Location Awareness

Dominic Duggan
Stevens Institute of Technology

Location Awareness

- Satellites and GPS
- Infrared and Ultrasonic
- LAN-based
- Cell-based

Based on materials by Charlie Leonard

SATELLITES AND GPS

3

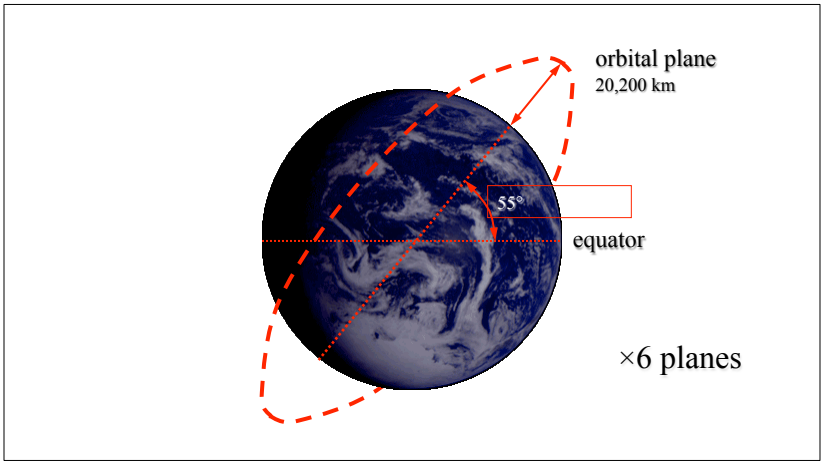
NAVSTAR GPS

Navigation Satellite Timing and Ranging

- The American Department of Defense started development in 1973
- Six orbital planes
  - plane = orbit containing multiple satellites
- 21 active satellites, plus 3 spares
  - Four per plane

4

NAVSTAR GPS



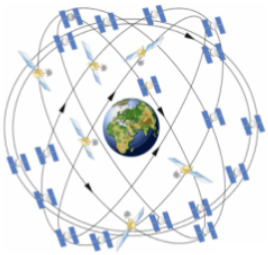
5

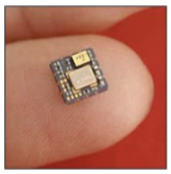
NAVSTAR GPS

- Why not geostationary at 36,000 km?
  - Transmitter
  - Launcher
  - Polar regions
- Compromise: 20,200 km so period is 12h
- However, many satellites needed
  - At least 17 satellites required
  - Today: 31 satellites = five to twelve in range!

6

NAVSTAR GPS

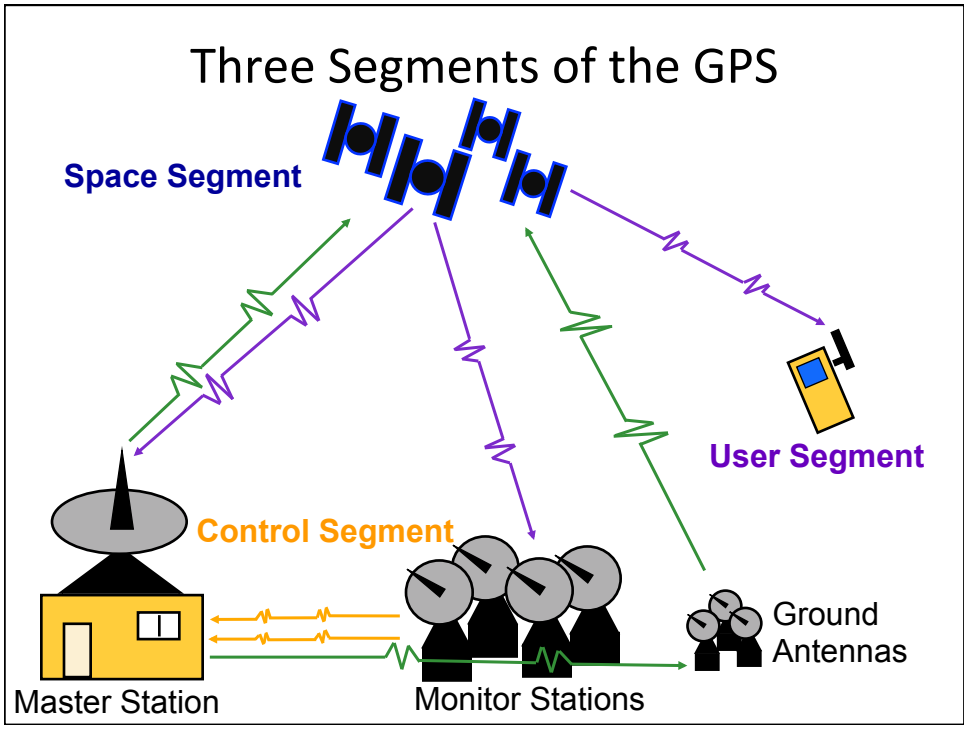


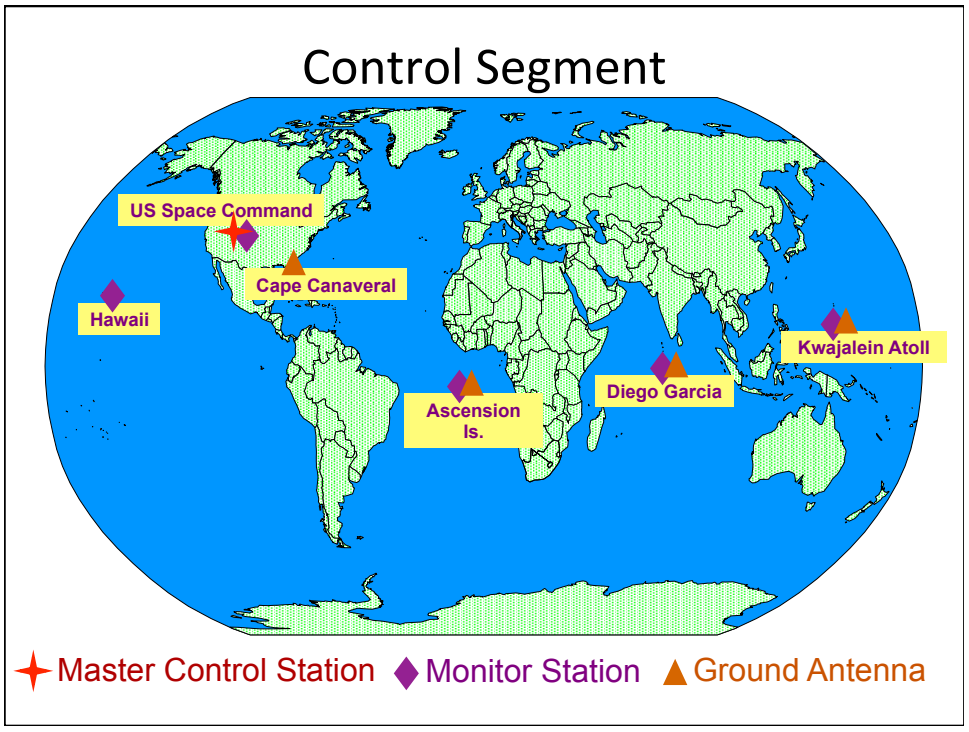


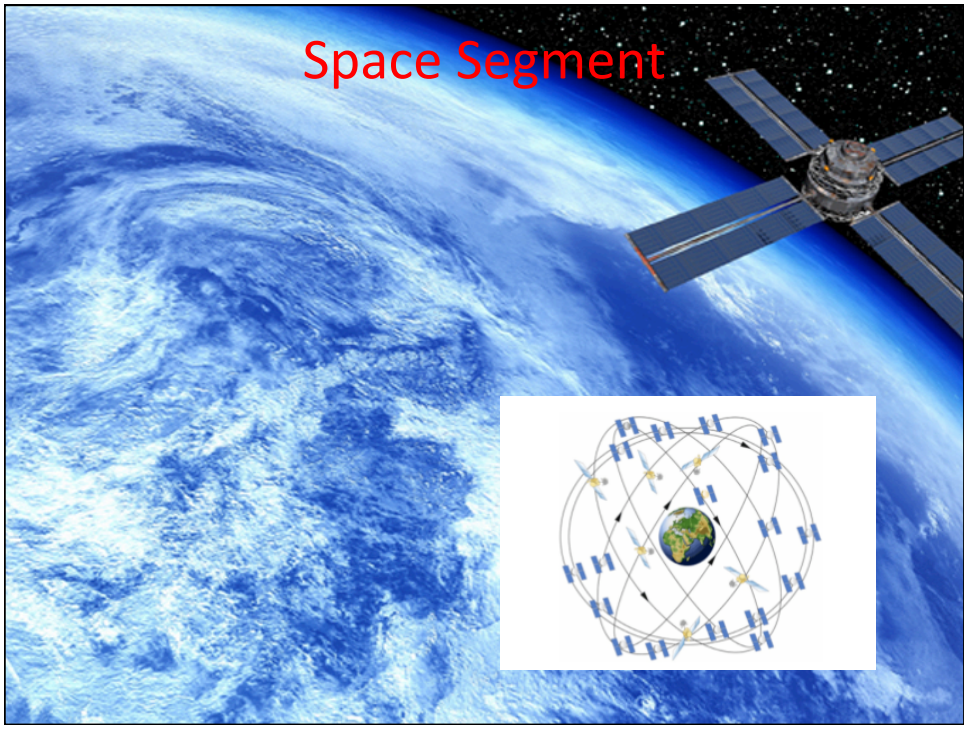
7

GPS ARCHITECTURE

8


Three Segments of the GPS
Space Segment
User Segment
Control Segment
Ground
Antennas
Master Station
Monitor Stations


Control Segment
US Space Command
Cape Canaveral
Hawaii
Kwajalein Atoll
Diego Garcia
Ascension
Is.
Master Control Station
Monitor Station
Ground Antenna



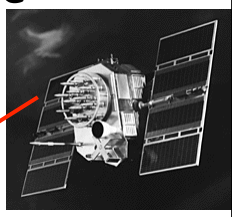
User Segment

- Military.
- Search and rescue.
- Disaster relief.
- Surveying.
- Marine, aeronautical and terrestrial navigation.
- Remote controlled vehicle and robot guidance.
- Satellite positioning and tracking.
- Shipping.
- Geographic Information Systems (GIS).
- Recreation.

Four Basic Functions of GPS

- Position and coordinates.
- The distance and direction between any two waypoints, or a position and a waypoint.
- Travel progress reports.
- Accurate time measurement.

Position is Based on Time

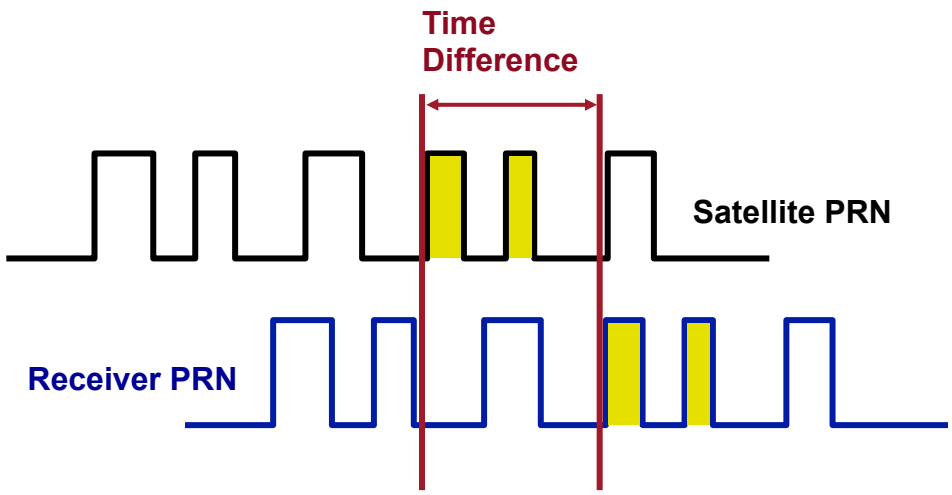


T

T + 3

Distance = "3 times the speed of light"

Pseudo Random Noise Code

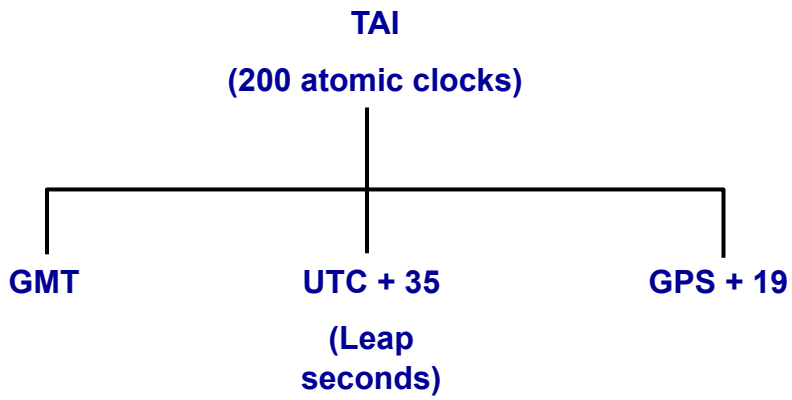


NAVSTAR GPS

- Satellites broadcast over two reserved frequencies
  - L1 frequency, at 1575.42 MHz
  - L2 frequency, at 1227.6 MHz
- L1 PRN code (civil)
  - Repeats every second
  - 3m accuracy (theoretical)
- L1 & L2 PRN code (U.S. military)
  - 10x bandwidth
  - Repeats every 38 weeks
  - 0.3m accuracy (theoretical)

16

What Time is It?



NAVSTAR GPS

- 63 milliseconds latency (1-way)
- Inaccuracy of 1 millisecond => error of 300 kilometers!
- Four atomic clocks (per satellite)
- Receiver clock?

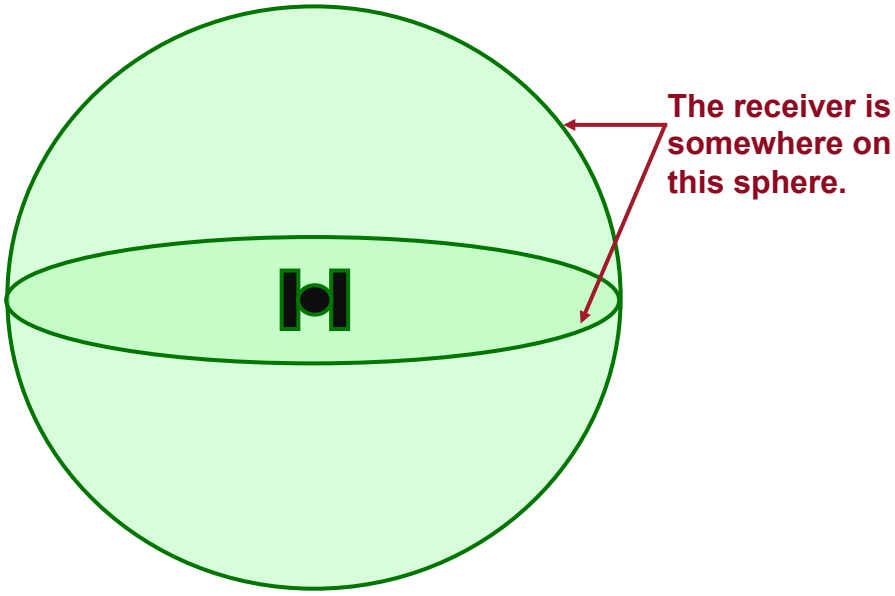
18

NAVSTAR GPS

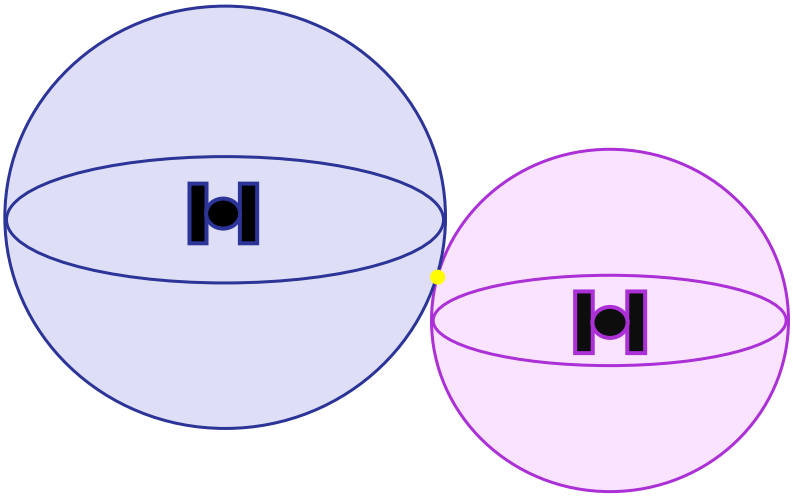
- Receiver: digital clock
- Precision vs accuracy
- Requires fourth satellite

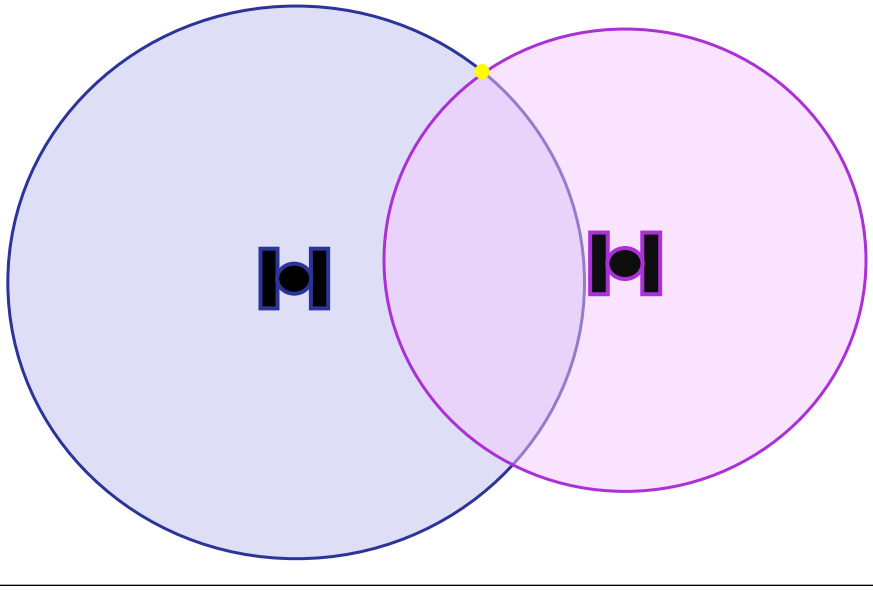
19

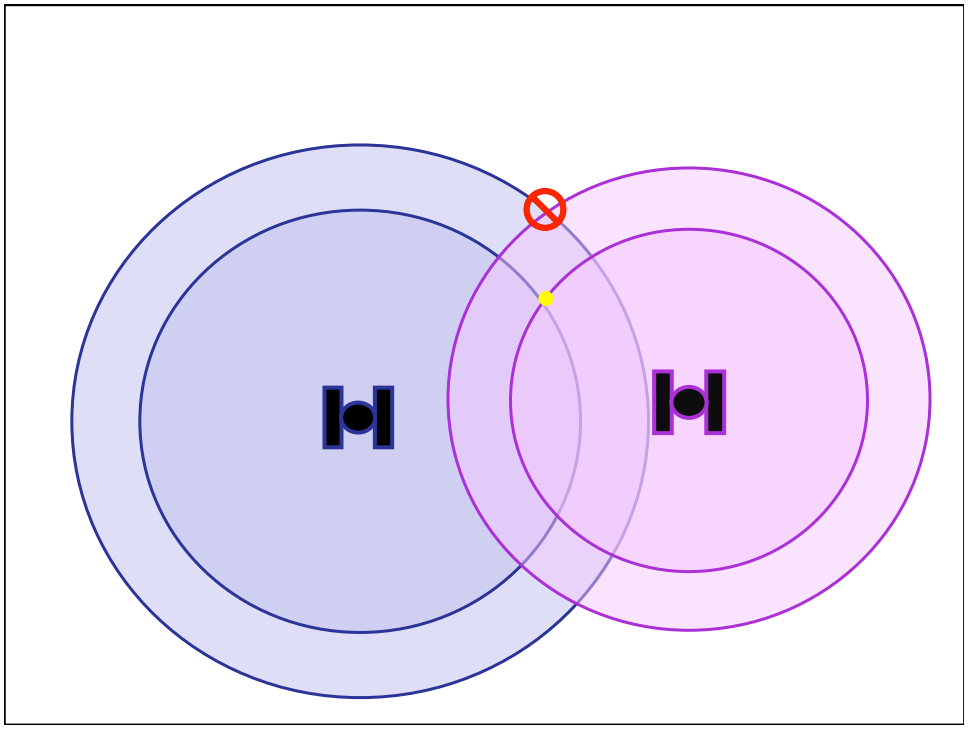
Signal From One Satellite

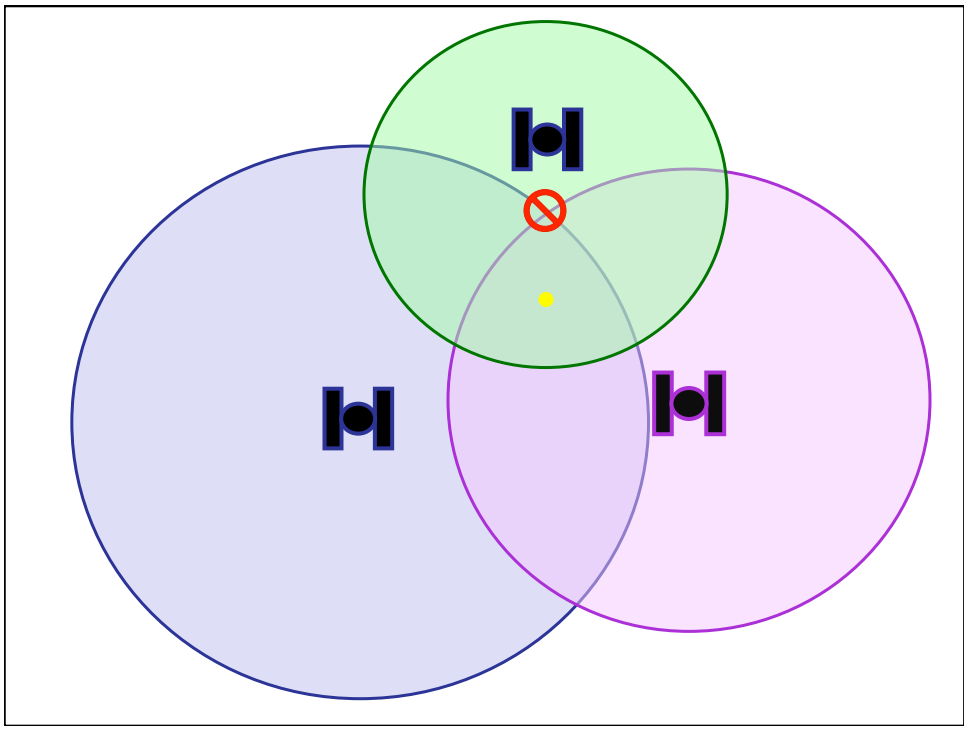


Signals From Two Satellites

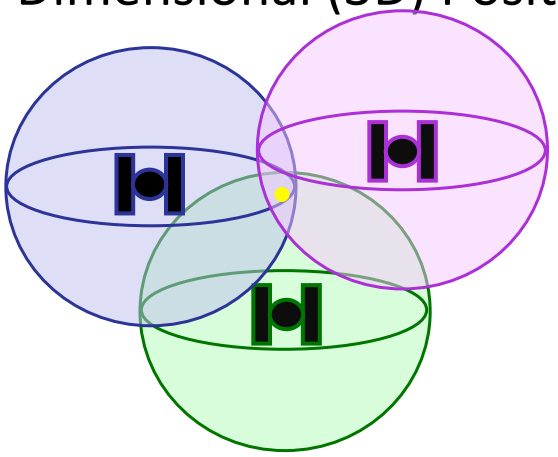




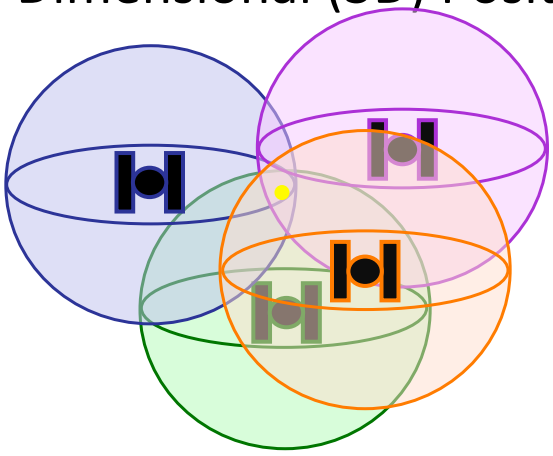




Three Dimensional (3D) Positioning

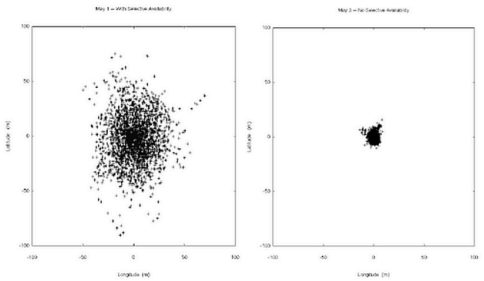


Three Dimensional (3D) Positioning



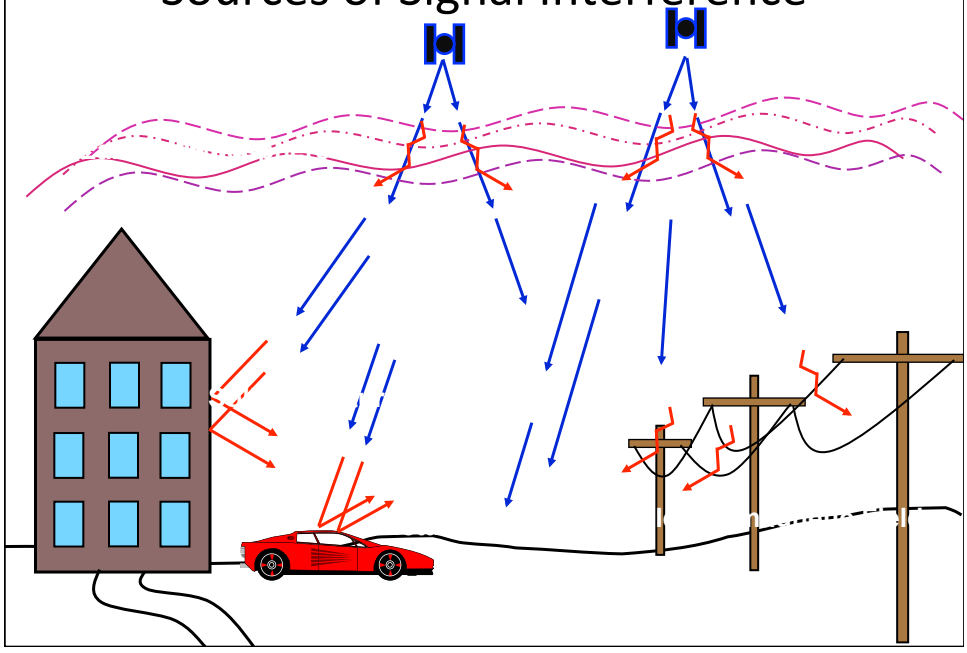
Selective Availability (S/A)

- DoD dithered the satellite time message.
- In May 2000 the Pentagon reduced S/A to zero meters error.



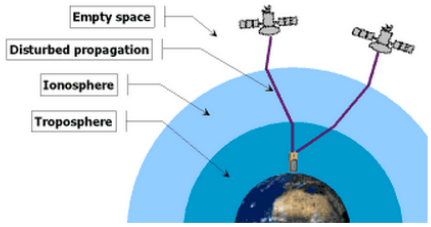
Plot of the position determination with and without SA
(Diagram from http://www.igeb.gov/sa/diagram.shtml (page no longer available)
With friendly permission of Dr. Milbert (NOAA))

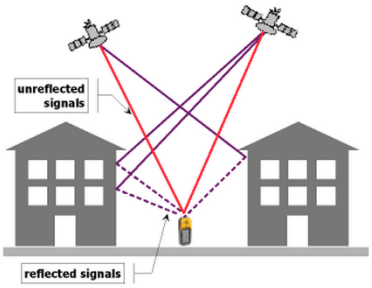
Sources of Signal Interference



Ranging Errors

- Ionosphere: 5m
  - Varies with signal frequency
  - Military: factor out w/L1,L2
- Troposphere: 0.5m
- Satellite position data: 2.5m
  - Use precise orbit data
- Satellite clock: 2m
  - Periodic corrections by ground stations
- Multipath: 1m
  - Obstacles
- Average Error: 3-5m
  - potentially 15m)





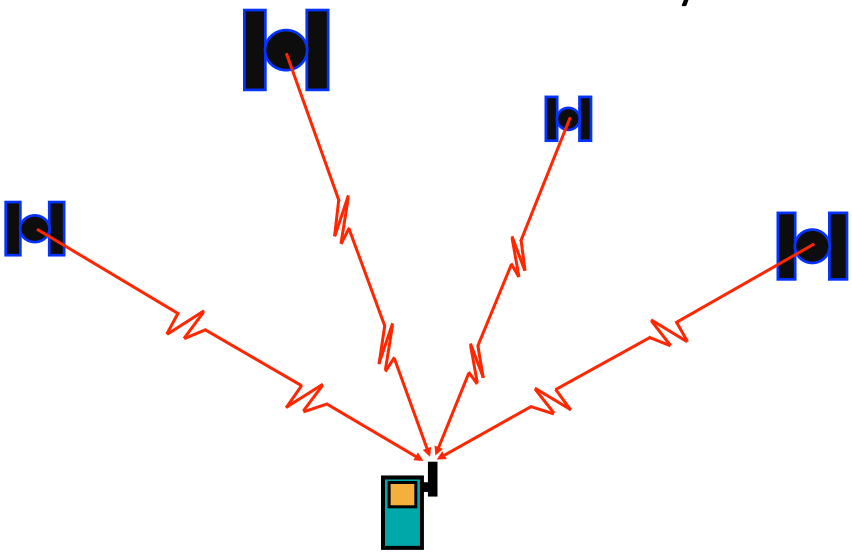
29

DILUTION OF PRECISION AND DIFFERENTIAL GPS

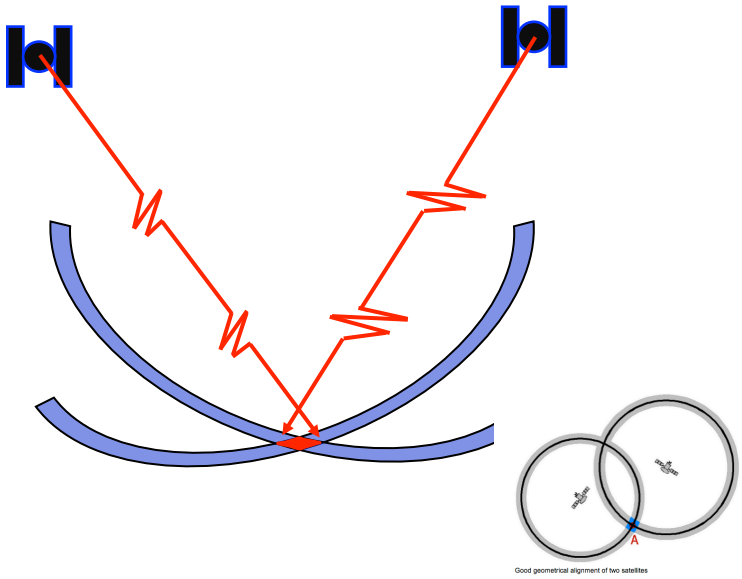
GPS Satellite Geometry

- Quality of GPS signals
- Accuracy of receiver trilateration.
- **Dilution of Precision (DOP):** each satellite's position relative to the other satellites being accessed by a receiver.
  - Position Dilution of Precision (PDOP)
- Up to the GPS receiver to pick satellites
  - best position triangulation

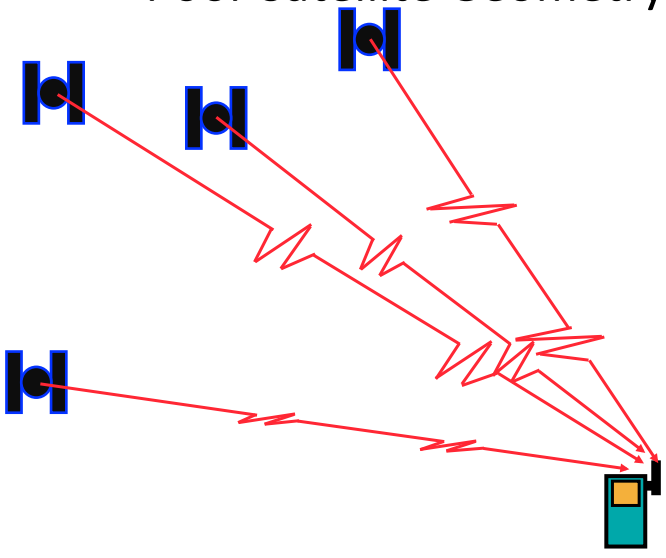
Good Satellite Geometry



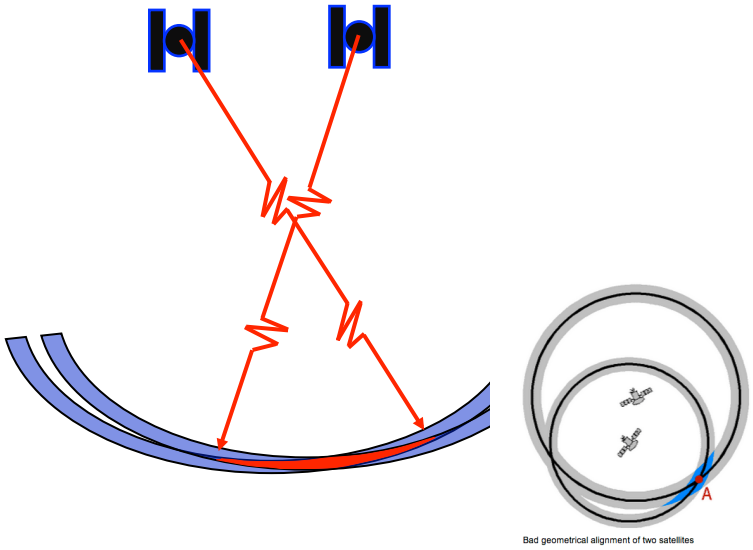
Good Satellite Geometry



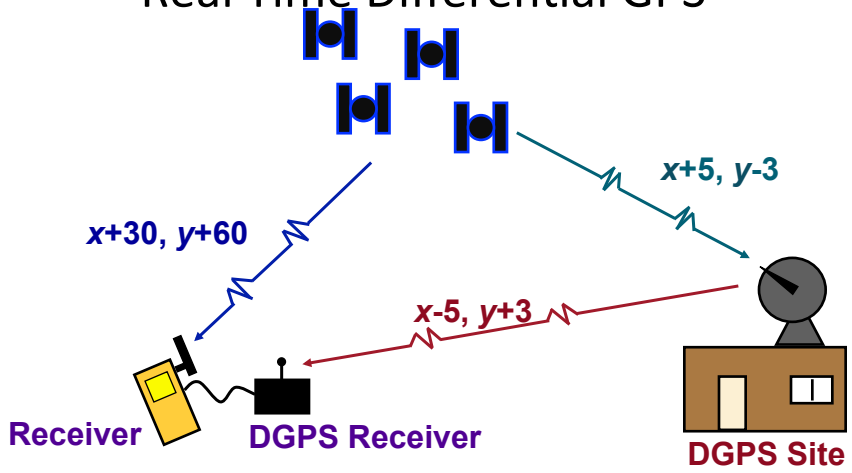
Poor Satellite Geometry



Poor Satellite Geometry



Real Time Differential GPS



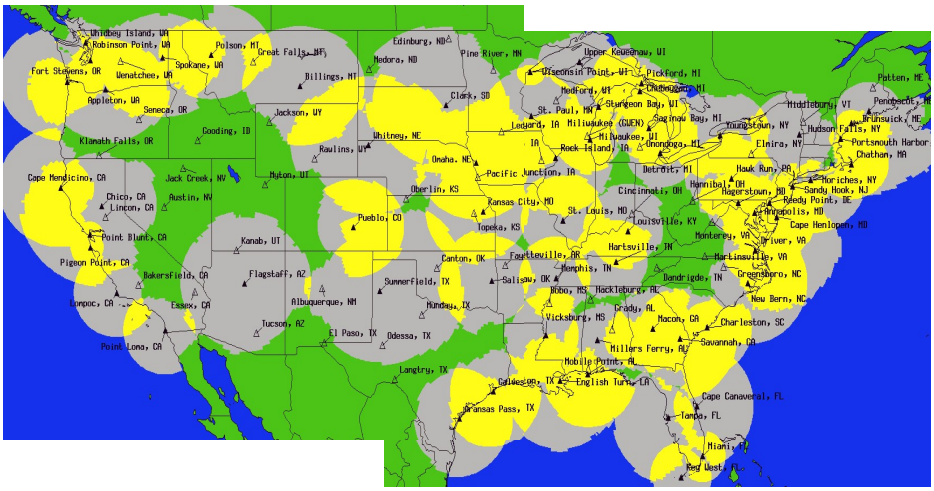
DGPS correction = *x*+(30-5) and *y*+(60+3)

True coordinates = *x*+25, *y*+63

True coordinates = *x*+0, *y*+0

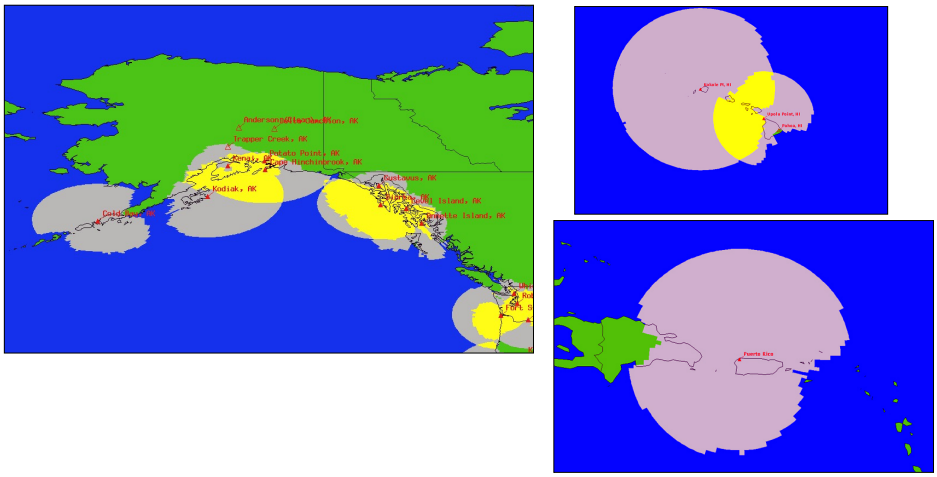
Correction = *x*-5, *y*+3

NDGPS Ground Stations



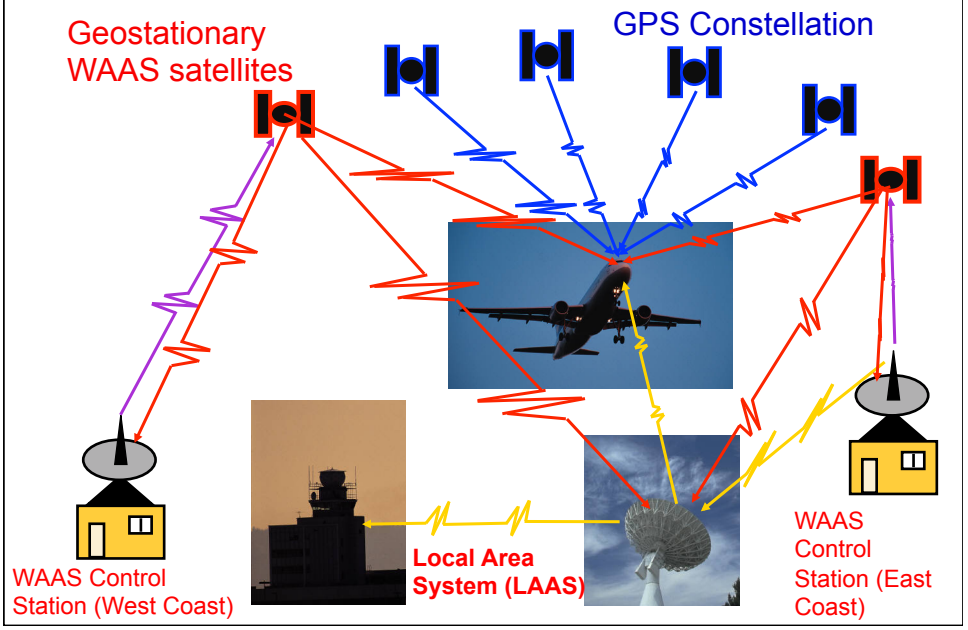
Yellow areas show overlap between NDGPS stations. Green areas are little to no coverage. Topography may also limit some areas of coverage depicted here.

NDGPS Ground Stations



Yellow areas show overlap between NDGPS stations. Green areas are little to no coverage. Topography may also limit some areas of coverage depicted here.

Wide Area Augmentation System



Differential GPS

- Use known errors at fixed GPS receivers to correct for errors at rovers
- Maritime GPS
  - Lighthouses
- Wide Area Augmentation System (WAAS)
  - 25 ground base stations
  - 4 geostationary satellites
- Real-time Kinematic (RTK)
  - Sub-centimeter accuracy
  - Expensive equipment
  - Requires line of sight between coordinating receivers

40

Technology	Basic GPS	GPS + WAAS	Real-time Kinematic GPS
Accuracy	★★☆☆ 3D coordinates with 10 m median accuracy	★★★☆ 3D coordinates with 2 m median accuracy	★★★★ 3D coordinates with 10 cm accuracy
Coverage	★★★☆ Outdoors with clear view of 4+ GPS satellites	★★☆☆ Outdoors in USA — Requires clear view of 4+ GPS satellites and a WAAS satellite	★☆☆☆ Outdoors with 4+ GPS satellites and requires line-sight between mobile unit and surveyed unit
Infrastructure cost	★★★★ $14 B US initial cost + $500 M US yearly for global coverage	★★★★ WAAS satellites needed in addition to GPS constellation	★☆☆☆ Beyond GPS constellation, requires calibrated, surveyed ground unit
Per-client cost	★★★☆ GPS antenna and chipset required	★★★☆ GPS antenna and WAAS-capable chipset required	★☆☆☆ Special RTK unit required
Privacy	★★★★ Location is estimated passively on the GPS unit	★★★★ Location is estimated passively on the GPS unit	★★★★ Location is estimated passively on the GPS unit
Well-matched use cases	Outdoor navigation for land, sea and air, emergency response, turn-by-turn driving directions, outdoor mapping/information/tour guide services, personell/pet tracking, fitness/activity tracking, gaming	Outdoor navigation for land, sea and air, emergency response, turn-by-turn driving directions, outdoor mapping/information/tour guide services, personnel/pet tracking, fitness/activity tracking, gaming	Outdoor navigation requiring extreme accuracy, surveying, aircraft landing, maritime construction

41

Others

- GLONASS (Russia)
- Galileo (EU)
- BeiDou (China)
- Future: multi-system receivers

42

INFRARED AND ULTRASONIC

43

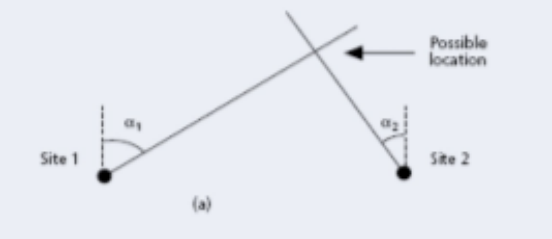
Estimating Position based on signals

- Angle of Arrival (AOA)
- Time of Arrival (TOA)
- Time Difference of Arrival (TDOA)

44

Angle of Arrival (AOA)

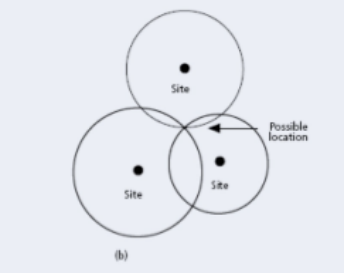
- Electronically steering a directional antenna or antenna array
- Triangulation
- Robust to multipath effects
- Need special hardware



45

Time of Arrival (TOA)

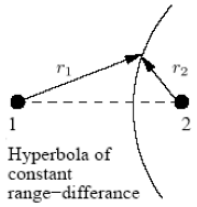
- Time of arrival (TOA)
  - Intersecting range intervals
  - Estimate response delay?
  - Timing errors in absence of LOS

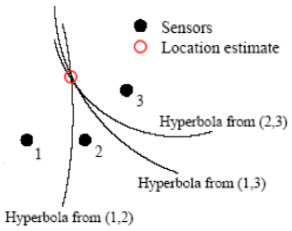


46

Time Distance of Arrival (TDOA)

- Hyperbola is set of points at constant time-difference from two base stations
- NB: Don't need TOA, just difference!
- Each pair of base stations gives hyperbola on which mobile device lies
- Intersect hyperbolic curves for location
- In network or device

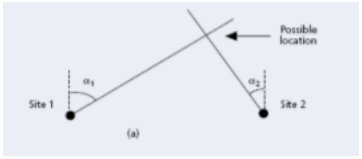


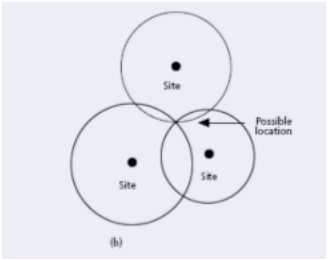


47

Infrared and Ultrasonic

- GPS signal cannot penetrate buildings
- Accuracy:
  - Absolute location
    - Use AOA, TOA to achieve 5-10cm
  - Symbolic location
    - Conference room
    - 0.5m: room vs hallway

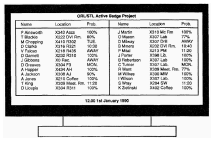




48

Active Badge: IR Proximity

- Characteristic
  - Office personnel wear badges (IR signal)
  - Rooms detect signals
  - Location server
- Applications
  - Call Forwarding
  - Privacy?

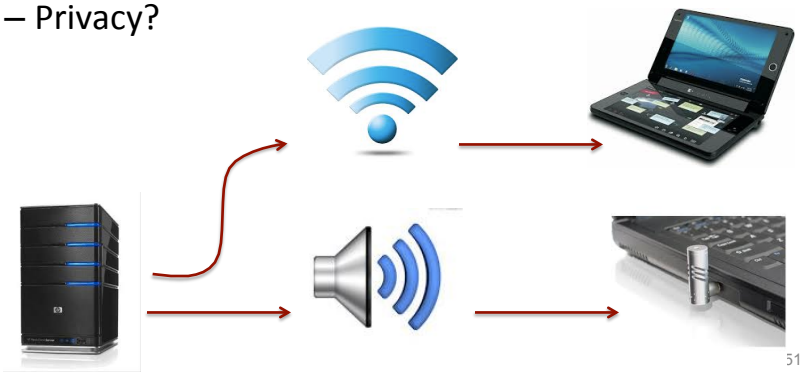


Ultrasonic Proximity

- Have rooms emit ultrasonic identifying pulses
  - Do not propagate outside room
  - Devices decode pulses
  - Glacial due to encoding signal in sound
  - Privacy?

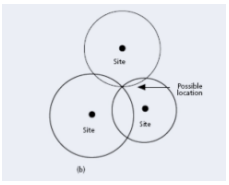
WALRUS: US Proximity

- Beacons emit identifying datagram over 802.11
  - Devices then listen for audio signal
  - Privacy?

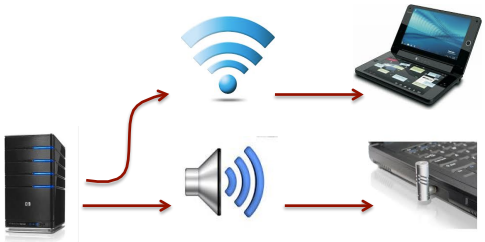


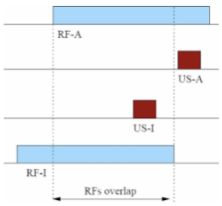
51

Cricket



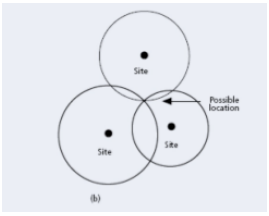
- Locate people in room
  - Emit radio (RF) and US
  - Receivers measure time difference
  - Estimate distance to beacon
  - Overlap RF, US to detect and ignore US collisions
  - Privacy?



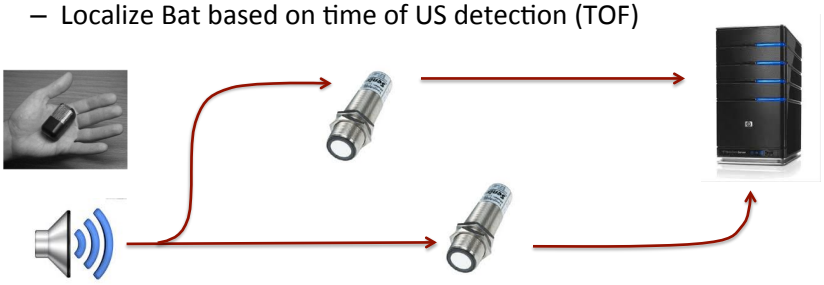


52

Active Bat

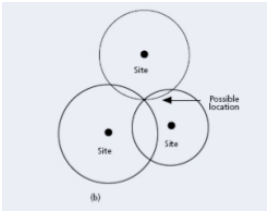


- Accuracy to within centimeters
- Use US receivers to localize pager-like devices (Bats)
  - Devices emit US
  - Environment listens for signals
  - Localize Bat based on time of US detection (TOF)

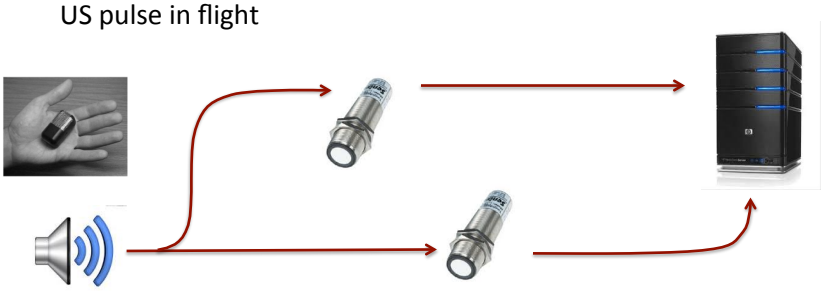


53

Active Bat

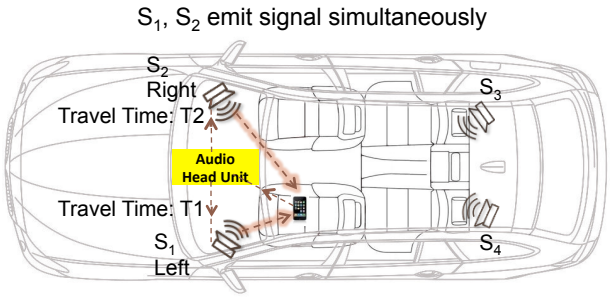


- Accuracy to within centimeters
- Accuracy due to dense deployment of receivers
  - 100 receivers in 100 m^2 office space
  - Slotted schedule to ensure no more than one Bat at a time has US pulse in flight



54

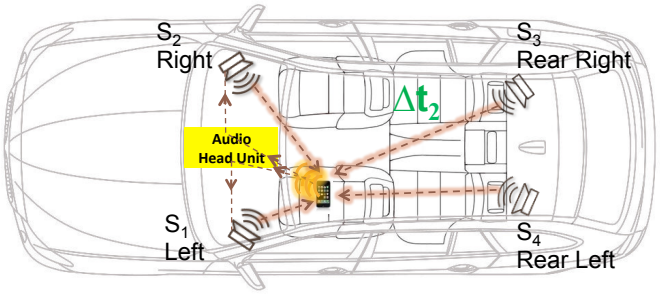
Car Phone Location: TOA

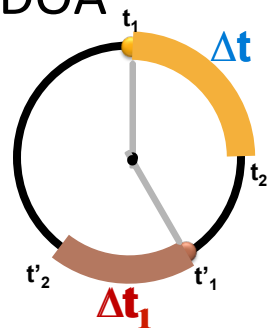


- Phone sends Bluetooth signal to audio
- Audio speakers emit US
- Phone detects US signals
- Problem: response delay

5

Car Phone Location: TDOA



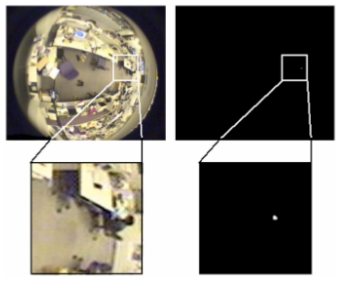


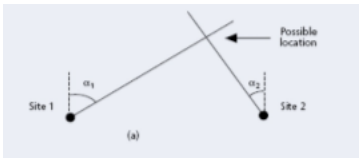
- Phone sends BT signal to S_1, then S_2
- Fixed interval Δt between signals
- Audio speakers send US signals
- $\Delta t_1 - \Delta t > 0$ closer to left signal (S_1)
- $\Delta t_1 - \Delta t < 0$ closer to right signal (S_2)

56

Altair: IR Angle of Arrival

- Wide-angle CCD video cameras
- IR filters to observe blinking IR-LEDs
- Angle of IR light's arrival estimated based on distance of lit pixel from center of image
- With 2 or more cameras, can estimate location of device
  - 1 camera and height assumption





57

Technology	Infrared proximity (e.g., Active Badge)	Ultrasound proximity (e.g., WALRUS)	Ultrasound TOF (e.g., Active Bat)	Infrared triangulation (e.g., ALTAIR)
Accuracy	★★★☆ Room ID with high accuracy	★★★☆ Room ID with high accuracy	★★★★ 3D location with 5 cm accuracy	★★★★ 3D location with 9 cm accuracy
Coverage	★★☆☆ Indoor only in room fit with IR receiver/beacon	★★☆☆ Indoor only in room fit with ultrasonic beacons	★★☆☆ Indoor only in room fit with ultrasonic infrastructure	★★☆☆ Indoor only in room fit with infrared cameras
Infrastructure cost	★★☆☆ Infrared receiver or beacon required for each room	★★☆☆ 1 or more ultrasonic receiver or beacon required for each room	★☆☆☆ Requires dense array of ultrasonic receivers	★☆☆☆ Requires 2+ calibrated IR cameras per room
Per-client cost	★★★☆ Inexpensive IR badge/dongle required	★★★★ Software only solution on device with microphone	★★★☆ Inexpensive ultrasonic badge/dongle required	★★★☆ Inexpensive IR badge/dongle required
Privacy	★★★★ If localization is performed on the client. Otherwise ★★☆☆ Opt-out easy by removing badge	★★★★ Localization is performed by mobile client	★★☆☆ Localization is performed by infrastructure. Opt-out easy by removing badge	★★☆☆ Localization is performed by infrastructure. Opt-out easy by removing badge/dongle
Well-matched use cases	Asset and personnel tracking, indoor mapping/ navigation/tour guides	Asset and personnel tracking, indoor mapping/ navigation/tour guides	Asset and personnel tracking, tangible UIs, fine-grained info services	Asset and personnel tracking, tangible UIs, fine-grained info services

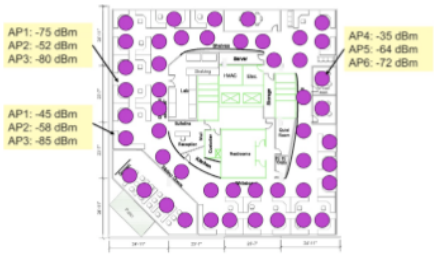
58

LAN-BASED LOCATION AWARENESS

59

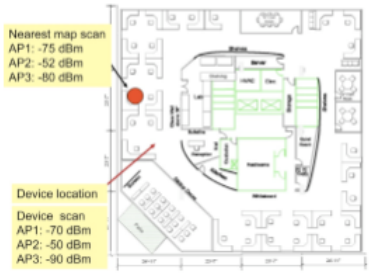
Signal Strength Fingerprinting

- RADAR
- Based on two properties:
  - Spatial variability
  - Temporal consistency
- Two phases
  - Mapping phase
    - Signal strength
    - Orientation of receiver
  - Location estimation phase
    - Client estimates its position based on scan



60

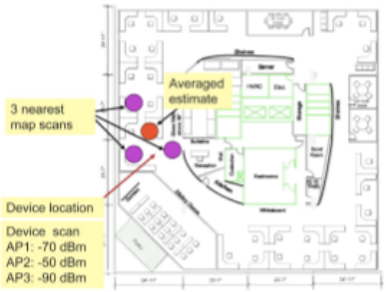
Signal Strength Fingerprinting



Nearest neighbor

AP	Scan 1	Scan 2	Difference
AP1	-70 db	-75 db	5 db
AP2	-50 db	-52 db	2 db
AP3	-90 db	-80 db	10 db

$$\sqrt{5 \cdot 5 + 2 \cdot 2 + 10 \cdot 10} = 11.4.$$



K-Nearest neighbor

61

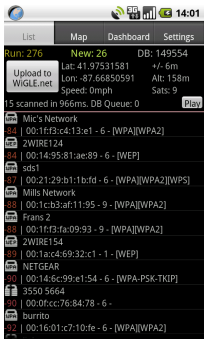
Signal Strength Modeling: Active Campus

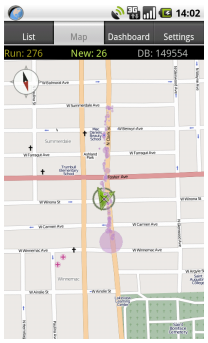
- AP database contains latitude, longitude, floor number for each AP
- Estimate client location from **signal strength**
- Distance estimates heavily weight strongest AP
- Higher accuracy indoors
  - Obstructions reducing AP ranges and likely locations

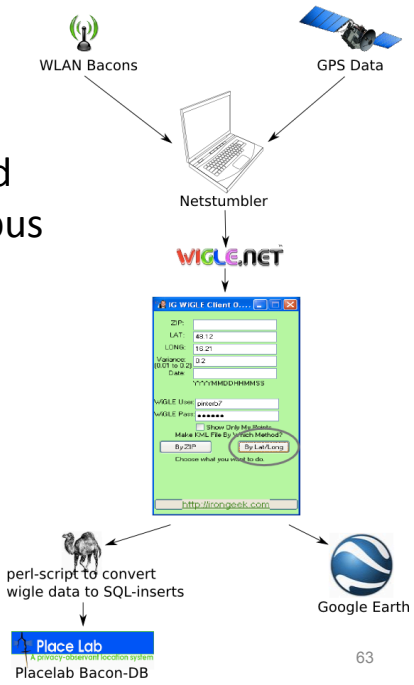
MAC	Latitude	Longitude
00:0f:34:ab:0c:e0	43°39′39.95″N	79°23′44.36″W
00:0f:f7:0c:e9:c0	43°39′44.32″N	79°23′47.33″W
00:0f:f7:0c:4f:03	43°39′43.50″N	79°23′37.86″W

Signal Strength Modeling: PlaceLab

- Use “war driving” to build AP database for off-campus
- See wigle.net







63

Technology	802.11 signal-strength fingerprinting (e.g., RADAR)	802.11 signal-strength modeling (e.g., Place Lab)	802.11 proximity (e.g., GUIDE)
Accuracy	★★★☆ 2D coordinates with 1-3 m median accuracy	★★☆☆ 2D coordinates with 10-20m median accuracy	★☆☆☆ Location accuracy dependant on AP density
Coverage	★★☆☆ Building to campus scale. Requires 802.11 coverage and radio map. Best accuracy achieved when 3+ APs are visible.	★★★☆ Areas with 802.11 coverage and radio map. Best accuracy achieved when 3+ APs are visible.	★★★☆ Anywhere with 802.11 coverage and an AP, location map.
Infrastructure cost	★★☆☆ No additional infrastructure is needed beyond 802.11 APs. Creating radio map is time intensive and new/moved APs require remap.	★★★☆ No additional infrastructure is needed beyond 802.11 APs. Creating radio maps is less work than for fingerprinting.	★★★★ No additional infrastructure is needed beyond 802.11 APs.
Per-client cost	★★★★ Software-only solution for devices with 802.11 NICs.	★★★★ Software-only solution for devices with 802.11 NICs.	★★★★ Software-only solution for devices with 802.11 NICs.
Privacy	★★★★ when localization is performed on the client. ★☆☆☆ when localization is performed in the infrastructure.	★★★★ when localization is performed on the client. ★☆☆☆ when localization is performed in the infrastructure.	★★★★ when localization is performed on the client. ★☆☆☆ when localization is performed in the infrastructure.
Well-matched use cases	Asset and personnel tracking in indoor environments, indoor mapping/navigation/tour guides	Social networking, tour guides, indoor/outdoor navigation/tour guides, fitness/activity tracking	Outdoor tour guides, nearby resource advertisement, activity tracking

64

CELL-BASED LOCATION AWARENESS

65

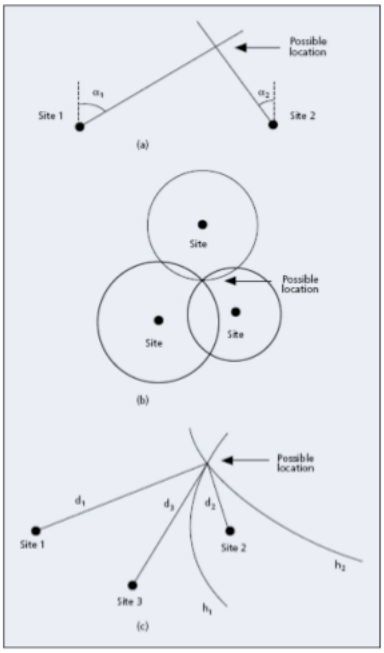
Cell ID-based

- Accuracy based on size of cell
  - 150m to 30km
  - Augment with round-trip time (RTT)
  - E911: locate handset within 50m with 66% accuracy

66

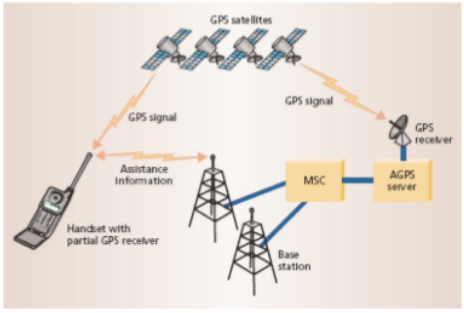
Radio Modeling

- GSM and CDMA based on time-of-flight measurements, rather than signal strength
- Time Difference of Arrival (TDOA)



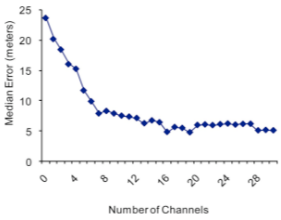
Assisted GPS

- GPS problems
  - Power
  - Unblocked view of sky
  - Signal acquisition
- A-GPS
  - Supplement GPS with measurements from nearby reference GPS receiver
  - Device only needs one satellite
  - Server can send constellation
  - QUALCOMM gpsOne: combine with base station range measurements



68

Signal Strength Fingerprinting



- Similar to 802.11 fingerprinting
  - "training phase" building radio map
  - Location estimation based on current signal measurements
- Advantage: cellular base stations more "stable" than 802.11 base stations
- Radio signal information more readily available on GSM than CDMA
- Wide fingerprints: 6 strongest GSM cells and readings of up to 29 additional GSM channels
  - Too weak for communication, good enough for location

69

Standards

- 3GPP and 3GPP2
  - Call for implementation of:
    - Cell-ID: inaccurate but "free"
    - TDOA: high accuracy, but requires changes to handset or network (or both)
    - A-GPS: requires handset changes

70

Technology	GSM signal-strength fingerprinting	GSM TOF and signal-strength modeling	GSM/CDMA proximity	Assisted GPS (A-GPS)
Accuracy	★★★☆ 2D coordinates with 4 m median accuracy in dense cell environment	★★☆☆ 3D coordinates with 100 - 200 m accuracy	★☆☆☆ Accuracy dependant on cell tower density (150 m – 30 km)	★★☆☆ 3D coordinates with 10 -150 m accuracy depending on number of GPS satellites visible
Coverage	★★☆☆ Building to campus scale. Requires cell network coverage and radio map. Best accuracy when 3+ cells are visible	★★★☆ Areas with GSM coverage and radio map. Best accuracy when 3+ cells are visible	★★★★ Anywhere with cell coverage and cell-to-location map	★★★☆ Outdoors with 4+ GPS satellites or indoors with cell network support + view of 1+ GPS satellite
Infrastructure cost	★★★☆ No additional infrastructure is needed beyond cell network. Creating radio map is time intensive	★★★★ No additional infrastructure is needed beyond cell network and map of tower locations	★★★★ No additional infrastructure is needed beyond cell network and map of tower locations	★★★☆ Beyond GPS constellation, requires deployment of fixed GPS receivers
Per-client cost	★★★★ Software only solution	★★★★ Software only solution	★★★★ Software only solution	★★★☆ GPS antenna and chipset required for handset
Privacy	★★☆☆ Even if location is computed on client device, the network still tracks a handset's associated cell	★★☆☆ Even if location is computed on client device, the network still tracks a handset's associated cell	★★☆☆ Even if location is computed on client device, the network still tracks a handset's associated cell	★★☆☆ Even if location is computed on client device, the network still tracks a handset's associated cell
Well-matched use cases	Asset and personnel tracking in indoor environments, indoor mapping/navigation/tour guides	Social networking, emergency response, neighorhood-scale information access, fitness tracking, outdoor mapping / navigation	Regional information access (weather, traffic, etc.)	Emergency response, indoor/outdoor information/tour guide services, personnel/pet tracking, activity tracking, gaming

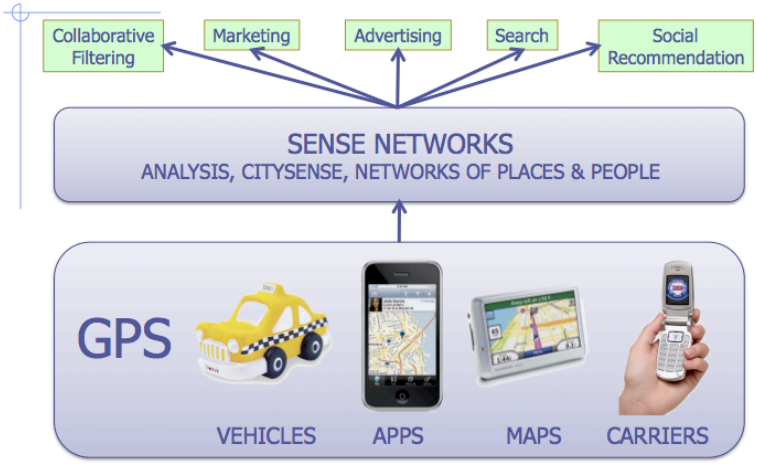
71

Based on materials by Tony Jebara

LOCATION-BASED SERVICES: CITY SENSE

72

GPS and location data

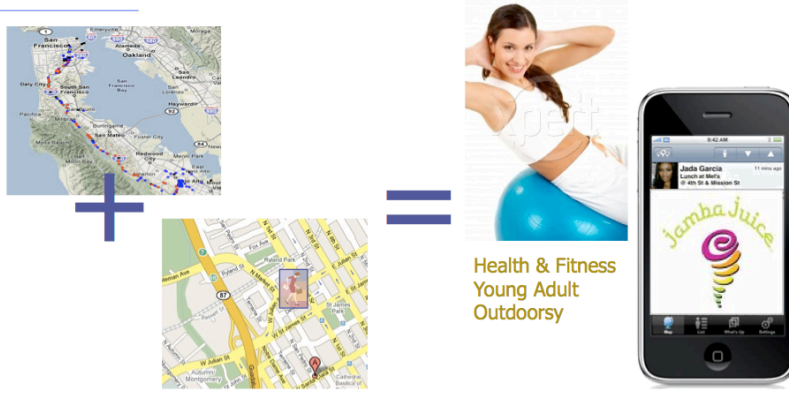


73

The Old View of LBS Data

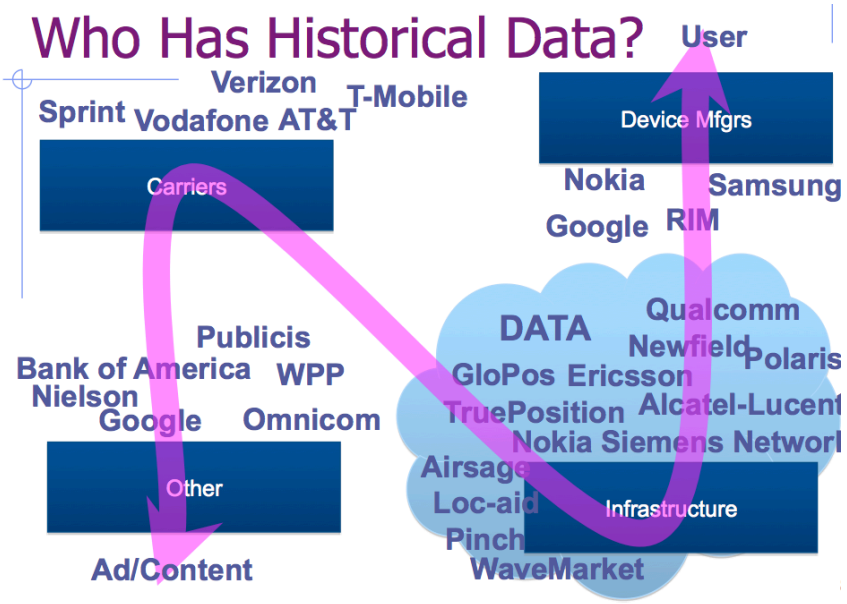
Single Ping @ Starbucks:
No personalization, no targeting
Can't use a single ping, too much error in space & time…
It's not just about *when* and *where* but also about ***who***

The New View of LBS Data



Location *history* for understanding & personalization
Store data over space and time to overcome accuracy issues
Lesson: save your LBS data to segment your customers!

Who Has Historical Data?



Network of People

Hard to say if User A is like User B…

User A

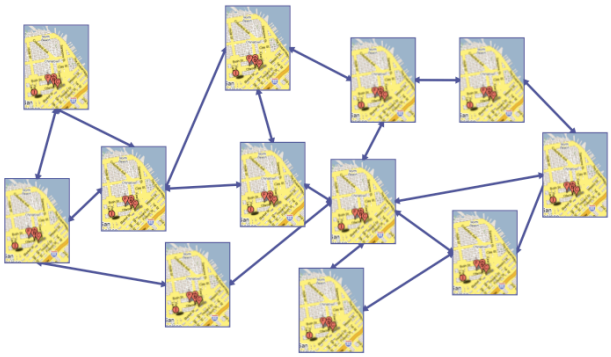
User B

… don't just see if they colocate physically
… do they overlap semantically (network of places)

77

Network of Places

Is place A like place B?
Look at each place's Flow, Commerce & Demographics



78

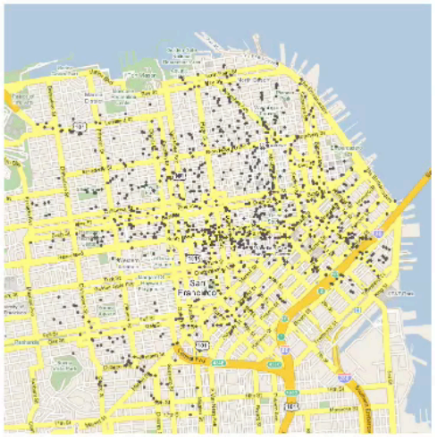
Network of Places: Flow

Look at flow A to B

Markov transition

Minimum Volume Embedding (MVE)

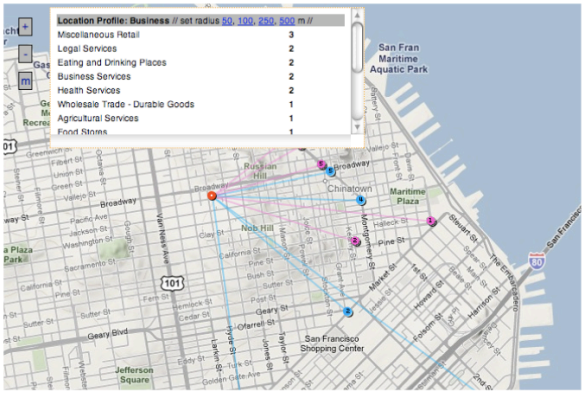
Color code clusters



79

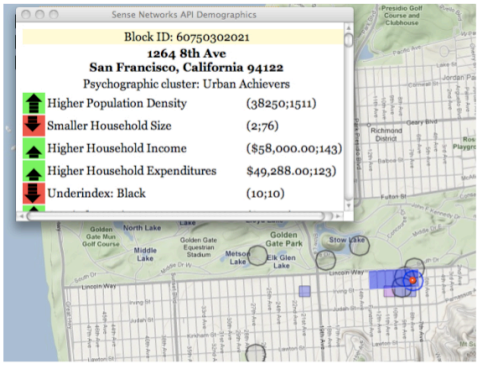
Network of Places: Commerce

Get each block's SIC (standard industrial categorization) code & cluster



Network of Places: Demographics

Get each block's census demographic data & cluster



81

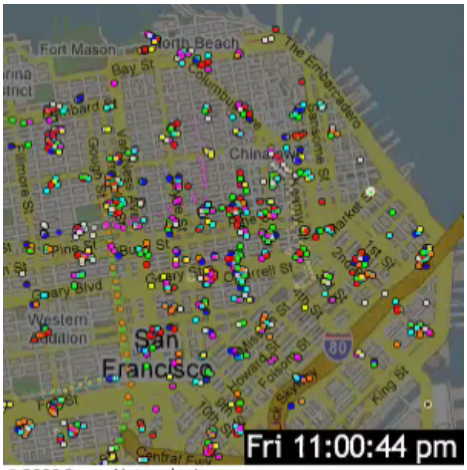
Encoding people

For each user, convert GPS trail into matrix of probabilities for week hour probability of being in

1) flow cluster
2) sic cluster
3) demographic cluster

Week Hour	FLO 1	FLO 2	…	FLO 20	SIC 1	SIC 2	…	SIC 97	DEM 1	DEM 2	…	DEM 78
1	.03	.31		.14	.03	.05		.41	.11	.04		.01
2	.14	.34		.02	.04	.05		.52	.01	.01		.00
…												
168	.07	.34		.51	.02	.06		.48	.02	.01		.00

82



Network of People → Tribes

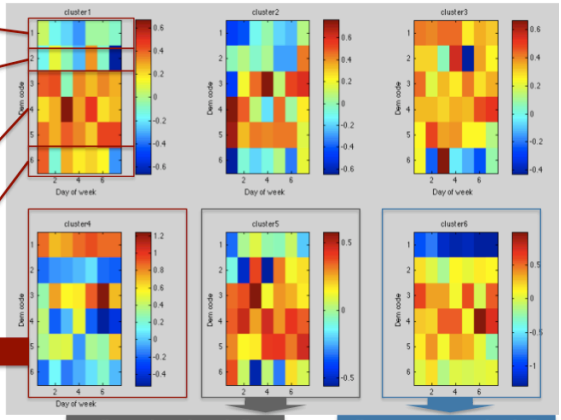
How often do they go out each day of the week?

Where do they hang out?

What is the avg age of most people in the neighborhoods they spend time in?

How racially diverse are the neighborhoods they spend time in?

Are the places they spend time in rich neighborhoods or poor neighborhoods?



"Young & Edgy"
- Out every night in young, racially diverse, low income neighborhoods

"Weekend Mole"
- Out occasionally on weeknights, typically middle-aged, Latino, middle-income neighborhoods

"Mature Homebody"
- Rarely goes out, typically spends nights in mature, white, higher income neighborhoods

84

The Next Net

(Stephen Baker, BusinessWeek)

85

Based on materials by John Krumm

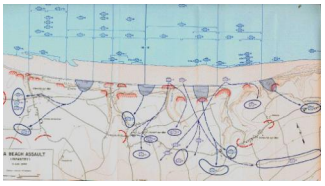
LOCATION PRIVACY

86

Subtleties of Location Privacy

"… a special type of information privacy which concerns the claim of individuals to determine for themselves when, how, and to what extent location information about them is communicated to others."

Duckham, M. and L. Kulik, *Location privacy and location-aware computing*, in *Dynamic & Mobile GIS: Investigating Change in Space and Time*, J. Drummond, et al., Editors. 2006, CRC Press: Boca Raton, FL USA. p. 34-51.



When: For D-Day attack, troop location privacy not important 60 years later

How: Alert fires to tell your family whenever you stop for pancakes

"Michael Mischers Chocolates"
"Weight Watchers"
To what extent: Accuracy high enough to distinguish?

87

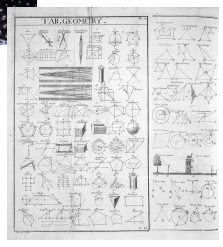
Computational Location Privacy

Law – Privacy regulations enforced by government

Policy – Trust-based, often from institutions

Encryption – Applies to any type of data.

Computational Location Privacy – Exploits geometric nature of data with algorithms

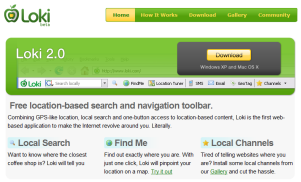


88

WHY REVEAL YOUR LOCATION?

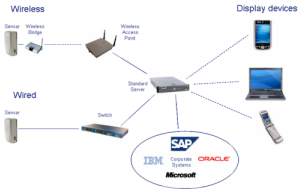
Why Reveal Your Location?

If you want to know your location, sometimes have to tell someone else.



Loki Wi-Fi locator – send your Wi-Fi fingerprint and get back (lat,long)

Quova Reverse IP – send your IP address and get back (lat,long)

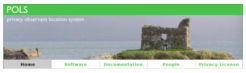


UbiSense – static sensors receive UWB to compute (x,y,z)

Exceptions

The Cricket Indoor Location System

Cricket – MIT



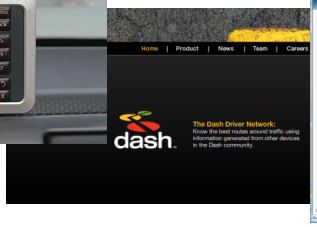
POLS – Intel Research

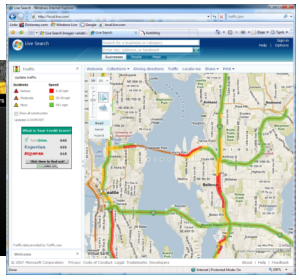
Variable Pricing

Congestion Pricing

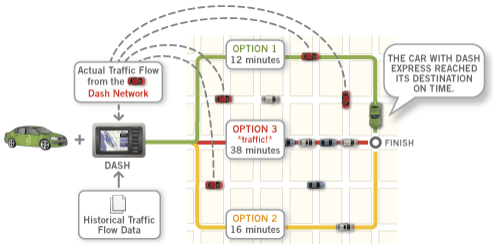
Pay As You Drive (PAYD) Insurance

Traffic Probes





http://dash.net/

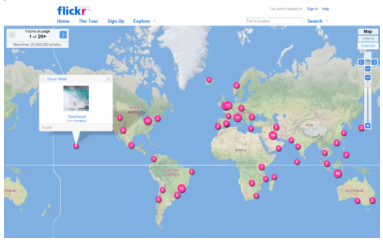


Routing The Dash Way

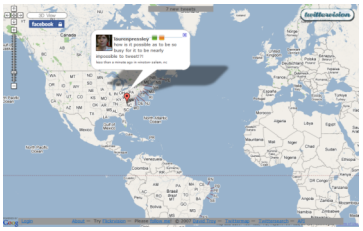
Dash was built from the ground up to be smart about traffic. Each Dash unit comes loaded with a database of historical traffic conditions for major metropolitan areas. Dash knows how fast traffic flows for every time and day of the year. Dash leverages this historical information along with traffic flow information sent from the network of other Dash drivers to provide users up to 3 routing options. Finally a solution that gives you routing options based on real world conditions!

Social Applications

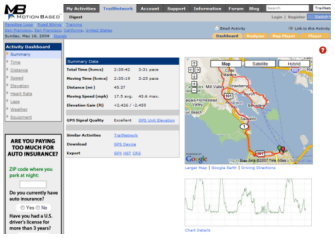
Dodgeball



Geotagged Flickr



Geotagged Twitter



MotionBased

93

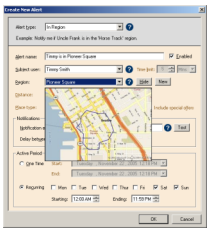
Location-Based Services

Navigation

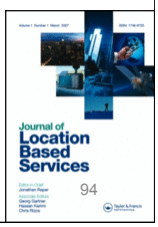
Local Information

Tracking

Games

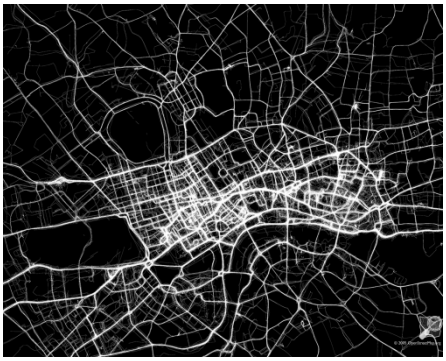


Location Alerts

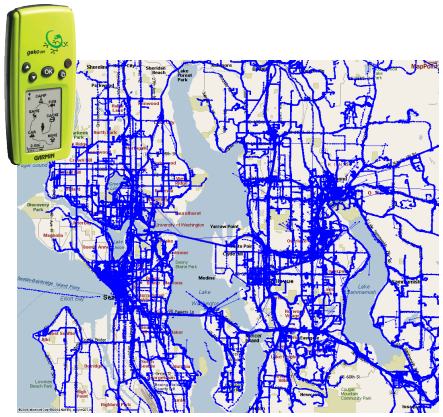


94

Research



OpenStreetMap (London)



MSMLS (Seattle)

95

DOES ANYONE CARE?

96

People Don't Care about Location Privacy

- 74 U. Cambridge CS students
- Would accept £10 to reveal 28 days of measured locations (£20 for commercial use) [1]

- 226 Microsoft employees
- 14 days of GPS tracks in return for 1 in 100 chance for $200 MP3 player

- 62 Microsoft employees
- Only 21% insisted on not sharing GPS data outside

- 11 with location-sensitive message service in Seattle
- Privacy concerns fairly light [2]

- 55 Finland interviews on location-aware services
- "It did not occur to most of the interviewees that they could be located while using the service." [3]

[1] Danezis, G., S. Lewis, and R. Anderson. *How Much is Location Privacy Worth? in Fourth Workshop on the Economics of Information Security.* 2005. Harvard University.

[2] Iachello, G., et al. *Control, Deception, and Communication: Evaluating the Deployment of a Location-Enhanced Messaging Service.* in *UbiComp 2005: Ubiquitous Computing.* 2005. Tokyo, Japan.

[3] Kaasinen, E., *User Needs for Location-Aware Mobile Services. Personal and Ubiquitous Computing*, 2003. 7(1): p. 70-79.

97

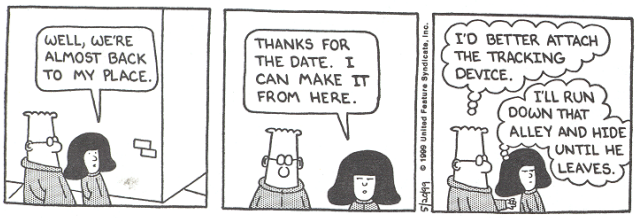
Documented Privacy Leaks

How Cell Phone Helped Cops Nail Key Murder Suspect – Secret "Pings" that Gave Bouncer Away
New York, NY, March 15, 2006

Stalker Victims Should Check For GPS
Milwaukee, WI, February 6, 2003

Real time celebrity sightings
http://www.gawker.com/stalker/

A Face Is Exposed for AOL Searcher No. 4417749
New York, NY, August 9, 2006



98

Subtleties of Location Privacy

- Interviews of location based services users
- Less worry about location privacy in closed campus [1]

- Interviews in 5 EU countries
- Price for location varied depending on intended use [2]

- Greeks significantly more concerned about location privacy
- Study two months after wiretapping of Greek politicians [2]

[1] Barkhuus, L., *Privacy in Location-Based Services, Concern vs. Coolness*, in *Workshop on Location System Privacy and Control, Mobile HCI 2004*. 2004: Glasgow, UK.

[2] Cvrček, D., et al., *A Study on The Value of Location Privacy*, in *Fifth ACM Workshop on Privacy in the Electronic Society*. 2006, ACM: Alexandria, Virginia, USA. p. 109-118.

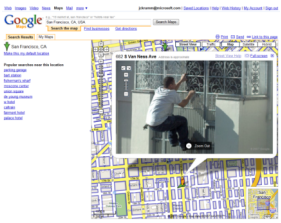
99

COMPUTATIONAL LOCATION PRIVACY THREATS

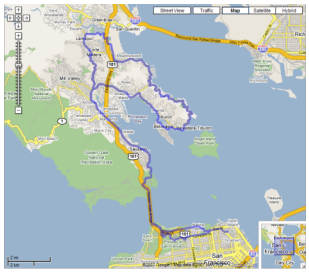
100

Computational Location Privacy Threats

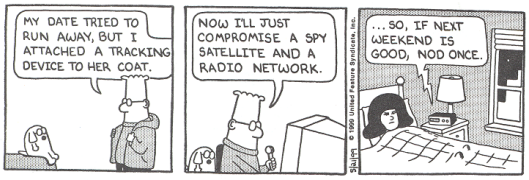
Not computational: stalking, spying, peeping



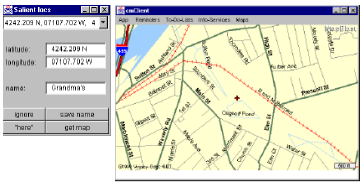
Not computational: browsing geocoded images



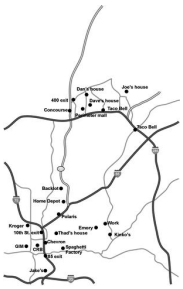
Not computational: browsing GPS tracks



Significant Locations From GPS Traces



comMotion (Marmasse & Schmandt, 2000)
- consistent loss of GPS signal → salient location
- user gives label (e.g. "Grandma's")

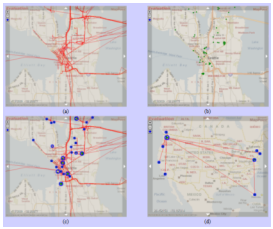


Ashbrook & Starner, 2003
- cluster places with lost GPS signal
- user gives label

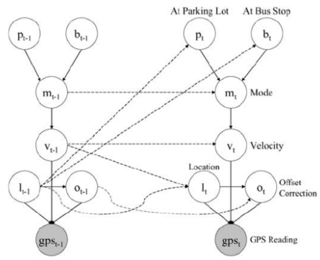
Common aim: find user's significant locations, e.g. home, work

Kang, Welbourne, Stewart, & Borriello, 2004
- time-based clustering of GPS (lat,long)

Project Lachesis (Hariharan & Toyama, 2004)
- time/space clustering
- hierarchical



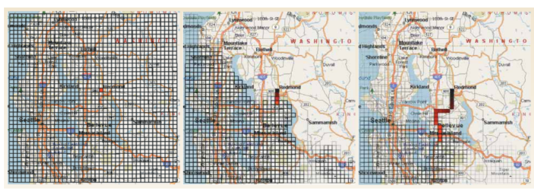
Context Inference



Patterson, Liao, Fox & Kautz, 2003
• GPS traces
• Infer mode of transportation (bus, foot, car)
• Route prediction

Location says a lot about you

Krumm, Letchner & Horvitz, 2006
• Noisy GPS matched to road driven
• Constraints from speed & road connectivity



Predestination (Krumm & Horvitz, 2006)
• Predict destination
• Extends privacy attack into future

Context Inference - Wow

Inferring Long-term User Properties based on Users' Location History

Yutaka Matsuo and **Naoaki Okazaki** and **Kiyoshi Izumi** and **Yoshiyuki Nakamura** and **Takuichi Nishimura** and **Kôiti Hasida**
National Institute of Advanced Industrial Science and Technology
y.matsuo@aist.go.jp, {kiyoshi, nmura, taku}@ni.aist.go.jp, hasida.k@aist.go.jp

Hideyuki Nakashima
Future University, Hakodate
h.nakashima@fun.ac.jp

Abstract

Recent development of location technologies enables us to obtain the location history of users. This paper proposes a new method to infer users' long-term properties from their respective location histories. Counting the instances of sensor detection for every user, we can obtain a *sensor-user matrix*. After generating features from the matrix, a machine learning approach is taken to automatically classify users into different categories for each user property. Inspired by information retrieval research, the problem to infer user properties is reduced to a text categorization problem. We compare weightings of several features and also propose sensor weighting. Our algorithms are evaluated using experimental location data in an office environment.

in a laboratory might imply working behavior for laboratory members. While for different types of users such as guests for a campus tour, being in a laboratory implies sightseeing behavior. Therefore, user modeling and behavior detection are mutually complementary: if we have a more precise user model, we can guess more precisely the user behavior, and vice versa. Automatically obtaining a user model will bootstrap activity recognition in a ubiquitous environment to enable context-aware information services.

IJCAI 2007

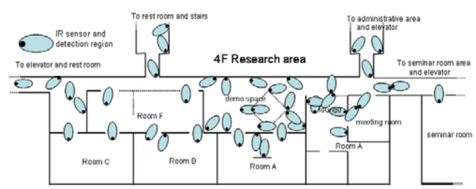


Figure 3: Sensor allocation map for a part of the fourth floor.

Indoor location sensors

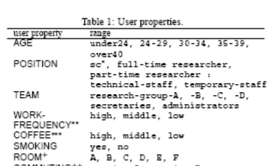


Table 1: User properties.

user property	range
AGE	under24, 24-29, 30-34, 35-39, over40
POSITION	sc*, full-time researcher, part-time researcher : technical-staff, temporary-staff
TEAM	research-group-A, -B, -C, -D, secretaries, administrators
WORK-FREQUENCY**	high, middle, low
COFFEE***	high, middle, low
SMOKING	yes, no
ROOM+	A, B, C, D, E, F
COMMUTING++	station-A, station-B

Machine learning to infer these properties based only on time-stamped location history

Good: TEAM, ROOM
OK: AGE, COFFEE, SMOKING
Bad: POSITION, WORK FREQUENCY

Location is Quasi-Identifier

Protecting Privacy Against Location-based Personal Identification*

Claudio Bettini[1], X. Sean Wang[2], and Sushil Jajodia[3]

[1] DSI, University of Milan, Italy. bettini@dsi.unimi.it
[2] Dept of CS, University of Vermont, Vermont. xywang@cs.uvm.edu
[3] CSIS, George Mason University, Virginia. jajodia@gmu.edu

Abstract. This paper presents a preliminary investigation on the privacy issues involved in the use of location-based services. It is argued that even if the user identity is not explicitly released to the service provider, the geo-localized history of user-requests can act as a quasi-identifier and may be used to access sensitive information about specific individuals. The paper formally defines a framework to evaluate the risk in revealing a user identity via location information and presents preliminary ideas about algorithms to prevent this to happen.

1 Introduction

There are currently over 1.5 billion mobile phone users worldwide and the numbers are still growing very fast. Location technologies can be currently used by wireless carrier operators to provide a good estimate of the user location. These techniques are being refined in order to meet the requirements imposed by federal institutions both in US and Europe for location-enhanced emergency services. Significantly more precise positioning is obtained by GPS technology which is already integrated in some mobile phones and it is likely to become a common feature of most product phones. Indoor positioning is also available based on a wide range of enabling technologies including ultrasound, Wi-Fi, and Bluetooth. Considering that mobile phones are rapidly evolving into multipurpose devices that can access a wide range of services, there is a general concern about how positioning information is stored, managed and released to possibly untrusted service providers.

This paper considers the privacy issues involved in accessing location-based services, i.e., services that, based on the user current position, can provide location-aware information. Typical examples are map and navigation services, services that provide information on close-by public resources (e.g., gas stations, pharmacies, ATM machines, ...), services that provide localized news (e.g., weather forecasts, road constructions, etc.), as well as more personalized services like proximity marketing or friend-finder.

* This work is partially supported by NSF under grants IIS-0430402 and IIS-0242237. The work of Bettini is also partially supported by the Italian MIUR (FIRB "WebMinds" project N. RBNE05WEXT_005)

Quasi-Identifier – "their values, in combination, can be linked with external information to reidentify the respondents to whom the information refers. A typical example of a single-attribute quasi-identifier is the Social Security Number, since knowing its value and having access to external sources it is possible to identify a specific individual."

Secure Data Management, VLDB workshop, 2005

Simulated Location Privacy Attack 1

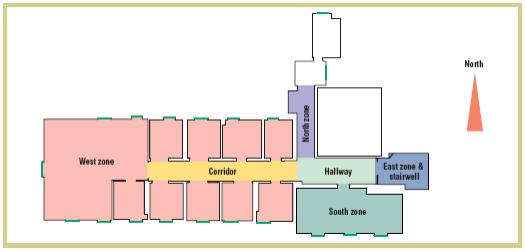
SECURITY & PRIVACY

Location Privacy in Pervasive Computing

As location-aware applications begin to track our movements in the name of convenience, how can we protect our privacy? This article introduces the mix zone—a new construction inspired by anonymous communication techniques—together with metrics for assessing user anonymity.

Many countries recognize privacy as a right and have attempted to codify it in law. The first known piece of privacy legislation was the English 1361 *Justices of the Peace Act*, which legislated for the arrest of eavesdroppers and nuisances. The Fourth Amendment to the US Constitution provides citizens' right to privacy, and in 1998 US Supreme Court Justice Louis Brandeis stated that privacy is the most cherished of the fundamental rights of a democracy.[1] The 1948 *Universal Declaration of Human Rights*[2] declares that everyone has a right to privacy at home, with family, and in correspondence. Other pieces of international legislation echo these rights. Although many people closely consider their privacy a fundamental right, comparatively few can give a precise definition of the term. The Global Internet Liberty Campaign[3] has produced an extensive report that discusses personal privacy at length and identifies four broad categories: information privacy, bodily privacy, privacy of communications, and territorial privacy.

Alastair R. Beresford and Frank Stajano*, University of Cambridge



Active BAT indoor location system

IEEE Pervasive Computing Magazine, Jan/March 2003

Experiment
- Attach pseudonym to each person's location history
- Check
  - Where does person spend majority of time?
  - Who spends most time at any given desk?
- Found correct name of *all* participants

Simulated Location Privacy Attack 2

INTELLIGENT TRANSPORTATION SYSTEMS

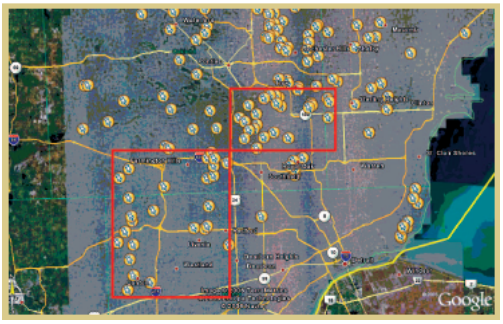
Enhancing Security and Privacy in Traffic-Monitoring Systems

This architecture separates data from identities by splitting communication from data analysis. Data suppression techniques can help prevent data mining algorithms from reconstructing private information from anonymous database samples.

Baik Hoh, Marco Gruteser, and Hui Xiong
Rutgers University

Ansaf Alrabady
General Motors

IEEE Pervasive Computing Magazine, Oct/Dec 2006



Experiment
- GPS histories from 65 drivers
- Cluster points at stops
- Homes are clusters 4 p.m. – midnight
- Found plausible homes of 85%

Simulated Location Privacy Attack 3

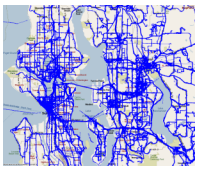
Fifth International Conference on Pervasive Computing (Pervasive 2007), May 13-16, Toronto, Ontario, Canada

Inference Attacks on Location Tracks

John Krumm

Microsoft Research
One Microsoft Way
Redmond, WA, USA
jckrumm@microsoft.com

Pervasive 2007



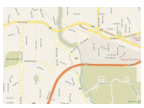
GPS Tracks (172 people)

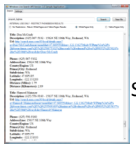
Home Location (61 meters)

Home Address (12%)

Identity (5%)

MapPoint Web Service reverse geocoding





Windows Live Search reverse white pages

Simulated Location Privacy Attack 4

On the Anonymity of Periodic Location Samples

Marco Gruteser and Baik Hoh

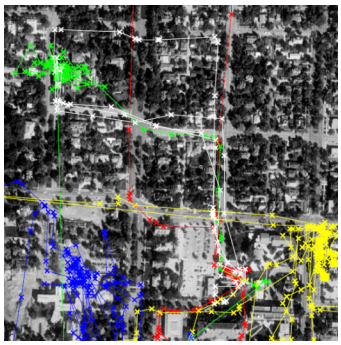
WinLab - Electrical and Computer Engineering Department
Rutgers, The State University of New Jersey
94 Brett Rd
Piscataway, NJ 08854
{gruteser,baikhoh}@winlab.rutgers.edu

Abstract. As Global Positioning System (GPS) receivers become a common feature in cell phones, personal digital assistants, and automobiles, there is a growing interest in tracking larger user populations, rather than individual users. Unfortunately, anonymous location samples do not fully solve the privacy problem. An adversary could link multiple samples (i.e., follow the footsteps) to accumulate path information and eventually identify a user.
This paper reports on our ongoing work to analyze privacy risks in such applications. We observe that linking anonymous location samples is related to the data association problem in tracking systems. We then propose to use such tracking algorithms to characterize the level of privacy and to derive disclosure control algorithms.

1 Introduction

The continuous improvements in accuracy and cost of Global Positioning System (GPS) receivers are driving new location tracking applications with a massive user base. For example, in the United States, cell phone providers can determine the positions of emergency callers through Assisted GPS, and the German federal government is funding the development of a GPS-based highway toll collection system for trucks. These systems are capable of sampling location information from a large numbers of users.
We anticipate great demand for this data, going far beyond the original applications of emergency positioning or toll collection. Aside from hotly debated uses such as in law enforcement and targeted marketing, there are also clearly benevolent uses. For example, vehicles could report the location of abrupt braking activity to improve road safety, navigation systems that optimize traffic flows could alleviate congestion and pollution, or movement models collected from cell phones may help predicting the spread of infectious diseases.
Sharing location information, however, raises privacy concerns [1, 2]. For example, frequent visits to clinics signal medical problems, attending meetings may reveal political preferences, and meetings of influential business managers could indicate pending business deals. As such, the problem of sharing location information is analogous to hospitals publishing medical records to epidemiologists and other medical researchers—it can be beneficial to society but invades on privacy.
Anonymizing data provides a solution that enables data access while maintaining privacy. Sweeney [3, 4] pointed out, however, that naive anonymization strategies, such

Security in Pervasive Computing, 2005



From "multi-target tracking" algorithms originally designed for military tracking

- Three GPS traces with no ID or pseudonym
- Successful data association from physical constraints

Simulated Location Privacy Attack 5

Simultaneous Tracking & Activity Recognition (STAR) Using Many Anonymous, Binary Sensors

Daniel Wilson & Chris Atkeson

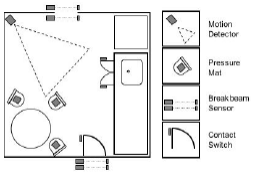
Robotics Institute
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
[dan.wilson, cga]@cs.cmu.edu

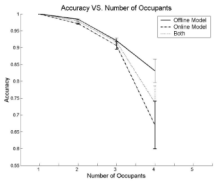
Abstract. In this paper we introduce the simultaneous tracking and activity recognition (STAR) problem, which exploits the synergy between location and activity to provide the information necessary for automatic health monitoring. Automatic health monitoring can potentially help the elderly population live safely and independently in their own homes by providing key information to caregivers. Our goal is to perform accurate tracking and activity recognition for multiple people in a home environment. We observe a "bottom-up" approach that primarily uses information gathered by many minimally invasive sensors commonly found in home security systems. We describe a Rao-Blackwellised particle filter for room-level tracking, rudimentary activity recognition (i.e., whether or not an occupant is moving), and data association. We evaluate our approach with experiments in a simulated environment and in a real instrumented home.

1 Introduction

Advances in modern health care are helping millions of people live longer, healthier lives. As a result, people aged 65 and older are the fastest growing segment of the US population (set to double in the next two decades) [5]. Current health-care infrastructure is inadequate to meet the growing needs of an increasingly older population. Clearly, there is a growing need to develop technological solutions.
One solution is to use automatic health monitoring to enable *aging in place*, in which elders live independently and safely in their own homes for as long as possible – without transitioning to a care facility. Automatic health monitoring uses data from ubiquitous sensors to infer location and activity information about at-risk occupants. Studies have shown that pervasive monitoring of the elderly and those with disabilities can improve the accuracy of pharmacologic interventions, track illness progression, and lower caregiver stress levels [13]. Additionally, [30] has shown that movement patterns alone are an important indicator of cognitive function, depression, and social involvement among people with Alzheimer's disease.
In this paper we introduce the simultaneous tracking and activity recognition (STAR) problem. The key idea is that people tracking can be improved by activity recognition and vice versa. Location and activity are the *context* for one another and knowledge of one is highly predictive of the other. We seek to provide the information that is vital

Pervasive, 2005





- Home with three occupants
- Two-state sensors
- Continuity analysis on thousands of sensor readings
- 85% correct data association

110

Simulated Location Privacy Attack 6

A spatiotemporal model of strategies and counter strategies for location privacy protection

Matt Duckham*, Lars Kulik†, and Athol Birtley†

*Department of Geomatics
University of Melbourne, Victoria, 3010, Australia
mduckham@unimelb.edu.au

†Department of Computer Science and Software Engineering
University of Melbourne, Victoria, 3010, Australia
lars@csse.unimelb.edu.au
a.birtley@ugrad.unimelb.edu.au

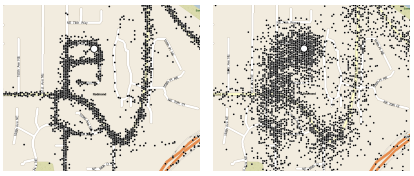
Abstract. Safeguarding location privacy is becoming a critical issue in location-based services and location-aware computing generally. Two drawbacks of many previous models of location privacy are: 1) the models only consider a person's location privacy protection, but not the invasion of location privacy by external agents; and 2) the models are static and do not consider the spatiotemporal aspects of movement. We argue that, to be complete, any model of location privacy needs to enable the analysis and identification of techniques both to protect and to invade an individual's location privacy over time. One way to protect an individual's location privacy is to minimize the information revealed about a person's location, termed *obfuscation*. This paper presents an explicitly spatiotemporal model of location privacy that models a third party's limited knowledge of a mobile individual's location. We identify two core strategies that a third party can use to refine its knowledge, so potentially invading that mobile individual's location privacy. A global refinement strategy uses the entire history of knowledge about an agent's location in a single step. A local refinement strategy iteratively constructs refined knowledge over time. We present a formal model of global and local refinement operators, and show how this formal model can be translated into a computational model in a simulation environment.

1 Introduction

Location privacy can be defined as a special type of *information privacy* that concerns the claim of individuals to determine for themselves when, how, and to what extent location information about them is communicated to others [6], cf. [21]. The emergence of low-cost location-aware computing, which combines powerful mobile computing platforms, wireless communication capabilities, and ubiquitous integrated positioning systems, has led to location privacy being acknowledged as a key challenge in information science (e.g., [17]). A failure to safeguard location privacy has been linked to a range of undesirable effects, including unsolicited marketing and *location-based "spam"*; decreased *personal safety*, such as might result from stalking or assault; and

GIScience 2006

Refinement operators for working around obfuscated location data



original σ= 50 meters noise added

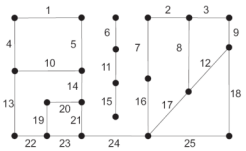


Fig. 2. Example geographic environment graph

Example refinement sources
- Must stay on connected graph of locations
- Movements are goal-directed
- Maximum speed constraint

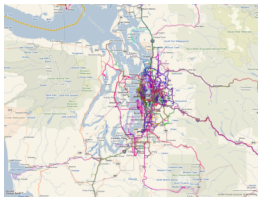
111

LOCATION PREDICTION

112

Where Do You Want to Go Today?

We already know, more or less.

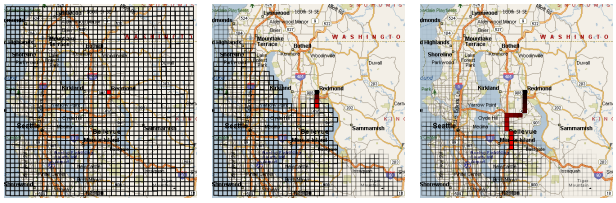


UbiComp 2006: The Eighth International Conference on Ubiquitous Computing, September 17-21, Orange County, CA, USA

Predestination:
Inferring Destinations from Partial Trajectories

John Krumm and Eric Horvitz

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA USA 98052
{jckrumm, horvitz}@microsoft.com



Efficient driving likelihood

Clues to destination
- Previous destinations
- Ground cover
- Efficient driving
- Trip time

Accuracy = 2 km median error at halfway point of trip

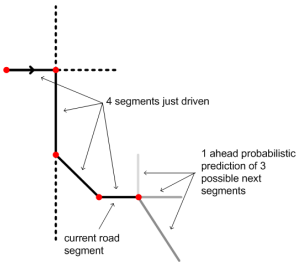
113

How Do You Want to Get There?

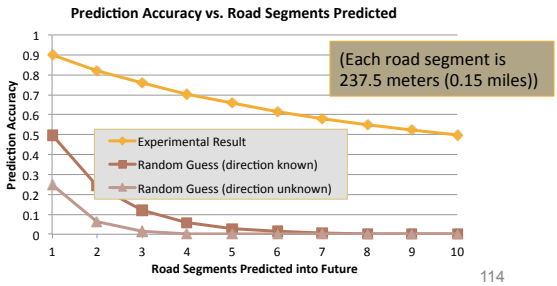
Paper Number 08AE-101

A Markov Model for Driver Turn Prediction

John Krumm
Microsoft Research



Predict next road segments based on past road segments (Markov model)



114

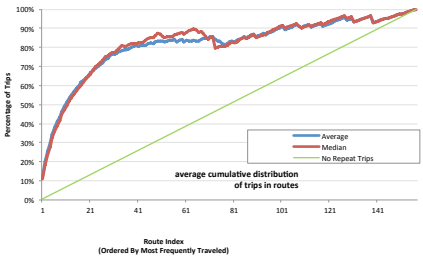
Full Route Prediction

Paper Number 08AE-283

Route Prediction from Trip Observations

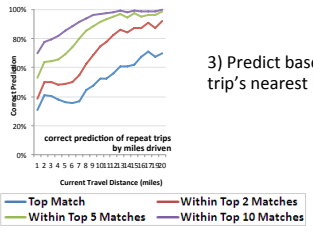
Jon Froehlich
University of Washington

John Krumm
Microsoft Research



1) Relatively small number of routes make up large fraction of drivers' trips

2) Cluster observed trips into repeated routes



3) Predict based on current trip's nearest historical route

115

COMPUTATIONAL COUNTERMEASURES

116

Computational Countermeasures

3

Location privacy and location-aware computing

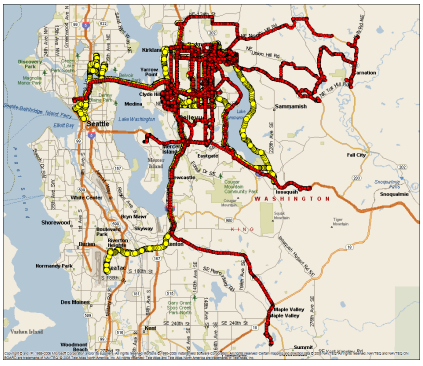
Matt Duckham & Lars Kulik
University of Melbourne, Australia

CONTENTS

Dynamic & Mobile GIS: Investigating Change in Space and Time, CRC Press, 2006

Four ways to enhance location privacy
1.Regulations – govt. enforced
2.Policies – trust-based agreements
3.Anonymity – pseudonyms and/or ambiguity
4.Obfuscation – reduce quality of data

Computational Countermeasures: Pseudonyms



Pseudonimity
• Replace owner name of each point with untraceable ID
• One unique ID for each owner

Example
• "Larry Page" → "yellow"
• "Bill Gates" → "red"

• Beresford & Stajano (2003) propose frequently changing pseudonym
• Gruteser & Hoh (2005) showed "multi-target tracking" techniques defeat complete anonymity

118

Computational Countermeasures: k-Anonymity

I'm chicken # 341, and I'm in this building (along with k-1 other chickens).

I'm chicken # 341, and I visited this place in the past 21 minutes (along with k-1 other chickens).

- k-anonymity introduced for location privacy by Gruteser & Grunwald, 2003
- They note that temporal ambiguity also gives k-anonymity
- Pattern of service requests could break k-anonymity (Bettini, Wang, Jajodia 2005)

Computational Countermeasures: Mix Zones

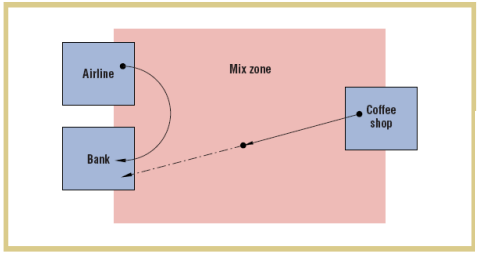
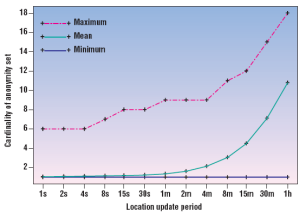


Figure 1. A sample mix zone arrangement with three application zones. The airline agency (A) is much closer to the bank (B) than the coffee shop (C). Users leaving A and C at the same time might be distinguishable on arrival at B.



Beresford & Stajano, 2003

- New, unused pseudonym given when user is between "application zones"
- "k-anonymous" when you can be confused with k-1 other people
- Anonymity (*i.e.* k) varies with busyness of mix zone
- Attack by trying to list all pseudonyms given to a person
- Can use probabilistic paths to associate pseudonyms

120

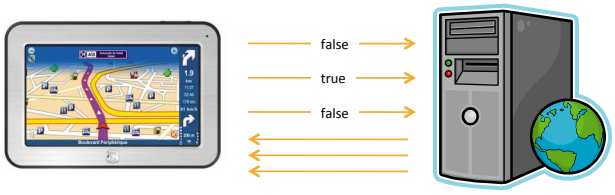
Computational Countermeasures: False Reports

An Anonymous Communication Technique using Dummies for Location-based Services

Hidetoshi Kido† Yutaka Yanagisawa†† Tetsuji Satoh†,††
†Graduate School of Information Science and Technology, Osaka University
††NTT Communication Science Laboratories, NTT Corporation
h-kido@ist.osaka-u.ac.jp yutaka@cslab.kecl.ntt.co.jp satoh.tetsuji@lab.ntt.co.jp

- Mix true location report with multiple false reports
- Act only on response from true report

Pervasive Services, 2005



- Communication overhead (addressed in paper)
- Attack by finding most sensible sequence of location reports
- Counter by making false sequences sensible (addressed in paper) (fun research project)

121

Computational Countermeasures: Obfuscation

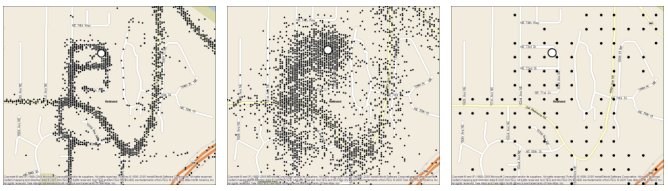
A Formal Model of Obfuscation and Negotiation for Location Privacy

Matt Duckham[1] and Lars Kulik[2]

[1] Department of Geomatics,
University of Melbourne, Victoria, 3010, Australia
mduckham@unimelb.edu.au
[2] Department of Computer Science and Software Engineering,
University of Melbourne, Victoria, 3010, Australia
lkulik@cs.mu.oz.au

- Formalizes obfuscation techniques
- Client & server can negotiate what needs to be revealed for successful location based service

Pervasive 2005

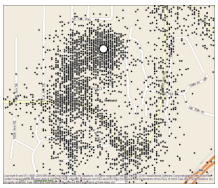


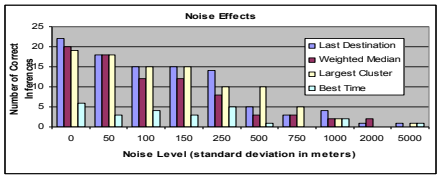
original low accuracy low precision

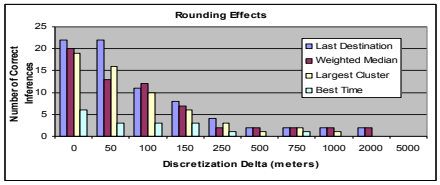
(from Krumm 2007)

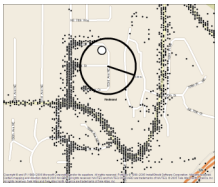
122

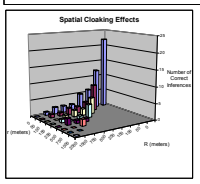
Computational Countermeasures: Obfuscation











Conclusion: need lots of obfuscation to counter privacy attack

Fifth International Conference on Pervasive Computing (Pervasive 2007), May 13-16, Toronto, Ontario, Canada

Inference Attacks on Location Tracks

John Krumm

Microsoft Research
One Microsoft Way
Redmond, WA, USA
jckrumm@microsoft.com

123

Computational Countermeasures: Obfuscation

Protecting Location Privacy Through Path Confusion

Baik Hoh
WINLAB
ECE Department
Rutgers, The State University of New Jersey
Email: baikhoh@winlab.rutgers.edu

Marco Gruteser
WINLAB
ECE Department
Rutgers, The State University of New Jersey
Email: gruteser@winlab.rutgers.edu

SECURECOMM 2005

Confuse the multi-target tracker by perturbing paths so they cross

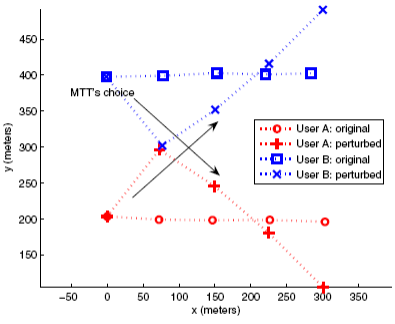


Figure 2. Two users move in parallel. The Path Perturbation algorithm perturbs the parallel segment into a crossing segment.

124

Conclusion

- Why reveal your location?
  - Lots of good reasons
  - Including just to know your own location
- Do people care about location privacy?
  - Not as much as we might expect
- Computational location privacy threats
  - Lots of sophisticated threats
- Location prediction
  - Even possible to infer your future locations
- Computational countermeasures
  - Much work on countermeasures
  - More work necessary as more threats come