AsiaFIF workshop
24, Aug, 2009
Cheju, Korea

Measurements and Modelling Research of Internet at BJTU

Yuchun Guo
Network & Communication Lab
Beijing JiaoTong University (BJTU)
http://netcomm.bjtu.edu.cn
Email: ychguo@bjtu.edu.cn

Overview

- Why measurement?
- Measurements and modelling
  - on network level
  - on application level
- A vision of the future

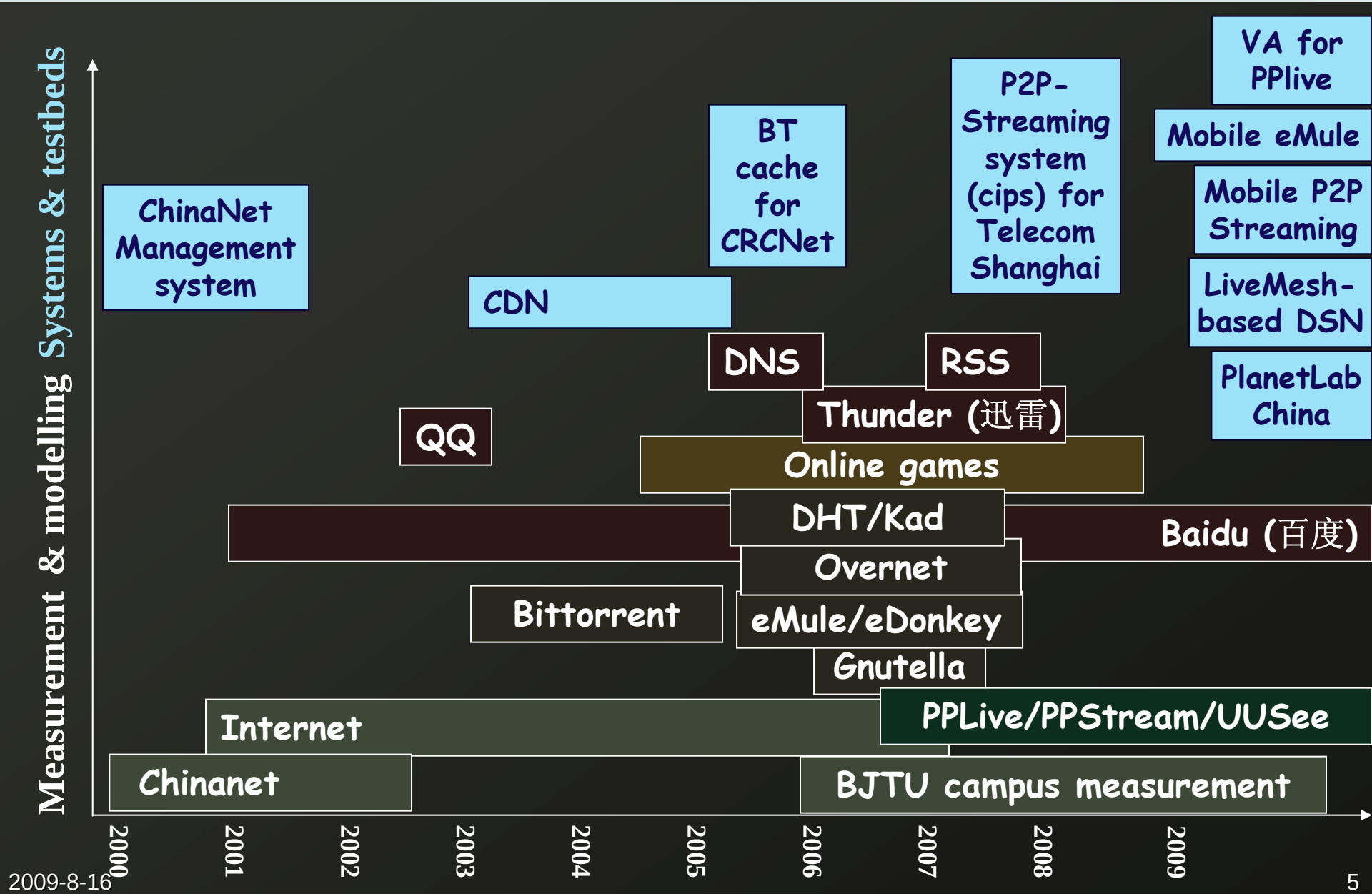
Why measurement?

- Internet
  - far from a totally engineered network like the telephone network
  - a complex network in the dark
- Academics, Administrators and Industries
  - little situational awareness regarding global dynamics and operational threats
  - traditional modelling/simulation methods not fit
- Real host of the Internet
  - Common people who use the Internet just for FUN
    - Human nature, not much engineering targets
- Finding what it is like via measurement is the base for research of the Internet
- Dilemma of measurement
  - The application owners do not like their networks to be probed by others, even the academics

Why measurement?

- What have been found about Internet via measurement?
  - Power-laws and small world phenomenon in topology and applications
  - Self-similarity and long range dependence in traffic
  - Problems in BGP routing
  - …
- What have we learned from our experience?
  - Targets selection: Popular applications
  - Measurement methodologies: by all means that may help
    - Cracking application protocols is critical but not the only way
  - A model is necessary before the measurement but the expectation for data collection via measurement is always not fulfilled, modelling level needs to be adjusted based on the achievable granularity of measurement
  - Measurement-analysis-modelling-measurement-…

Measurements and modelling



Network-level

- Measurements based on the accessibility to network management
  - Netflow / MRTG traces
  - Router records
  - Server logs
- ChinaNet
  - Netflow / MRTG
- Public datasets
  - Routeviews, Whois, CAIDA, NLANR

Work based on ChinaNet datasets

- Model AS traffic properties and AS sizes, eps the observed sizes
- Infer the traffic demands and routes between nodes from the link traffic and topology
  - MRTG data is not as irrelevant but may leak critical network configuration and application characteristics
- Identify different applications (e.g. P2P, QQ/MSN) from traffic traces of coarse granularity
  - IP address / port # inference combined with the understanding of protocols or behavior properties
- Find that network sampling underestimates unfairness in bandwidth sharing, and validate by theoretical models and simulation results
- Model the ‘fetch at most once’ property of the content popularity in P2P networks based on ChinaNet traces and BT traces locally measured

Work based on the public datasets

- Design a large-scale topology visualization tool BOSAM to compare networks
- Infer the geographical coverage strategies of ISPs in network deployment based on mining of AS topology geographical coverage properties
- Analyze the temporal evolution of AS topology in 10yrs
- Identify hijacks and attacks in BGP routing, investigate the effectiveness of routing policies in ISPs' competition, model the effective range of routing policies based on the AS relationship
- Evaluate the possibility of flat routing (VRR/ROFL) on AS level, evaluate several flat routing schemes and the critical factors (distance between ASes, distance of an AS to the core, and AS degree)
- Study the convergence property of BGP routing, propose the "Anti Loop" routing algorithm with good route convergence

Application level: measurements

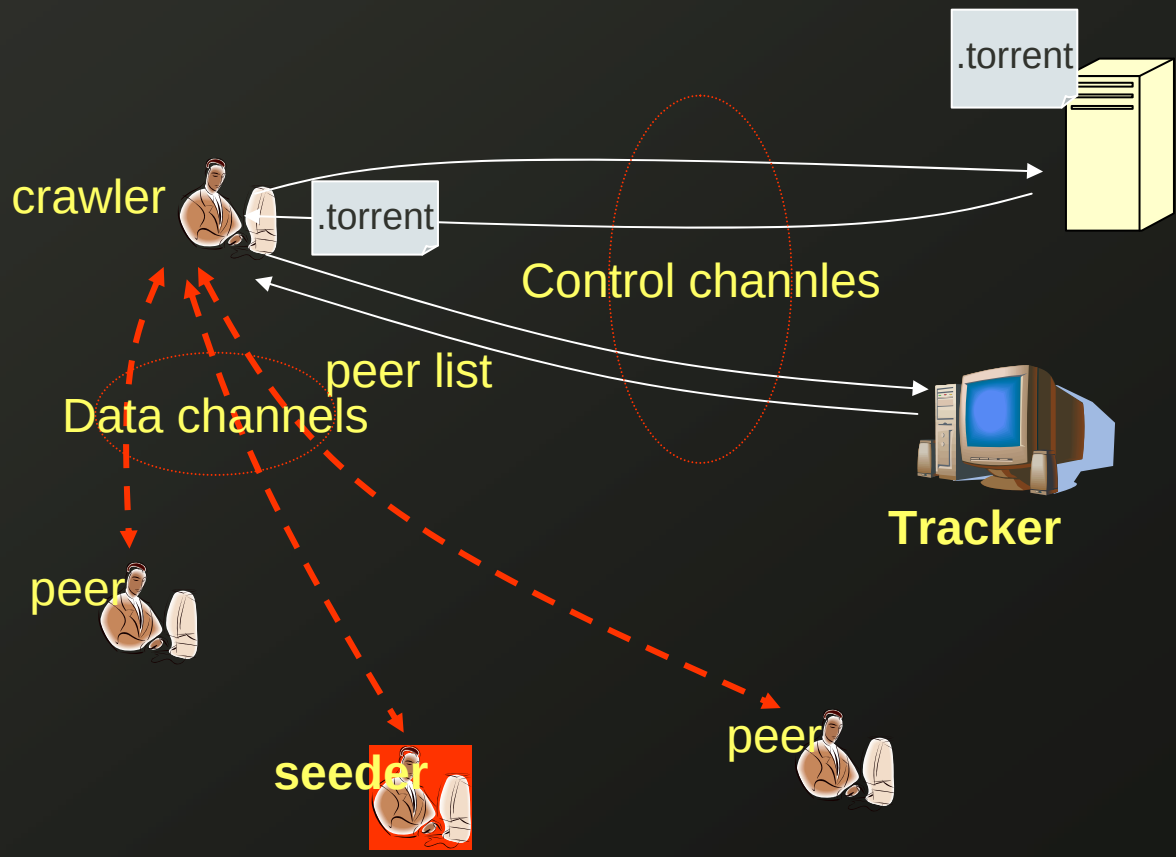
- File sharing systems
  - BitTorrent
  - Gnutella
  - eMule/eDonkey
  - Overnet
  - Kad
  - Thunder (迅雷 Xunlei)
- P2P Streaming
  - PPLive, PPStream, UUSee
- Online games
  - Crazyracing
  - Crazytank
  - World Of Warcraft (WOW)
- Others
  - Instant messaging
    - QQ
  - RSS
  - Blog
  - Search engine
    - Baidu
    - Korea
- Systems
  - BT cache
  - CIPS P2P streaming
  - CDN ?

Application-level: methodology

- BJTU campus network
  - GIGEMON System (DAG 4.3GE/Endace: Gigabyte Ethernet Monitor
  - Closed measurement
    - Single vantage point, NOT enough for a global view of applications
- Protocol analyzing via reverse engineering: Almost all P2P protocols
- Constructing the crawler measurement platform
  - Single PC / distributed platform based on Linux-boxs
- Identifying what can and cannot be probed with crawlers
  - Constructing a website and conducting loop-back tests to make up for the limits of the crawler
- Identifying the possibility of large-scale network measurement and the advantages and disadvantages for the network

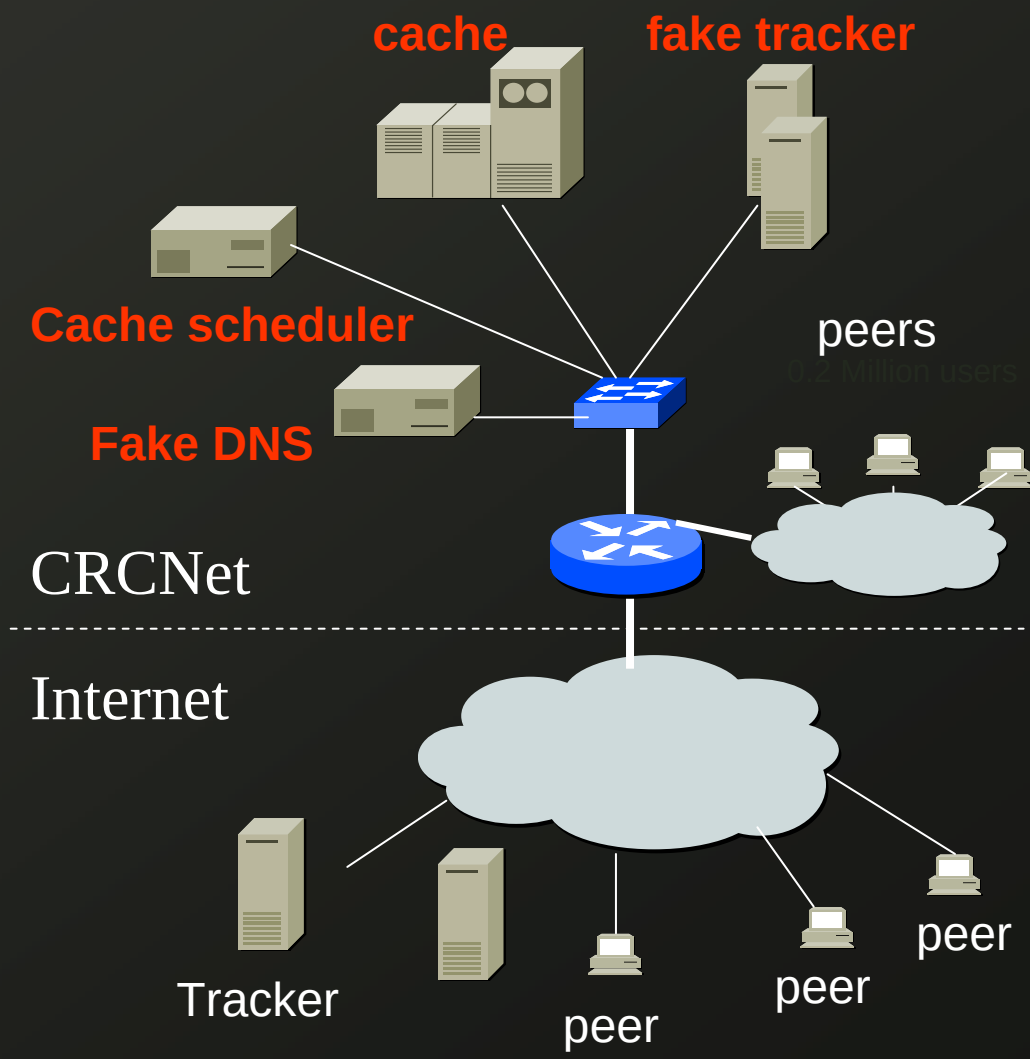
Bittorrent

- Measurements
  - Built the BT crawler platform
  - Obtained the BJTU BT website ‘Chenguang(晨光)BT’ Tracker log
  - Crawled the .torrent web (ImageGarden)
  - built the BT website ‘北京大学生网站’ (Beijing Undergraduates)
  - Linux RedHat9 downloading trace



Bittorrent

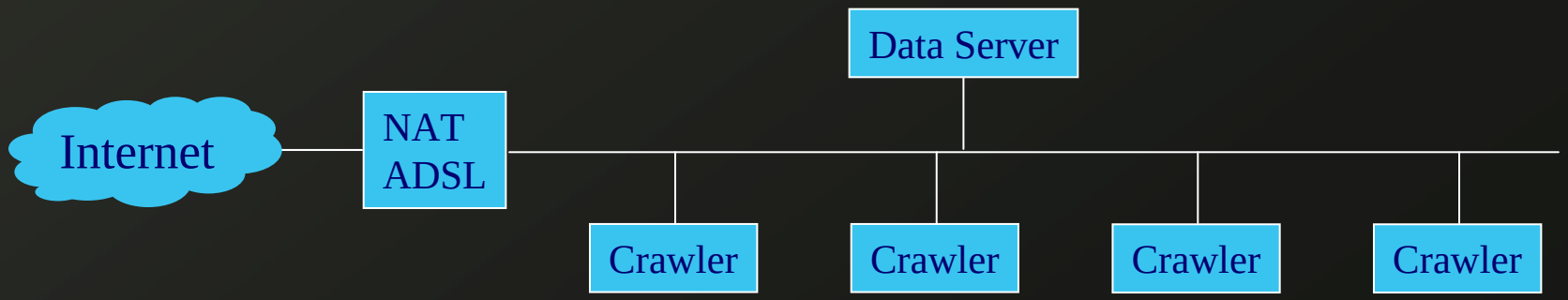
- System dynamics
  - Peer/seeder dynamics
    - Downloading dynamics
  - Content dynamics
    - The distribution process of each piece
  - Topology dynamics
    - The dynamic state of connectivity
- User behavior
  - Torrent lifetime
  - Seeding period, share factor
  - Content contribution: show-off, power-law
- P4P
  - BT cache system
    - ISP faking trackers / limiting inter-ISP traffic



BT cache system

Gnutella

- Methodology study: **Crawling Platform optimization**
  - Topology crawling
    - Recursive (BFS) / **Periodical long-term tracking**
    - snapshot / **dynamics (churns)**
    - Ultras / **all**
    - # Crawlers (LinuxBox) / **Bandwith limitation (ADSL)**
  - Fileinfo crawling
    - Webpage downloading / Recursive probes
- Analysis and modelling
  - Topology structure and churn features based on dynamic crawling
  - Files sharing and distribution properties



Kadmelia / Overnet, eMule, Azureus

- The Kadmelia algorithm supports the p2p applications greatly in practice but receives not much reports in measurement research
- Topology inference
  - sniffer the route-table maintenance per user
    - Distribution and dynamic of k-bucket
  - Infer the topology
    - Fetching others' route-table (not trivial on Overnet)
- Sensing network activities
  - Fake ID, one million
    - Practical 400,000 users, 24 of them having about 10,000 IDs

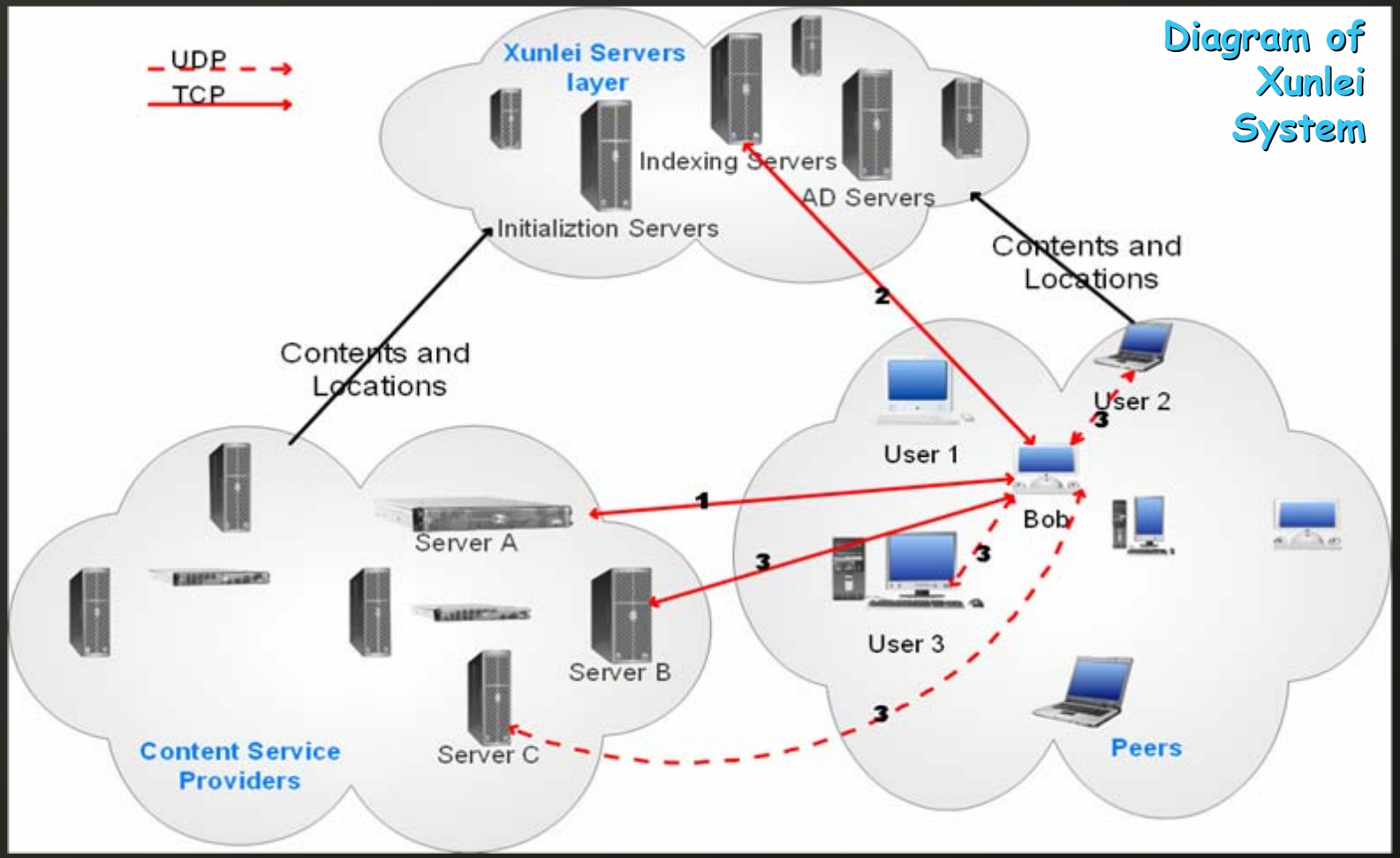
eMule/eDonkey

- Identify the eMule traffic from the passive measurement traffic traces
- Find the existence of Fake servers and the interaction between users and the fake servers, Trace and record users' server state
  - Prove that effects of fake servers are limited
- Harvest the filelists over eMule servers
  - For future P2P search
- Compare the keywords semantics in P2P networks and that in non-P2P context
  - Keywords contained Not in the text/document but in the file-names
  - Content unreadable (audio, video, …)

Xunlei (迅雷, Thunder)

- Most popular downloading platform in China
  - But nobody knows how XunLei works
  - denounced because of its behavior of downloading files from unclicked servers.
- Measurement
  - Challenges: Private protocol, failed in hacking protocol
  - Tools: Wireshark (sniffering)
  - Behavior backward inferring
    - source and destination IP addresses, protocol, bytes on packet level
  - Software downloading
    - ownership support
  - P2P content
    - Acceleration, the pioneer of VA

Xunlei


UDP
TCP
Xunlei Servers layer
Indexing Servers
AD Servers
Initializtion Servers
Diagram of Xunlei System
Contents and Locations
Contents and Locations
2
User 2
3
User 1
1
Bob
Server A
3
3
Server B
User 3
3
Content Service Providers
Server C
Peers

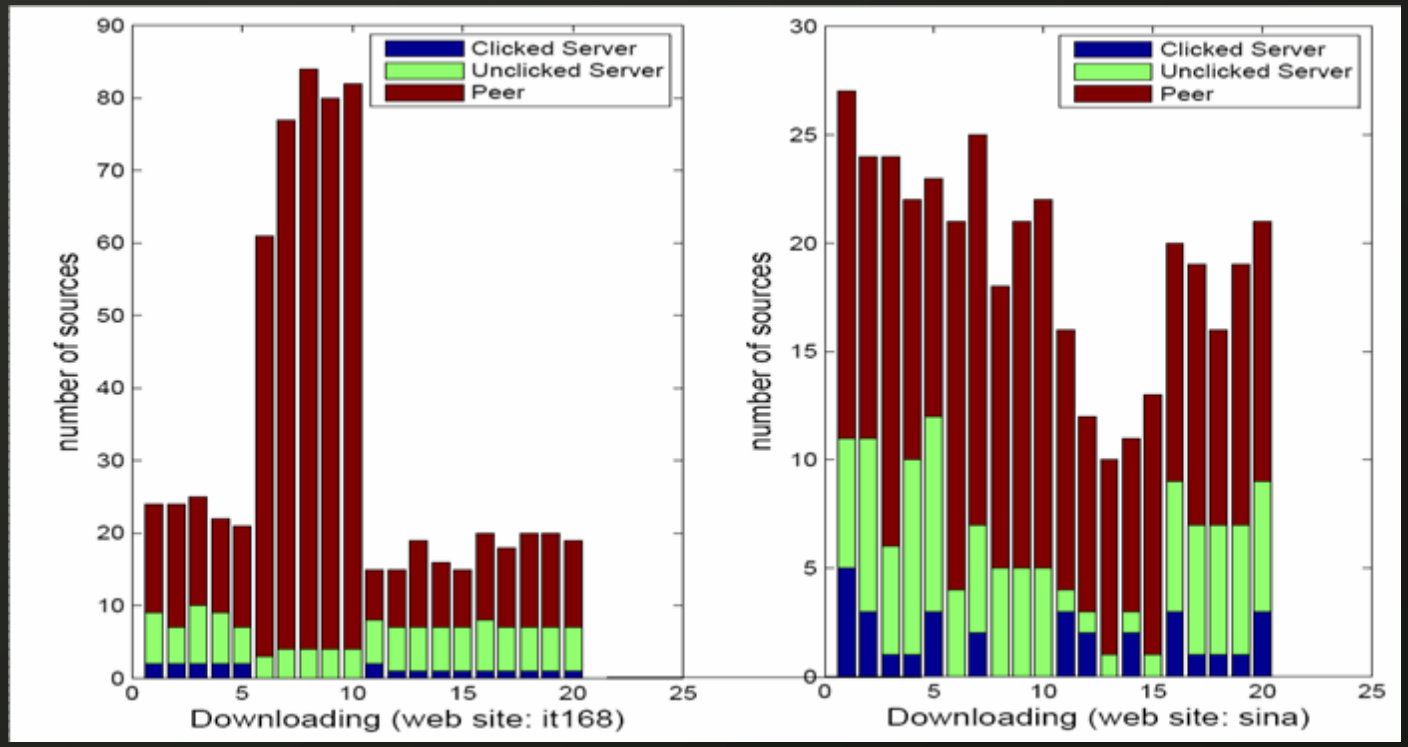
Xunlei

Click a link from site A → **Xunlei servers provided sources** → **Download files from multi-sources**

Which kind of sources? Who provide data?

sources

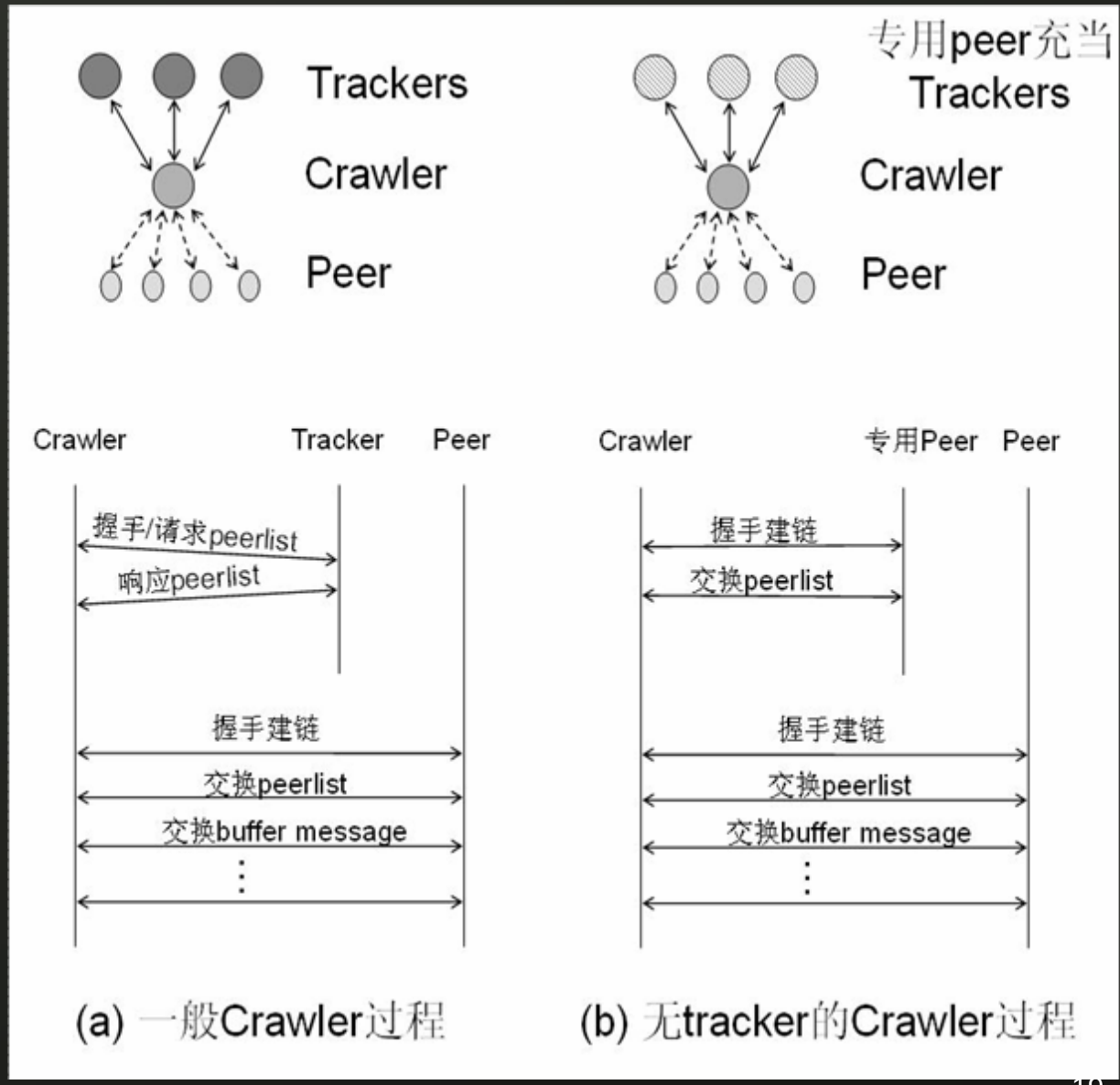
- Clicked servers
- Unclicked servers
- peers



Number and types of sources Xunlei clients use to download data (forty experiments each to two portal sites)

P2P Streaming

- PPLive/PPStream/UU See
- Live/VOD
- reverse-engineering / automatic client sniffer
- encrypted tracker
  - crawler bypassing tracker
- Key protocol messages' parsing
- ADSL limitation
- Flush-crowd event measurement (such as Spring Fest., NBA)



P2P Streaming: Live

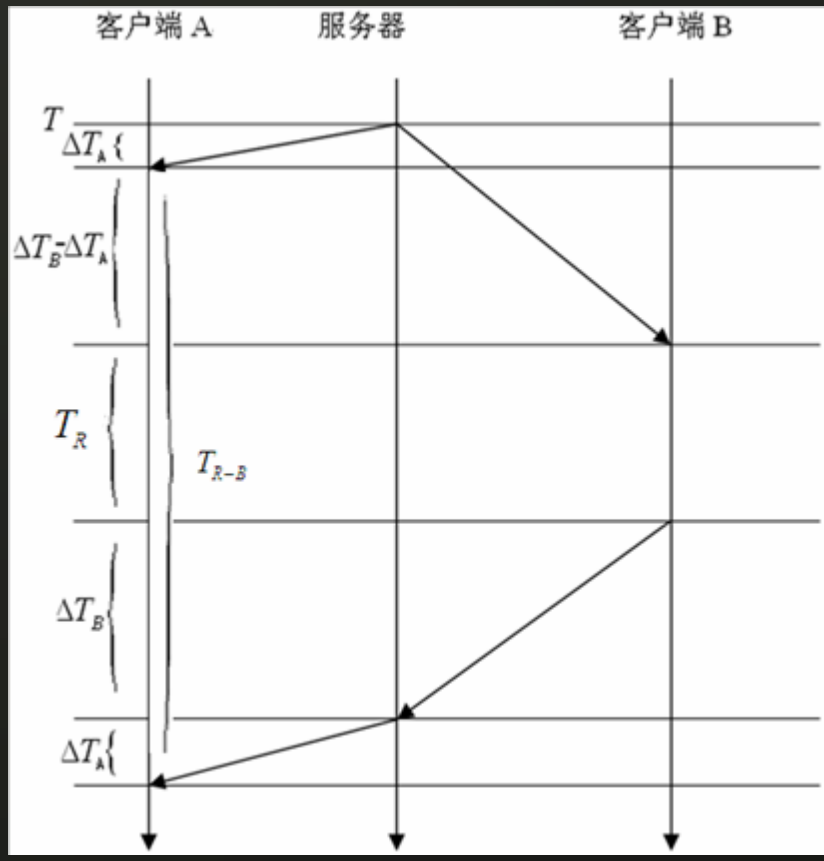
- User behavior
  - The traditional queueing model is not applicable
    - User's arrival and leaving Not depend on service time but on program-schedule
    - distribution of service time per user are not same
- System design of PPL, PPS, UUsee
  - Rate and timing
  - Streaming segmentation and index
  - Buffering strategies
  - Reference model for p2p live streaming
    - Critical parameters (e.g. set-up time, initial buffer offset,…)
- System dynamics
  - Sharing environment
  - Buffer occupation: temporal distribution and assemble distribution
  - Flush-crowd and Rate changes: effects on the system
  - Transport protocol selection: TCP or UDP

P2P Streaming: VOD

- Single channel
  - 1/3 watching, 2/3 only contributing
  - 2/3 continuously watch, 1/3 jump
  - User watching behavior affect the network sharing of pieces
- Multiple channel
  - Evaluate the conflict between provision slow response and highly dynamically changed demand of resources
    - for each channel at a time
      - # watching users
      - # available copies
  - Managing the availability of content according to the number of watching users: P2P storage of replications
    - Improving the performance based on the measurement of replication strategies in PPLive

P2P Online games

- Wow
  - API provided by the game provider: Who
  - Propose an algorithm to
    - find groups from the measured admixture of all groups in a COPY
    - Maintain a history of group-ship between user pairs
  - Design a reputation system based the group-ship for Anti-cheating
- crazyracing (跑跑卡丁车)
  - snapshot provided by the game provider
  - Estimate the network delay of remote player to the server based on
    - observation of an event on local terminal
    - the remote players' response time to this event



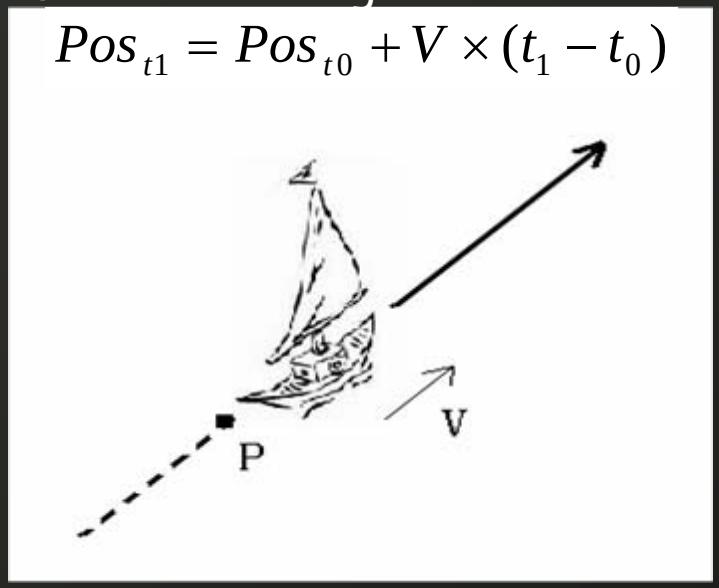
$$T_{R-B} = 2\Delta T_B + T_R$$
$$\Delta T_B = (T_{R-B} - T_R)/2$$

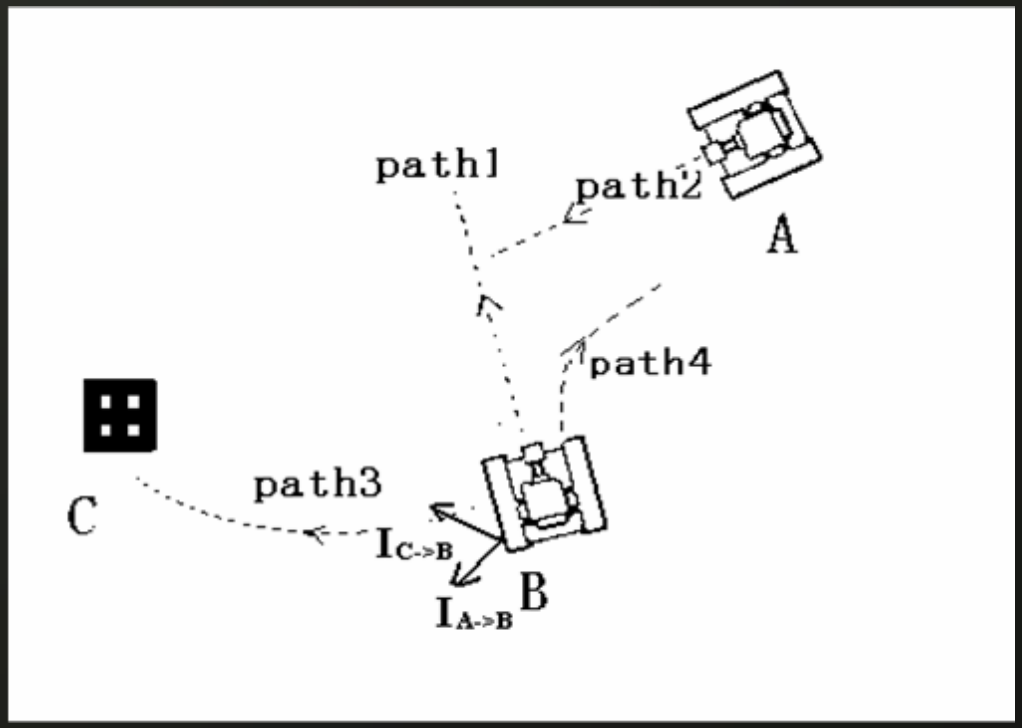
P2P Online games

- Crazytank
  - Set up a server and modify the protocol
  - Prediction of game paths
    - Original: Speed and accelerate
    - Modification: environment sensing, including scenario, counter players, game script

Dead Reckoning

$$Pos_{t1} = Pos_{t0} + V \times (t_1 - t_0)$$





Others

- QQ/MSN
  - Passive traffic traces
    - Traffic extraction / Protocol analysis
    - Temporal periodical properties
    - User behaviors
- Search engine
  - Korea musical top hits
  - Baidu video / audio search
- RSS
  - Repetition of top-lines
- Baidu Zhidao

Korean Musical Search Charts

- The web site www.naver.com covers than 70% of the Korean market,more than 35M Users.
- The naver-charts show the top-10 search keywords (singer, song,…)
  - refresh time ~ 6 seconds.
- We collected the naver-charts in one week, from July 9th to July 17th, 2008, kept records of top-10 keywords per 20s
- Observations:
  - Gaps between the singer rank and the song rank
  - Special users' influence on the ranking
    - Special searching pattern "singer+song" by special users

Gaps between the singer rank and the song rank.

SINGER RANK	SINGERS	SONGS	SONG RANK
1	A	a=g'	7
2	B	b	NONE
3	C	c	NONE
4	D	d=f'	6
5	E	e=e'	5
6	F	f=b'	2
7	G	g=c'	3
8	H	h	NONE
9	I	i=a'	1
10	J	j	NONE
NONE	DD	d'	4
NONE	HH	h'	8
NONE	II	i'	9
NONE	JJ	j'	10

A= Lee Hyoli, a= Lee's third album
E=SG Wannabe, e=the song of "lalala".
I=MC Mong, i="circus"

Vision for the future

- Applications to be measured (continuously)
  - P2P streaming system
  - Social network
  - Online multiplayer games
- Some trends
  - Wireless network
  - Evolution of application styles

Applications to be measured continuously

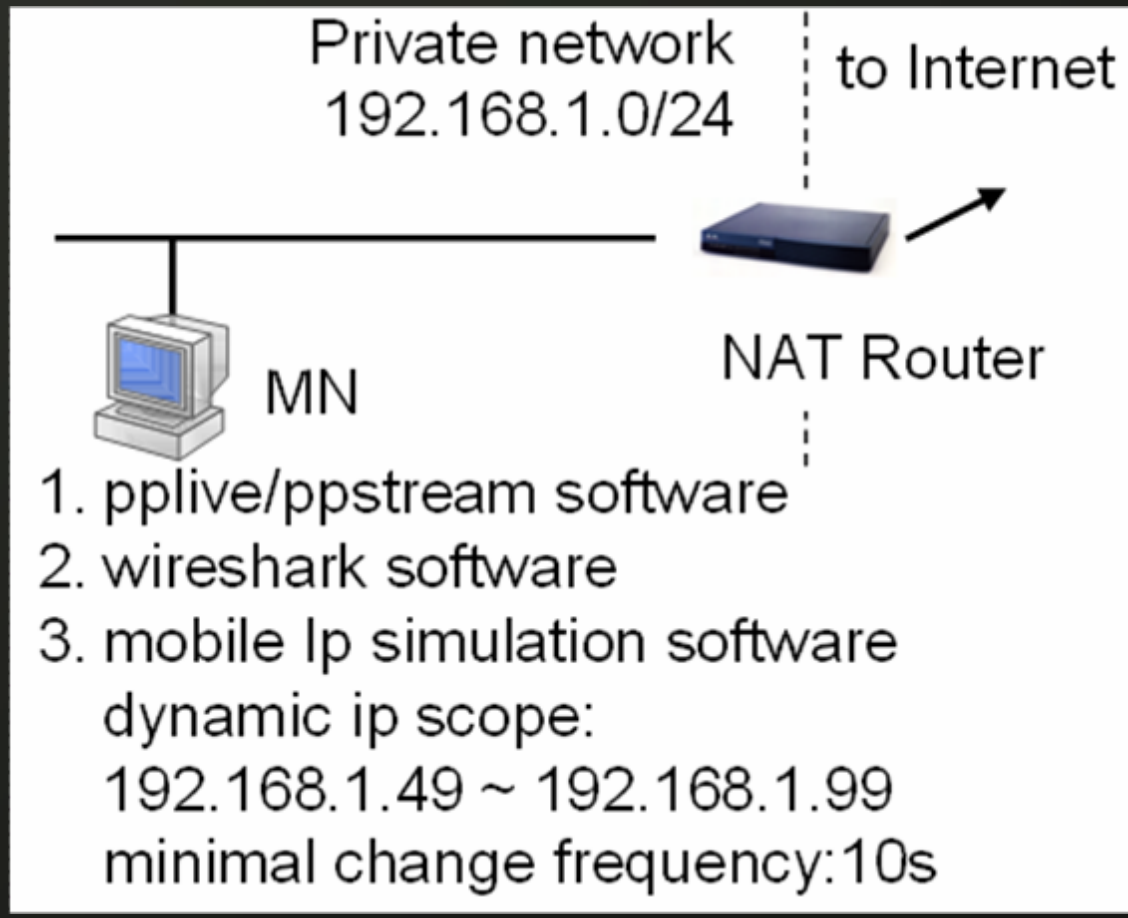
- P2P streaming system
  - User behavior
  - Cache / replication strategies
  - Performance improvement and evaluation
- Social network
  - Measurements
  - Sampling techniques
  - Data mining techniques adapt to the huge size of online social networks
  - Effects on the performance of other applications and mechanisms
- Online multiplayer games
  - Behavior
  - Virtual economy

Wireless network

- Two features
  - Mobility
  - Broadcast environment
- Two targets
  - Evaluate the present P2P in wireless environment
    - Simulated network experiment
    - Performance of present P2P
  - Desired P2P in wireless environment
    - Identify the key factors of desired wireless P2P
      - Leader per cell

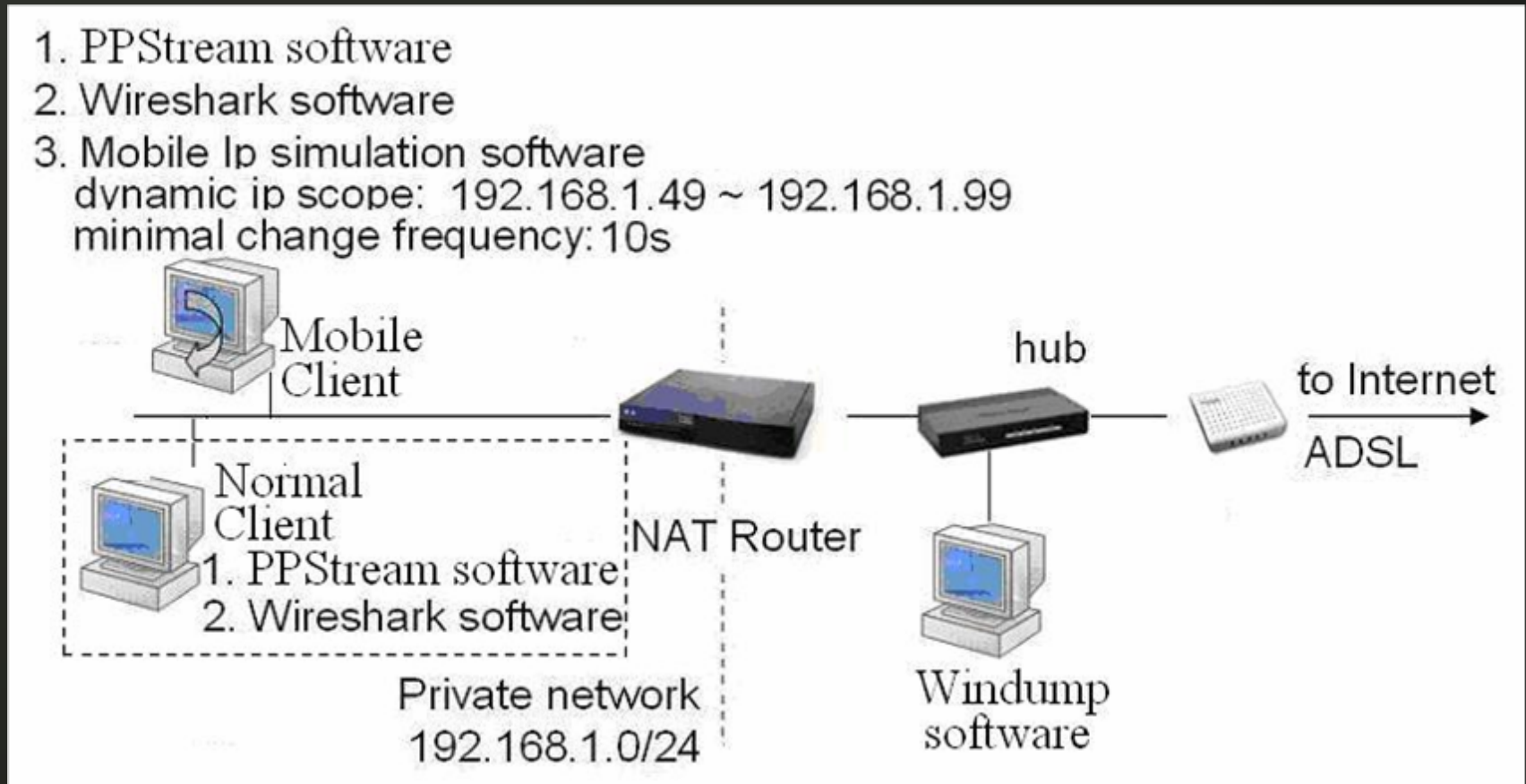
Mobility of P2P Systems

- Mobility emulation
  - real network based mobile IP emulation platform
  - Change the IP address of a PC automatically, periodically



Mobility of P2P Systems

- Deployment
  - **sniffers: PPLive / PPStream /** eMule
  - **possibility of lost response**



Mobility of P2P Systems

- Analysis and observations
  - TCP / UDP protocol availability during handover
    - TCP: extremely poor, 6s
    - UDP: neglectable, ~20ms
  - p2p application service availability during handover (PPStream)
    - Mobile terminal: not much degradation
      - Few overhead:
        - 400Kb/s, 5% retransmission , 30 chunks per handover
        - The number of chunks = RTT * Playback-rate
    - Other connected peers seems hurt very much
      - 40% reduce in upload
      - Connect time duration significantly reduced
  - Few mobile terminals hardly hurt the performance of P2P application
- PPLive/PPstream
  - Porting from TCP to UDP during above measurements
  - data encryption of PPLive: not understandable

Evolution of Application Styles

- Get Content from others
  - Web server, client browser, search for content
  - R&D
    - search engine
    - Webpage crawler
    - Webgraph: power law
- Share Content that I have
  - P2P distribution and P2P search
  - R&D
    - large-scale distribution system
    - search
    - Dynamic of systems
    - Analysis of swarm
- Publicize Content that I create
  - Server: indexing rendezvous point and search
  - R&D
    - Single media Publishing platform: human relationship /
    - Accelerating platform: distribution of content
    - Search engine
- co-create and co-play Content with others
  - From publishing platform to rich media community
    - Feature facebook?

co-create and co-play content with others

- Engineering human relationship?
  - Present view: Engineering the system, observing the human nature
  - A New Vision: engineering the Social relationship
- Research
  - What is it ?
    - not clear
    - not simply
      - A webpage crawler
      - A graph study
      - A community finding

Measurements and Modelling Research of Internet at BJTU

- Thank you!
- Q & A

Yuchun Guo

Network & Communication Lab
Beijing JiaoTong University (BJTU)
http://netcomm.bjtu.edu.cn

Email: ychguo@bjtu.edu.cn