



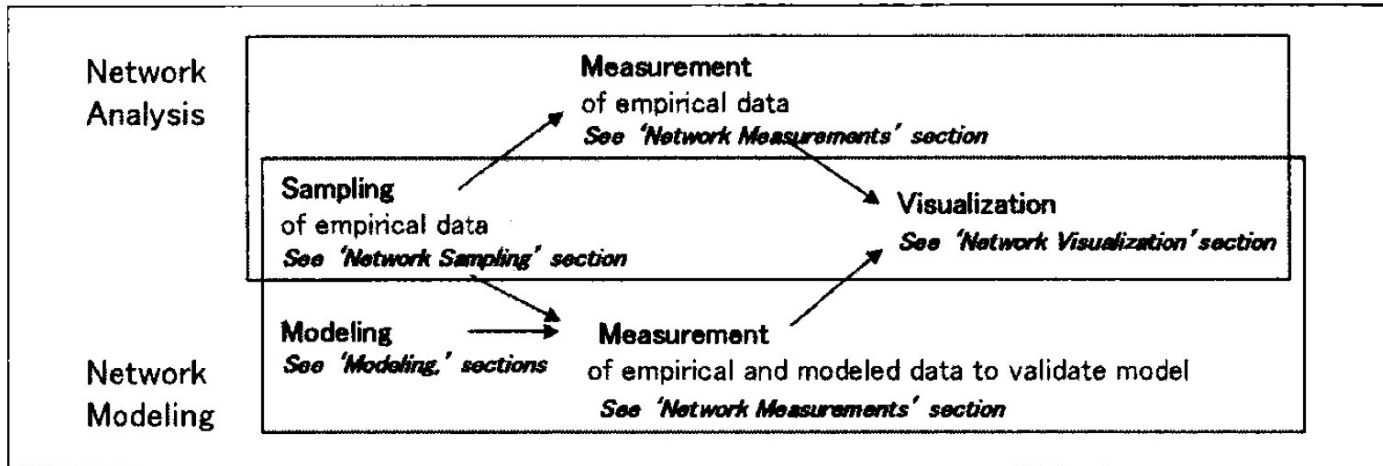
Network science



Věda o sítích

- interdisciplinární obor
- studium sítí – biologických, technologických, vědeckých atd.
- techniky a algoritmy vyvinuté v různých oborech – matematika, statistika, fyzika, sociologie (analýza sociálních sítí), informační věda (bibliometrické sítě), počítačová věda
- teorie – původ v matematice (teorie grafů a diskrétní matematika), relevantní pro efektivní design technologických sítí, vědeckých sítí, komunikačních sítí apod.

Věda o sítích



- analýza sítí – tvorba popisných modelů
- modelování sítí design procesních modelů umožňujících predikce – vzhled do struktury a dynamiky sítí



Relevantní otázky informační vědy

- Jak zajistíme stabilitu a zabezpečení technologické infrastruktury? (internet, www)
- Jaké vlastnosti sítí podporují/brání šíření a přístupu k informacím?
- Jaká je struktura vědeckých sítí, jak se vyvíjejí a jak mohou být využity k zefektivnění komunikace vědeckých poznatků?



Náhodné sítě

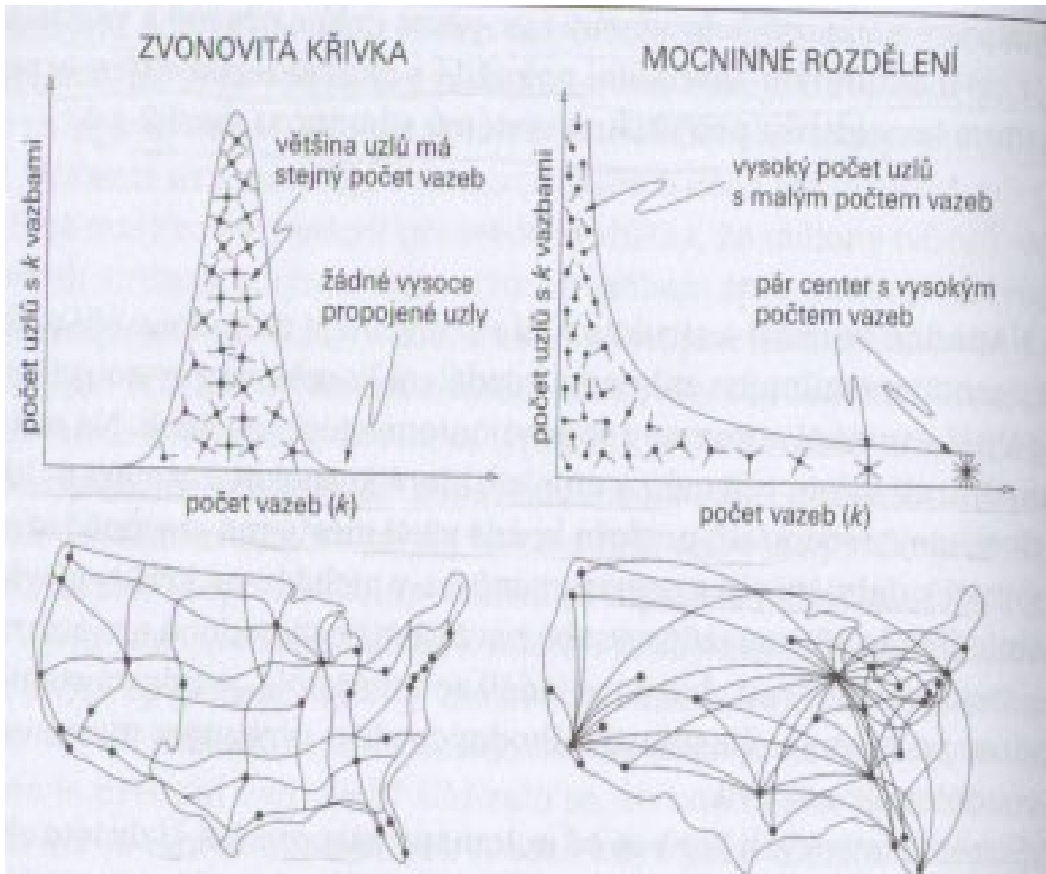
- zkoumá přírodní zákonitosti v různých sítích
- pokud každý uzel má 1 hranu – všechny spojeny do sítě
- **náhodné sítě** – spojují uzly náhodně, každý má stejnou pravděpodobnost spojení s novým uzlem
- většina uzlů má stejný počet hran – charakteristická škála



Bezškálové sítě

- **slabé (překlenující) vazby** mezi vzdálenými uzly – můstky zkracující vzdálenost
- **centrální uzly** – zapojeny do mnoha shluků, spojují menší shluky, předurčují topologii sítě, ovlivňují stabilitu a odolnost vůči chybám
- v reálných sítích většina uzlů málo hran, centrální uzly – anomální počet hran, vzácné → neexistuje typický uzel, chybí škála

Mocninný zákon



matematický
výraz

určuje rozdělení
počtu vazeb mezi
uzly



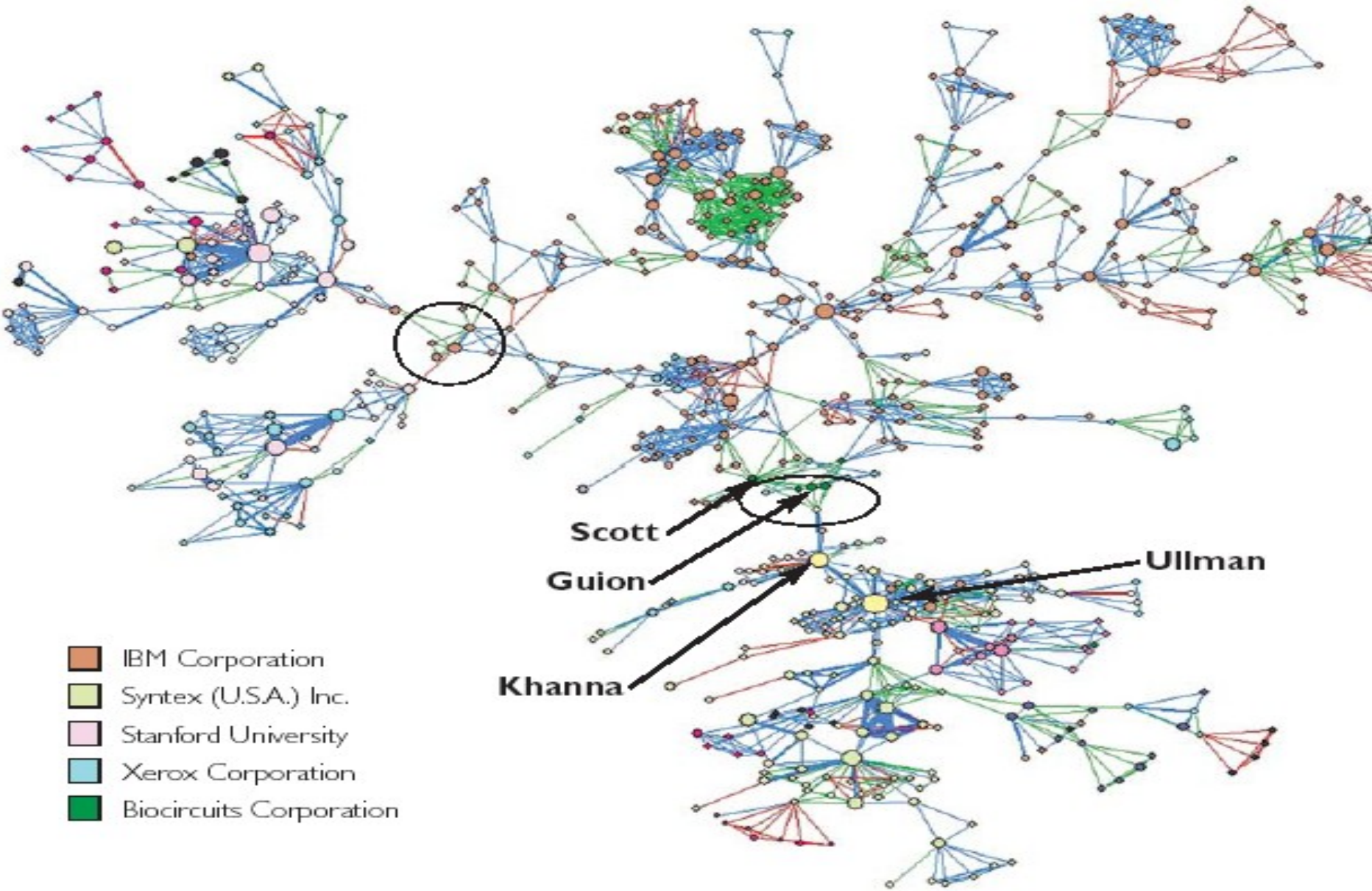
Sítě malého světa

- **struktura malého světa** – průměr hran uzlu roste nad kritickou hranicí 1 → hustá síť vazeb utváří gigantickou pavučinu, shluk
- obrovské množství uzlů – malé vzdálenosti mezi nimi
- síť s vysokým klastrovacím koeficientem a krátkou charakteristickou cestou mezi uzly





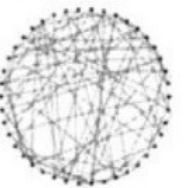

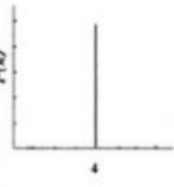
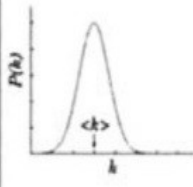
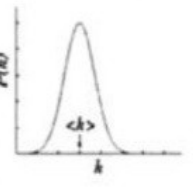
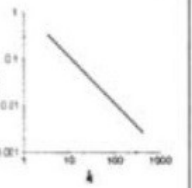

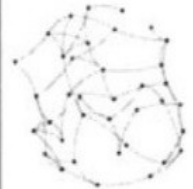
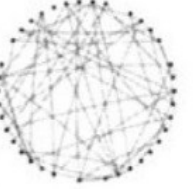

Sítě malého světa

- sociální sítě – princip šesti kroků
- kolik kroků mezi webovými stránkami?
- jaké další sítě mají strukturu malého světa?
- v nekonečně velkých sítích ani nekonečná nahodilost nevytvoří síť malého světa



○ **FLEMING, Lee – MARX, Matt. Managing Creativity in Small Worlds. *California management review*. Vol. 48, No. 4, 2006. p. 13.**



Network type	Regular lattice	Small-world	Random	Scale-free / Heavy-tail
Layout				
Number of nodes	42	42	42	42
Number of edges	84	84	84	81
Diameter	11	5	5	5
Characteristic path length	5.6	2.9	2.76	2.6
Clustering coeff. (Eq.4)	1	0.31	0.21	0.26
Clustering coeff. (Eq.6)	0.5	0.16	0.13	0.22
Average degree	4	4	4	3.86
General degree distribution				
After removal of the five most highly connected nodes				



Spojování sítí

- **stupeň uzlu** (node degree) – počet hran připojených k uzlu, tj. počet nejbližších sousedů, součet přicházejících a odcházejících hran
- **cesta** (path) – počet uzlů, které jsou mezi zvolenými uzly, které chceme spojit. Měří se nejkratší cesta, cykly, dosažitelnost, průměr – měří velikost sítě (kolik hran musí být překonáno, abychom spojili nejvzdálenější pár uzlů)
- **hustota** – poměr počtu hran v grafu k celkovému počtu uzlů na druhou



Vlastnosti uzlů a hran

- **Bonacichův mocniný index** – počítá i stupně uzlů spojených s vybraným uzlem
- **blížkost** (centrality) – závisí na stupni uzlu, vzdálenost uzlu ke všem ostatním
- **Eigenvector** – snaha najít nejcentrálnější uzel v celé síti
- **Betweenness centrality** – pozice uzlu v síti, tok který je schopný kontrolovat, počet nejkratších cest mezi párovými uzly, které procházejí daným uzlem



Topologie sítí

- **stromy** – hierarchický graf, uzel potomek má jeden uzel rodiče, kořenový uzel – povstává z něj celá struktura
- **multigraf** – grafy obsahující smyčky (hrana spojuje uzel sám se sebou), více hran – dva uzly spojené více vazbami)
- **sítě s váhami** – zachycují sílu a frekvenci interakcí (př. velikost traffic mezi internetovými routery, počet pasažérů používajících různé aerolinky)



Topologie sítí - stabilita

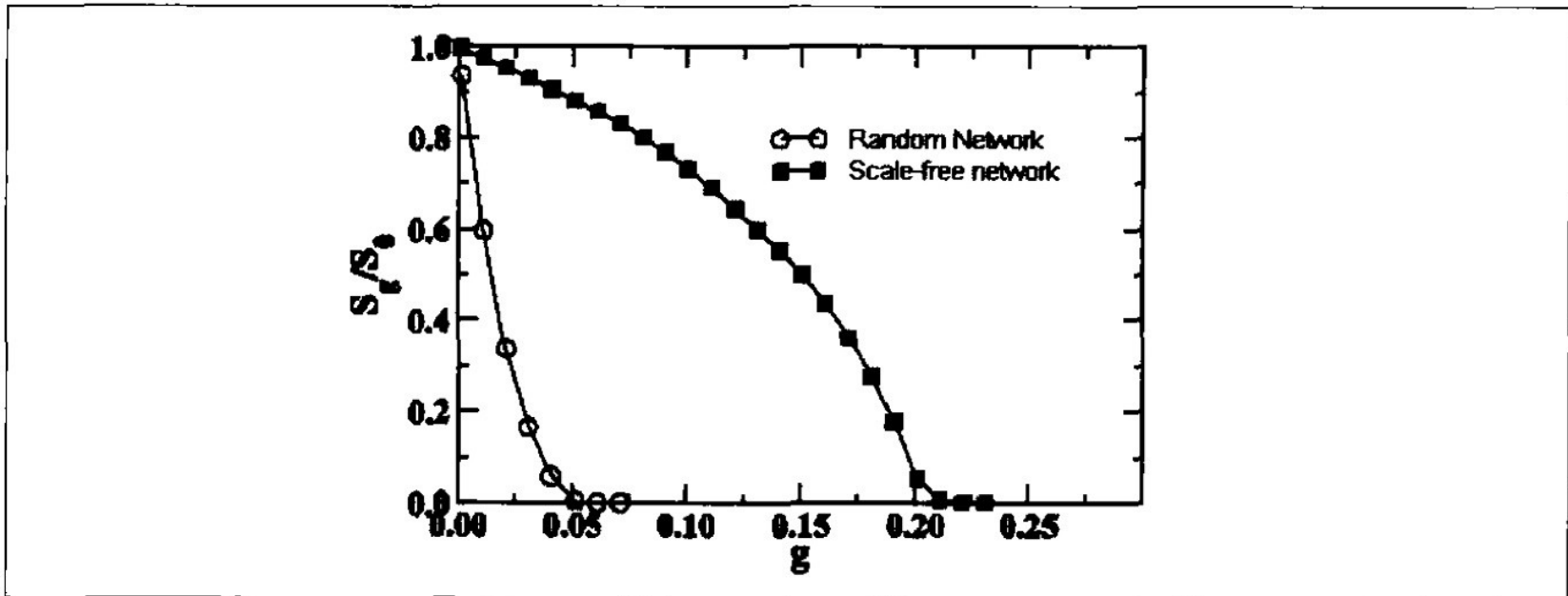
- většina komplexních sítí v přírodě - bezškálové
 - poruchy v bezškálových sítích – pravděpodobnost postihu stejná pro všechny uzly → většinou postiženy malé uzly, kterých je víc než centrálních
- sítě se nerozpadají, ani při poruše jednoho centra, ostatní centra drží síť pohromadě



Topologie sítí - stabilita

- možnost poškození – cílená likvidace center → fragmentace sítě → kolaps sítě
- náhodné sítě – nutno odstranit velké množství uzlů
- bezškálové sítě – nutno odstranit malé množství center → necitlivé k poruchám, zranitelné cíleným útokem

Topologie sítí - stabilita



průběh cílených útoků na internet



Topologie sítí - evoluce

- sítě nejsou **statické**, ale **dynamické** – rozrůstají se o nové uzly
- připojování uzlů se neřídí *strukturním* narůstáním (stejná pravděpodobností připojení), ale **preferenčním připojováním** (preferovány uzly s více vazbami)
- **samoorganizované sítě** – nezávislá činnost miliónů uzlů vytváří emergentní, složitě organizované chování → evoluce



Topologie sítí - konkurence

- **necentralizované sítě** – v bezškálových sítích konkurence → hierarchie uzlů (mocninný zákon), vítěz nevede příliš výrazně
- **centralizované sítě** – hvězdicová typologie, vytrácí se bezškálovost. Jediné centrum a malé uzly, vítěz bere vše (př. Microsoft na trhu s operačními systémy)



Topologie sítí

- **orientované sítě** – jednosměrné vazby (web, potravinové řetězce) → ovlivňuje topologii, síť je nehomogenní
- **neorientované sítě** – obousměrné vazby, nejkratší cesta tam i zpět shodná

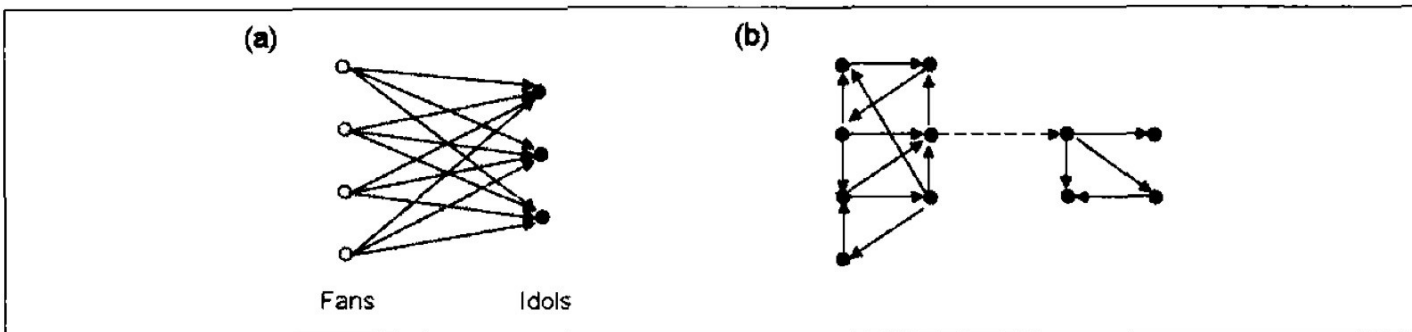


Lokální struktury

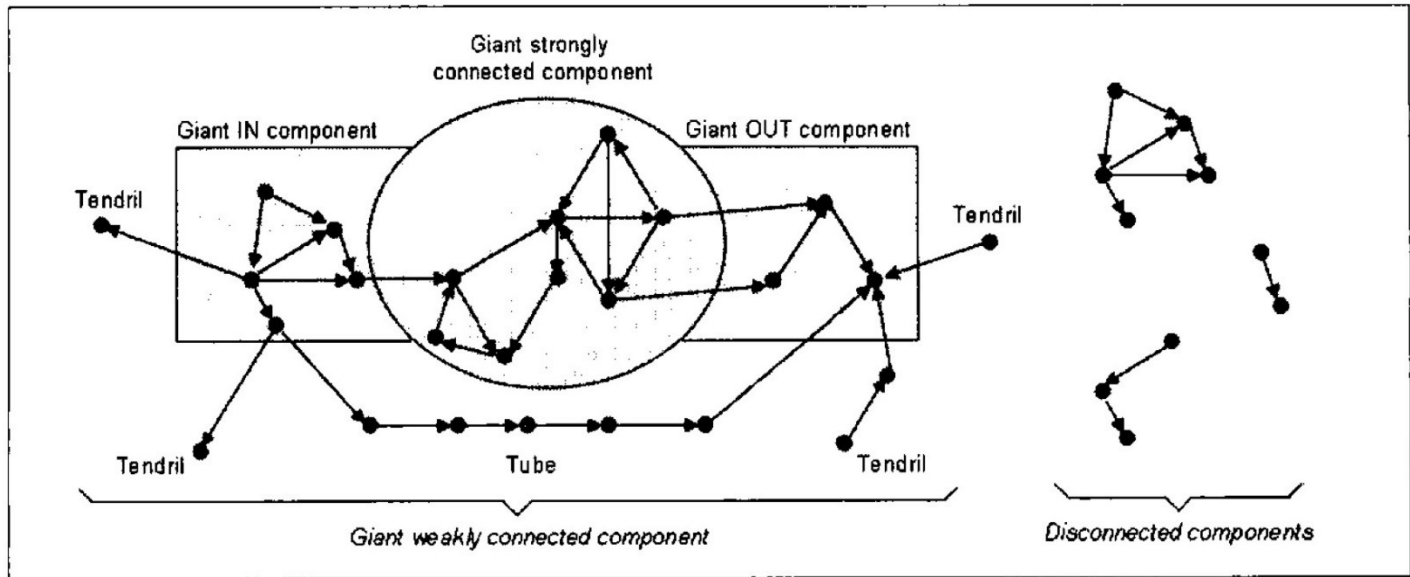
- **klastrovací koeficient C** – stupeň, jakým jsou mezi sebou spojeny uzly sousedící s daným uzlem – různé typy sítí – mřížka, malý svět, náhodný, bezškálový
- **kliky** – subgraf plně propojených tří uzlů
- **frakce** – celkový počet spojených trojic uzlů

Lokální struktury

- **motivy** – malá vzory spojení v síti:
- **n-kliky** - dlouhá vlákna
- **n-grafy** - nové uzly v klanech již spojeny s existujícími uzly
- **n-plexy** - spojení s k členy subgrafu



Komponenty sítě





Virtuální sítě spolupráce

- **kolaboratoř** – kolaborace + laboratoř
termín Williama Wulfa
- laboratoř beze zdí, umožňující výzkum bez geografických omezení
- interakce s kolegy, vzdálený přístup k přístrojové technice, sdílení dat a výpočetních zdrojů, zpřístupnění informací v digitálních knihovnách
- vědec který chce používat → porozumění způsobu použití přístrojové techniky → pomoc zodpovědného vědce-technika → předchází uživateli, nutno budovat sociální vztahy



Virtuální sítě spolupráce

- sociální vztahy tvoří kolaboratoř stejně, jako její nástroje
- interaktivní pracovní prostor –eTable pro face to face konzultace
- klíčník (gatekeeper) – osoba s překlenujícími vazbami, spojuje firmy, týmy, oddělení organizací...
 - významná role při inovacích: důvěrné vztahy, kredibilita/exkluze, sdílení novinek, informací, větší podíl na rozhodování a moci, tj. tvořivém řešení



?

- Proč je těžké zaindexovat všechny dokumenty na www?
- Jaké výhody má centrální uzel při sdílení znalostí?

Senzorické sítě ve vědě

vestavěné
senzorické sítě –
studium růstu
rostlin, vzorce
větru a počasí,
chemické a
biologické
působení, aktivity
zvířat, zdroje
kontaminace půdy
a vody

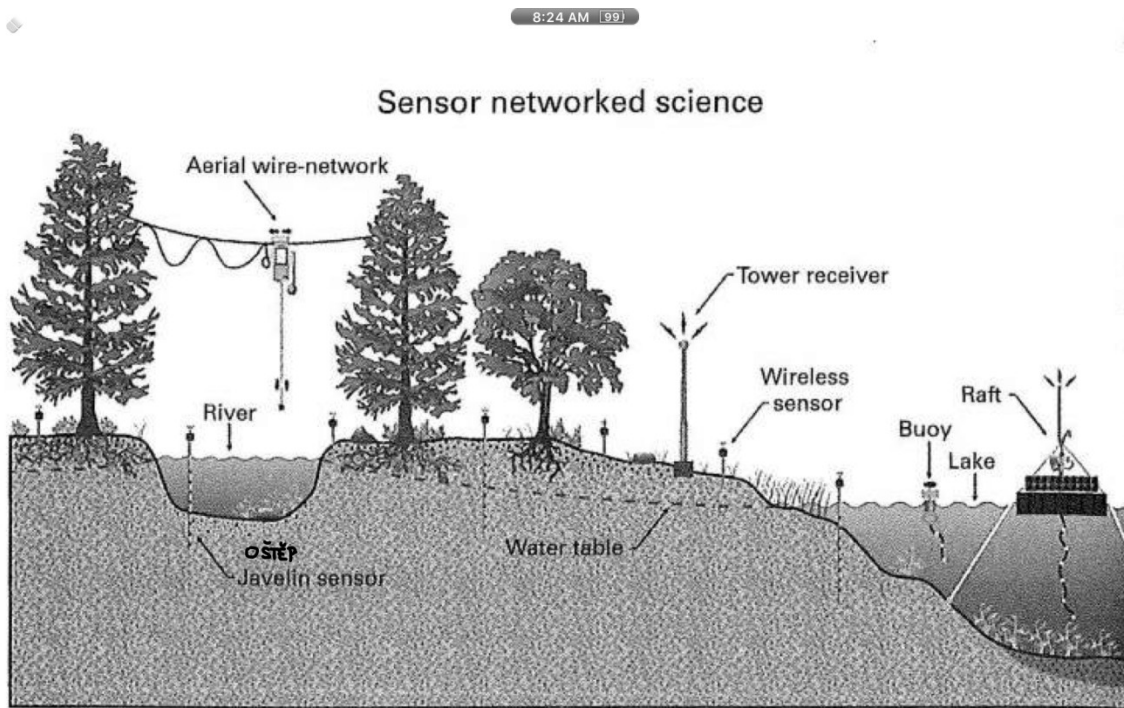


Figure 5.1

Sensor network technologies used in CENS. Credit: Jason Fisher.

Senzorické sítě ve vědě

- případ malé vědy – inkonzistence výzkumných metod – adaptivní. vysoká validita dat, nízká replikovatelnost
- spojení inženýrů a techniků (vývoj aplikací pro vědu) s vědci (nové technologie pro výzkumné metody)
- velké množství dat přesahující možnosti ručního zpracování – př. růst kořenů – kamera v plastové trubici v zemi odesílá data do senzorické sítě, až 60 000 obrázků za den (10 GB), mořská biologie – lidé sběr vzorků vody 4 x za 24 hodin, senzory každých pět minut
- teorie, co jsou data se vyvíjí spolu vývojem technologií a dozráváním spolupráce a metod. Vliv rozhodnutí při designu hardwaru a softwaru – vliv jaká data mohou být získána



Senzorické sítě ve vědě

- senzory neměří vítr, arsen, nitráty a další vědecké proměnné přímo, měří napětí a další zjistitelné indikátory – binární signály interpretované statickými modely
- ground truting – používání známých metod měření k testování validity nových měřících metod – smysluplnost vědeckých dat vyžaduje vědecké modely
- chybí standardy dat, archivy, klasifikační mechanismy usnadňující a integrující výměnu zdrojů dat.
- Data management starostí výzkumníků, chybí vhodný aparát. Sdílení osobními kontakty, neexistují formální prostředky na propojení dat s publikacemi a dalšími výzkumnými objekty.
- data držena soukromě, sdílený fond požadován pouze specializovanými obory jako seismologie či genomika