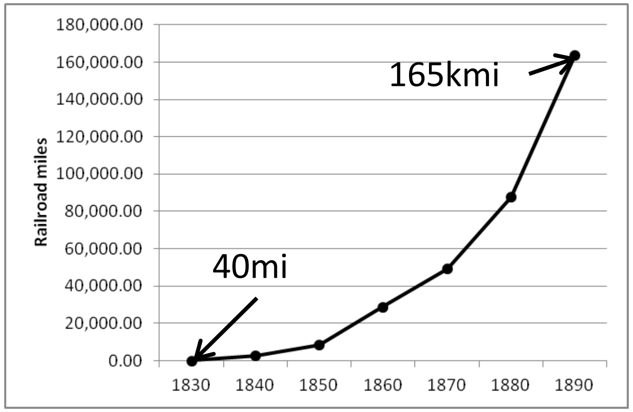
Source: http://www.isc.org/solutions/survey/history

- From less than 100 route entries in core routing tables to close to half a million routes in 25 years (a growth of close to **four** orders of magnitude)

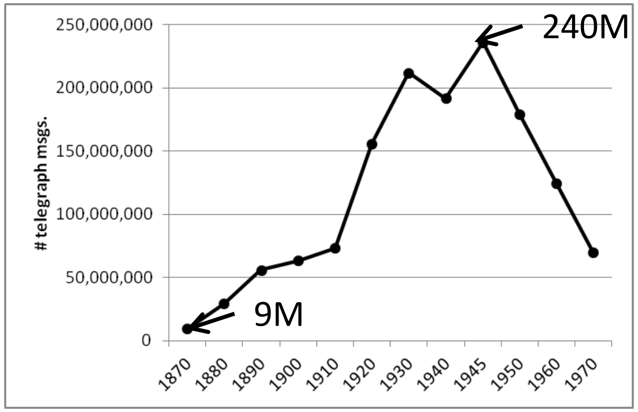
Source: http://bgp.potaroo.net

Putting Things in Perspective

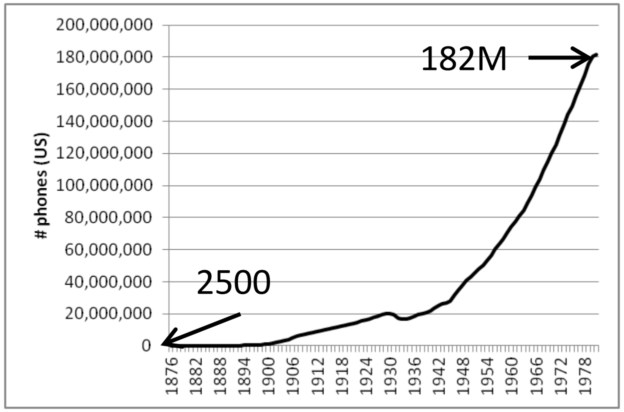
Railroad (US)
United States Census Bureau



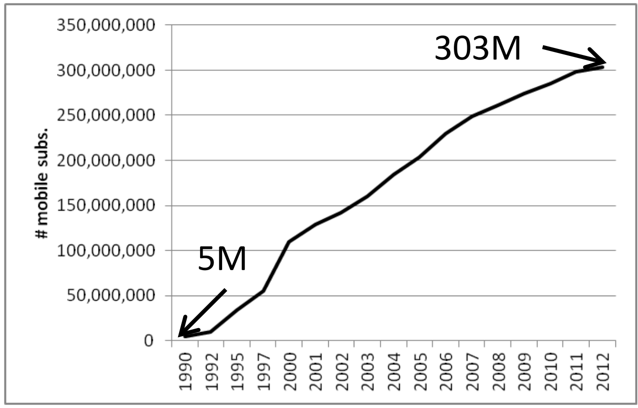
Telegraph (US)
United States Census Bureau



Telephone (US)
http://galbithink.org/telcos/telephones-1876-1981.xls



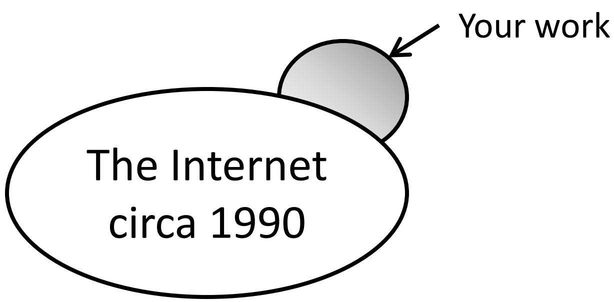
Mobile (US)
http://www.ctia.org



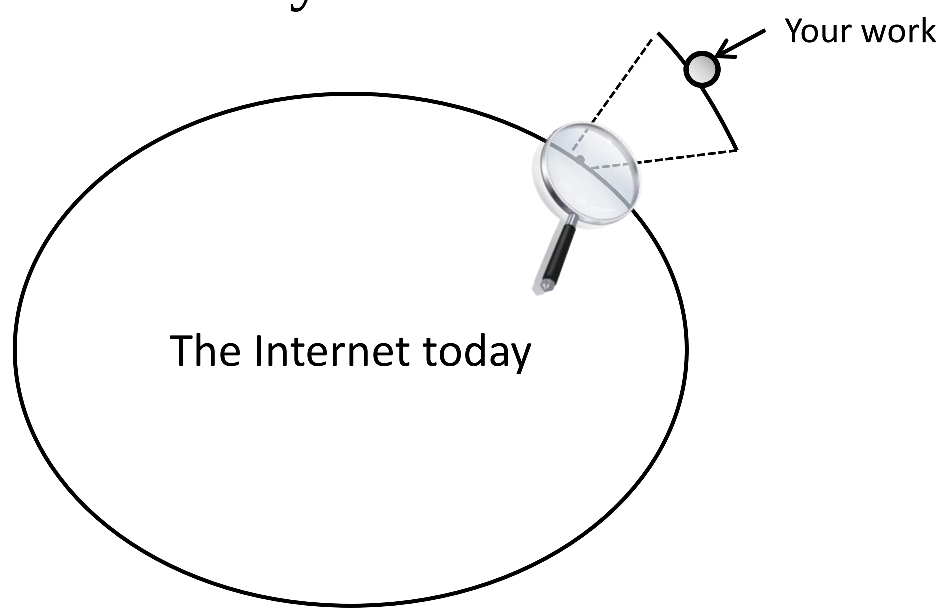
- These were arguably some of the most transformative technologies of their time, but they pale in comparison to the Internet
  - **Railroad**: About four orders of magnitude growth in 60 years, and then flat
  - **Telegraph**: Less than two orders of magnitudes in 75 years, and then a precipitous drop
  - **Telephone**: Five orders of magnitude growth in 100 years, and then mostly flat
  - **Mobile**: About two orders of magnitude growth in 20 years, but starting to saturate (even worldwide)

And What It Means for Internet Research

- The impact of your work in the early 1990s



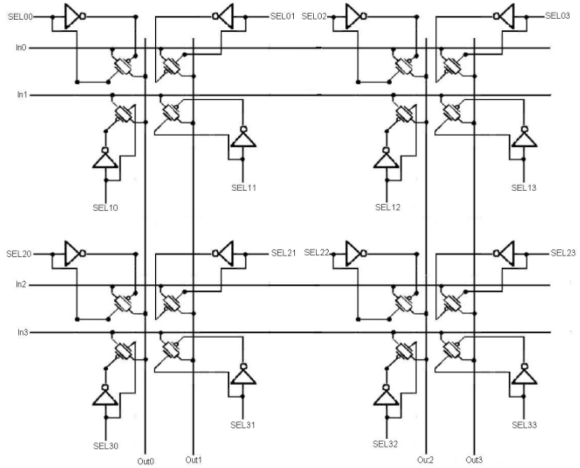
- Scaling things up to today's environment



- There are still lots of fun things to do in Internet research, but the glory years are gone and unlikely to come back

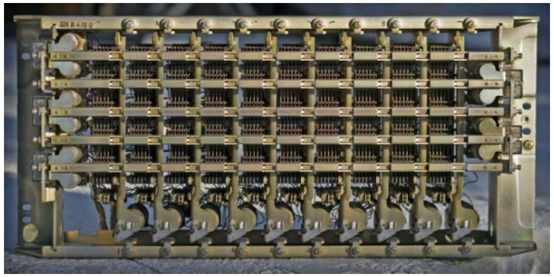
A Unique Combination Unlikely to Repeat Any Time Soon

- The Internet revolution is really the semi-conductor revolution
- As is common in such instances*, the new technology is first used to make existing designs better, before realizing that it allows entirely new designs, *i.e.*, the network as the computer or rather a computing network

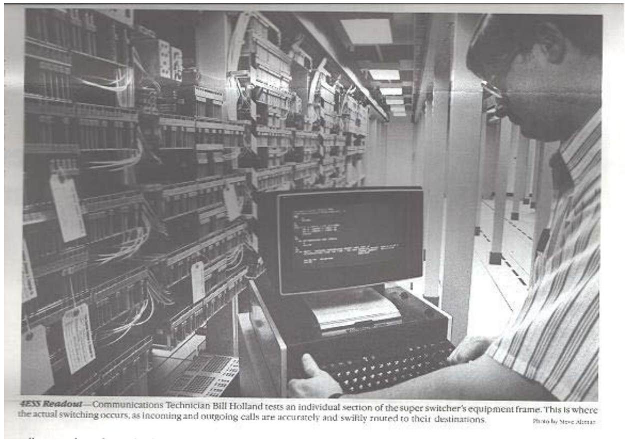


***** Brian Arthur, *"The Nature of Technology: What it is and how it evolves."*

From http://atomictoasters.com/wp-content/uploads/2010/12/Telephone-Operators.jpg

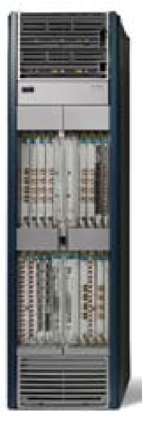


From http://en.wikipedia.org/wiki/Crossbar_switch

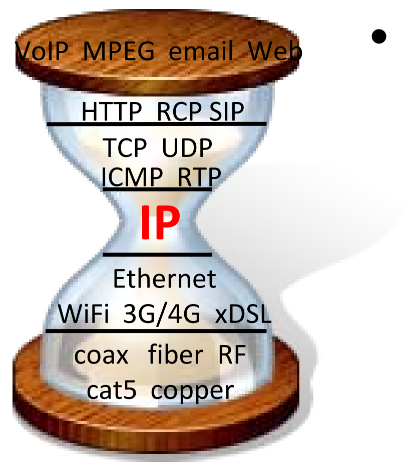


http://www.phworld.org/switch/4ess.jpg

The Internet Paradigm Shift



- Routers "compute" on packets rather than just transmitting them



- And arguably, IP's success is largely due to picking "just the right level of computation", *i.e.*, neither too little, nor too much

- Some may argue that the next logical steps is to ride the computation wave further and increase the level of computation the network performs, but the jury is still out

- More importantly though, irrespective of the outcome, we are looking at <u>*evolution* and not *revolution*</u>, so that barring another technology discontinuity that will usher in the potential for drastically new designs, we are now stuck in an age of progressive changes

Networking Research
From Revolution to Evolution

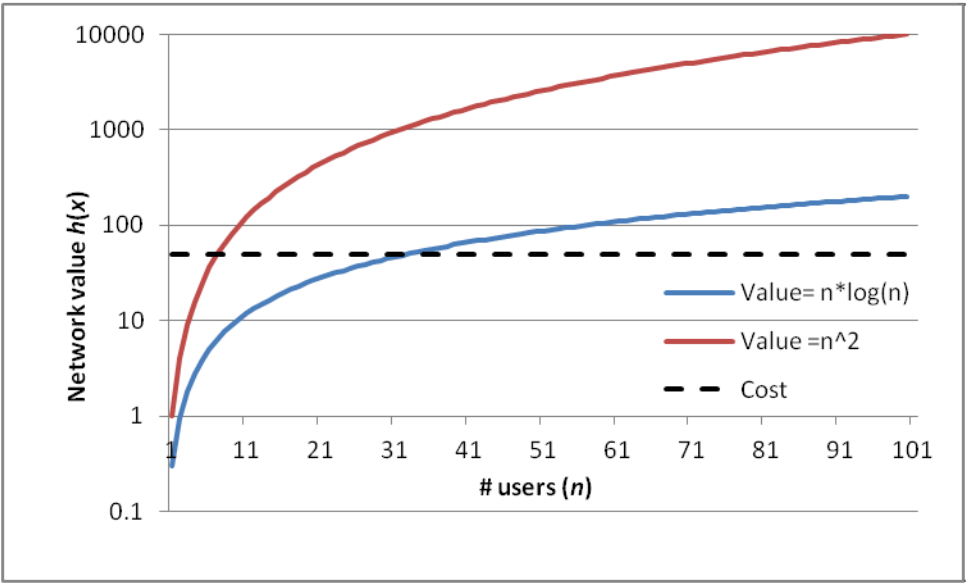
- There are clearly many interesting networking problems that remain to be solved, but
  - Barring a new technology paradigm shift, most efforts are likely to be improvements to the functionality or performance of existing solutions
  - Getting recognition or adoption of your work is likely to be increasingly difficult, unless you target a specific and active sub-community

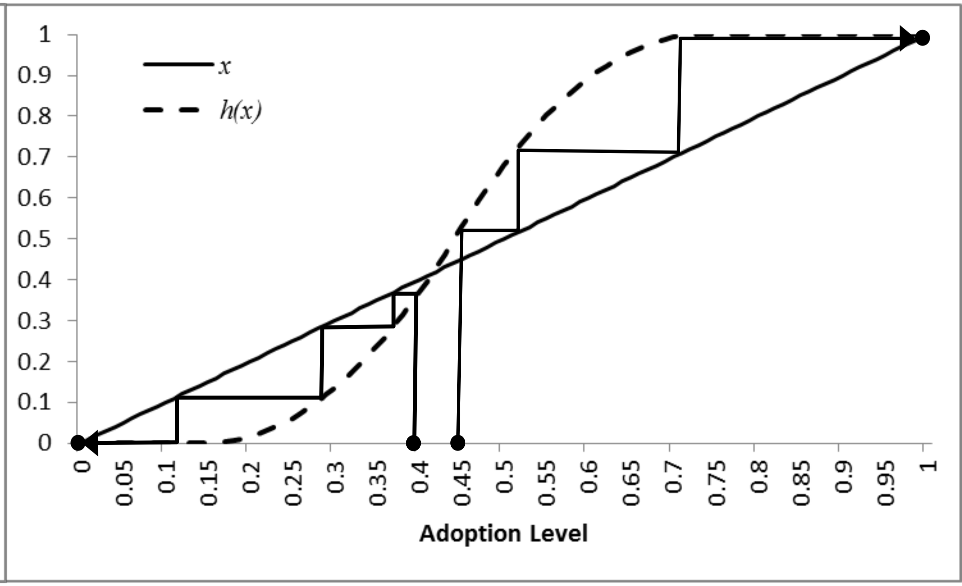
- Fortunately, the now ubiquitous nature of the Internet has also created many new and challenging research areas. I'll mention two.
  1. How do you get things done/adopted in large-scale networks?
  2. Can we predict what outcomes emerge from networked interactions and why?

TECHNOLOGY ADOPTION IN LARGE-SCALE NETWORKS

The Adoption Conundrum of Network Technologies

- Useful above a certain adoption threshold, but how to get there?
  - See, *e.g.*, A. Ozment and S. E. Schechter, "Bootstrapping the adoption of Internet security protocols." Proc. WEIS 2006, Cambridge, UK, for a relevant discussion

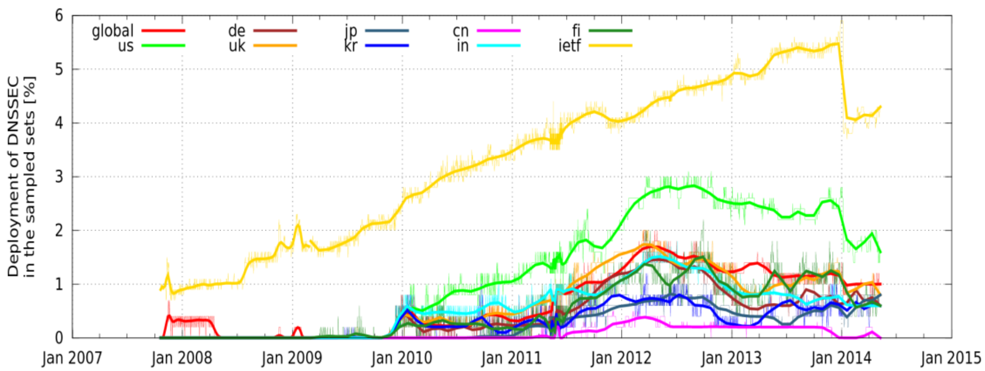


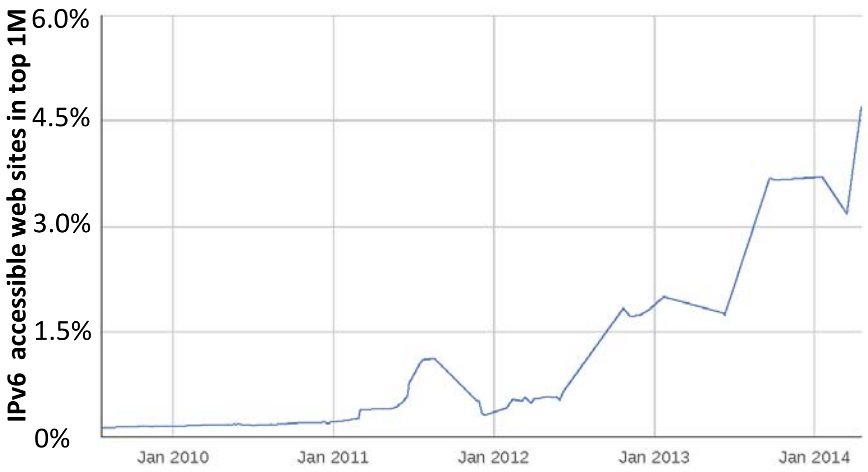


The Adoption Conundrum of Network Technologies

- And there are plenty of examples to illustrate the adoption challenges of network technologies & services

From https://eggert.org/meter/dnssec (sample of ~7300 sites)





- DNSSEC standard first published in 1999, but updated in 2005, and again in 2008
- Sweden deploys DNSSEC in 2005
- IANA signs the root zone of the DNS in 2010
- Still barely a few % of sites in 2014…

- IPv6 standardized circa 1998
- IANA allocates last block in February 2011
- World IPv6 Day in June 2011
- World IPv6 Launch in June 2012
- Still, it took IPv6 15 years to go from 0 to just over 45,000 websites (out of 1M)…

Framing the Problem

- How do we overcome the “chicken-and-egg” adoption dilemma faced by most network technologies and services?
- It is a serious problem that has affected or delayed the success of many network technologies
  - See IAB Workshop on Internet Technology Adoption and Transition (ITAT), Cambridge, UK, December 2013
- Bundling is a potentially useful mechanism to overcome initial adoption hurdles
  - I like A but don’t care too much for B, but will still adopt A+B and in the process help improve B’s eventual adoption
- Great idea, but when will it actually work (and not hurt)?
- Surprisingly, not much is known on the subject, and exploring the question turns out to be a lot of fun

What Do We Know About the Question?

- Two relevant bodies of work
  - Product and technology diffusion
  - Product and service bundling
- Much work in marketing research on product diffusion in the presence of externalities
  - Little or no work accounting for the impact of bundling
- Investigation of bundling strategies has focused on devising optimal pricing strategies
  - Models account for product demand correlation, and highlight the benefit of negative correlation
  - Until recently, externalities were absent from these models
  - Three recent works have explored bundling with externalities
    - All three focus on optimal pricing and assume independent demands, *i.e.*, no correlation in the values users assign to different products

Setting Things Up

- Modeling individual adoption decisions based on utility functions

$$V_i(x_i(t)) = U_i + e_i x_i(t) - c_i, \text{ where}$$

  - U_i is the user's (random) valuation for technology i (follows a certain distribution)
  - e_i is the strength of technology i's externality factor (how value increases with adoption)
  - $x_i(t)$ is the level of adoption of technology i at time t (varies from 0 to 1)
  - c_i is the adoption "cost" of technology i (resources, training, upgrades, acquisition, etc.)
- Adoption $\Leftrightarrow V_i(x_i(t))>0$, with equilibria such that $h_i(x_i^*)=x_i^*$, where $h_i(x)=P(U_i>c_i-e_ix_i)$
  - Rational users want to see positive utility from adopting
  - Equilibria when # adopters exactly matches # users with positive utility
- When bundling two technologies (1 and 2), the bundle's utility $V(x(t))$ is of the form

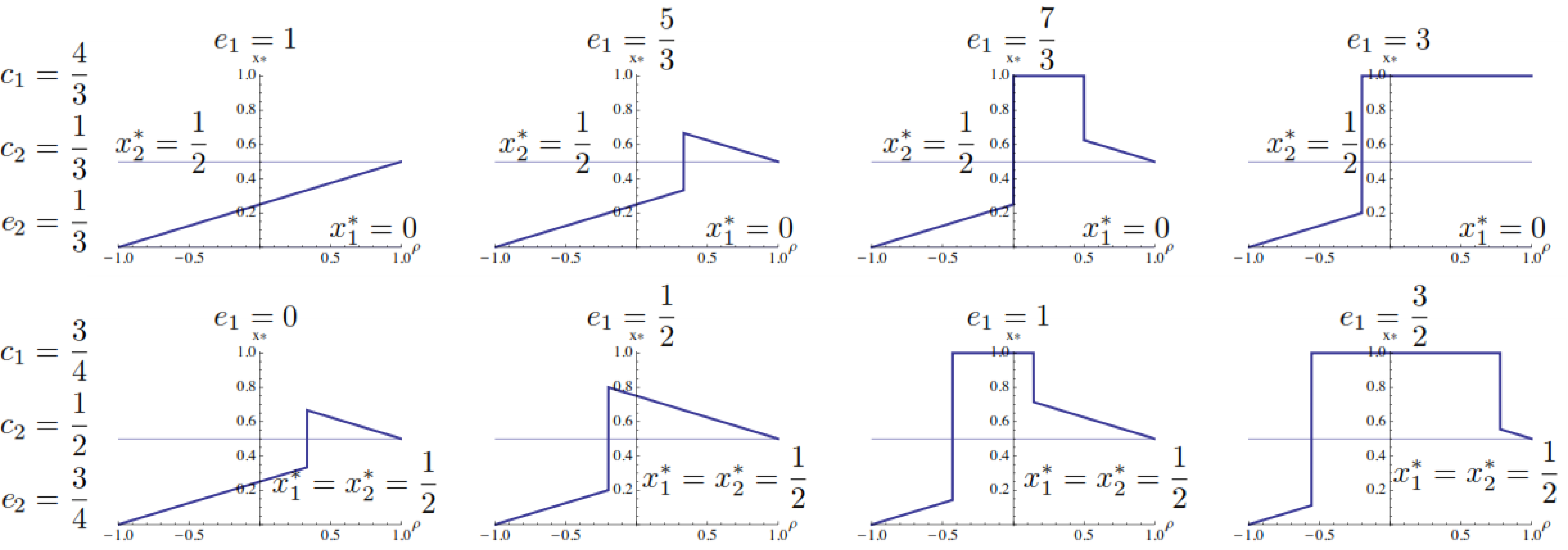
$$V(x(t)) = U + ex(t) - c$$

  - Where† $U = U_1+U_2$, $e = e_1+e_2$, $c = c_1+c_2$, and $x(t)$ is the bundle's adoption level at time t

The question we seek to answer is "**When is $x^* \geq \max\{x_1^*, x_2^*\}$?**" With a focus on the impact of the *<u>joint distribution</u>* $F(U_1,U_2)$; and in particular correlation

† Can be generalized to account for complements/substitutes and (dis)economies of scope

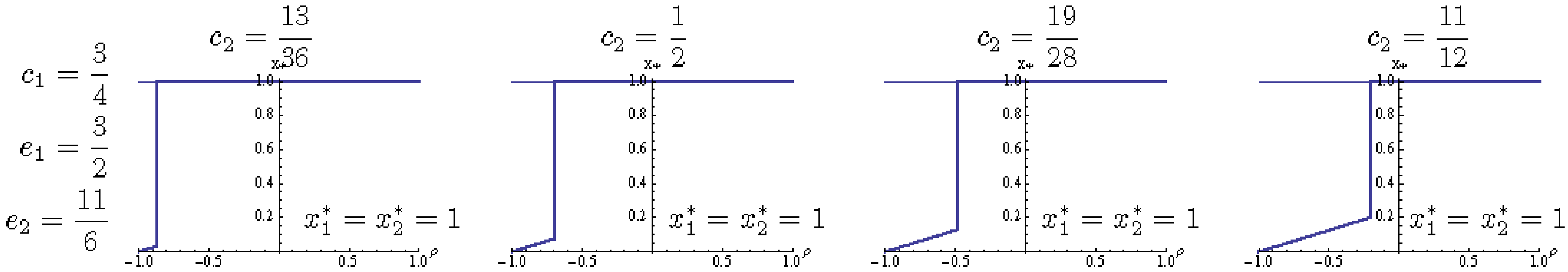
Win-Win Scenarios (simple distribution)



For WW outcomes: Choose technologies that are

1. (a) either heterogeneous in cost-benefit structure
   (b) or average (in cost & externality)
2. Sufficiently correlated in user valuation, but not too much

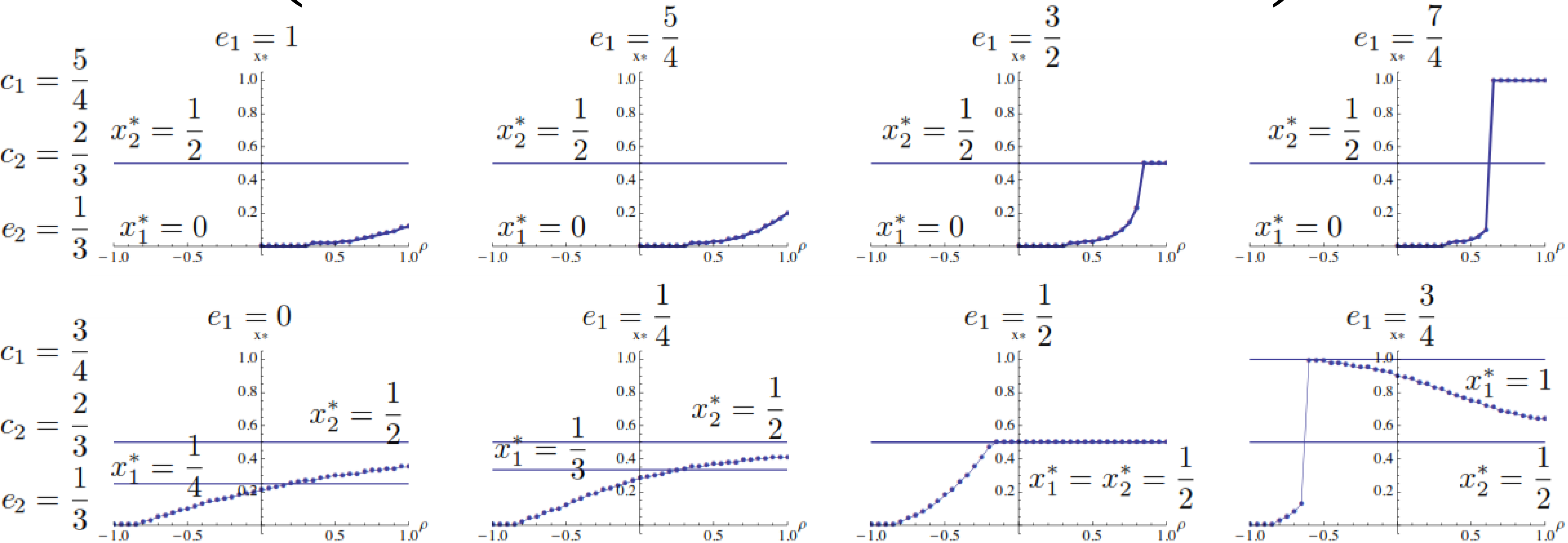
Lose-Lose Scenarios (simple distribution)



LL outcomes can arise when valuation correlation is negative enough

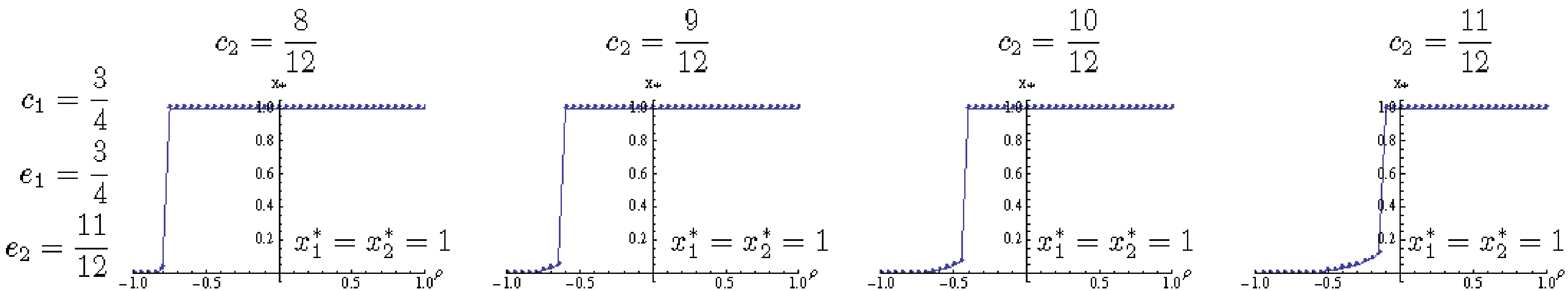
- Negative correlation means that few users like both services
- Can prevent early adoption phase to reach critical mass, *i.e.*, past the adoption level for which externality can start fueling continued adoption growth

WW Scenarios (a more realistic distribution)



- WW outcomes qualitatively similar in behavior
  - Correlation must exceed a threshold
  - Exceeding that threshold can be harmful

LL Scenarios
(a more realistic distribution)



- LL outcomes under similar conditions
  - Arise again mostly for negative correlation

Summary

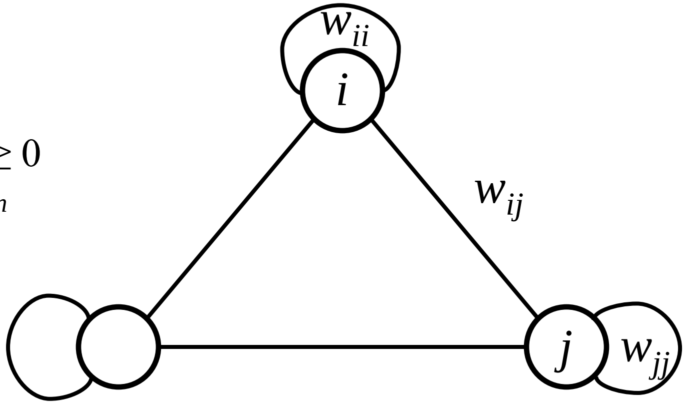
- Bundling as a tool (though clearly not the only one) to address the chicken-and-egg adoption problem faced by network technologies
- **Main finding**: Correlation in how users value technologies appears to play an important (enough positive correlation, but not too much)
  - Positive correlation helps attract early adopters beyond the critical mass needed for externalities to kick-in
  - Too much positive correlation creates a large pool of users that don't value either technology, hence potentially limiting overall adoption
  - Correlation creates "discontinuities" by affecting whether or not the critical mass threshold is crossed
- The results are obviously preliminary and call for further investigation and preferably empirical validation

DOES GREATER CONNECTIVITY MAKE FOR GREATER PARTISANSHIP?

The Role of Party Affiliation

- The ubiquity of modern communications means that we are constantly aware of, and possibly influenced by the opinions of our peers (or groups of peers)
- There is also no denying that we are living in an increasingly polarized world (stark divide between for/against with little in lieu of intermediate opinions)
- Are these two factors connected, and can we investigate this question in a principled manner?

A Generic Model of Opinion Formation in a Network



- A (fully-connected) network of n vertices
  - Symmetric interaction *weights* $\{w_{ij}\}$: $w_{ij} = w_{ji}$, $w_{ii} \geq 0$
  - Network *state* (opinions) $\boldsymbol{x} = (x_1,\ldots,x_n)$ in $\{-1,+1\}^n$
  - (State) *update* sums: $S_i(\boldsymbol{x}) = \Sigma_j w_{ij} x_j$
- Asynchronous state updates $\boldsymbol{x} \rightarrow \boldsymbol{x'}$: $x'_{i'} = \text{sgn } S_{i'}(\boldsymbol{x}) = \text{sgn } \Sigma_j w_{i'j} x_j$ (some i')
  - A weighted majority opinion $x'_i = x_i$ (for $i \neq i'$)
- **Given**:
  - Initial state $\boldsymbol{x}(0)$
  - "Honest" update schedule $\{i(k), k \geq 1\}$
  - Asynchronous update dynamics on $\{-1,+1\}^n$ $\boldsymbol{x}(0) \rightarrow \boldsymbol{x}(1) \rightarrow \boldsymbol{x}(2) \rightarrow \ldots \rightarrow \boldsymbol{x}^*$
- This basic system is known to always converge to one of (exponentially) *many* fixed points (equilibria) $\boldsymbol{x}^*$, s.t. $x^*_i = \text{sgn } S_i(\boldsymbol{x}^*)$ $(1 \leq i \leq n)$

Overall Approach

- Consider (initially) a full network, *i.e.*, all users communicate with (are aware of) each other and influence each other's opinions
- Introduce a two-party structure on the network
  - Users belong to one party or the other
- Investigate influence of *party* on opinion formation dynamics and equilibria
  - Starting from some distribution of initial opinions, how do opinions evolve and eventually settle?

- Two distinct models that account for party influence in different ways
  1. Party affiliation as the ***dominant*** factor in determining how users influence each other: **Random interactions** model
  2. Party affiliation as an ***indirect*** factor in determining how users influence each other: **Profile-based interactions** model

Random Interactions Model

- Influence weights between users are random, but biased as a function of party affiliation
  - Party affiliation ***directly*** affects users' influence on each other
    - Users in the same (different) party are more likely to influence each other positively (negatively)
  - In the context of a *two party* system: P_1 and P_2

$$\tilde{w}_{ij} = \begin{cases} w_{ij} & \text{if } (i, j) \in P_1 \times P_1 \text{ or } P_2 \times P_2 \\ -w_{ij} & \text{if } (i, j) \in P_1 \times P_2 \text{ or } P_2 \times P_1 \end{cases}$$

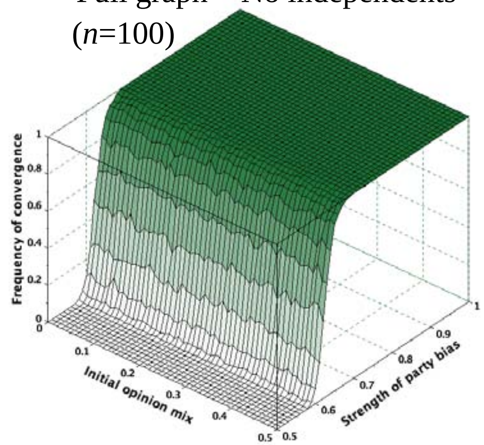
  - w_{ij} is a (+1,-1) Bernoulli r.v. with parameter $p \in (1/2, 1]$, where p captures strength of party bias
    - *Positive* bias for *intra-party* interactions
    - *Negative* bias for *inter-party* interactions

Main Findings

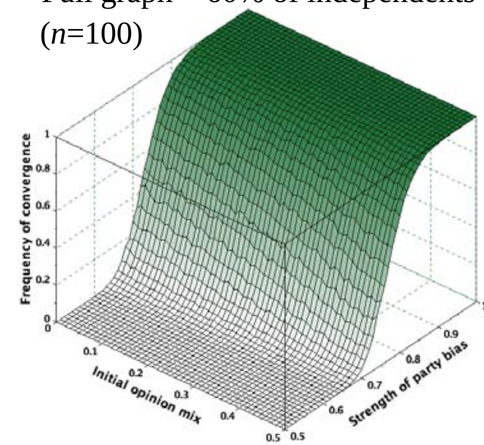
- Polarized outcomes arise with high probability (asymptotically approaching 1), even in the presence of only moderate party bias
  - This occurs even in relatively small populations (*n*=100)
- The result holds even under various relaxations
  - Presence of “independents”
  - Presence of “zealots”
  - Introduction of a network structure (Erdös-Rényi)

- In other words, when party affiliation is the dominant influence factor, <u>*polarized outcomes are now the norm*</u>

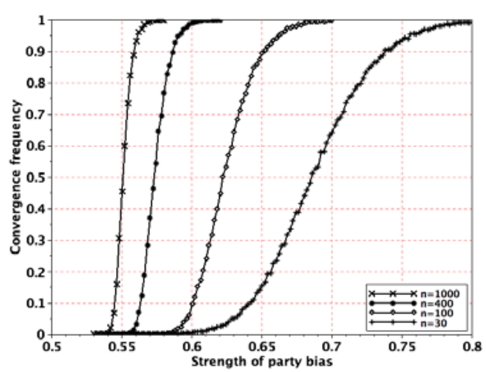
Convergence to polarized outcome
Full graph – No independents
(*n*=100)



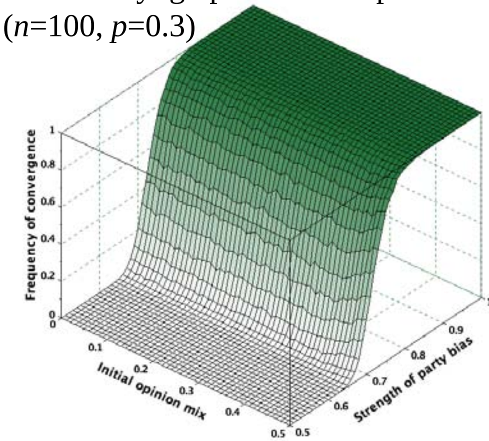
Convergence to polarized outcome
Full graph – 60% of independents
(*n*=100)



Impact of population size *n*
Full graph – No independents

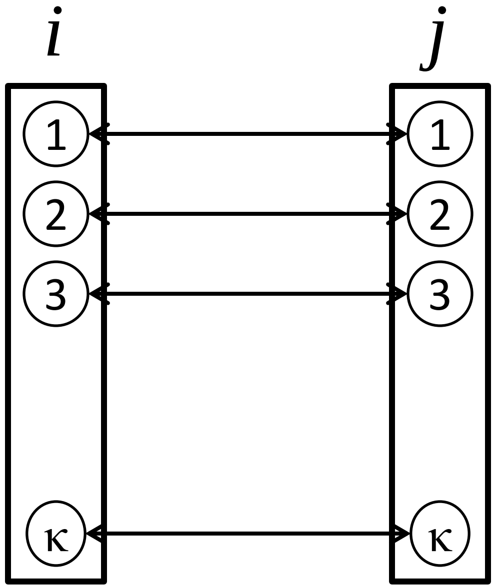


Convergence to polarized outcome
Erdös-Rényi graph – No independents
(*n*=100, *p*=0.3)



A Profile-Based Model

- Users are characterized by their *profile*
  - Voting record (for or against) on a set of κ previous independent issues
  - Node *i's* profile: $\pi_i = (\pi_{i1}, \ldots, \pi_{i\kappa})$



- Profiles are random, but party affiliation biases the odds of a for/against position on a given profile issue
  - i.i.d. signed Bernoulli random variables, each with a party induced bias

$$P\{\pi_{il} = +1\} = p > 1/2$$
$$P\{\pi_{il} = -1\} = 1-p < 1/2$$

Party P_1

- Users influence each other based on how well their profiles are aligned, *i.e.*, the number of issues on which they hold the same opinion
  - An ***indirect*** party influence model

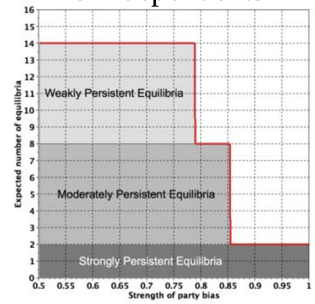
$$w_{ij} = \left\langle \pi_i, \pi_j \right\rangle = \sum_{l=1}^{\kappa} \pi_{il} \pi_{jl}$$

Main Findings

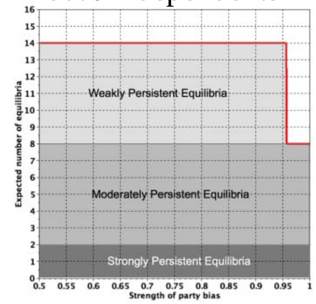
Profiles of size κ=3
Population size n=100

- Diversity of outcomes is preserved
  - Multiple possible equilibria (independent of population size, though dependent on strength of party bias and profile size)
  - Final equilibrium depends on initial opinions
- Heterogeneity of opinions within a party remains at equilibrium, even under heavy party bias
  - Heterogeneity declines in terms of both number of possible outcomes and level of dissent within a party, but does not disappear

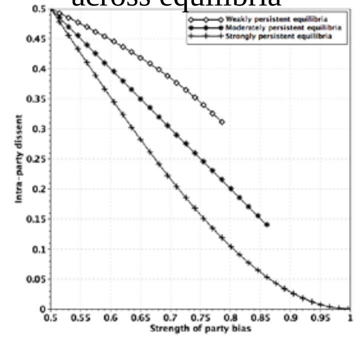
Possible equilibria
No independents



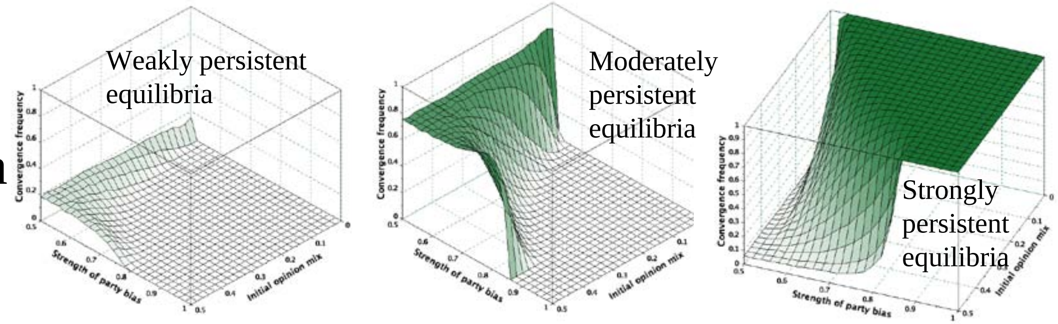
Possible equilibria
60% independents



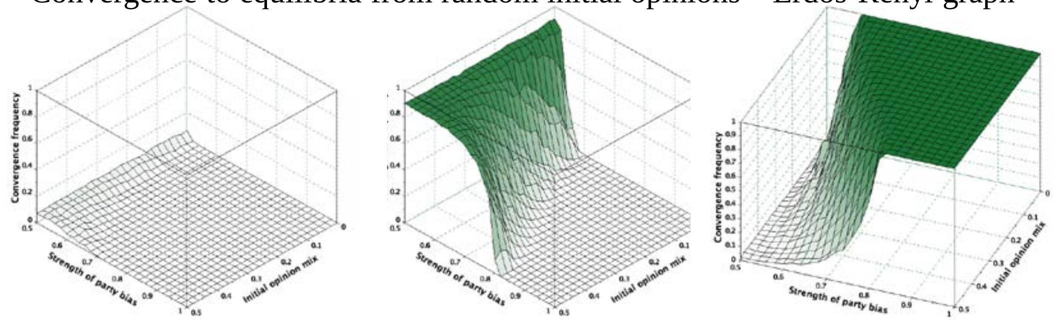
Opinion diversity
across equilibria



Convergence to equilibria from random initial opinions – Full graph



Convergence to equilibria from random initial opinions – Erdös-Rényi graph



Summary

- The glory days of Internet research are behind us
  - It is still possible to do interesting and fun work, but we are unlikely to again see the same level of visibility and excitement
- There is, however, a whole slew of new “networking” problems made possible by the ubiquitous connectivity the Internet affords
  - I offered two among many possible examples
    - Understanding how to effectively *upgrade* a large-scale network infrastructure
    - Opinion formation in partisan networks

References

[1] R. Guerin, J. C. de Oliveira, and S. Weber, "Adoption of bundled services with network externalities and correlated affinities." To appear in ACM Transactions on Internet Technology (early version available on ArXiv, October 2013).

[2] M. H. Afrasiabi, R. Guérin, and S. Venkatesh, "Opinion Formation in Ising Networks." Proc. ITA 2013 Workshop, San Diego, CA, February 2013.

[3] M. H. Afrasiabi, R. Guérin, and S. Venkatesh, "Spin glasses with attitude: opinion formation in a partisan Erdös-Rényi world." Proc. ITA 2014 Workshop, San Diego, CA, February 2013.