



Intelligent Systems

Petri Myllymäki
Complex Systems Computation Group
Department of Computer Science
University of Helsinki, Finland

<http://www.cs.helsinki.fi/petri.myllymaki/>

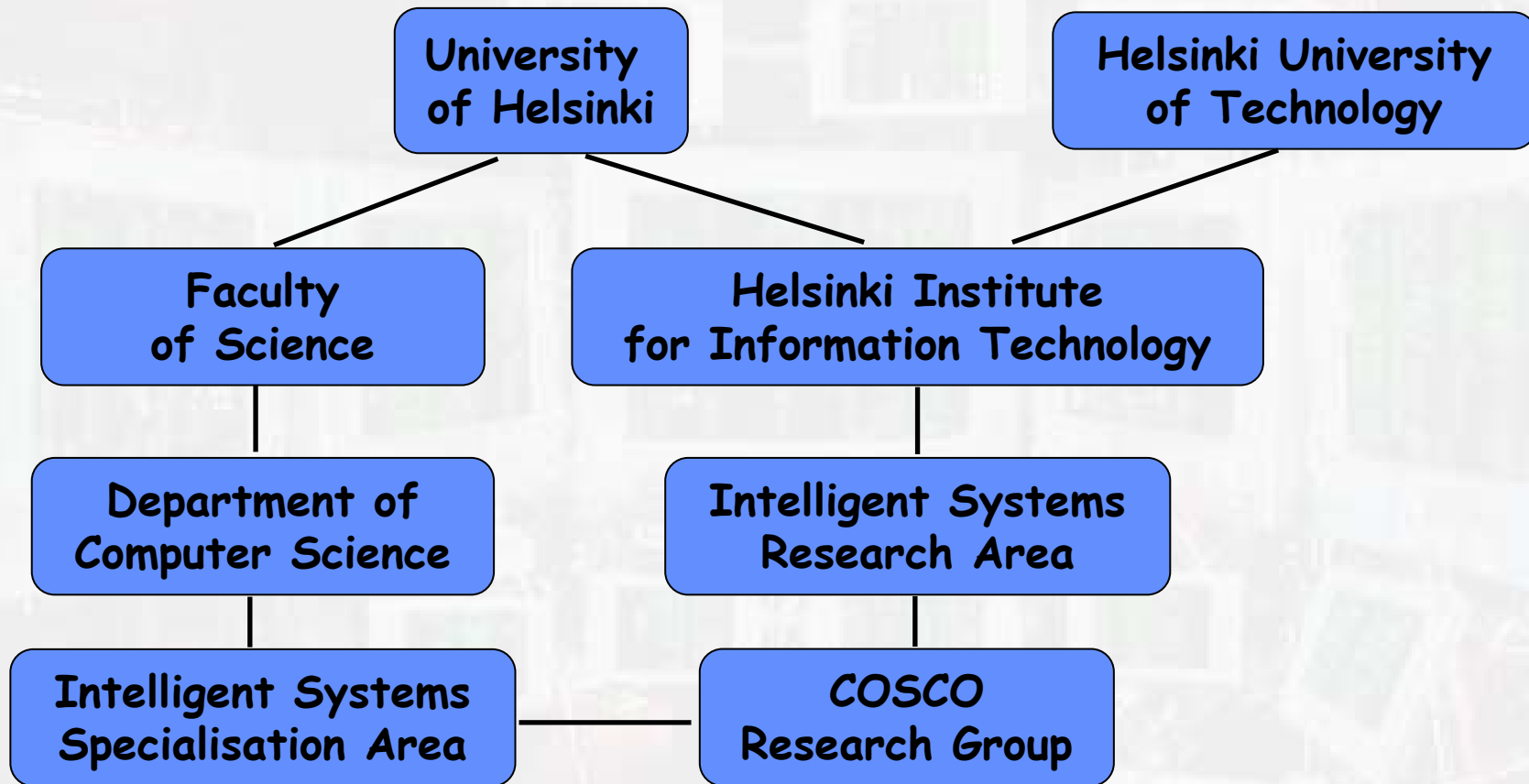
Intelligence?

- Computer science is considering ways to **make things automated**
- Intelligent Systems studies the question **how can intelligence be automated?**
- Intelligent **behavior** can be characterized by **learning/adaptation**
- In order to learn one needs to **remember** and **generalize** (world is never repeating itself!)
- In order to generalize one needs **models**

Name of the game?

- Artificial intelligence
- Science of uncertainty
- Adaptive and intelligent systems
- Computational intelligence
- Soft computing
- Real-world computing
- Complex systems computation
- Deep computing
- ...
- Keywords: knowledge representation, reasoning logic, expert systems, machine learning, data mining, Bayesian networks, neural networks, fuzzy systems, evolutionary computing, artificial life, robotics, planning, optimization...

Teaching and Researching Intelligent Systems



http://www.cs.helsinki.fi/alykkaat/





Complex Systems Computation Group CoSCo

27-
31.1.2003

Hannes Wettig presents the paper "Bayesian Analysis of Online Newspaper Log Data" at the 2003 International Symposium on Applications and the Internet (Orlando, USA, January 2003).

→ [SAINT 2003](#)

28.1.2003

Professor Henry Tirri gives talk "Adaptive modeling of Internet users" at the concluding seminar of USIX technology programme

→ [USIX 2003](#)

28.1.2003

Miikka Miettinen gives talk "User Profiling and Visualization in www-based learning environments" at the concluding seminar of USIX technology programme

→ [USIX 2003](#)

21.1.2003

Professor Henry Tirri gives invited talk "Computational World" at Europe 2020 seminar arranged by Finnish Society for Information Services

15.1.2003

Professor Henry Tirri gives an invited talk "Personalized Adaptive Interfaces" at Media and Everyday Life seminar arranged by StoraEnso and AlmaMedia

12.1.2003

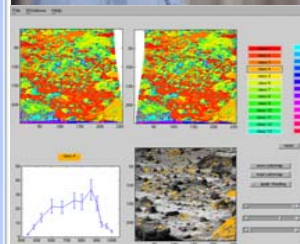
Professor Henry Tirri gives talk "Individual theory - Mathematics of Personalization in Computer Networks" at Finnish Science Days

→ [Tieteen päivät](#)



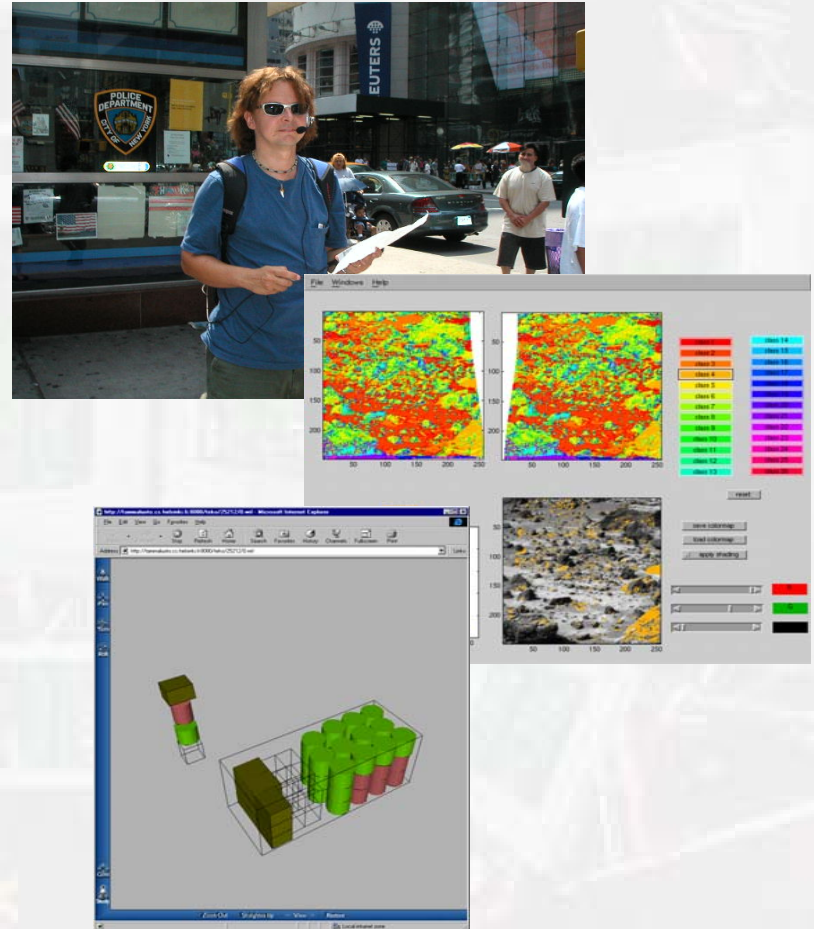
B-Course

Internet



Student: So what is it like to do research in AI?

- Great fun!
- Some hacking
- some math
- Some gaming
- Some technology "freaking"
- Some science fiction



"On the shoulders of Giants"



Andrey Nikolaevich Kolmogorov



Rev. Thomas Bayes

CoSCo Research areas

- Probabilistic and information-theoretic modeling in sciences and business
 - Information-theoretic modeling approaches
 - Bayesian and Causal Networks
 - Finding the position of mobile devices
 - Personalization
 - Next generation information search
 - Tools and theory for E-learning
- Stochastic optimization in complex domains



Bayes vs. MDL

Under regularity conditions $-\log P_{NML}(x^n | M) =$

$$-\log P(x^n | \hat{\theta}_i(x^n)) + \frac{k}{2} \log \frac{n}{2\pi} + \log \int \sqrt{\det I(\theta)} d\theta + o(1)$$

Under regularity conditions $-\log P_{Bayes}(x^n | M) \approx$

$$-\log P(x^n | \hat{\theta}_i(x^n)) + \frac{k}{2} \log \frac{n}{2\pi} - \log w(\hat{\theta}) + \log \int \sqrt{\det I(\theta)} d\theta + o(1)$$

If we take Jeffrey's prior

$$w(\theta) = \frac{\sqrt{\det I(\theta)}}{\int \sqrt{\det I(\theta)} d\theta} \dots \dots \dots \text{☺}$$

B-course Data Analysis Server (<http://b-course.hiit.fi>)

Final report of Housing [Bcourse] - Microsoft Internet Explorer

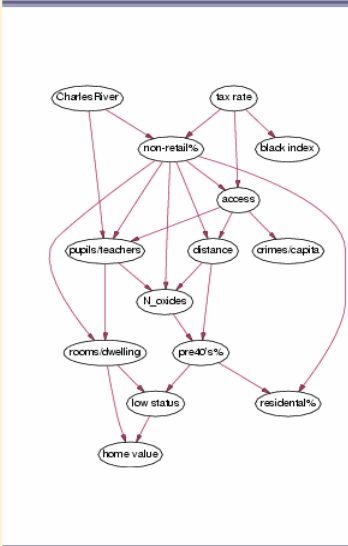
File Edit View Favorites Tools Help

Back Forward Stop Reload Search Favorites Media Print

Address <http://b-course.cs.helsinki.fi/cgi-bin/showreport.py?sid=8F76AC8DF292E0A4735A798A021F62C7&n=Final> Go Links »

Final report of Housing

The search is now over. Here you can find a picture of the final model, both in full-size viewing (PNG) and printable (Post-Script) formats. You can study the strengths of the arcs and also take a look at the two causal graphs that speculate about the possible causalities underlying the observable statistical dependencies. Finally, you can study the models interactively.



Dependency model

- » [Strength of dependencies \(arcs\)](#)

Causal models

- » [Naive causal model](#)
- » [Not-so-naive causal model](#)

Playgrounds

- » [Java playground](#) (Recommended!)
- » [Vanilla playground](#)

Pictures

- » [Natural size picture \(png-format\)](#)
- » [Picture for printing \(ps-format\)](#)

Hugin Lite export

- » [HUGIN Lite file](#)

Information about the search

During the search, 80408 candidate models were evaluated. Doing the extra 69685 evaluations since last report paid off, since the new model is over million times more probable than the last time best.

Internet

Quick Summary

- Designed for dependency analysis with graphical models
- ASP architecture (works with most browsers)
- inference of Bayesian networks (and elementary causal networks)
- "tutorial style" user interface
- no user modifiable parameters
- interactive tool for inference
- extendible platform (v 2.0 classification)

So what “science” was needed?

- Theory, heavy theory
- Empirical work with data sets ...
- Multidisciplinary co-operation
- Brilliant hacking (B-course had predecessors: D-Side, BAYDA)

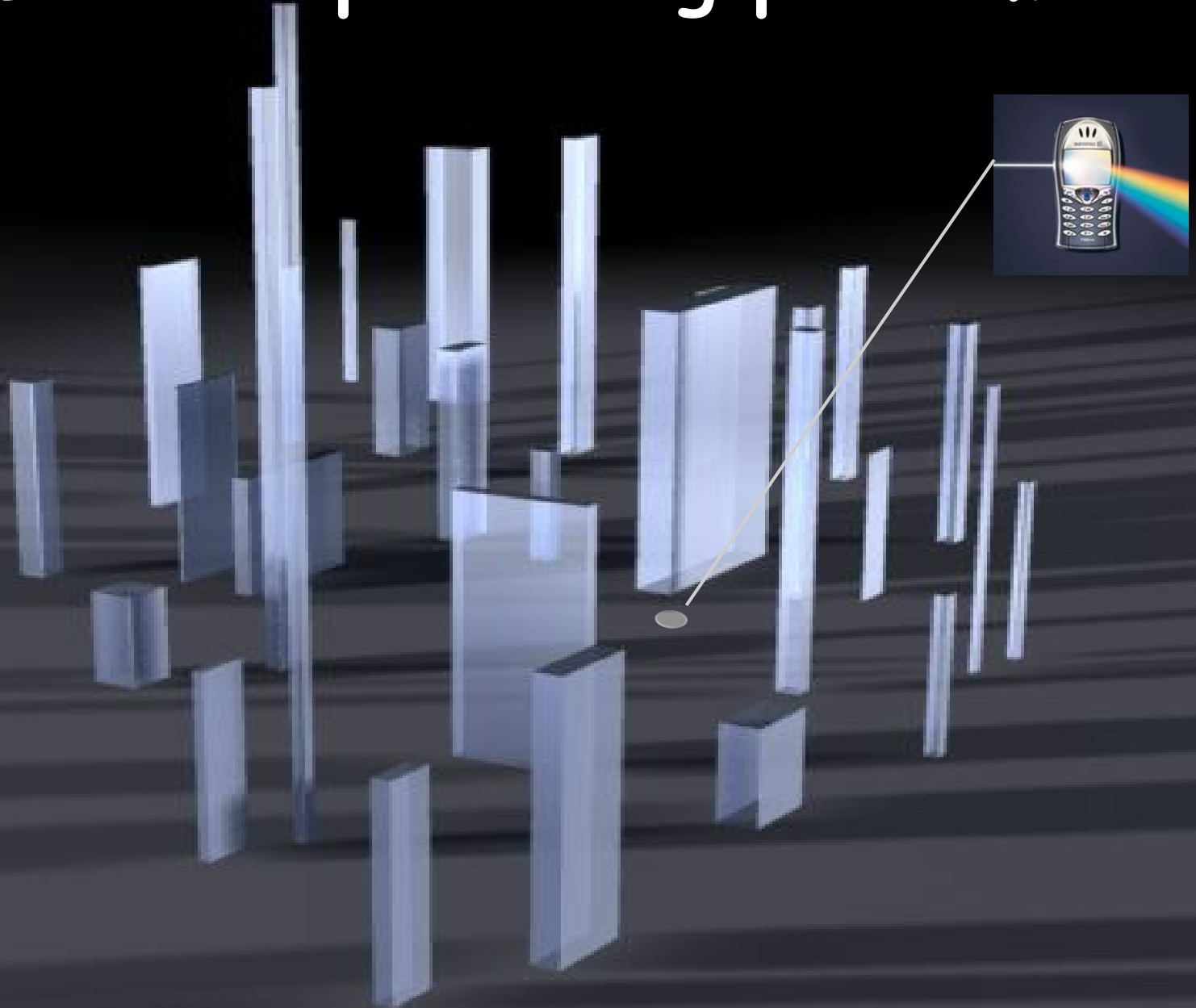


LocIt

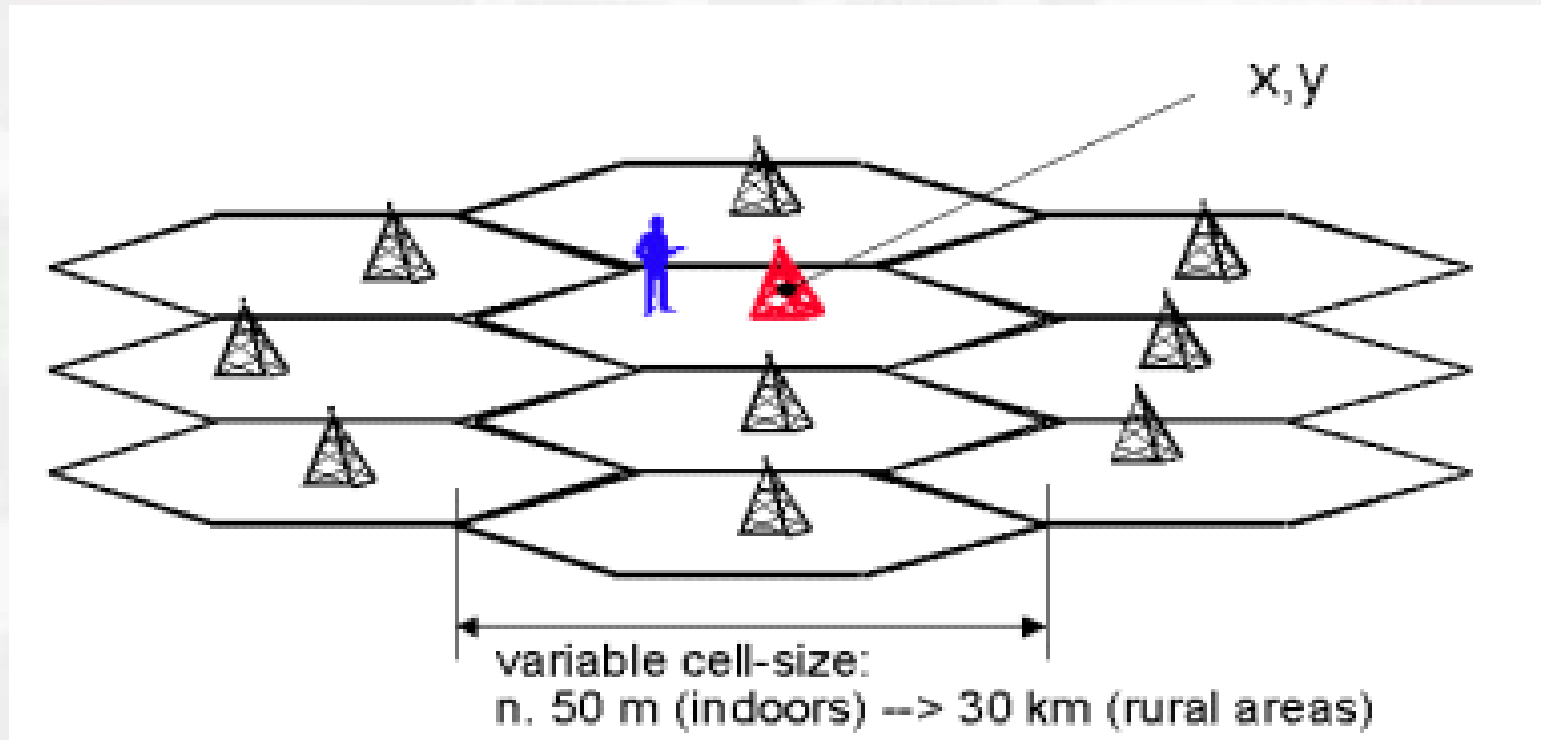
Mobile Device Positioning



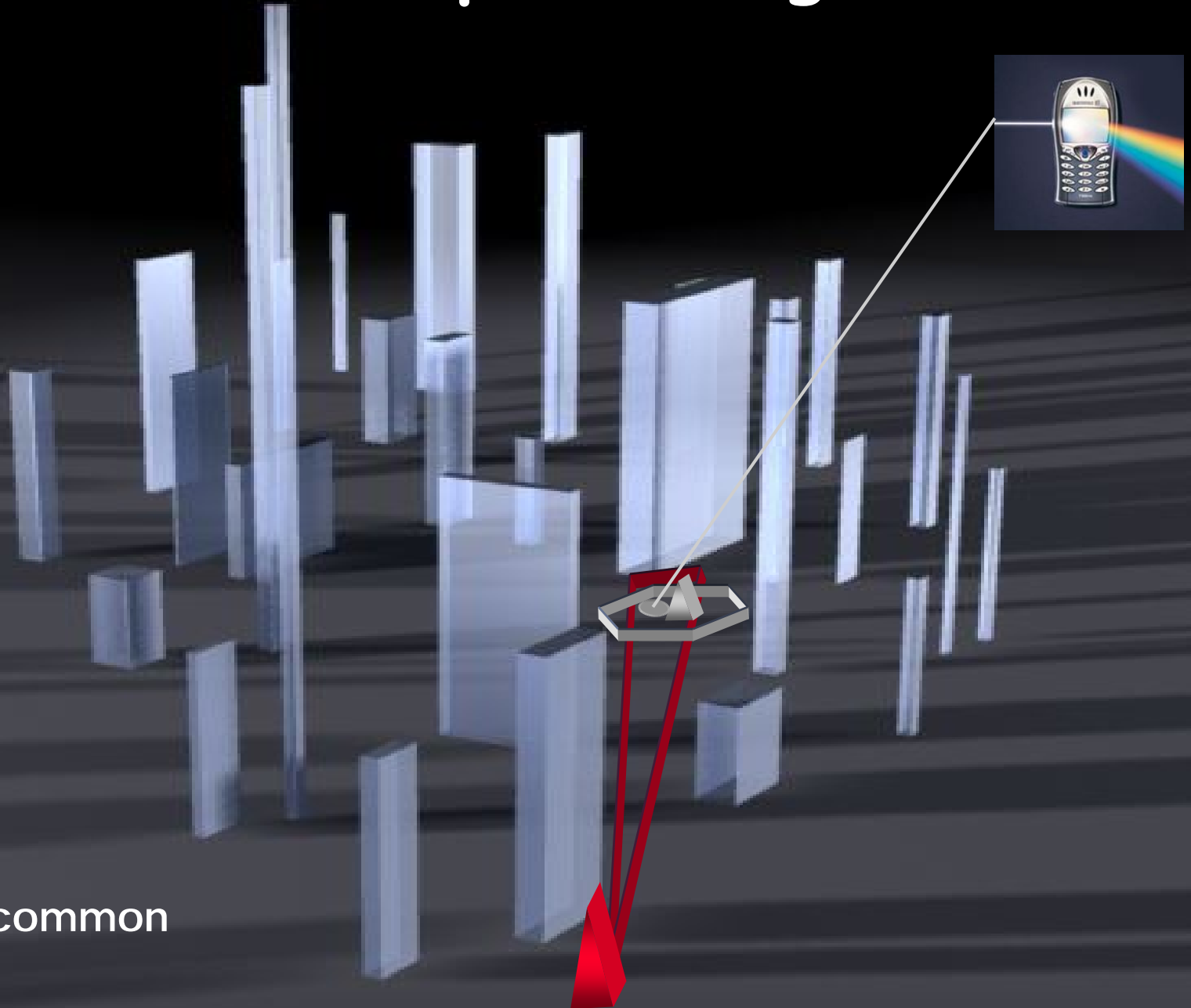
Location positioning problem



Cell ID

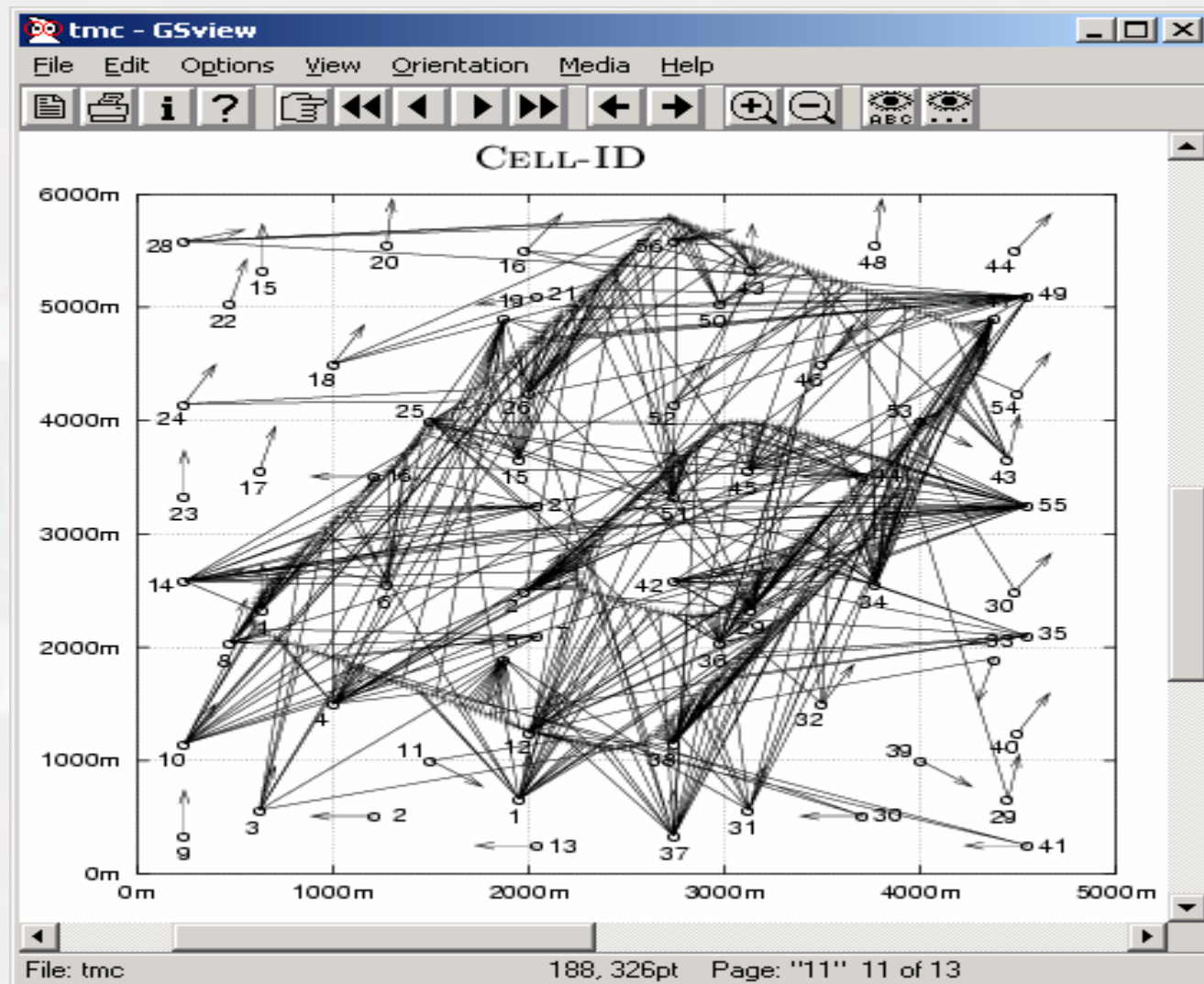


Urban positioning

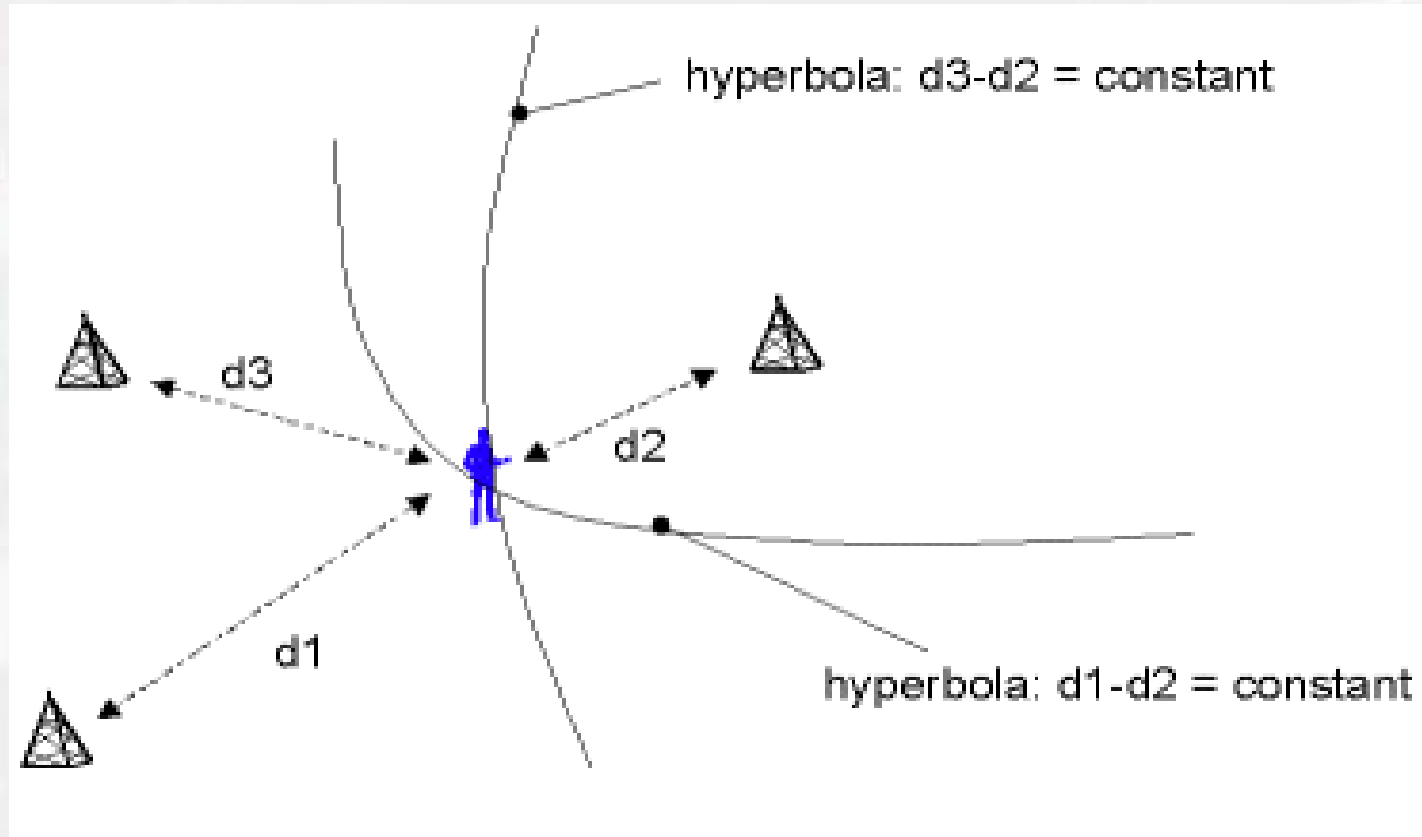


- errors > 500m common
+ simple

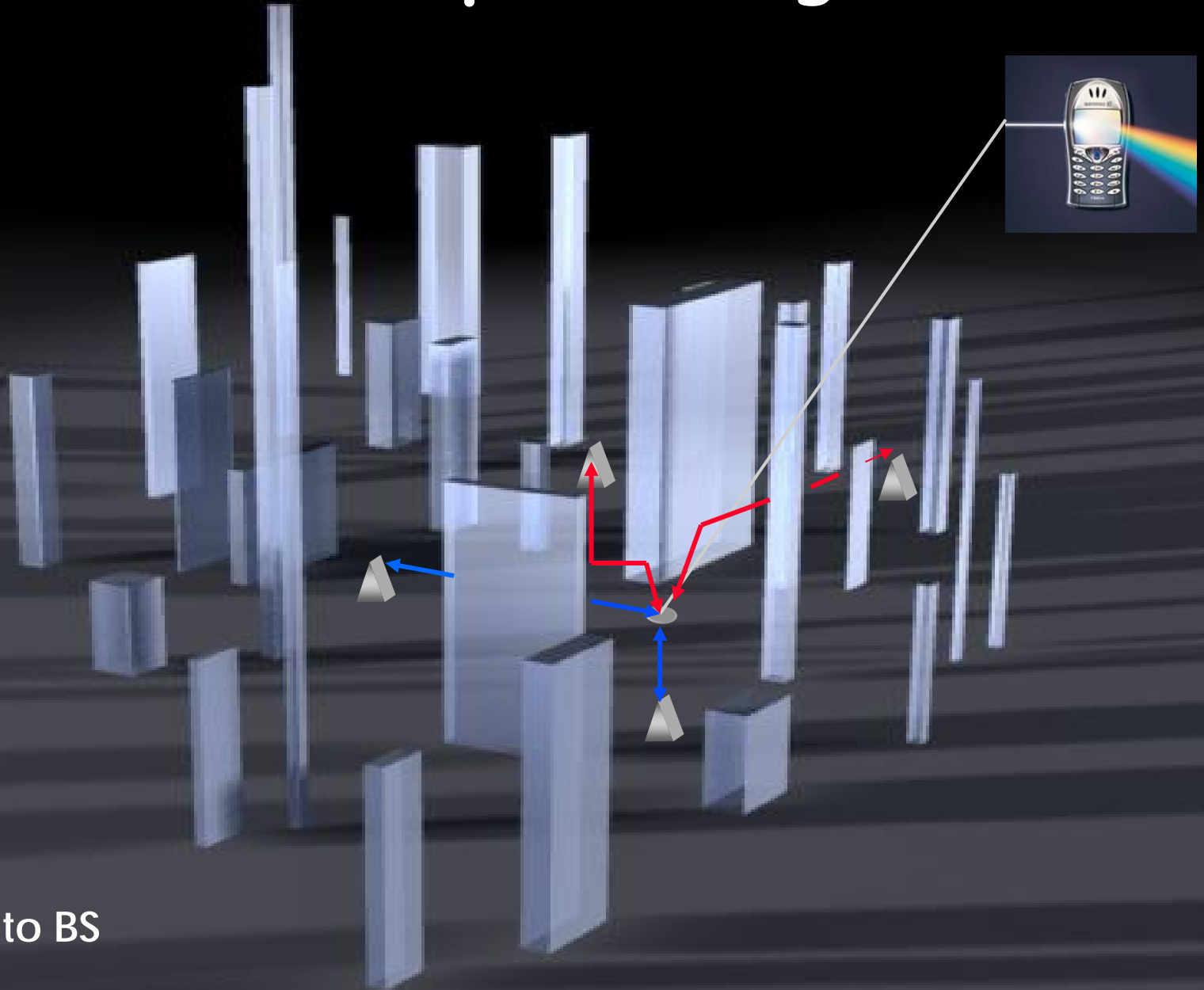
Cell ID errors



Enhanced Observed Time Difference (E-OTD)

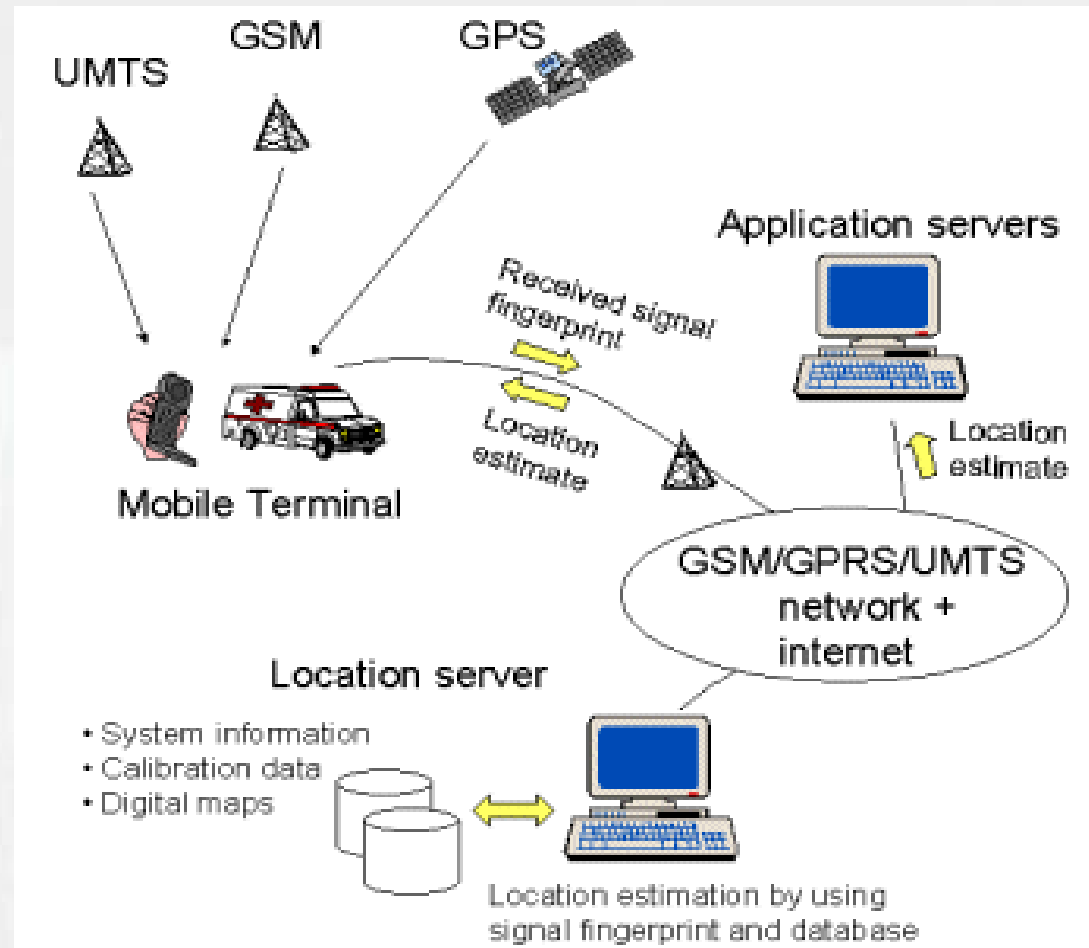


Urban positioning

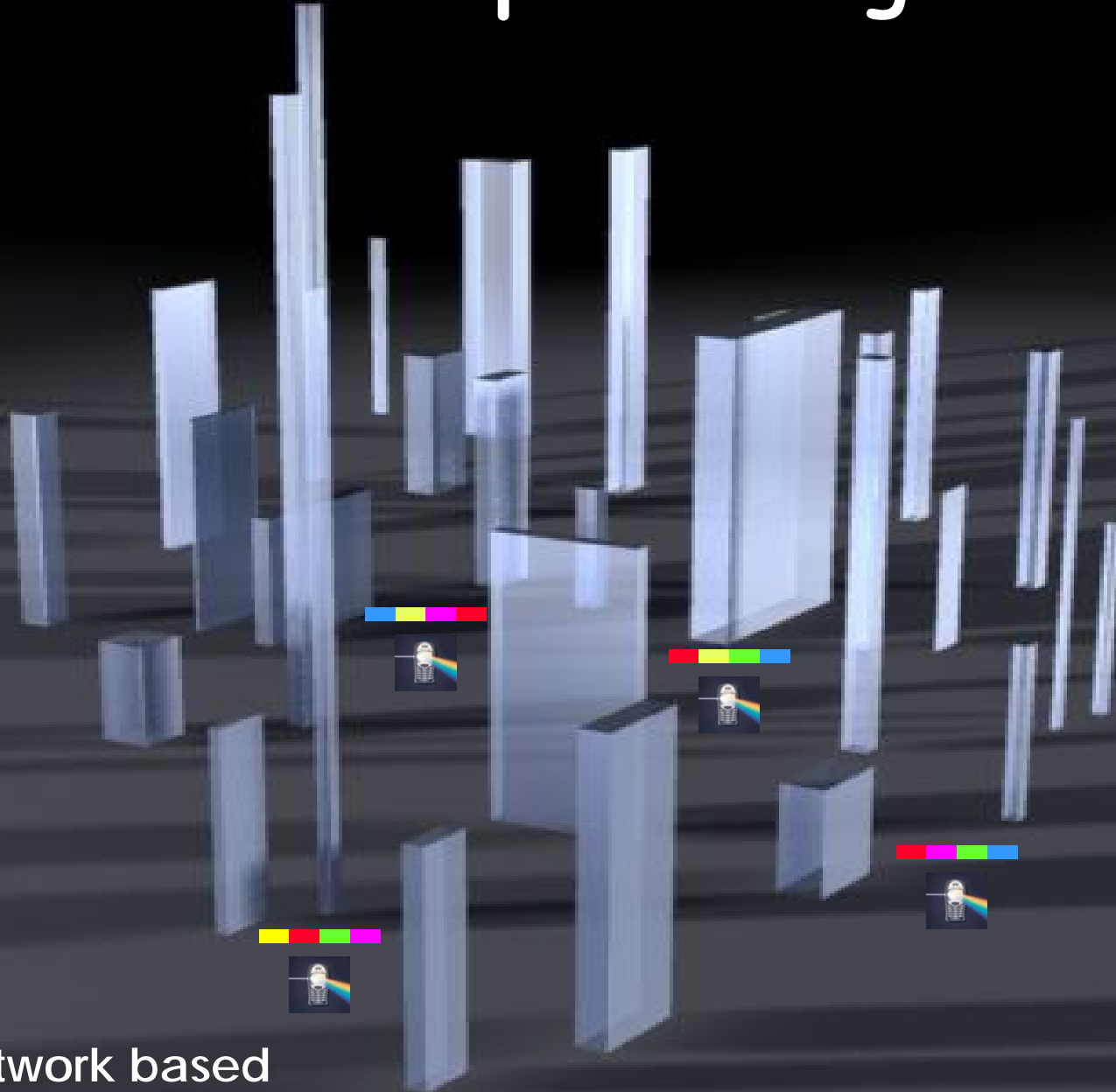


- multi-paths
- no line of sight to BS
- extra hardware

Modeling approach



Urban positioning

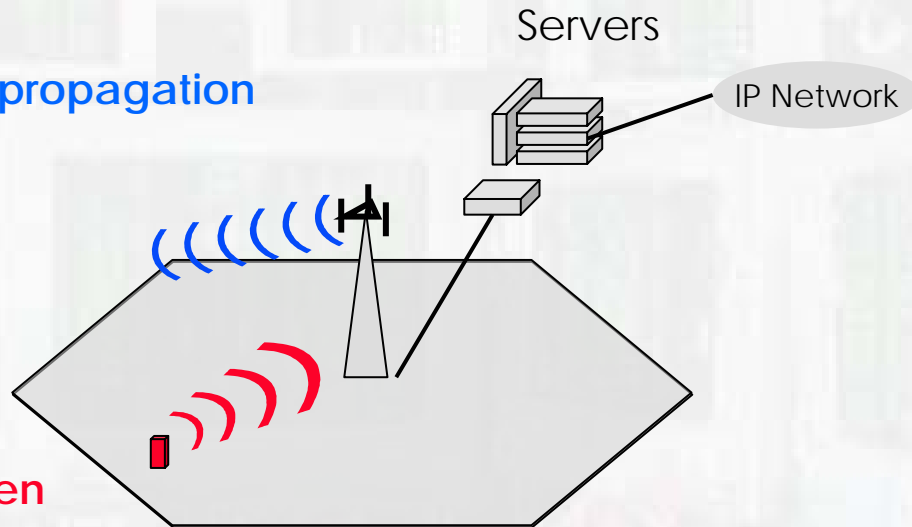


+accurate
+handset or network based
-modeling required

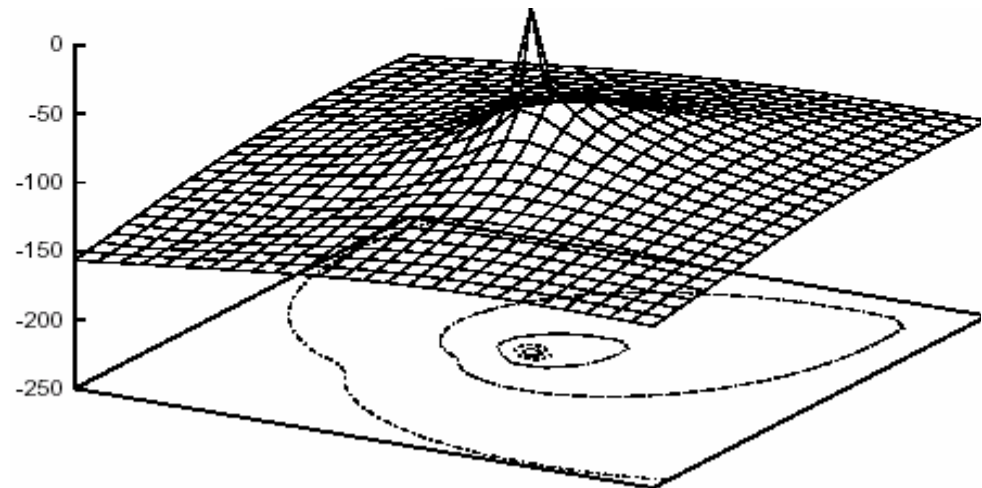
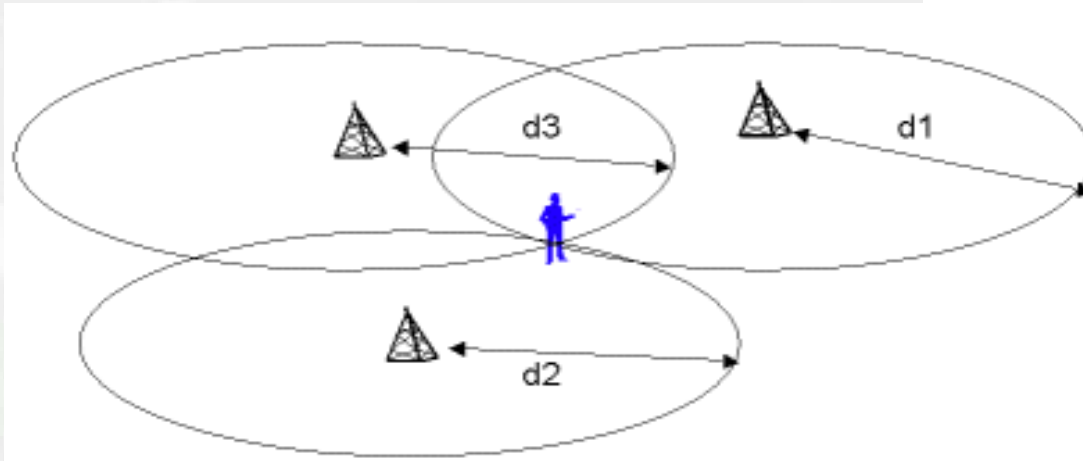
Modeling alternatives

"Dowlink modeling: the propagation models (GM level 1)"

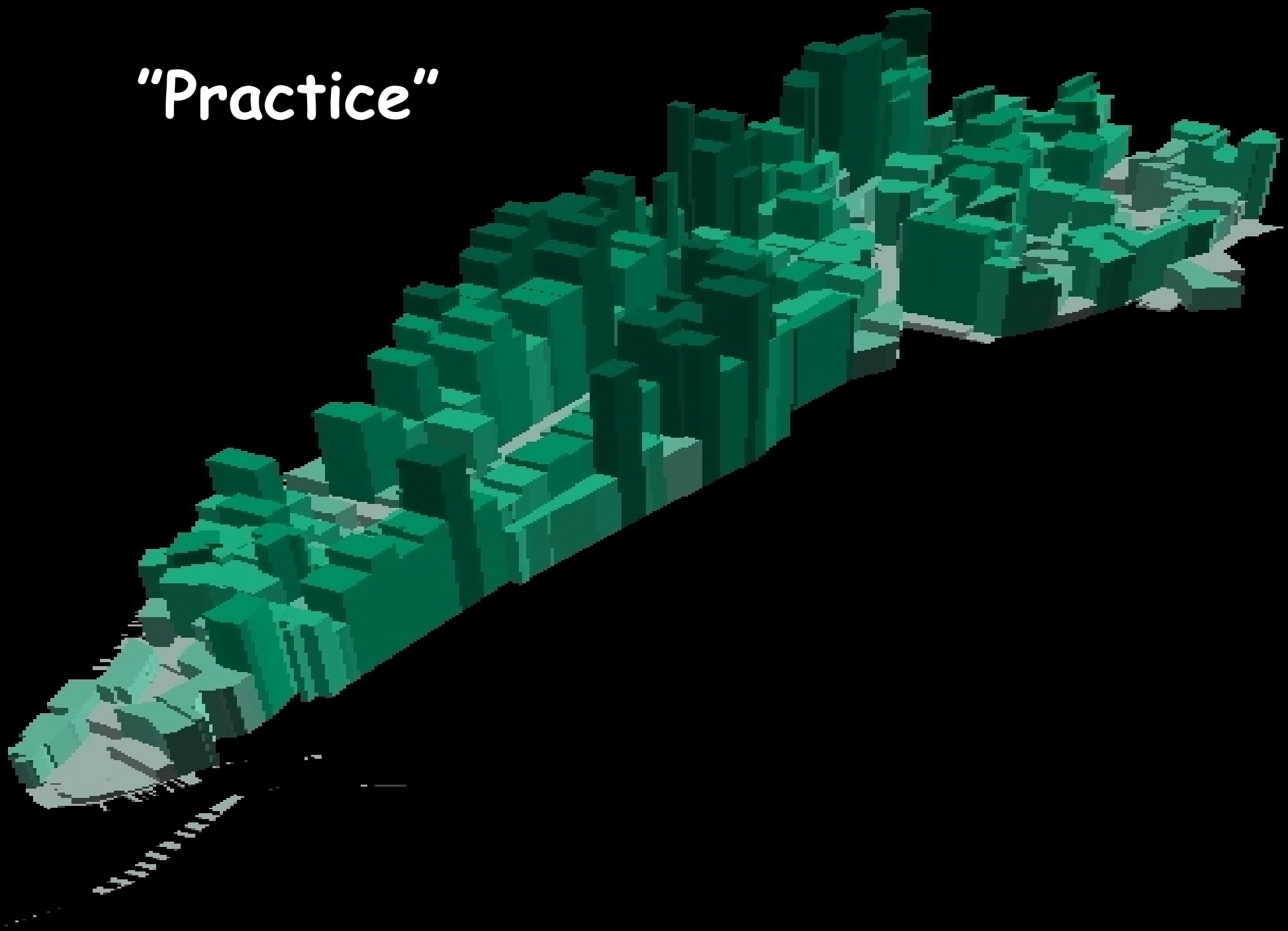
"Uplink modeling: data driven models (GM level 2)"



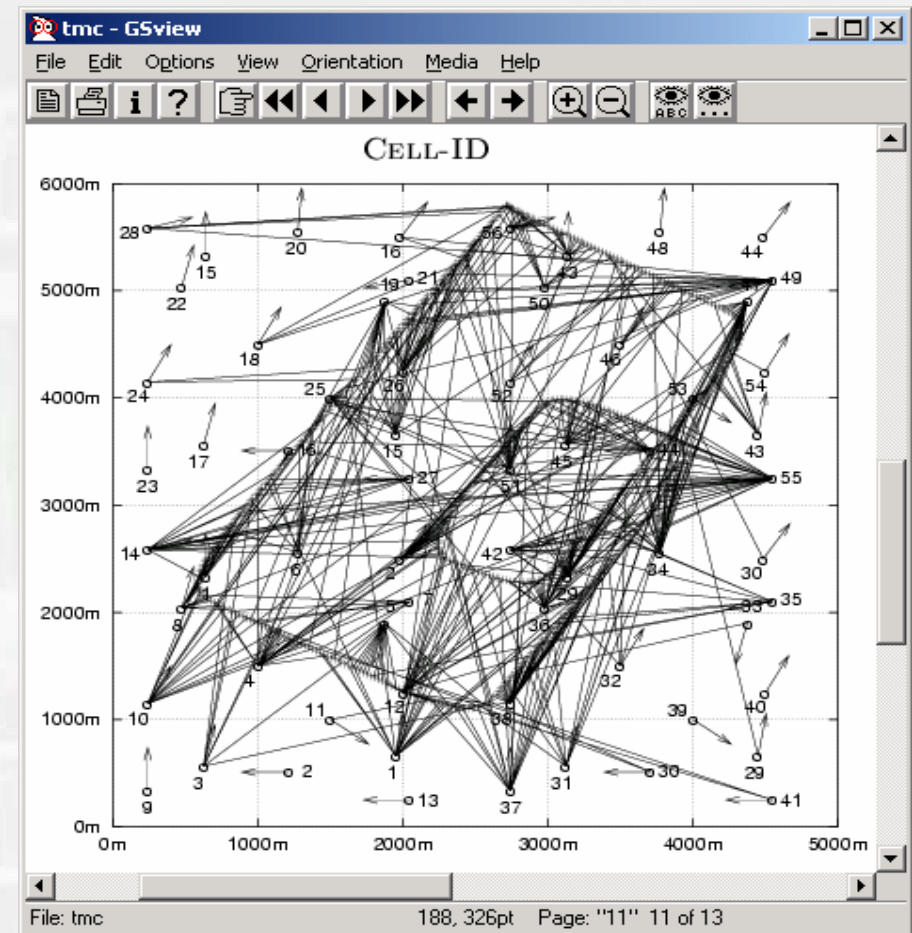
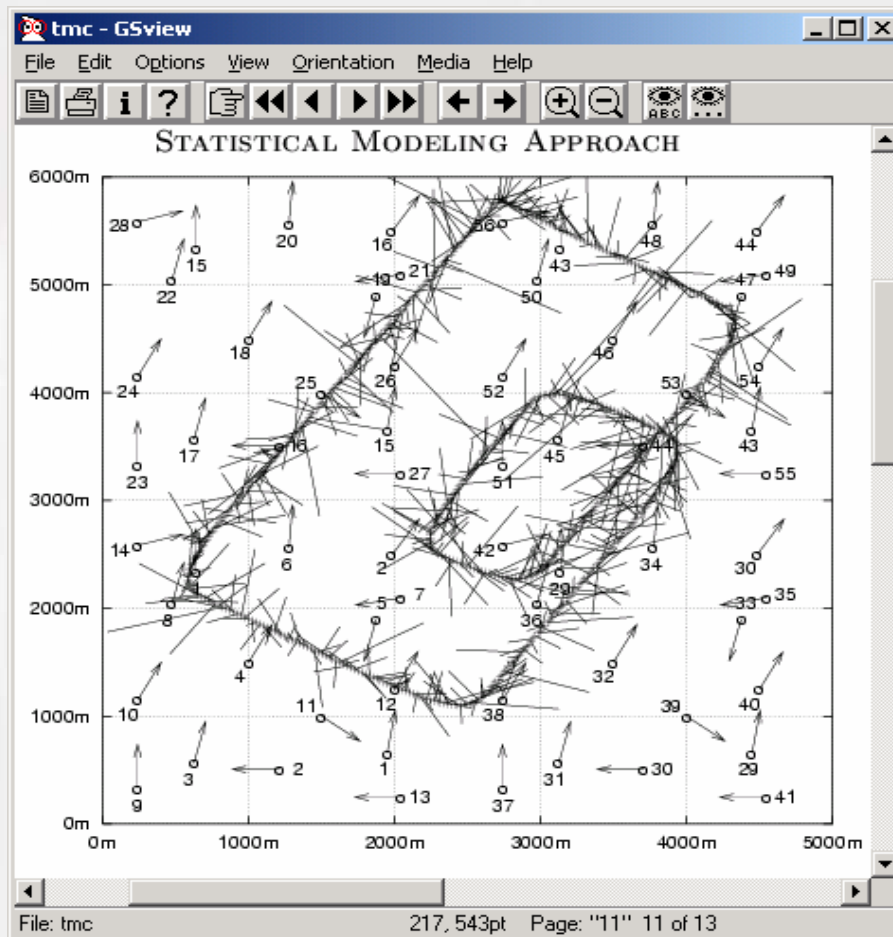
"Theory"



"Practice"



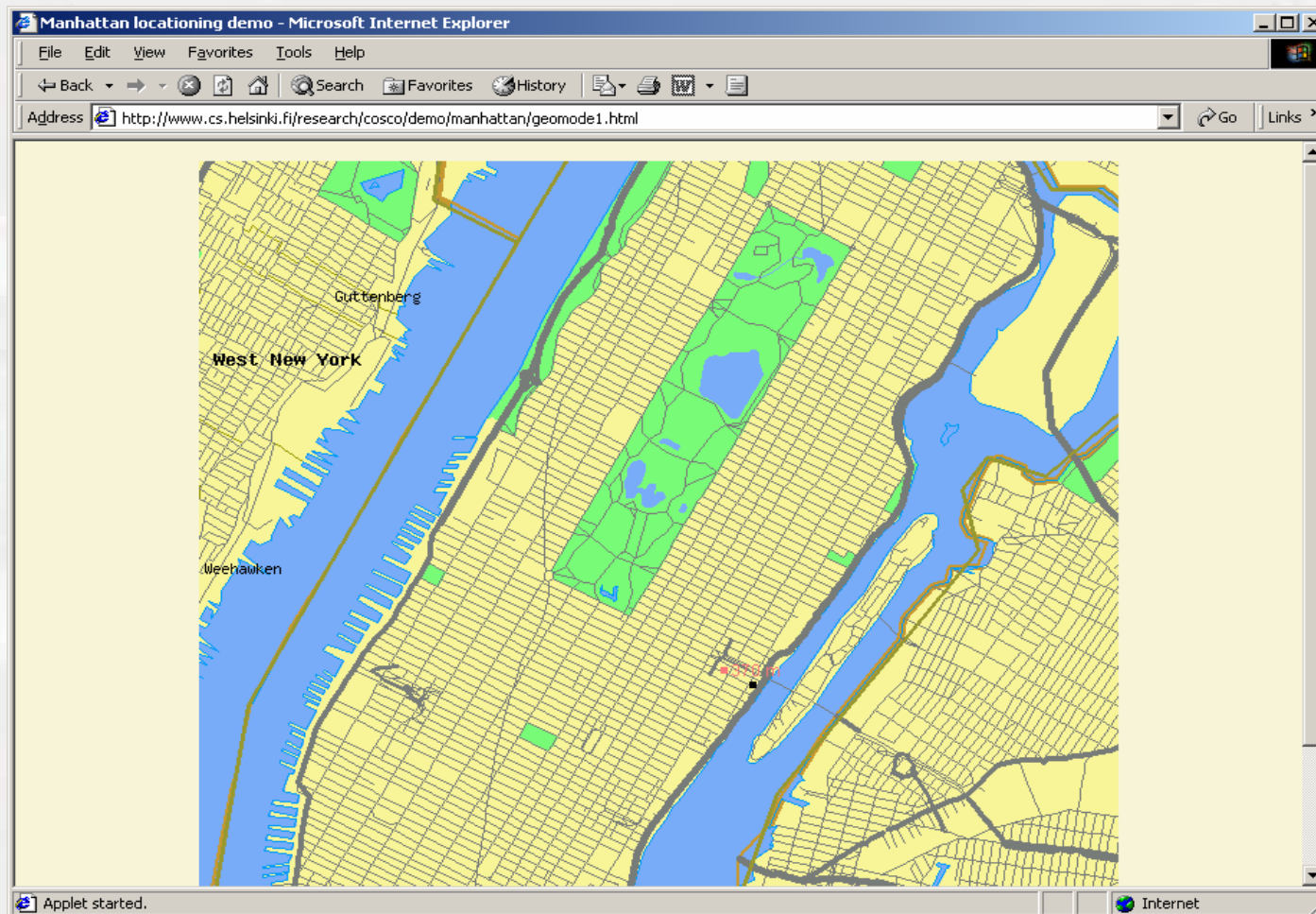
Model-based vs. Cell ID positioning errors visualization



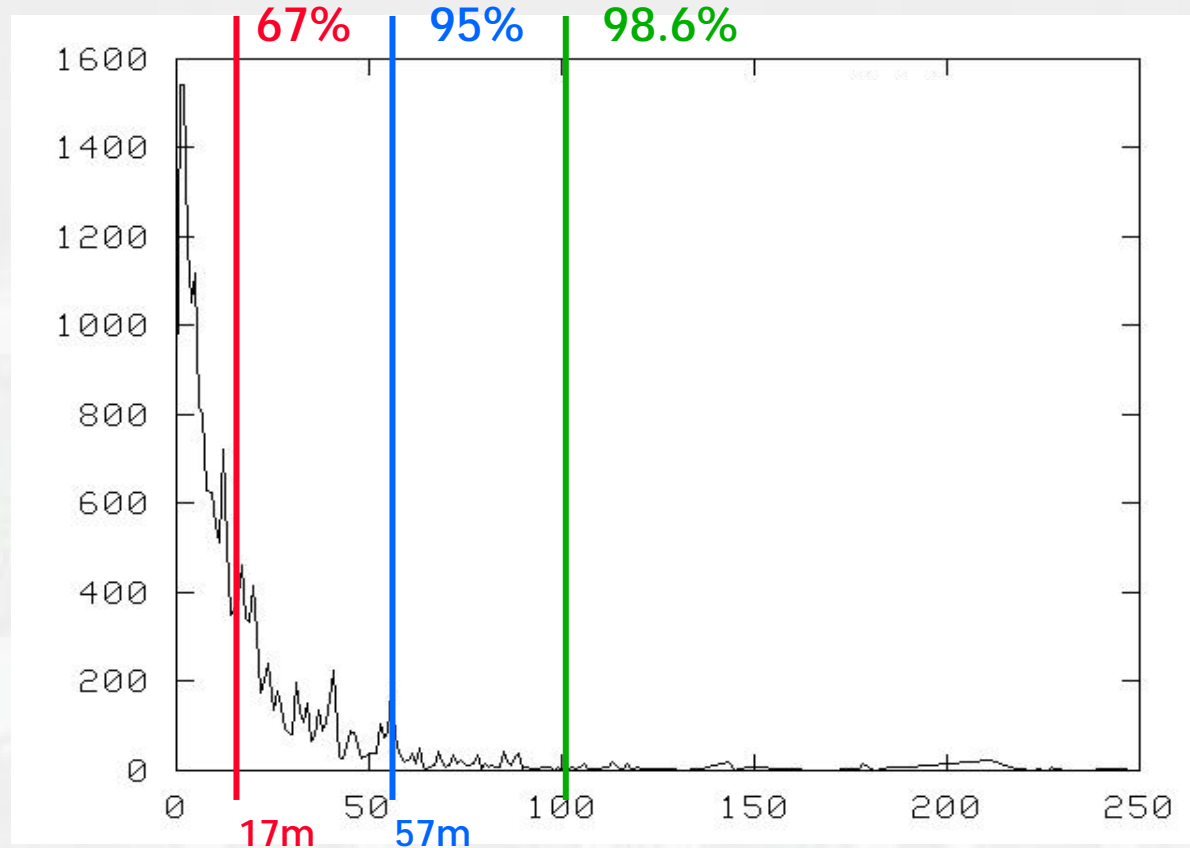
Mobile Positioning: Manhattan Trials



NYC Trial 2001

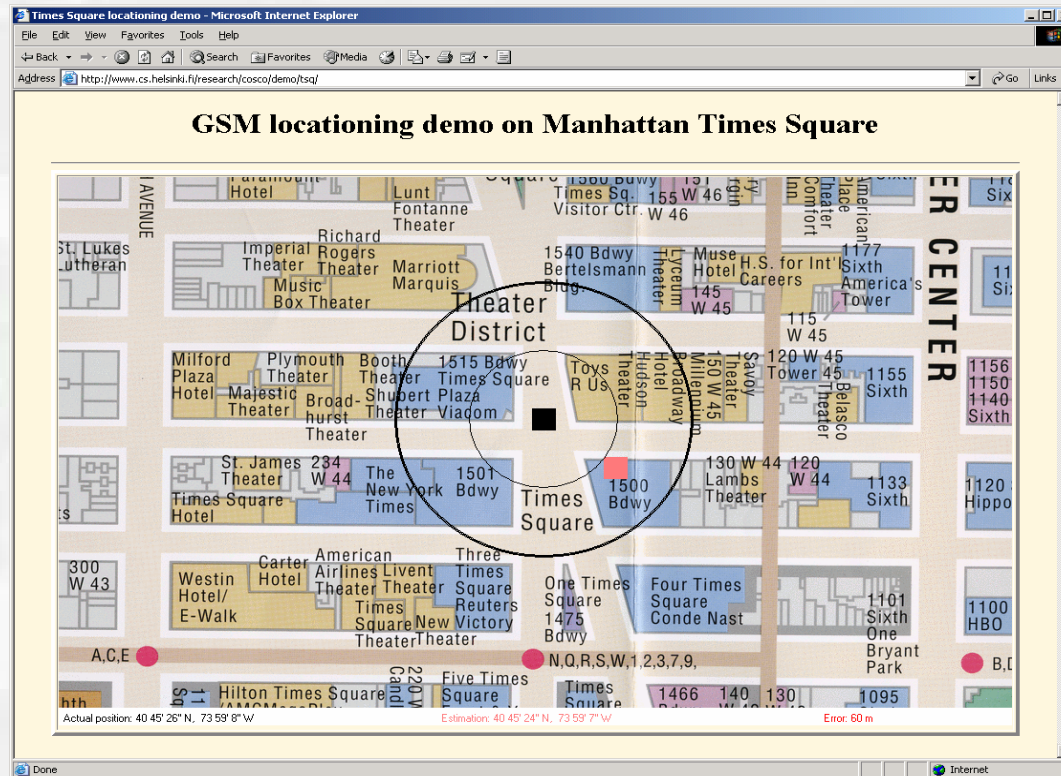


Accuracy NYC Trial 2001

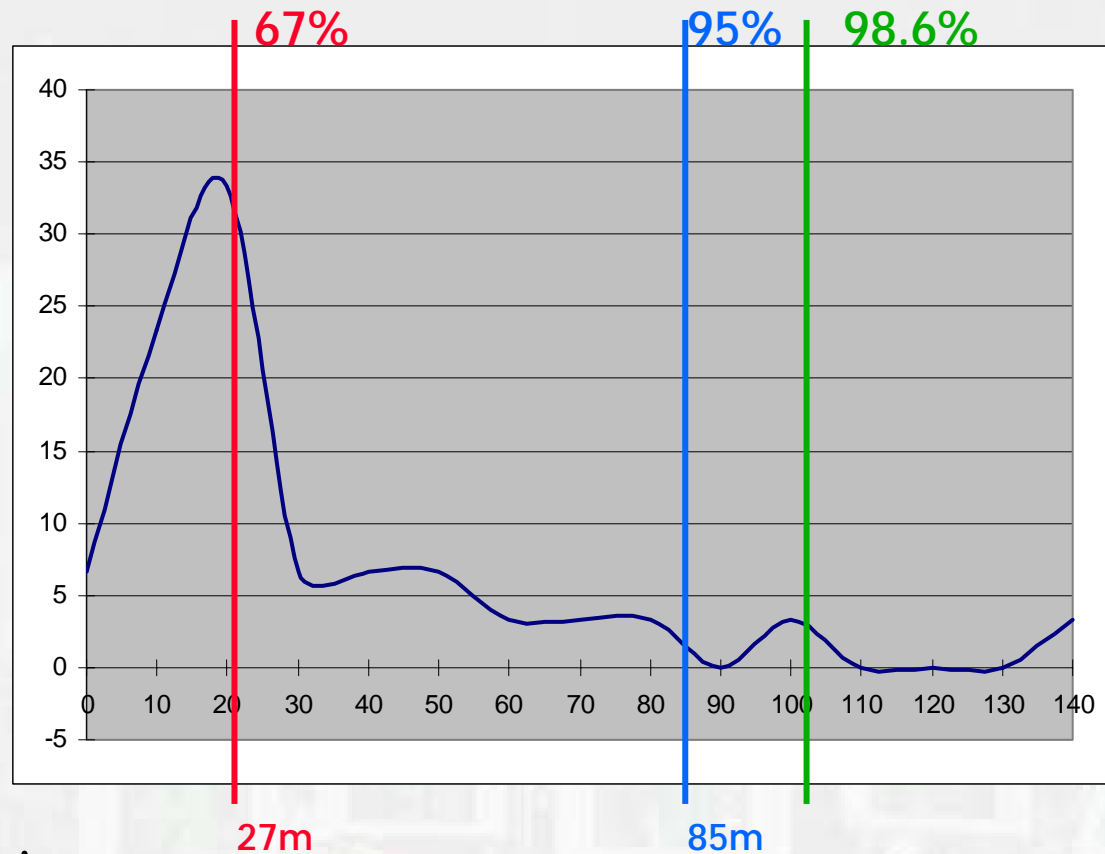


- 20166 points
- tracking; testing done in a car;

Trials: Manhattan 2002

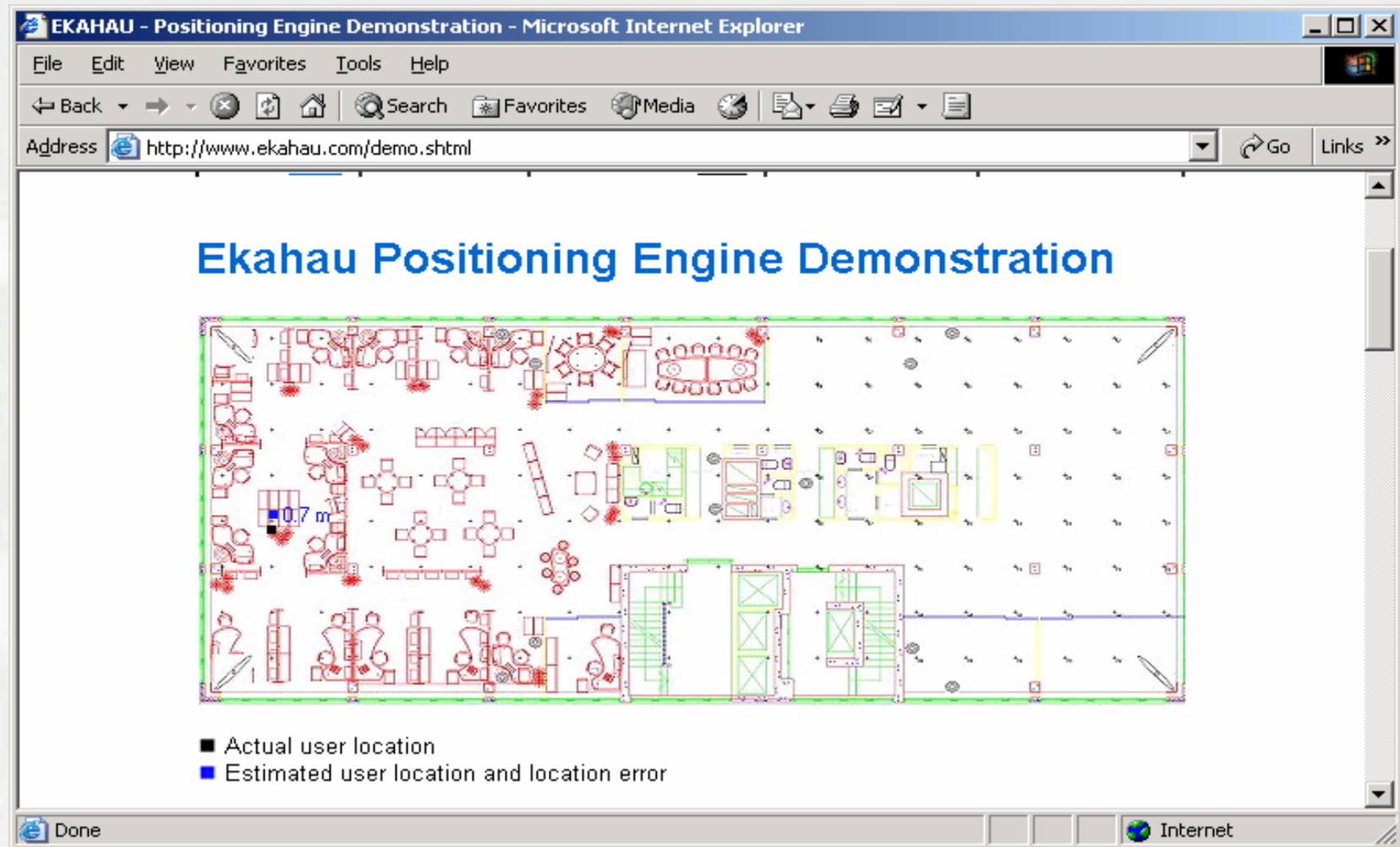


Accuracy NYC Trial 2002



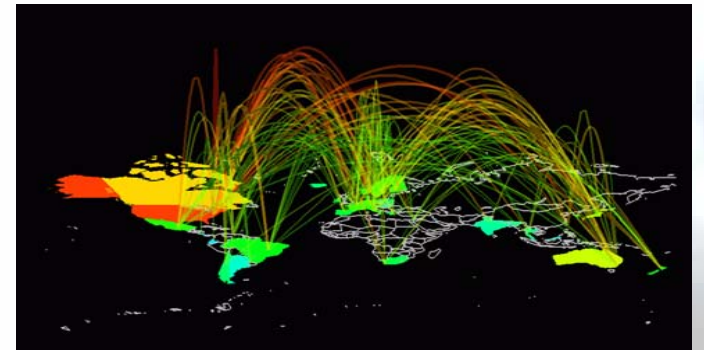
- 30 points
- static; testing done by walking;

Genaralization: WLAN

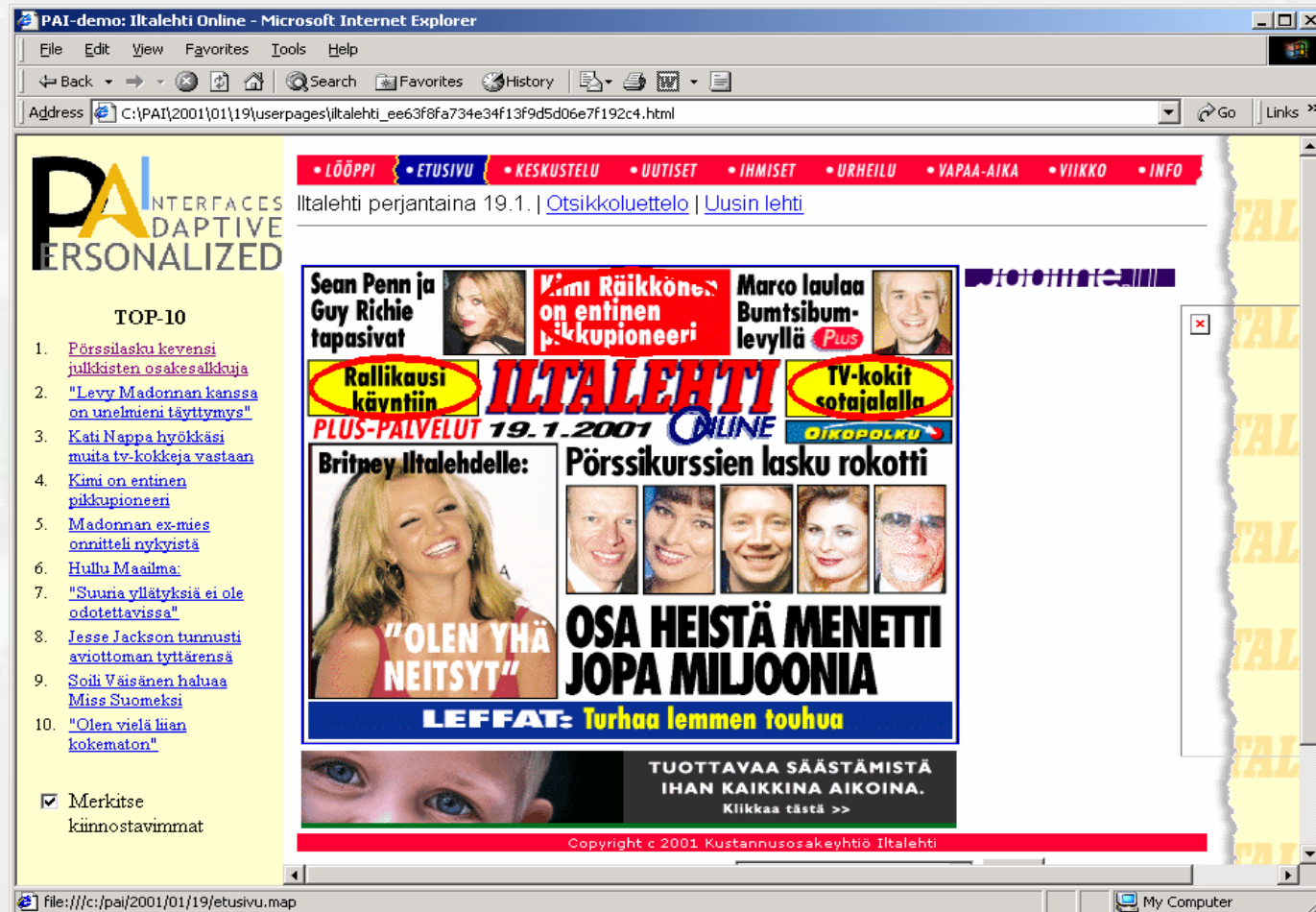


Personal WEB

Personalization and Intelligent Search



Personalized, adaptive interfaces



Future of Search



Prediction



"By 2010, the amount of data online in Internet and corporate intranets exceeds one **yottabyte** (10^{24})"

Knowledge workers spend 35% of their productive time searching for information online, while 40% of the corporate users report they cannot find the information they need to do their jobs on the intranet.

--Working council of CIOs, Business Wire, February 2001

Problems with

- Google's immortal cookie (2038)
- Google records all they can
- Google retains all data indefinitely
- Google ignores privacy policy questions
- Google hires spooks (NSA)
- Google's toolbar is spyware
- Google's cache is problematic
- Google's database as a data mining resource (150 million searches/day)

It's a war...



Next Generation Information Management Research

Collaborative search



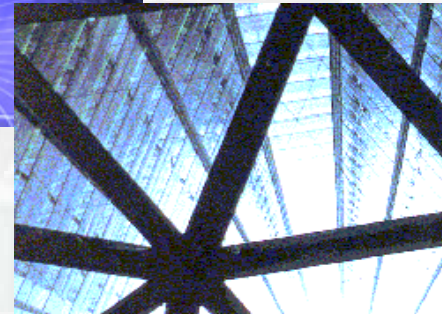
Search-Ina-BOX



YDIN



Mobile search



Topic-specific search

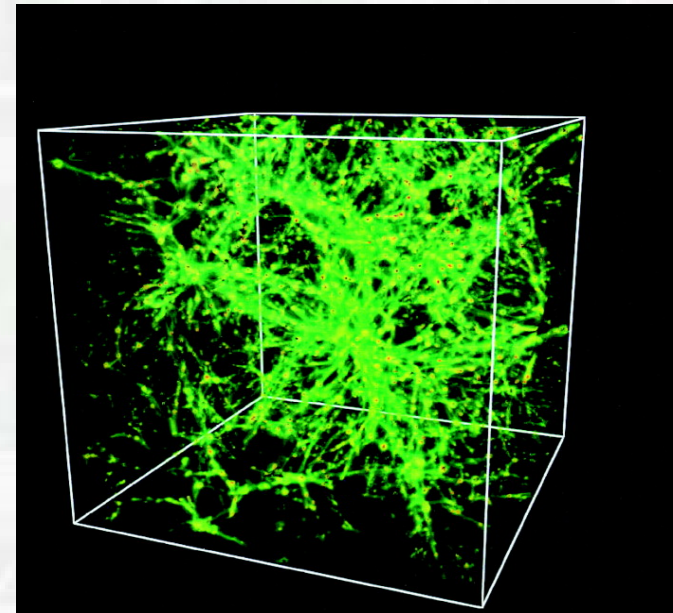
Emerging key technologies

***"Chelsea Football Club",
"Clinton DeLeon" and
"Chelsea Clinton"***

- language processing
- (probabilistic) statistical modeling
- personalization
- collaborative filtering

What is NOT there

- no major search engine performs any sophisticated language processing (scalability)
- no open source engines provide language processing
- non-interfering and (search) integrated personalization
- no quality corporate-wide intranet search systems (several sites)



Related projects

Scalable Probabilistic Methods
for the Next Generation Search Engine (PROSE)



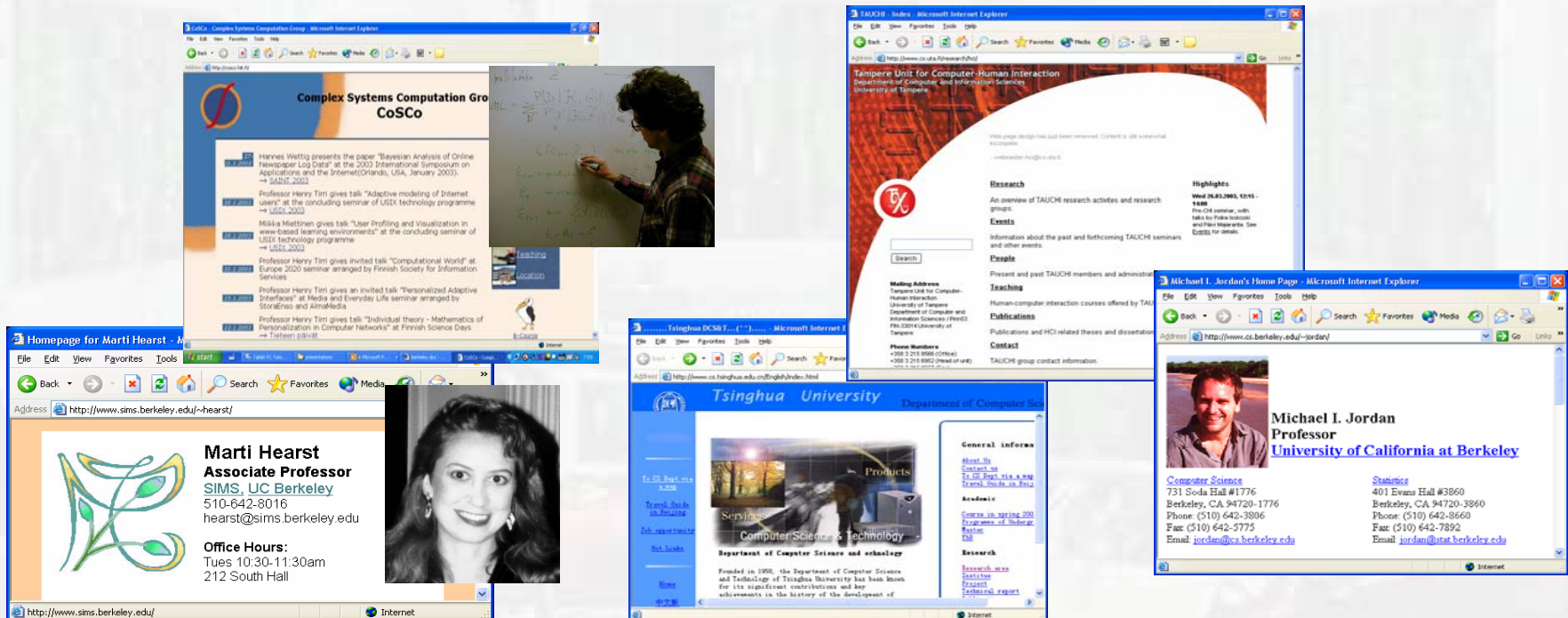
Search-Ina-BOX



ALVIS (FP6 STREP)

SIB research core competencies

- probabilistic modeling for embedded language models, automatic content analysis and personalization
- advanced user interface design

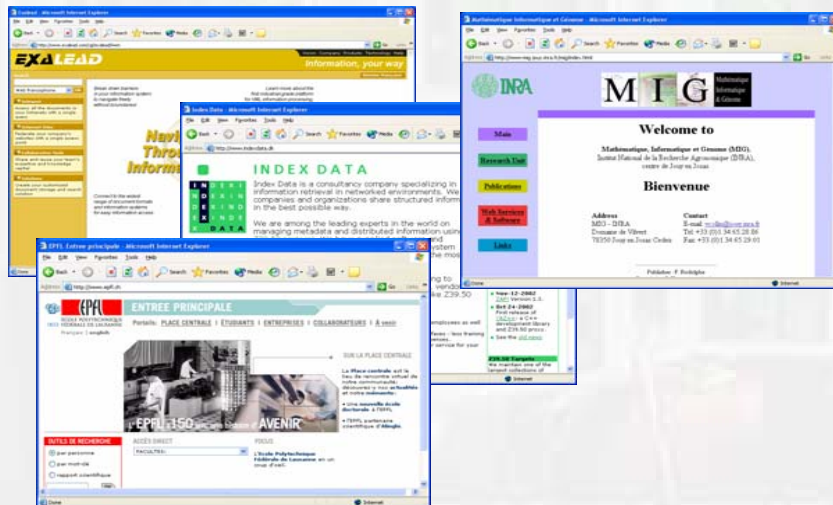


Superpeer Semantic Search (ALVIS)



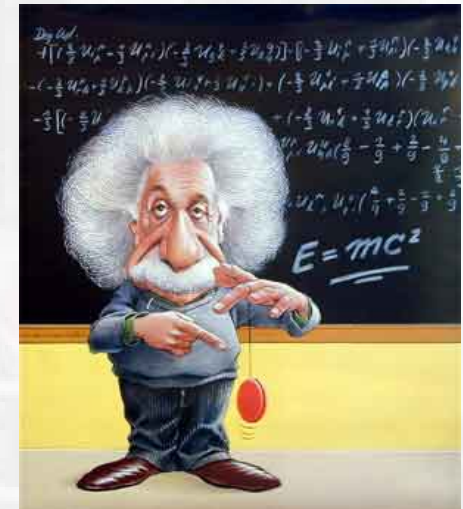
ALVIS (FP6 STREP)

- HIIT
- MIG-INRA (France)
- EPFL (Switzerland)
- Lund University (Sweden)
- JSI (Slovenia)
- Index Data (Denmark)
- Exalead (France)
- DTV (Denmark)
- Université Paris Nord

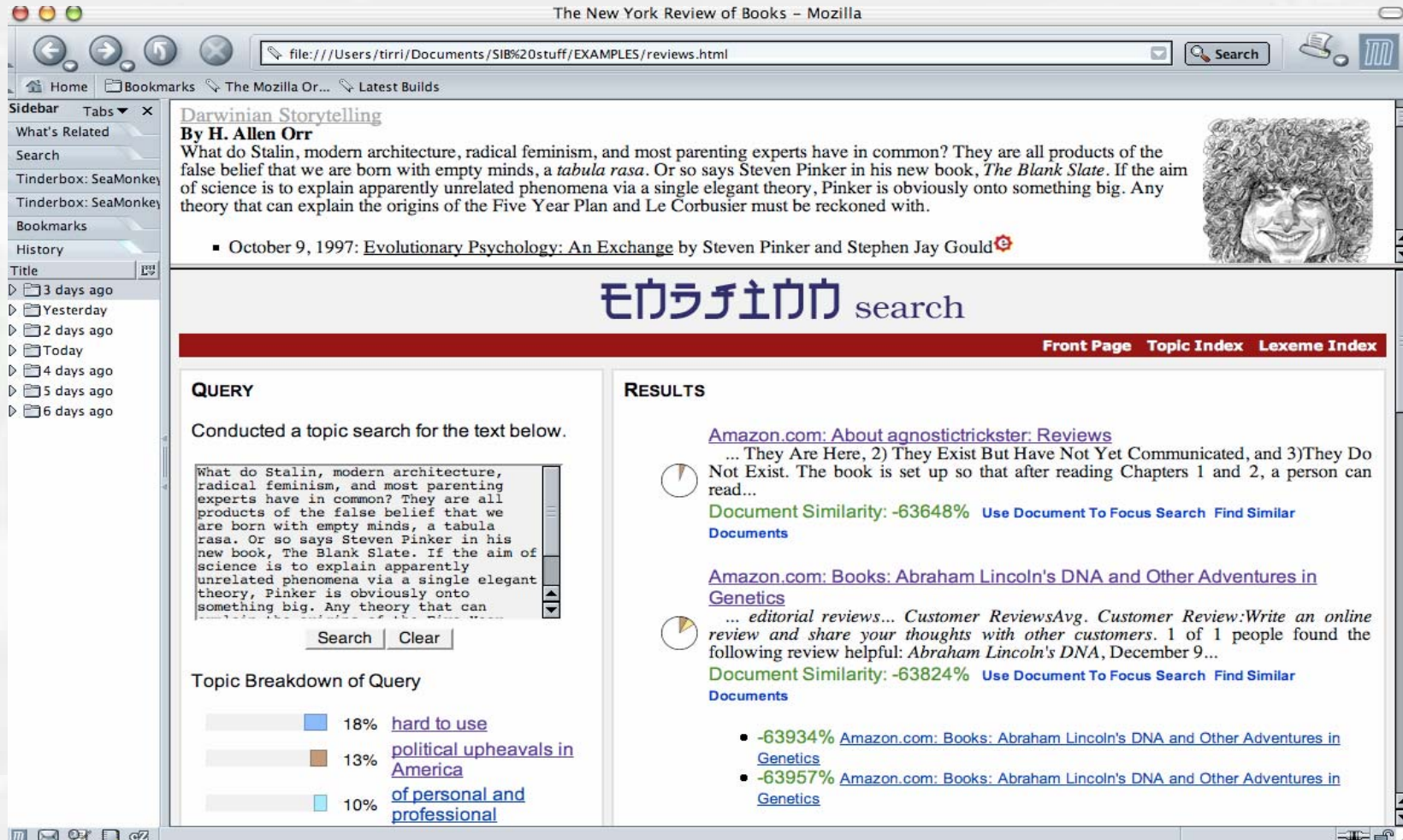


モラルチエール search

- Open Source
- Advanced language model (hidden from user)
- Probabilistic query models
- Integrated personalization
- Superpeer architecture



モラルチェンจ์ search



The New York Review of Books – Mozilla

file:///Users/tirri/Documents/SIB%20stuff/EXAMPLES/reviews.html

Home Bookmarks The Mozilla Or... Latest Builds

Search

Search

Tinderbox: SeaMonkey

Tinderbox: SeaMonkey

Bookmarks

History

Title

3 days ago

Yesterday

2 days ago

Today

4 days ago

5 days ago

6 days ago

モウラチ土の search

Front Page Topic Index Lexeme Index

Component '64' Summary

Component Number: 64

Empirical Proportion: 0.003218

Dirichlet Proportion: 0.001197

Expected No. Of Lexemes: 1285.36

Expected No. Of Documents: inf

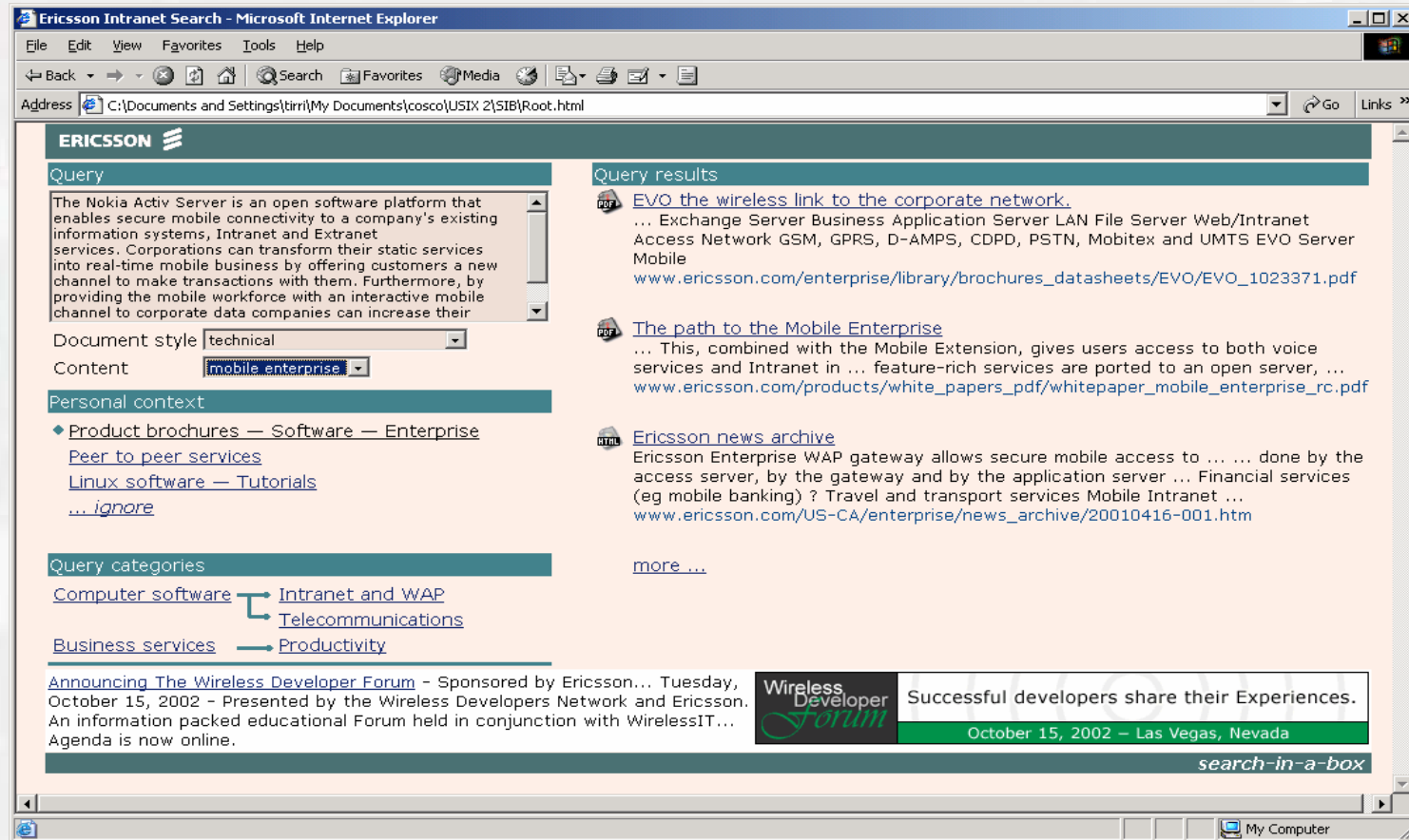
TYPICAL LEXEMES IN COMPONENT (ORDERED BY PROBABILITY)

[Anime](#) (0.050), [Japanese](#) (0.016), [Buffy](#) (0.016), [series](#) (0.014), [episode](#) (0.012), [Vol](#) (0.011), [characters](#) (0.011), [Pokemon](#) (0.011), [DVD](#) (0.010), [fan](#) (0.008), [by](#) (0.008), [movie](#) (0.007), [animation](#) (0.007), [Sailor Moon](#) (0.006), [English](#) (0.006), [Japan](#) (0.006), [Animated](#) (0.006), [Vampire Slayer](#) (0.005), [Otaku](#) (0.005), [manga](#) (0.005), [story](#) (0.004), [Ash](#) (0.004), [Gundam](#) (0.004), [Miyazaki](#) (0.004), [Tenchi](#) (0.004), [Ranma](#) (0.003), [Slayer](#) (0.003), [into](#) (0.003), [watch](#) (0.003), [list](#) (0.003), [battle](#) (0.003), [Akira](#) (0.003), [See](#) (0.003), [viewer](#) (0.003), [Save](#) (0.003), [volume](#) (0.003), [Angel](#) (0.003), [Pokémon](#) (0.003), [Collection](#) (0.003), [Dragon Ball Z](#) (0.003), [first](#) (0.003), [New](#) (0.003), [cartoon](#) (0.003), [plot](#) (0.003), [Sellers](#) (0.003), [get](#) (0.003), [Video](#) (0.002), [TV](#) (0.002), [show](#) (0.002), [Genres](#) (0.002), [original](#) (0.002), [dubbed](#) (0.002), [seen](#) (0.002), [Pikachu](#) (0.002), [Search](#) (0.002), [world](#) (0.002), [Tokyo](#) (0.002), [evil](#) (0.002), [Pioneer](#) (0.002), [Hayao](#) (0.002), [set](#) (0.002), [box](#) (0.002), [Goku](#) (0.002), [Advanced](#) (0.002), [color](#) (0.002), [Boxed Sets](#) (0.002), [Showtimes](#) (0.002), [team](#) (0.002), [errors](#) (0.002), [action](#) (0.002), [fight](#) (0.002), [voice](#) (0.002), [Powerpuff Girls](#) (0.002), [the](#) (0.002), [Spirited](#) (0.002), [Away](#) (0.002), [Pokémon](#) (0.002), [Correct](#) (0.002), [a](#) (0.002), [Neon Genesis Evangelion](#) (0.002), [end](#) (0.002), [Cowboy Bebop](#) (0.002), [Solomon](#) (0.002), [Uncut](#) (0.002), [best](#) (0.002), [DBZ](#) (0.002), [Love Hina](#) (0.002), [watching](#) (0.002), [comic](#) (0.002), [adventure](#) (0.002), [information](#) (0.002), [Rocket](#) (0.002), [Dragonball Z](#) (0.002), [Escaflowne](#) (0.002), [collector](#) (0.002), [dub](#) (0.002), [Sailor](#) (0.002), [Super](#) (0.001), [Master](#) (0.001), [Charles](#) (0.001), [Browse](#) (0.001), [violence](#) (0.001), [Earth](#) (0.001), [start](#) (0.001), [robot](#) (0.001), [features](#) (0.001), [B](#) (0.001), [subtitles](#) (0.001), [Top](#) (0.001), [Samurai](#) (0.001), [network](#) (0.001), [Spike](#) (0.001), [version](#) (0.001), [Transformers](#) (0.001), [Drama](#) (0.001), [Misty](#) (0.001), [vision](#) (0.001), [General](#) (0.001), [become](#) (0.001), [Gundam Wing](#) (0.001), [contains](#) (0.001), [must](#) (0.001), [Tenchi Muyo](#) (0.001), [freak](#) (0.001), [Ghost](#) (0.001), [Kiki](#) (0.001), [ending](#) (0.001), [Princess](#) (0.001), [Mononoke](#) (0.001), [VHS](#) (0.001), [Kids](#) (0.001), [Warrior](#) (0.001), [visit](#) (0.001), [Demon](#) (0.001), [mysterious](#) (0.001), [Evangelion](#) (0.001), [Anno](#) (0.001), [saga](#) (0.001), [creator](#) (0.001), [her](#) (0.001), [here](#) (0.001), [Totoro](#) (0.001), [fighting](#) (0.001), [Fanatic](#) (0.001), [language](#) (0.001),

"Ubisearch"



Search-In-a-Box concept



http://cosco.hiit.fi/search/





Intelligent Tools for E-learning

EDUFORM

EDUFORM - Microsoft Internet Explorer

EDUFORM
by
Co.S.Co

Learning Experiences and Motivation

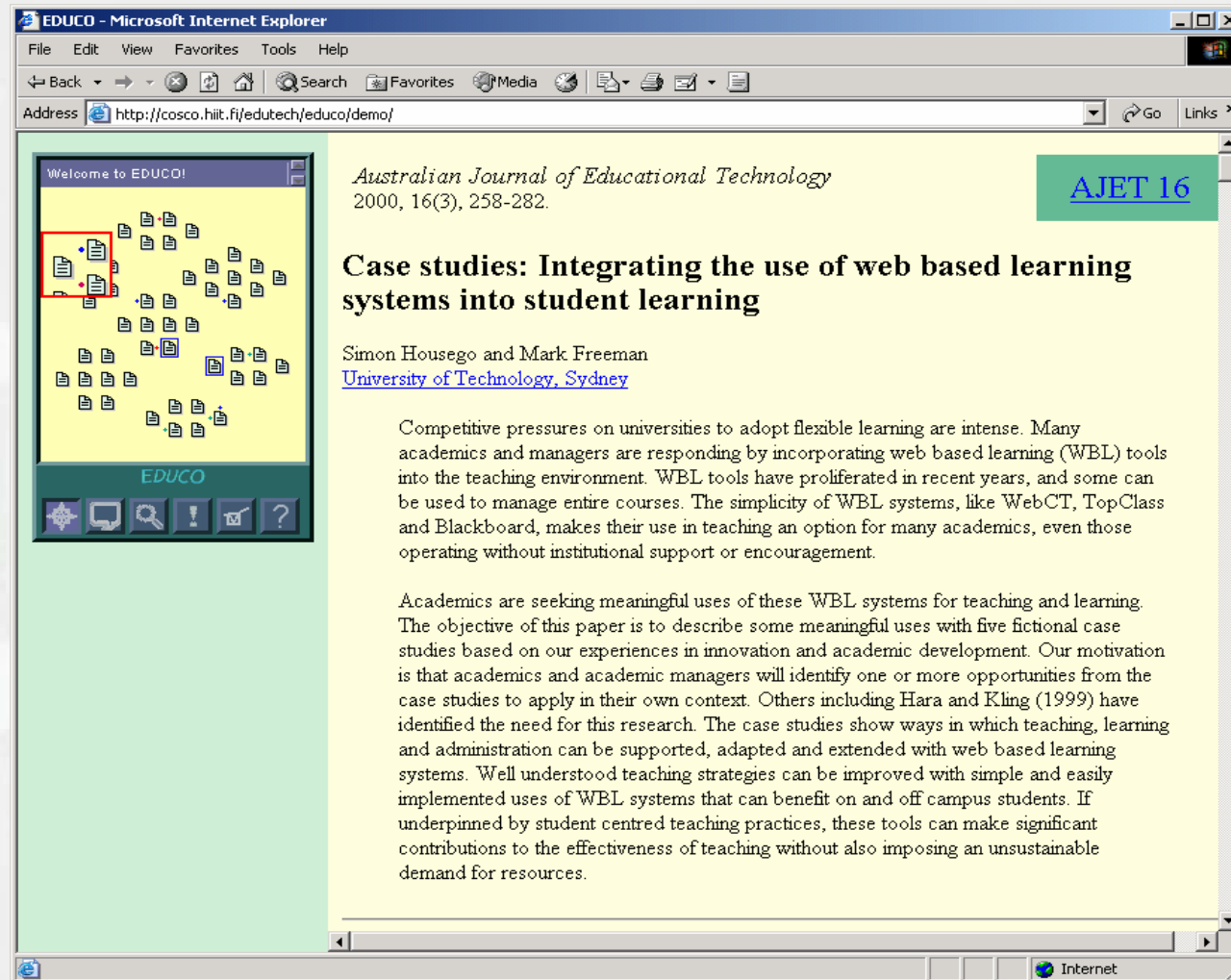
| | Disagree | 1 | 2 | 3 | 4 | Agree | 5 |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|---|
| 5. During an exam I wonder how I am performing in comparison to other students. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | |
| 6. When taking part in a practical examination I am concerned about failing and what will happen as a result. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | |
| 7. When answering essay questions I am also concerned about the other questions on the test that I cannot answer. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | |

42% left

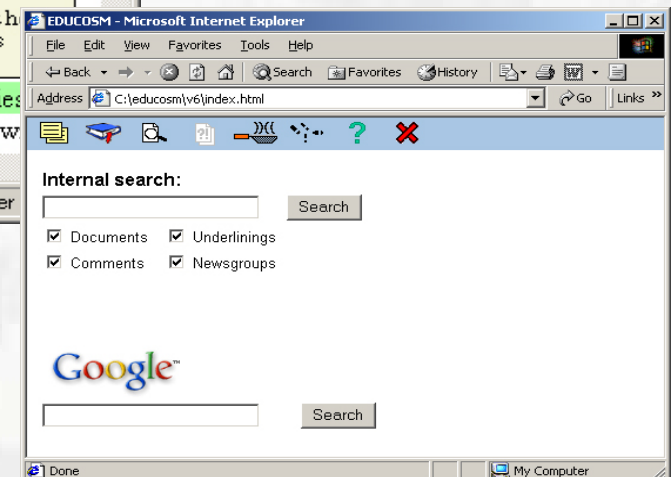
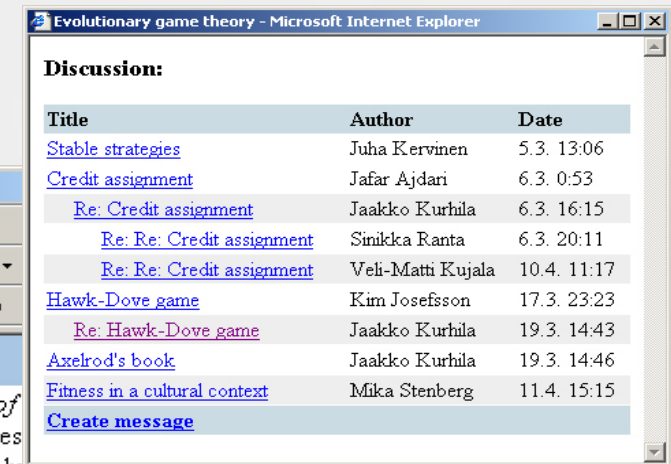
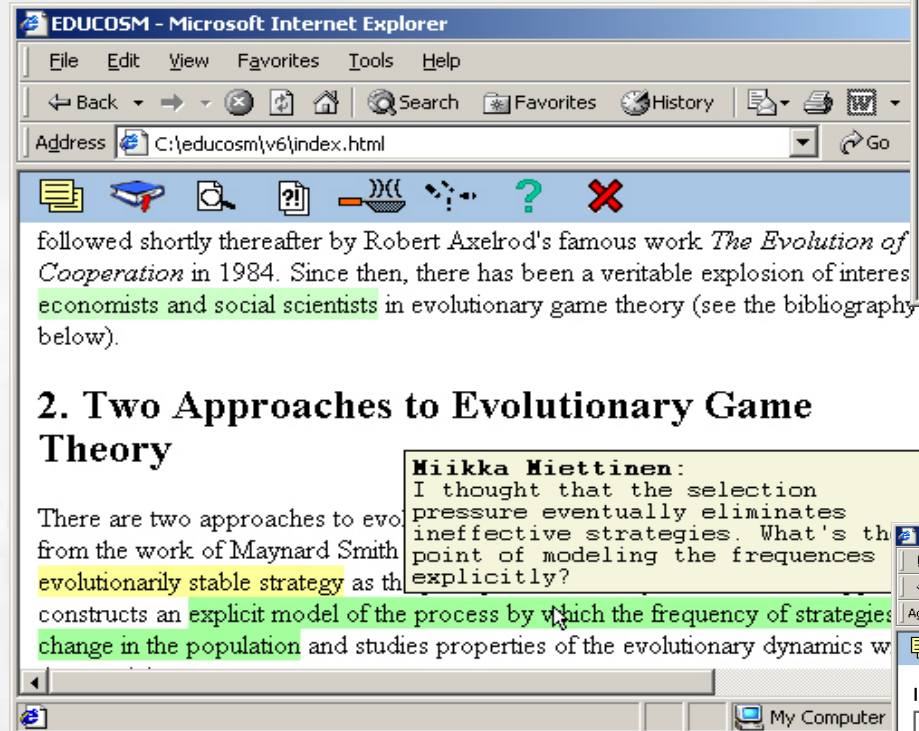
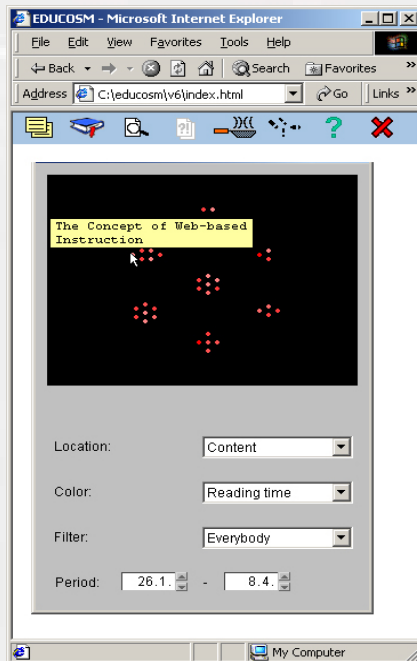
0% 100%

Internet

EDUCO



EduCosm



and other things

