Internet Science

Dr. Heiko Niedermayer

Cornelius Diekmann, M.Sc.
Prof. Dr.-Ing. Georg Carle

Lehrstuhl für Netzarchitekturen und Netzdienste
Institut für Informatik
Technische Universität München

Version: July 9, 2014

Agenda

Slides: 27

Introduction

Introduction

- Understanding the Internet as IP, HTTP, and HTML may not even nearly cover all aspects of the actual Internet.
- In particular with respect to all the processes that drive all kinds of activities in the Internet and also in its definition, operation, and management. Most of all also its usage, the things people create using it.
- In this chapter, we want to briefly sketch how human influence among other kinds of non-technical influences can look like.

Internet Science

- Internet Science is a activity of the EU to promote interdisciplinary Internet research.
- internet-science.eu

Mobility

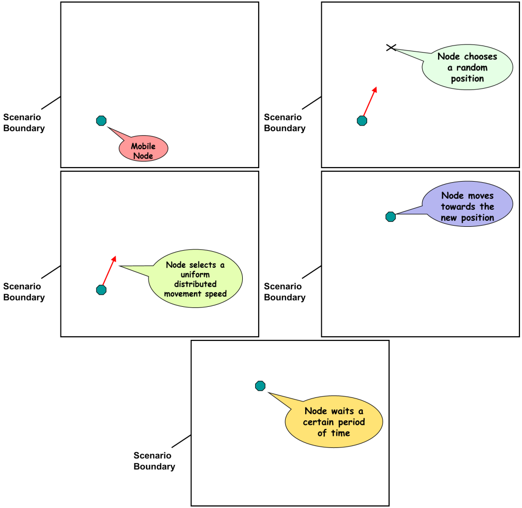
Mobility

- One aspect of human influence is user mobility. In a mobile network, the user moves around walking while using the network. In the beginning we have users move their home once their neighborhood did not meet their requirements anymore. In fixed networks, users might connect to the network from different places over time, yet they do not move while being connected.
- Mobility patterns can be influenced by what exists in the area of interest. Underground train lines and streets may lead to certain mobility or directional movement patterns. Interesting places attract people to move to the place and stay there for some time. This may be an entry of the underground, a lecture hall, . . .
- When it is not one particular area that is of interest for the research, researchers have generated random mobility patterns to study the impact of mobile nodes in a network.

Random Walk and Random Waypoint

- Random Walk
  - Each node walks in random direction with random speed for random time.
- Random Waypoint
  - Select random destination
  - Move to destination with random speed
  - When destination reached, stop.
  - Wait for random period of time. Then repeat process.
- Obstacles in the area and borders of the area are addition issues to consider when realizing these behaviors.

Random Waypoint



Popularity

Popularity

- Mobility was about the location of the user. Now, we are somewhat interested in the location of what is requested.
- Zipf's Law states that popularity of resources follows a power law distribution.
- Measurements in P2P networks (Gnutella, eDonkey) suggest that requests for popular videos were lower than a power law distribution would suggest.
  - Explanation via Fetch-Once: Popular videos gain their number of views also from repeated views. If the files has been downloaded already, the only get one download for multiple views. Unpopular content, however, may only get one view and one download.

Popularity over Time

- This Zipf's Law discussion misses the timing aspect of videos. Even the most popular video may not get a lot of views long after its release. So, the views for a particular day may differ a lot from the overall distribution.
  - This is important for studying caching where a constant distribution that follows Zipf's Law would suggest that you just cache the all-time most popular moves until your cache is full. And the people would forget to look the currently big movie because it might not be that big to make it to the all-time top.
- Models for popularity need, thus, to account for temporal locality (interest over time) when popularity of items might change during the time of study. One example is the Shot Noise Model by Ahmed, Traverso et al.

Popularity over Time - Classifying Request Behaviour

- For social media and YouTube videos, there have been proposal to classify popularity curves into:
  - Junk: most all-time views on one day
  - Quality: peak day with most views has a significant fraction (e.g. more than 20%) of all-time views
  - Viral: peak day does not account for a significant fraction of all-time views (e.g. less than 20%)
- Our own studies show that the classifications have to be adapted to the particular situation.

Game Theory

Game Theory

- From popularity we might now know what people might want. Now, we might start to model how they try to get there.
- We have seen Braess' Paradox in the introduction chapter of the lecture. People select their way to be optimal. Given all people selecting their optimal way, the way may be congested and less optimal. If some had chosen a less optimal way instead, all would have been better of.
- Game Theory adds a strategic dimension. This can be used to model economic aspects (best movie for 5 € or 2nd-best for 1 €?, invest and upgrade network in Garching to support better data rates or let is as it is?), other kind of decision making (privacy or sharing with friends on Facebook), and also security (cooperate or leach in file sharing networks?, attack your local network or remain friends with admin?).

Game Theory

- For simplicity, we now assume problems with 2 players, namely row player and column player and we write down the game as a table of actions and payoffs.
- Assumption: The players do not know what the other player plays and they do not have means to coordinate.
- The actions of the row player are given in the rows, for the column player in the columns.
- Each cell of the table contain an entry (row player payoff, column player payoff).

row player \ *col. player*	A	B
A	(0,0)	(1,2)
B	(2,1)	(3,3)

Table: Game 1

So, row player will get 2 or 3 in payoff if she choses B. The same is true for column player. In this game, the only rational choice is to play B.

Prisoner's Dilemma

row player / *col. player*	Silent	Betray
Silent	(-1,-1)	(-3,0)
Betray	(0,-3)	(-2,-2)

Table: Prisoner's Dilemma

- The negative payoffs refer to the years in prison.
- Actions are remain silent or betray the other and get benefit for cooperating with the authorities.
- What shall you do?
  - If you remain silent, the other's best strategy is to betray (free instead of 1 year).
  - If you betray, the other's best strategy is also to betray (2 years instead of 3 years).
  - Thus, the optimal strategy is to betray.

Prisoner's Dilemma

row player \ col. player	Silent	Betray
Silent	(-1,-1)	(-3,0)
Betray	(0,-3)	(-2,-2)

Table: Prisoner's Dilemma

- Note that while it is the best strategy for both to betray, the best cell in the table with respect to overall payoff is both remain silent.

row player \ col. player	Cooperate	Defect
Cooperate	(2,2)	(0,3)
Defect	(3,0)	(1,1)

Table: Prisoner's Dilemma with positive payoffs

Game Theory - Terminology

- As a player in a game, we have learned that we cannot simply take the action with the highest payoff in some cell. The other player may chose an action that leads to a different cell.
- In the games players follow strategies on how to pick their actions. Examples: pick action A, pick action B, randomize between A and B.
- So, we always have to think what the other might chose and then select an action with the highest payoff under this constraint (Best Response).
- A pair of strategies to select an action is a Nash Equilibrium if both strategies are best responses to the other strategy. Note, Nash Equilibria have the notion of being optimal strategies for the players, yet as we have seen for Prisoner's Dilemma, not necessarily the best outcome overall.

Iterated Prisoner's Dilemma

row player \ *col. player*	Cooperate	Defect
Cooperate	(2,2)	(0,3)
Defect	(3,0)	(1,1)

Table: Prisoner's Dilemma with positive payoffs

- Assume, we play the game twice.
  - If we first try to cooperate, we might get more payoff than the (defect, defect).
  - However, in the second game, we always are better of if we defect.
  - This means, the other will defect, this means, oh oh oh.... both will defect. First game become like last game, more gain if player defects also in first game (no loss later because both will defect in second game).
  - Again, the best strategy is to defect twice.

Iterated Prisoner's Dilemma

row player \ *col. player*	Cooperate	Defect
Cooperate	(2,2)	(0,3)
Defect	(3,0)	(1,1)

Table: Prisoner's Dilemma with positive payoffs

- Assume, we continue to play another round with probability p (say p=0.75).
  - To cooperate, means getting 1 payoff less.
  - To defect, means to get 1 payoff more once, but then the other might not cooperate anymore and also defect. So, we lose 1 payoff every game afterwards.
  - Expected number of additional games: GEOM(p), $E[X] = \frac{p}{1-p}$, here $E[X] = \frac{0.75}{0.25} = 3$
  - Defect means to gain 1 in one game and lose 1 in 3 games on average. Thus, better to cooperate.
  - Note, the uncertainty about the end makes cooperation beneficial.

Bounded Rationality

Bounded Rationality

- Game Theory studies strategic decision making where all actors make rational choices.
- In reality, people deviate. Bounded rationality tries to explain that by limitations that lead to deviation from optimality.
- Limitations = bound for resources used in human decision-making:
  - Information known (uncertainty)
  - Cognitive limitiations
  - Finite amount of time
- Note that while traits like altruism can be covered by rational game theoy (more pay-off if money given to others), some aspects of irrationality come into systems due to different goals and values (designer / reseacher vs actual human).
- Note that bounded rationality assumes a rational choice under limitations. Behaviorism assumes irrationality as driving force.

Application of Game Theory, Bounded Rationality, ...

- Game Theory, Bounded Rationality, Behavioursm and alike can clearly be used when human actors are directly involved.
  - Social networks, social media, peering of ISPs, ...
- But how to introduce them into lower technical layers of infrastructures where most is automated?
  - Fight deviations from intended protocol. People, Operators, etc. might get benefit from modifying protocols and values in their favor.
    - e.g. in BitTorrent do not upload to peers that do not upload (Tit-for-Tat)
  - It might help to model and understand how the infrastructure is developed, built, operated, and used.
    - Things may fail that are not considered to fail in system design. Typical assumptions: Bob's browser will only send correct TLS requests to Alice. The CA issues certificates only to the right and correct entity. The keys never get compromised. Thus, authentication always correct. Maybe assume all assumptions are only true probabilistically?
    - Expect and model potential deviations in entity behavior. Under what deviations will it still work?

Reader: A Case Study

Recommended Reading:

Recommended Reading:
Robert Axelrod, Rumen Iliev, Timing of cyber conflict.
`http://www.pnas.org/content/early/2014/01/08/1322638111`

Thanks for your attention!

Questions?