Heterogeneous Networks: a Big Data Perspective

Arash Behboodi

October 26, 2015

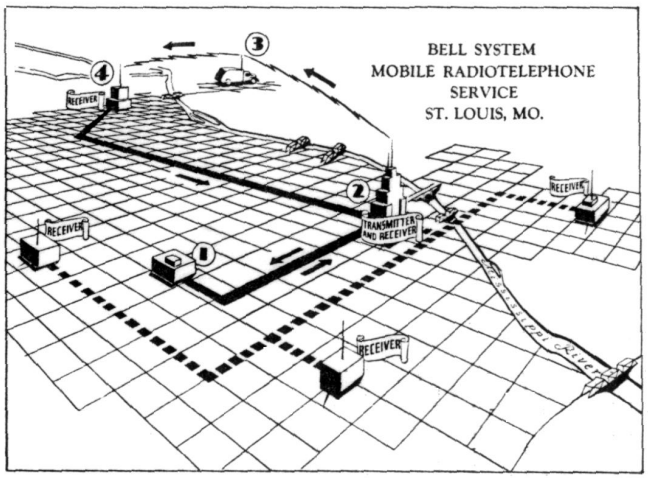
Institute for Theoretical Information Technology
Prof. Dr. Rudolf Mathar
RWTH Aachen University

Wireless Communication: History

- 1897: Guglielmo Marconi
- 1915: Transatlantic voice transmission
- 1921: Short wave radio (2.3MHz - 25.82MHz)
- 1935: Demonstration of FM
- 1946: AT&T First mobile telephone service (150 MHz band)

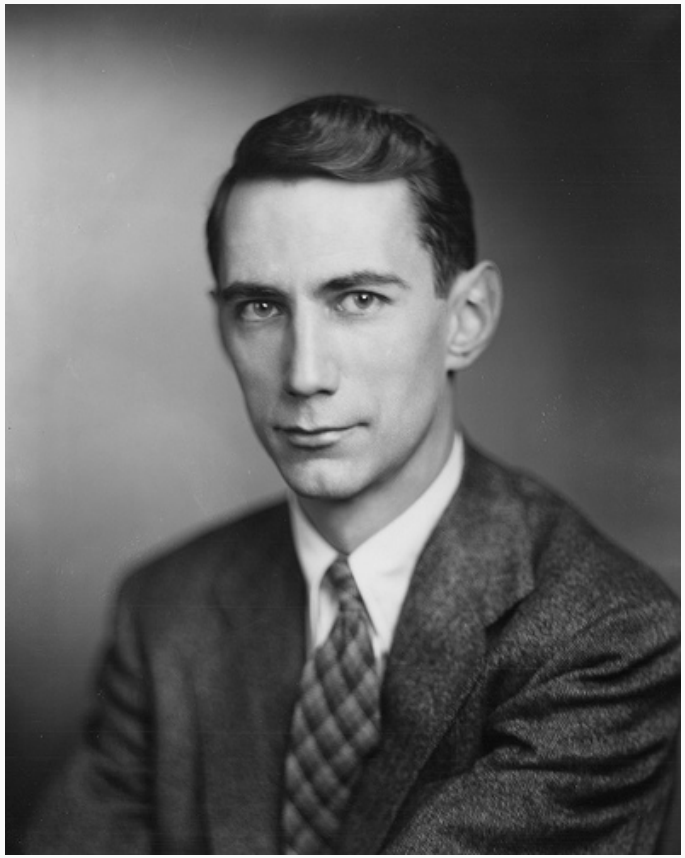
Wireless Communication: History

- 1897: Guglielmo Marconi
- 1915: Transatlantic voice transmission
- 1921: Short wave radio (2.3MHz - 25.82MHz)
- 1935: Demonstration of FM
- 1946: AT&T First mobile telephone service (150 MHz band)



Wireless Communication: History

- 1949: Claude Shannon
  - AWGN channel capacity: $C = W \log(1 + \frac{P}{NW})$ [bit/s]
- 1960 - 1970: Bell Labs developed cellular concept
- 1979: 1G cellular system deployed (Japan)
  - Analog modulation



Wireless Communication: History

- 1989: Qualcomm proposes CDMA
- 1991: 2G cellular system deployed (Finland)
  - Digital modulation (GSM, CDMA)
- 1995: First commercial launch of CDMA (Hong Kong)
- 2002: 3G cellular system deployed (South Korea)
- 2007: Apple iPhone launched

Cellular Networks: History

Challenges:

- The channel changes in time (mobility)
- The channel changes in frequency (multipath)
- The channel changes in space (path loss, shadowing)
- Resources are scarce: power, bandwidth, etc
- Interference

Performance metrics:

- Capacity, signal-to-noise-interference ratio
- Grade-of-Service/Quality-of-Service

Cellular Networks: History

Collaboration of engineers, physicists and mathematicians

- Probability, statistics, optimization, algebra, harmonic analysis, etc
- Integrated circuits, antenna design, etc
- Coding, modulation, equalization, etc

Some highlights:

- Multiple access (time, frequency, code)
- Diversity
  - Time: repetition coding
  - Frequency: multi-carrier systems (OFDM)
  - Space: multi-antenna systems (MIMO)
- Capacity approaching codes: LDPC, Turbo codes
- Interference (cooperation, management, alignment, avoidance)

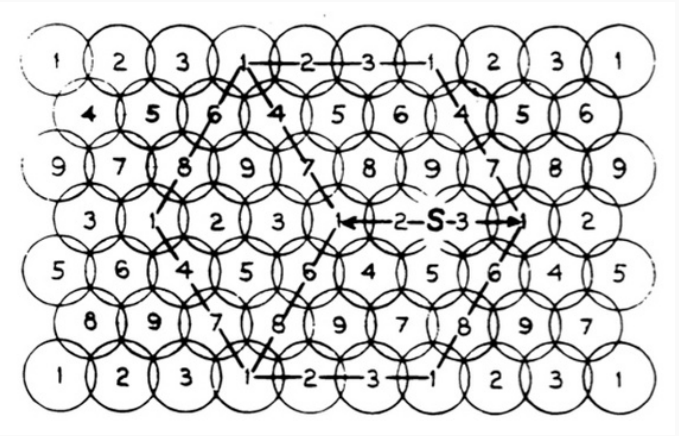
Cellular Networks: History

Some features:

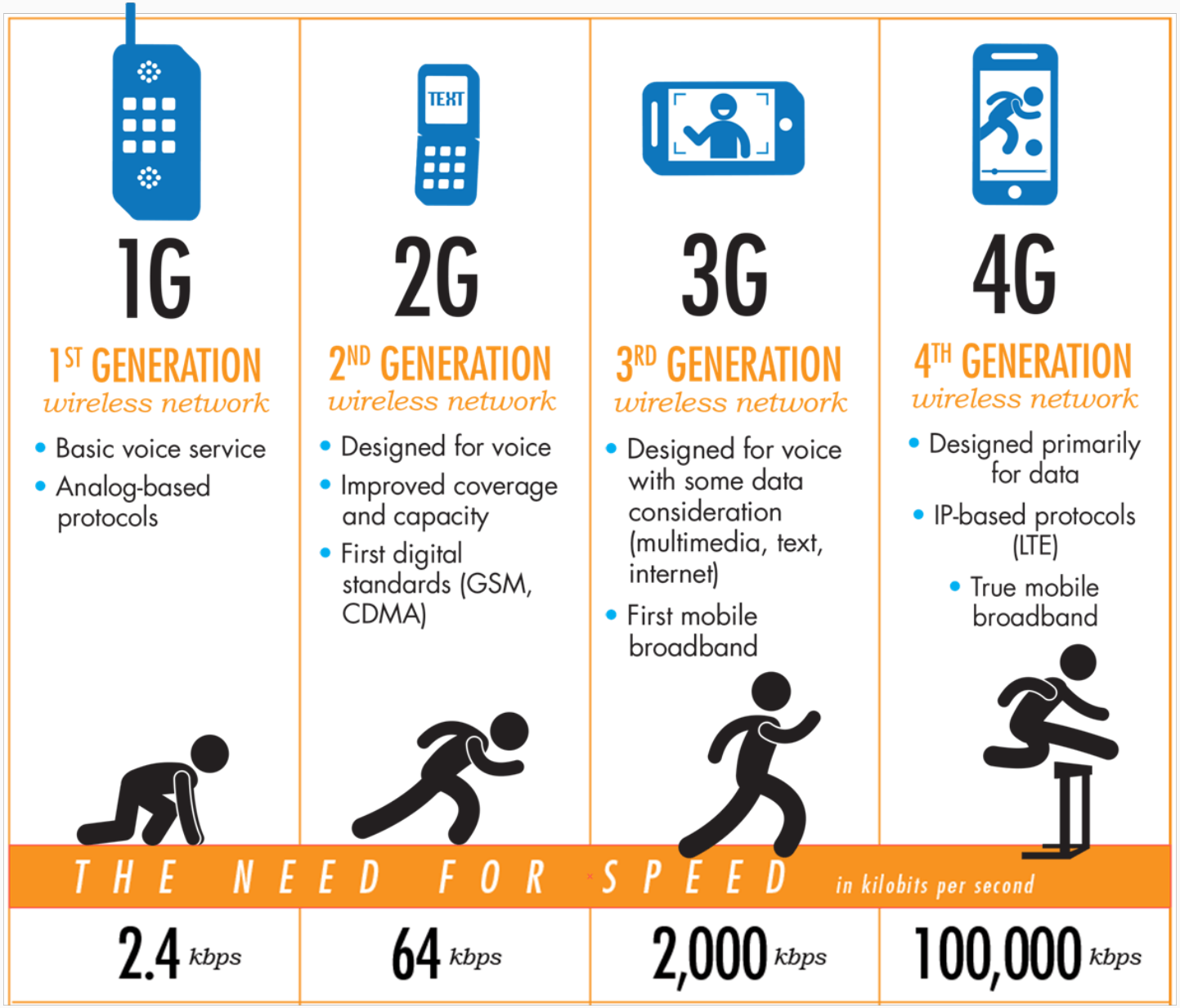
- Hexagonal regular planning
  - Frequency reuse
  - Handovers
- Voice driven
- Cellular traffic engineering
  - Queueing theory

Theoretical issues:

- Fundamental limits (network information theory)
  - Multiple access channel
  - Broadcast channel
  - Interference channel
- Design guidelines

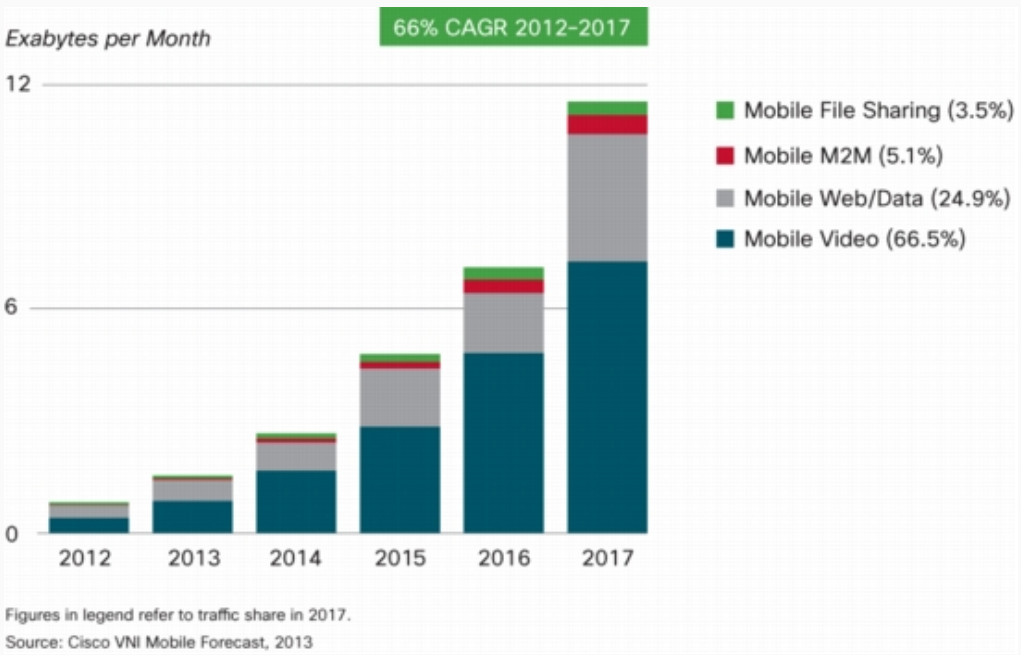


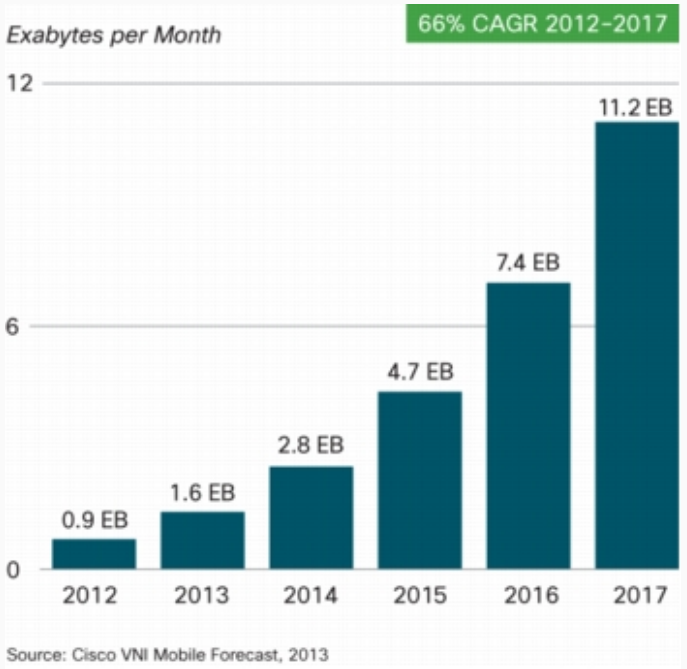
Evolution of Wireless Networks



Cellular Networks: Now

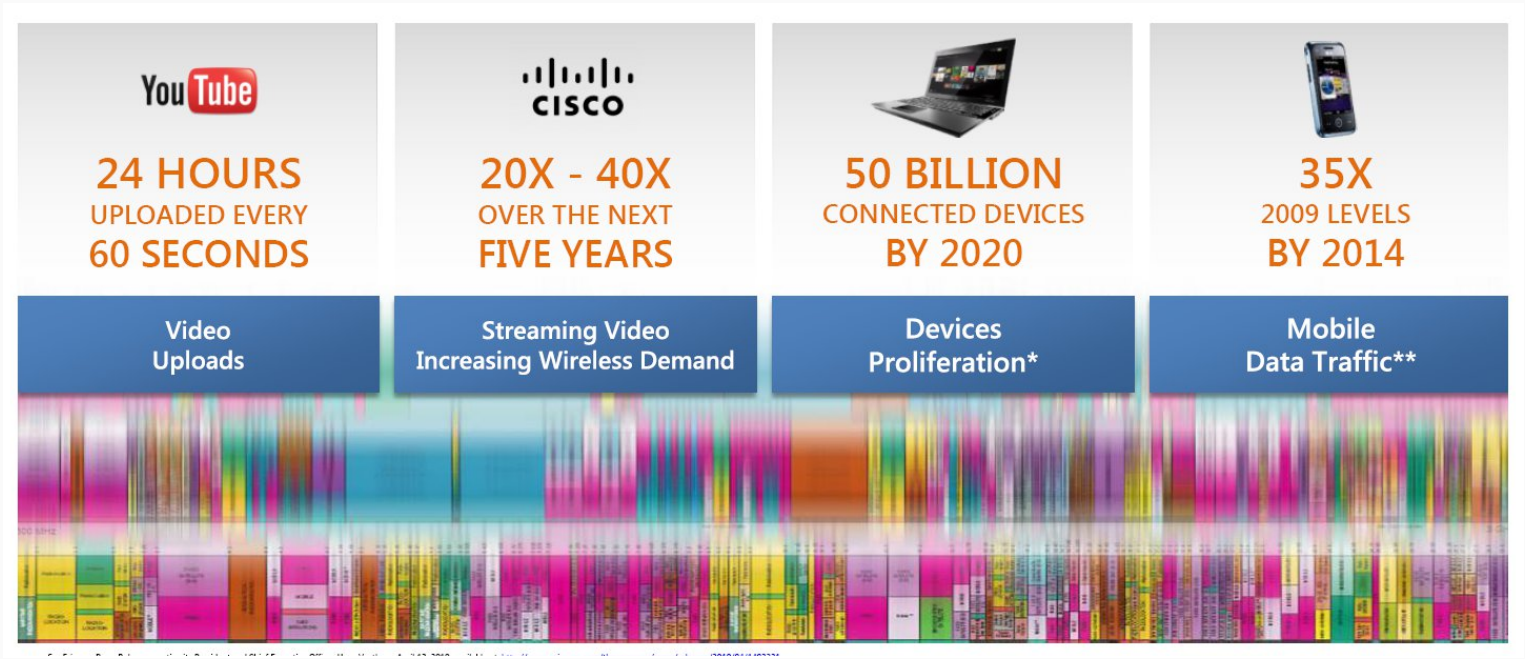
- Exponential growth of generated data
- Mobile video more than 66% of the whole data (1.5 GB per smart phone per month, 2015)
- 1.2 Billion smart phones, tablets sold in 2013
- Internet of things: *Always stay connected*





Cellular Networks: Now

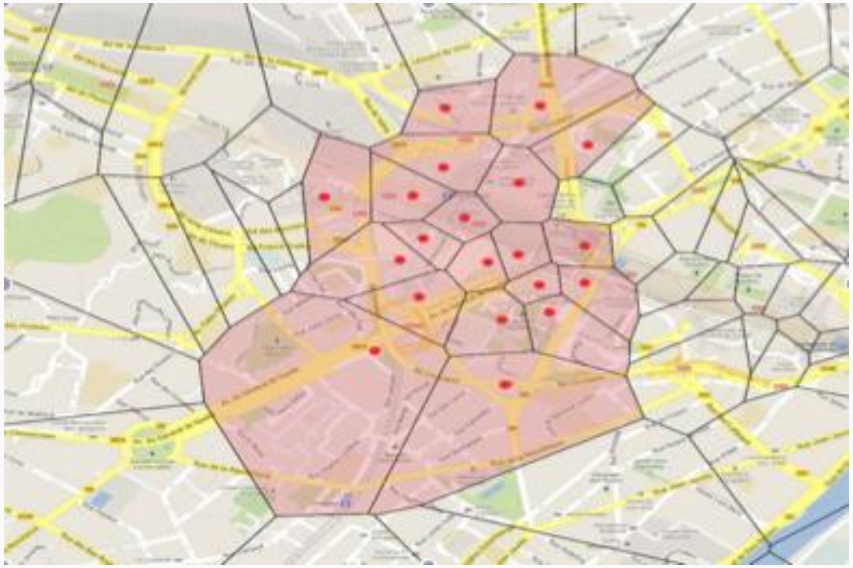
- Exponential growth of generated data
- Mobile video more than 66% of the whole data (1.5 GB per smart phone per month, 2015)
- 1.2 Billion smart phones, tablets sold in 2013
- Internet of things: *Always stay connected*



Cellular Networks: Now

Features:

- Large volume of data from users are stored
  - Call Detail Record (CDR)
  - Location
- Irregular base station deployment
- Proliferation of local wireless networks as well as newly deployed base stations



La Défense- Paris area

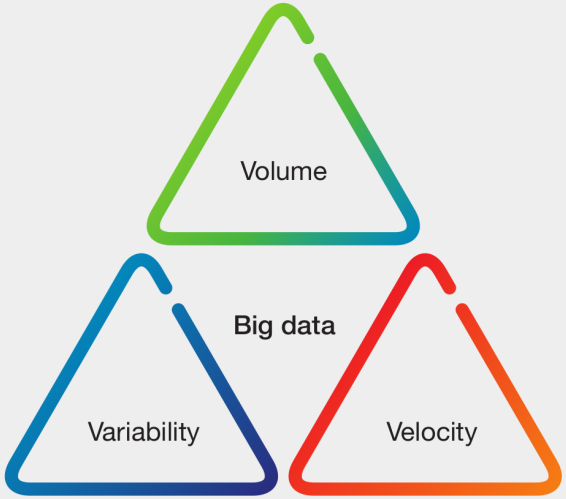
Cellular Networks: Now

Challenges:

- Exponential capacity demand
- Quality-of-Experience(QoE) driven: not always reducible to throughput
- Mobile video

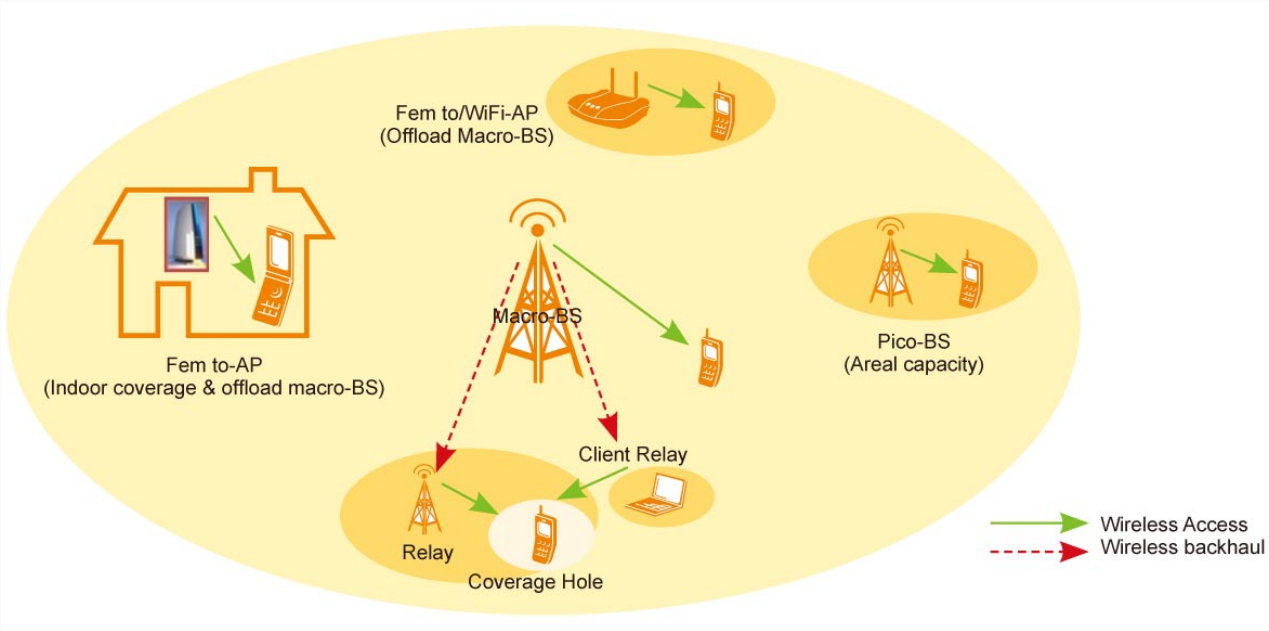
Problem

How can the users' datasets improve the performance of cellular networks? Big data paradigm ...



Heterogeneous Networks (HetNets)

- An architecture where different types of wireless networks coexist
  - Different RATs (LTE, WiFi, Zigbee, etc)
  - Different Tiers (femto-cells, pico-cells, etc)
- Densification as a solution: deployment of small cells
- Central notion: load balancing, offloading, traffic steering



Heterogeneous Networks (HetNets)

- Offloading: transfer the user from a fully loaded to a lighter load base station
  - Not a good idea to connect always to the strongest base stations

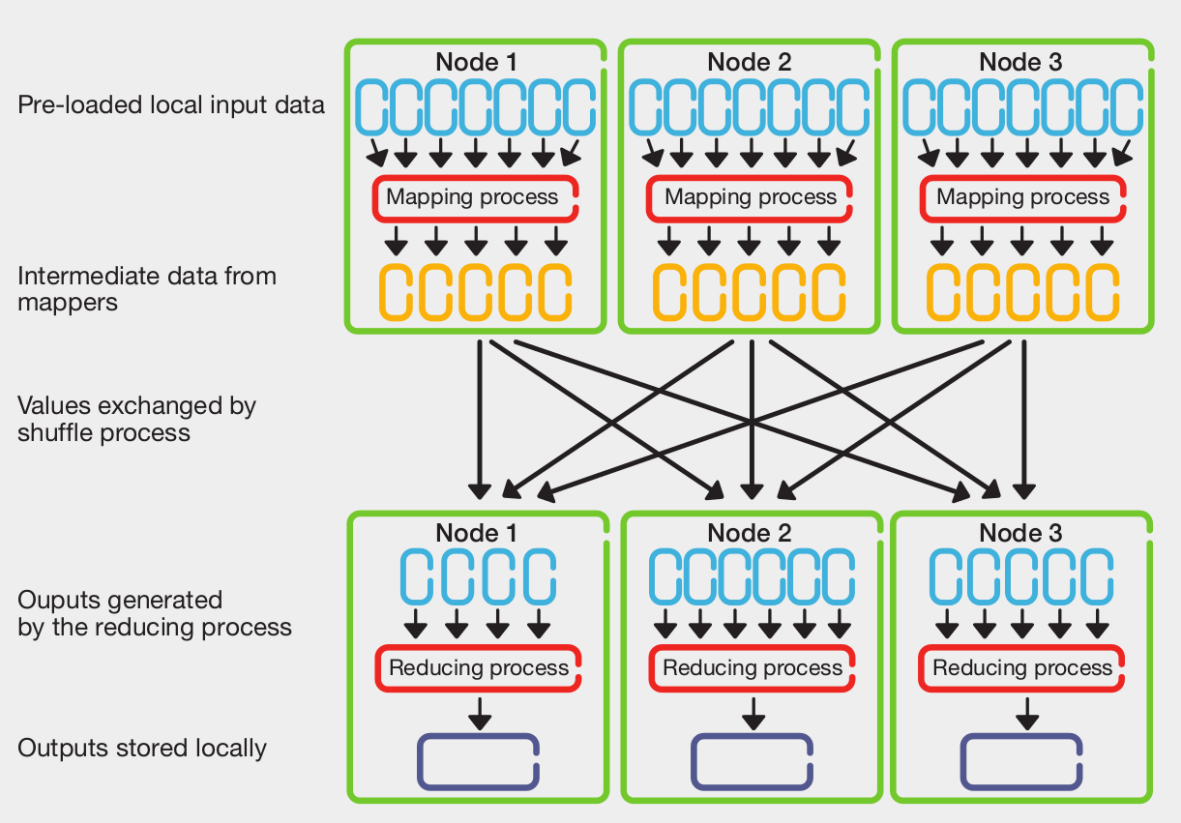
How big data analytics can be used?

- Capacity: predict when the base station is overloaded (a nearby event such as football)
  - Offload the traffic to other RATs/Tiers
- QoE: Predict when a call is dropped
  - Using user's trajectory choose the next base station
  - Providing smooth handover for instance by caching

Heterogeneous Networks (HetNets)

Example: Ericsson research and Ericsson's BSS portfolio management

- MapReduce (on Hadoop) based batch processing
- Analyzing 293,877 CDRs per second compared to 3,220 CDRs in the legacy system
- Might need to process up to 200 million CDRs per day



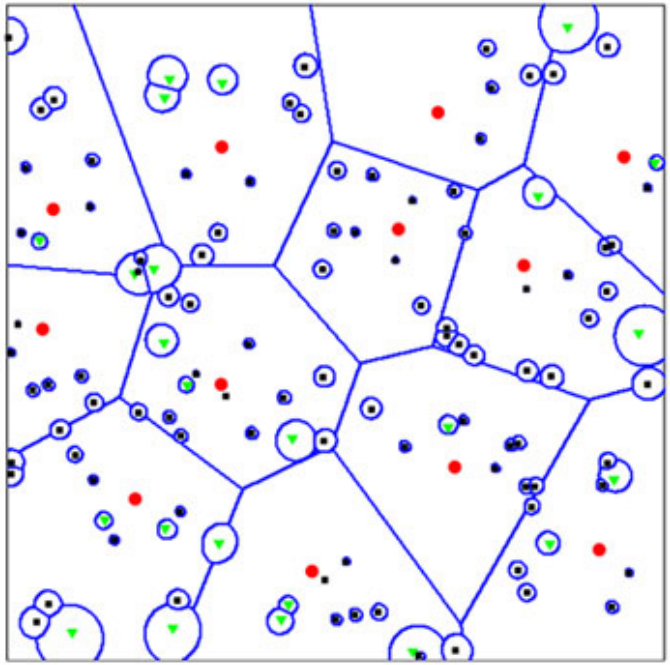
Heterogeneous Networks (HetNets)

Problems:

- Results are data specific currently and not repeatable
- No insights about the performance analysis of these methods

One side note: how to model HetNets

- Stochastic geometry: Poisson point processes with different densities in $\mathbb{R}^d$



Massive MIMO, mmWaves etc

Other solutions:

- Massive MIMO: very large number of antennas
- mmWaves: moving to higher frequencies
- Antenna tilting and distributed antenna

How big data analytics can be used?

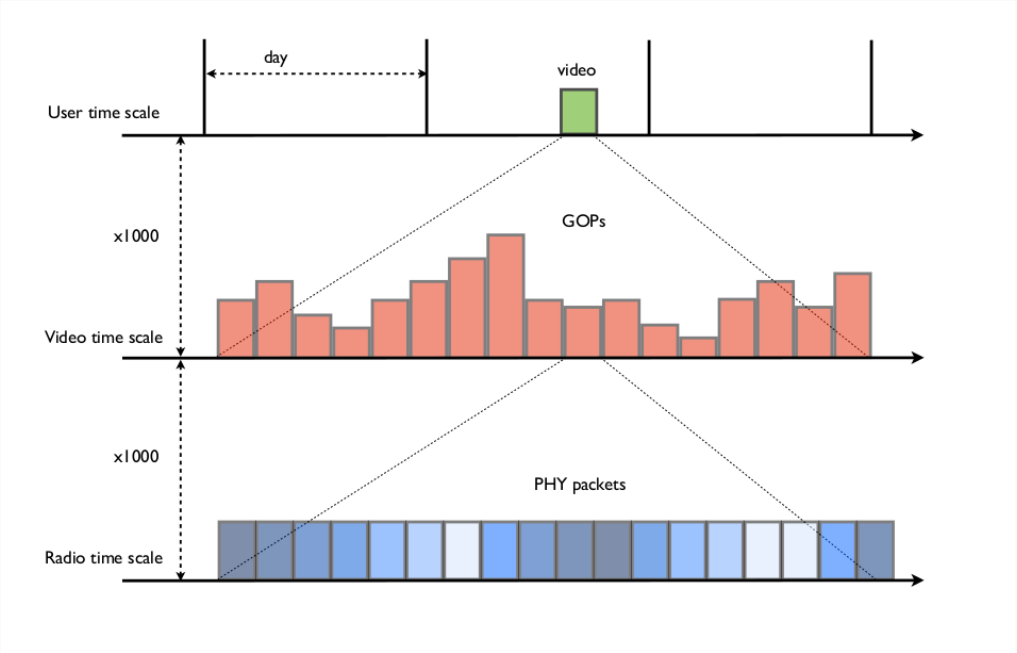
- Capacity: Find the trajectory of users
    - Beamforming and antenna tilting to provide better coverage

Video Streaming and Caching

Mobile Video:

- Preference: low but steady quality of the video
- Asynchronous content reuse: traffic generated by a few popular files accessed in a totally asynchronous way
- Predictable demand distribution

Common solutions: Cross layer design

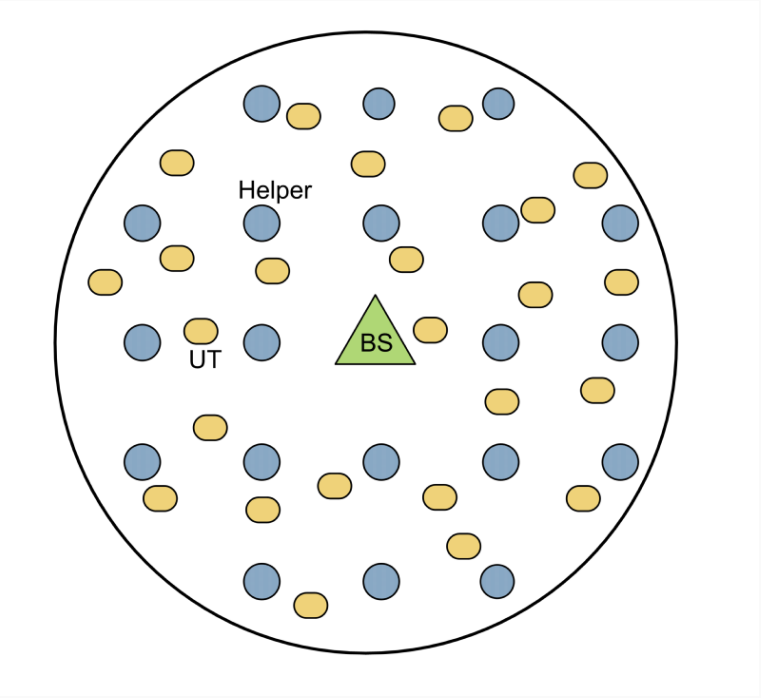


Video Streaming and Caching

Cross layer design:

- Source-Channel coding separation theorem in information theory
  - For point-to-point communication, a modular design works equally well
  - Not true for multi-user case as well as many other case
- Joint-design of video-scheduling and resource allocation

Common solutions: Caching



Video Streaming and Caching

Caching

- Cache different chunks of popular data on different caching nodes
- Distributed video streaming

How big data analytics can be used?

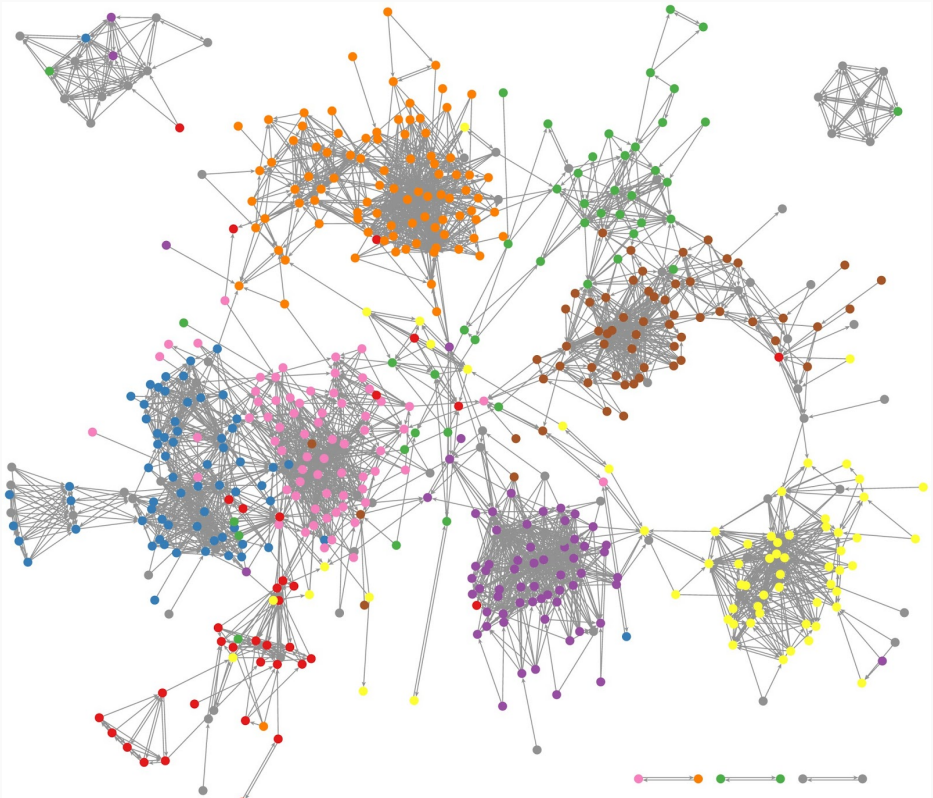
- Social network analysis
  - Find "influential" users and cache the video where it is most likely to be seen
- CDR analysis
  - Find the videos that are highly likely to be seen again and cache them based on CDR data

Side note: Interesting information theoretic studies of caching: coded caching, index coding

Community Detection

How community detection can be used?

- Intelligent video caching
- Efficient resource allocation
  - Clustering users into different communities (data-type, video-type, gamer, voice-type)
  - Employing resources correspondingly



Final Remarks

Big Data in Heterogeneous Networks

Big Data

- We need huge data to learn user's demand : Volume
- Data can vary from one user to another: Variability
- We have to decide fast: Velocity
- We have to make good decisions: Veracity

Big Data in Heterogeneous Networks

Datasets usage

- Extracting trends
  - Long-term trends, Seasonal trends, Short-term trends
- Base station deployment/activation
- Caching
- Fraud detection/ fault detection
- Load balancing
- Resource allocation
- Consistent connectivity (handover strategy, etc)

Big Data in Heterogeneous Networks

Final remarks:

- Industry is ahead of academia: datasets
- How to provide guidelines
  - Theoretical modeling
  - Data visualization
- Big data analytics can be used as business intelligence
  - Marketing campaigns such as user-customized publicities
  - Service providers feedback such as post-sale satisfactions

Dilemma

Why does it work at all? Is it worth the effort? Guidelines? Performance guarantees and benchmarking?

Questions?