Mobility data analysis to understand unknown diseases behavior

The case of facial paralysis

Jalel Akaichi

PhD, University of Sciences and Technologies of Lille

Associate Professor, University of Tunis

General introduction

Introduction (1/2)

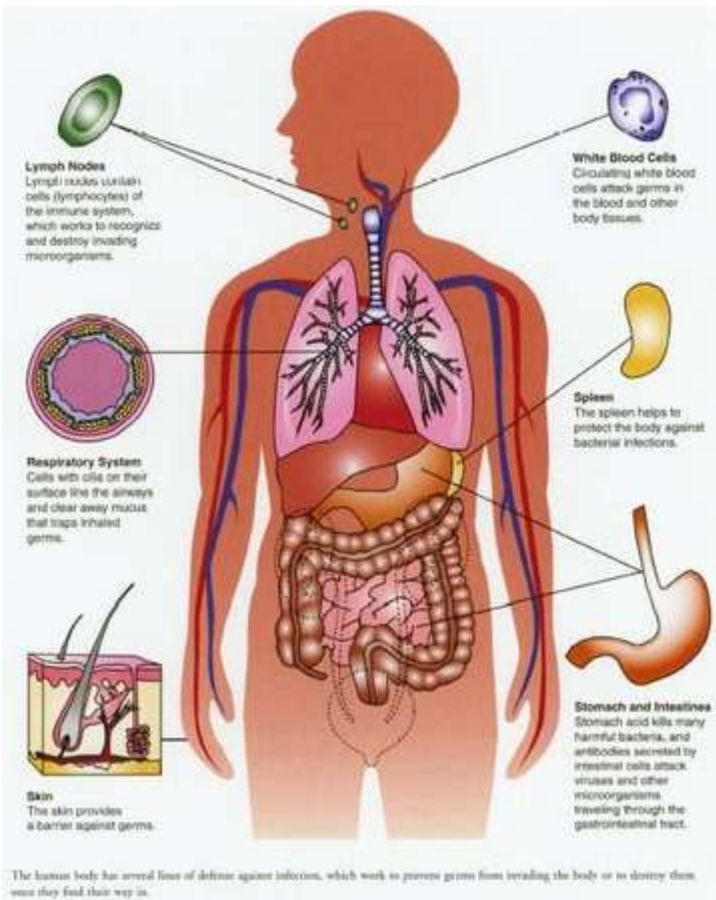
- Our **everyday** actions, as expressed by the **way** we live and **move**, leave digital **traces** in information systems.
  - Move in workplace, perform a surgery, etc.
- This is due to the use of **mobile location** aware devices.
  - That allow us to communicate.
  - That allow them to locate us ! Thanks to positioning technologies.
- Through these **traces** we can **sense** the objects **movements** in a space.
  - City, delegation, hospitals, human body, etc.
- Their potential value is high because of the increasing volume, pervasiveness and **positioning** accuracy of these **traces**.
  - Varied queries can be performed.

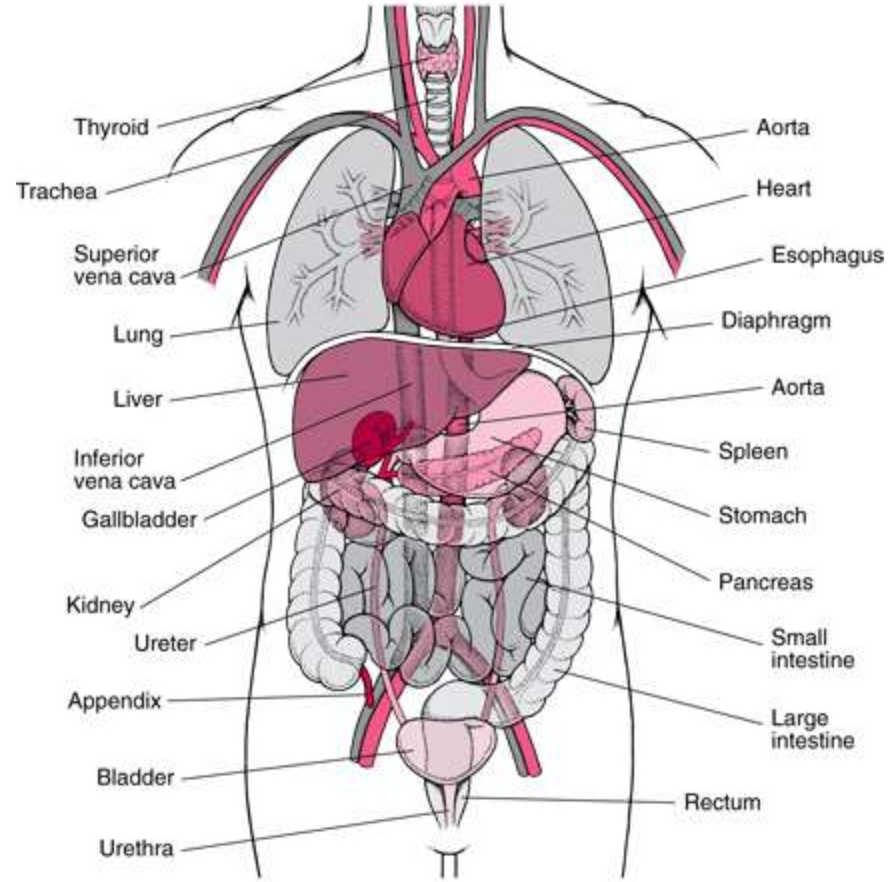
Introduction (2/2)

- Location technologies are capable of:
  - Providing a better estimation of a mobile object's **position**.
- Positioning system-equipped mobile devices can:
  - Transmit **location** information to some services provider.
- Latest advances like :
  - Wi-Fi and Bluetooth devices are becoming a source of data for indoor positioning.
  - Wi-Max can become an alternative for outdoor positioning.

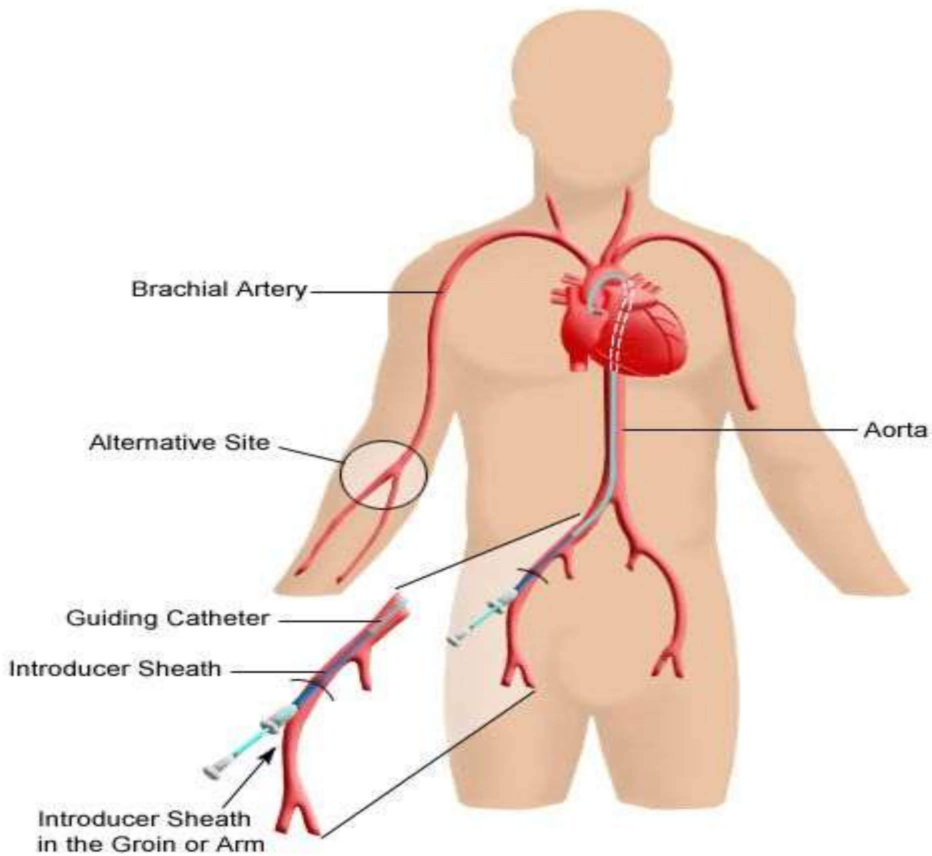
Some kinds of mobility data scenarios

Scenarios in health care (1/4)

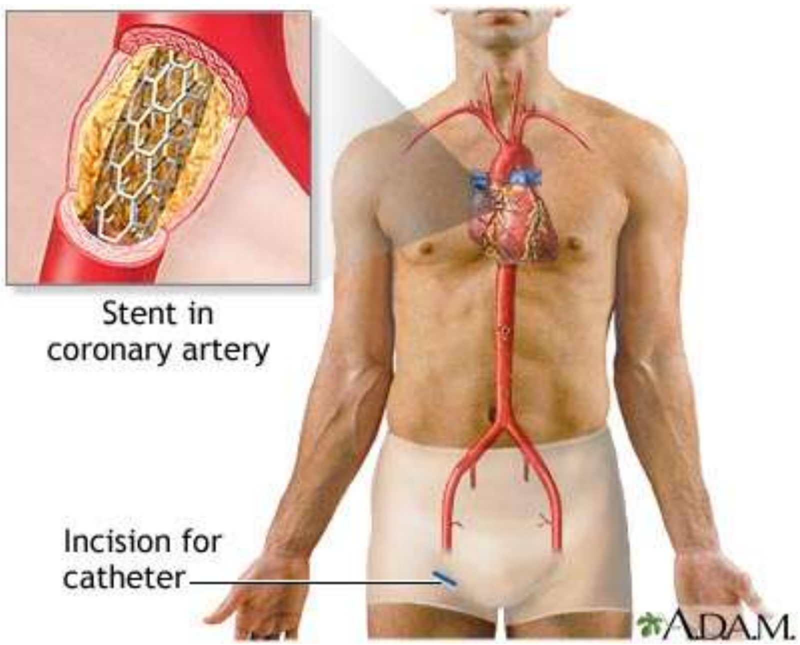


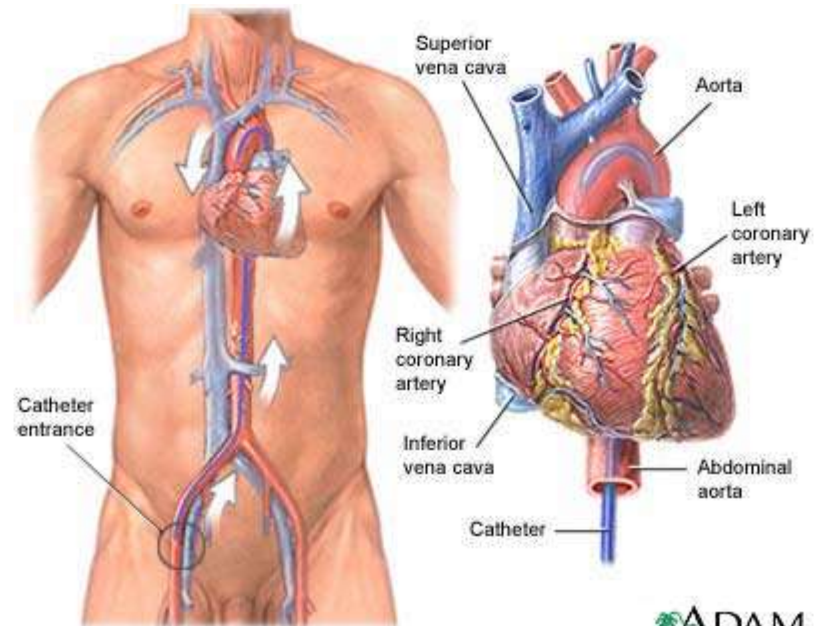


Scenarios in health care (2/4)

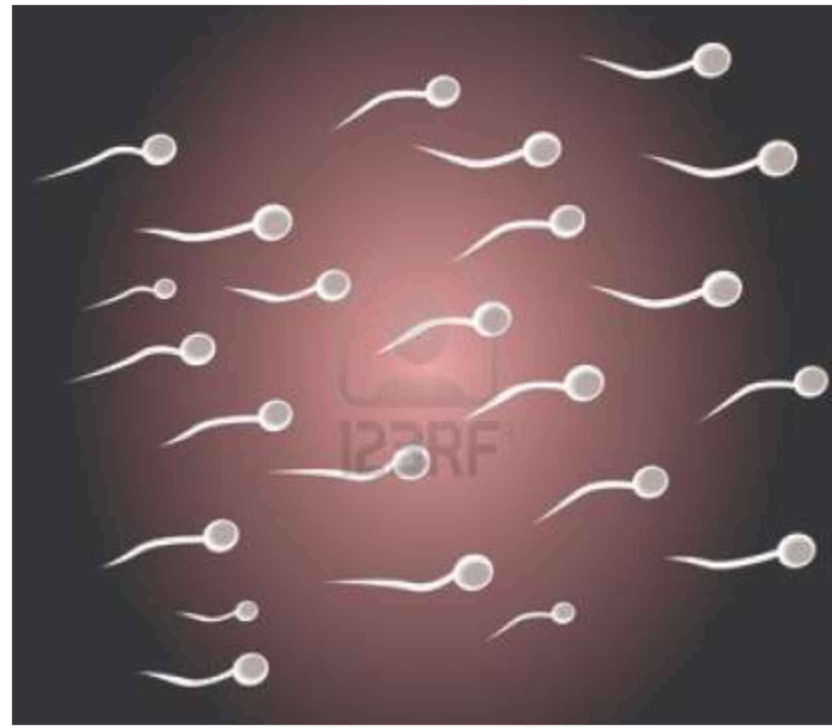


Scenarios in health care (3/4)





Scenarios in health care (4/4)

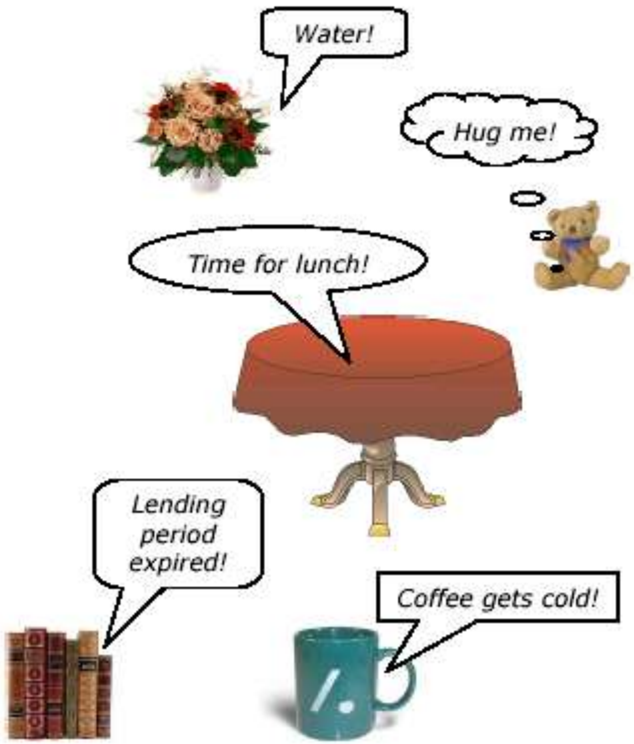


Pervasive systems

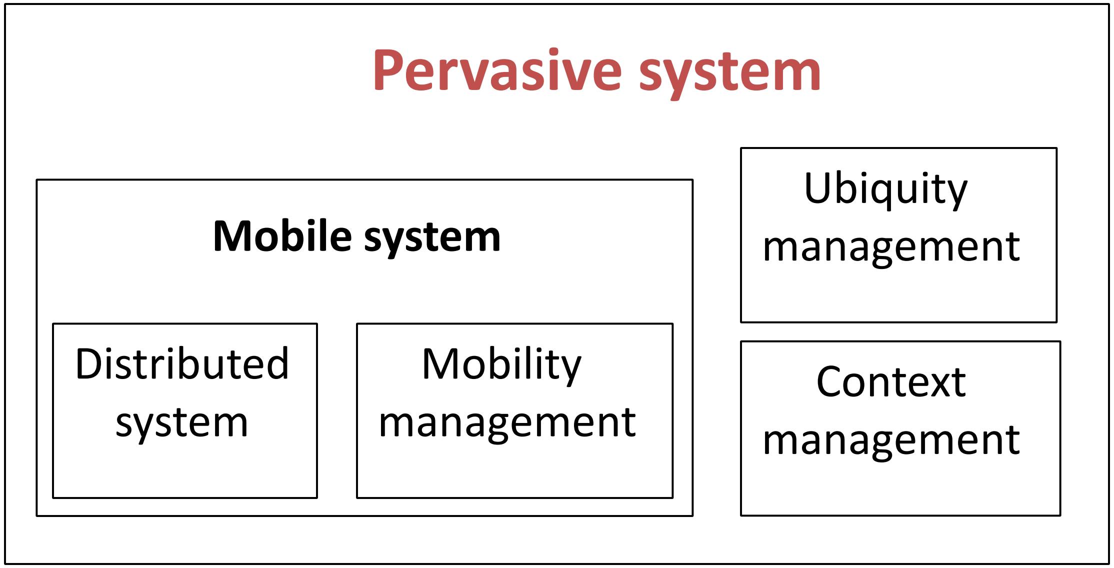
Pervasive computing

- Ubiquitous
  - Accessible from anywhere.
- Mobile
  - Integrate mobile devices.
- Context-aware
  - Take into account the execution context.
- Pervasive
  - Associate ubiquity, mobility and context-awareness.
- Ambient
  - Integrated into daily objects.

Ambient systems



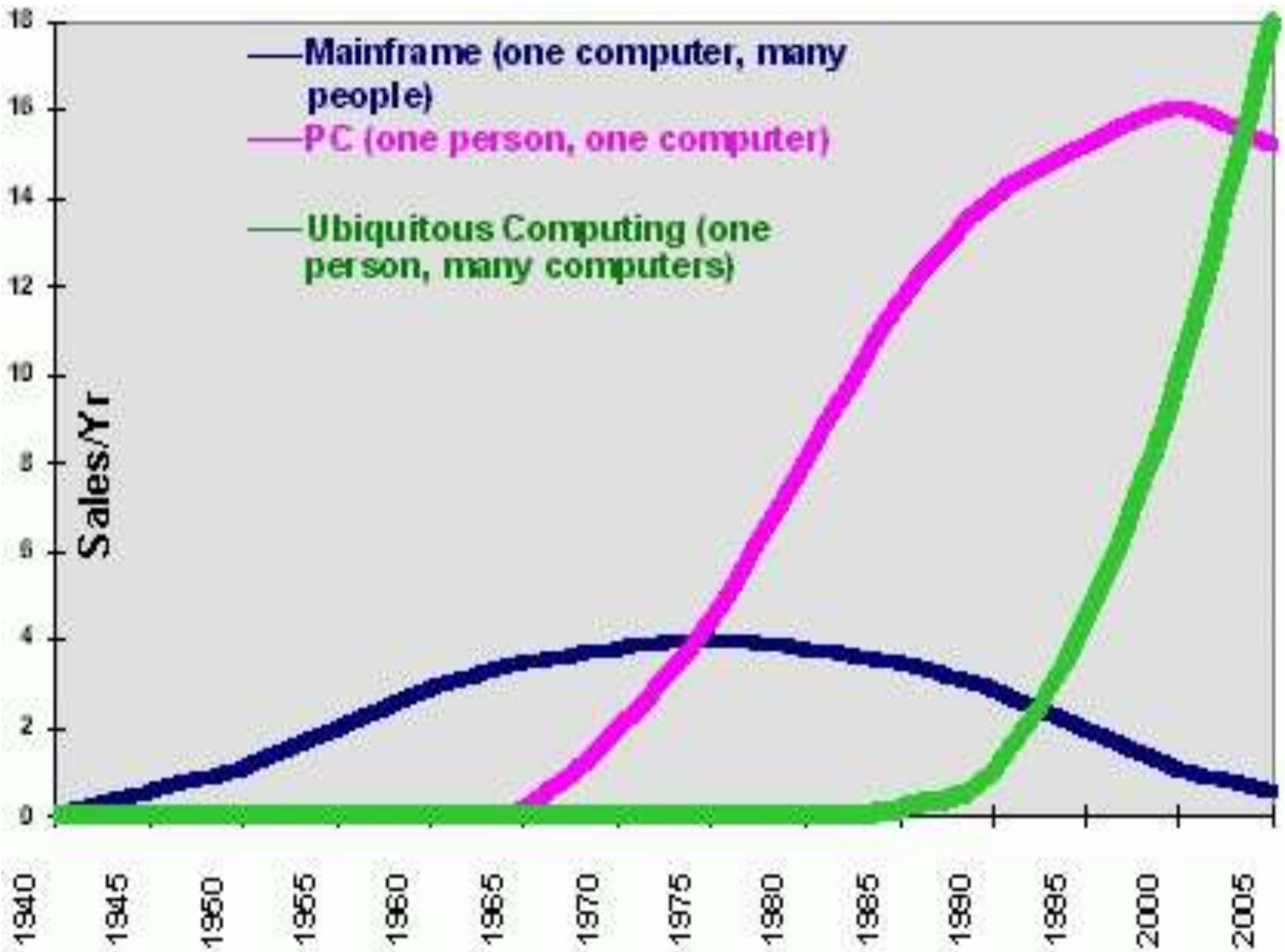
Pervasive system



Four waves - Four paradigms

- Mainframe computing (60’s-70’s)
  - Massive computers to execute big data processing applications.
  - Very few computers in the world.
- Desktop computing (80’s-90’s)
  - One computer at every desk to help in business-related activities.
  - Computers connected in intranets to a massive global network (internet), all wired.
- Mobile computing (90’s-00’s)
  - A few devices for every person, small enough to carry around.
  - Devices connected to cellular networks or WLANs.
- Ubiquitous computing (now)
  - Tens/hundreds of computing devices in every room/person, becoming “invisible” and part of the environment.
  - WANs, LANs, PANs – networking in small spaces.

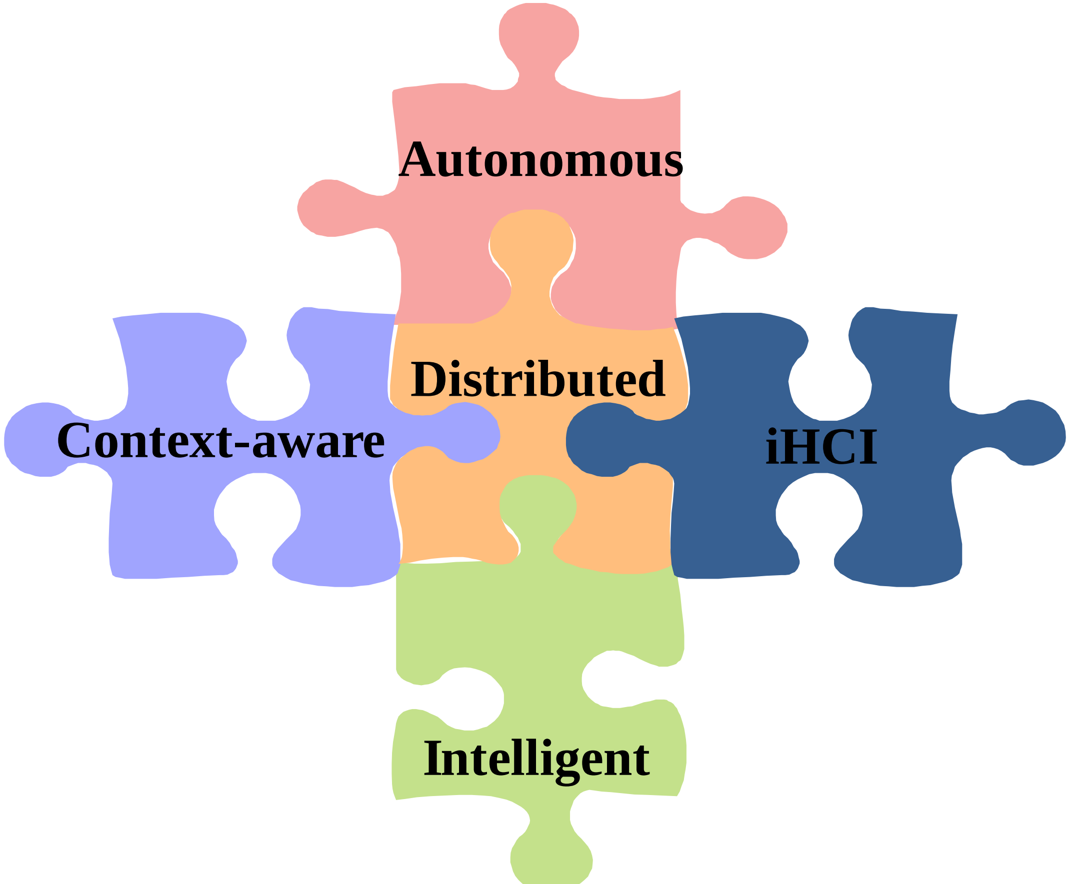
Trend: Weiser's 3 waves of computing



Ubiquitous computing

- Ubiquitous computing:
  - Activates the world.
  - Is invisible, everywhere computing that does not live on a personal device of any sort, but everywhere.
  - Makes a computer so embedded, so fitting, so natural, that we use it without even thinking about it.
- Also called: pervasive, deeply embedded, sentient computing, and ambient intelligence.

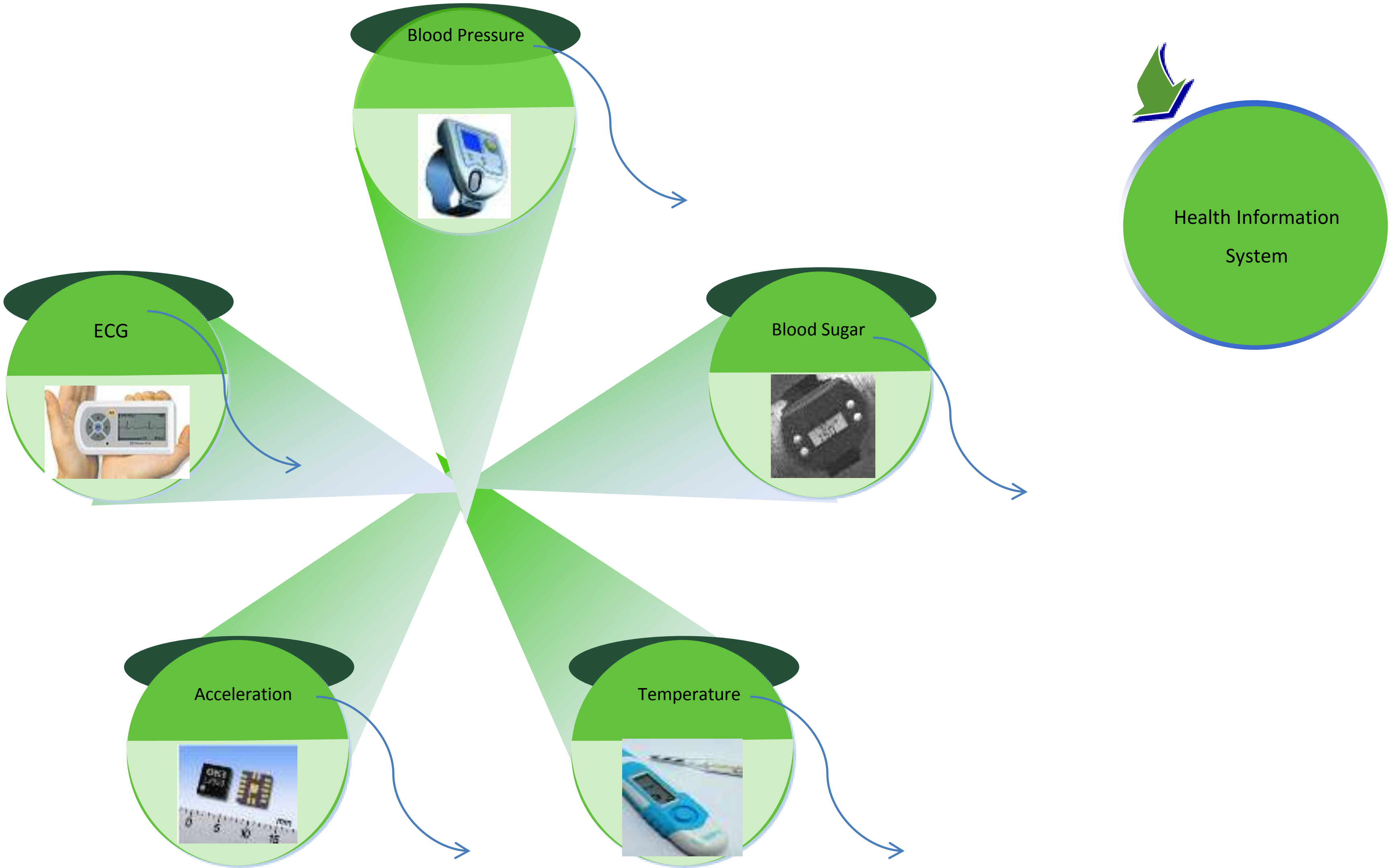
Five main properties for ubiquitous computing



Ubiquitous health care

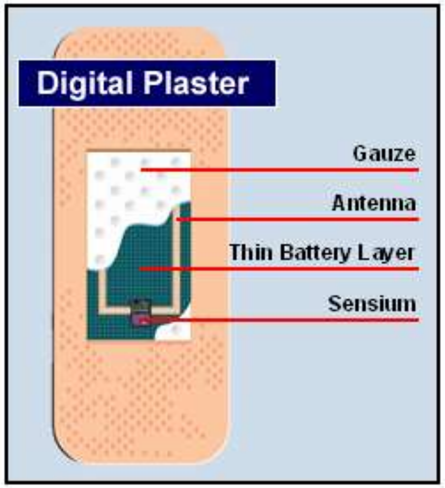
- Environment that collects the information by attaching sensors to medical “objects” and managing real-time information through the network.
- Providing health service of precaution, diagnosis, treatment, post management, etc.
- Everywhere at anytime.

Ubiquitous health care example

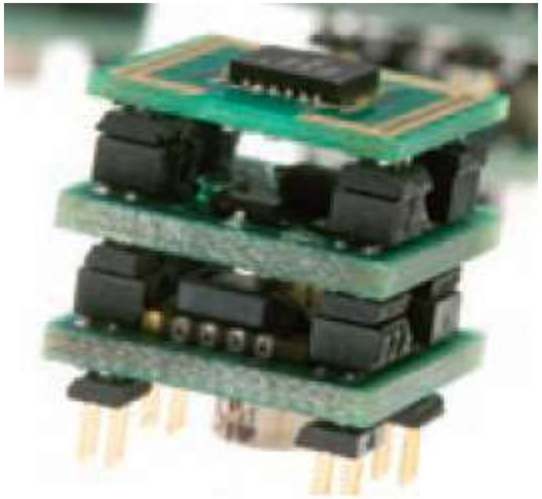


Digital bandage

Low power and small Size

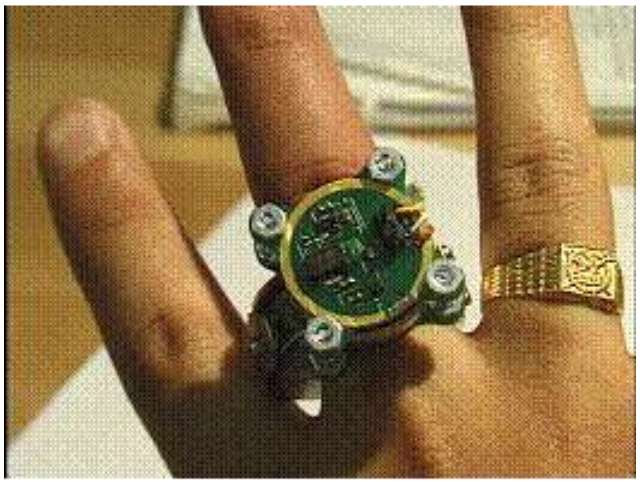


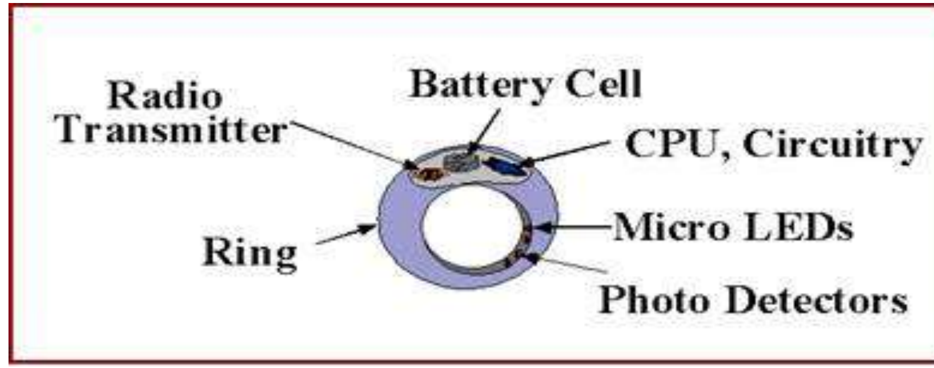
The 'digital' plaster has been designed to be as small as possible

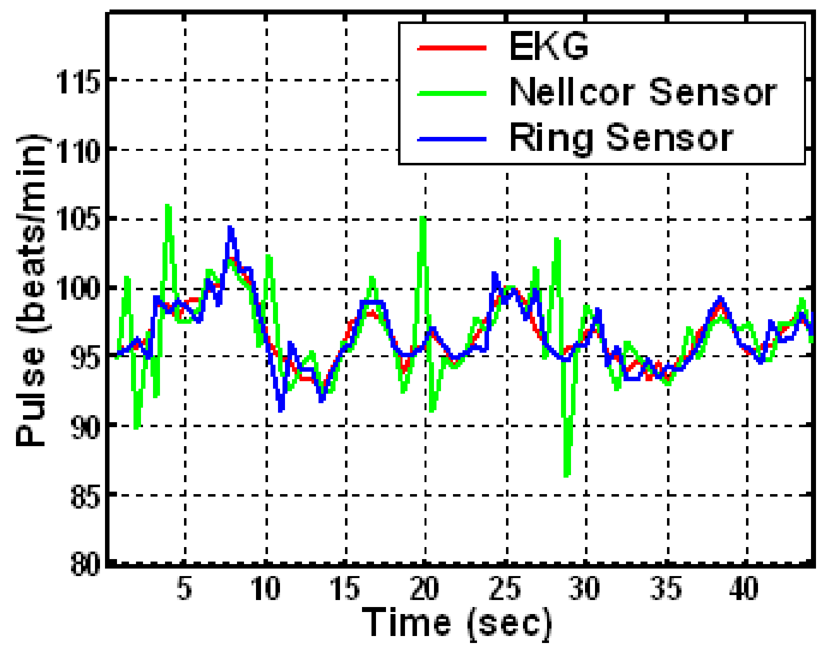


Toumaz company in England	Digital Plaster
IMEC in Belgium	Wireless Sensor Platform

MIT Wireless Ring Sensor





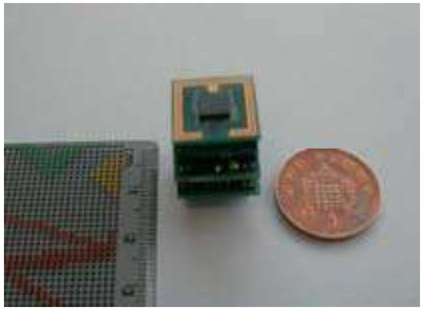


Measurments (wireless)

- Heart rate
- Heart rate variability
- Oxygen saturation
- Estimation of blood pressure

Enabling technologies

- Wireless (data) communication
  - Higher bandwidth
  - Lower power
  - Commodity (readily available and secure)
- Small form factor devices
  - Shrinking electronics
  - Better displays
  - New input methods
- Personalization
  - Machine learning
  - Inference



Extremely varied

- Embedding for smart control
  - Embedded systems for cars, airplanes, patients, etc.
- Creating new computing devices
- Connecting the existing physical world to a computational infrastructure
  - Ordinary objects and tasks re-evaluated and extended with computational/communication capabilities

Mobile computing

- The application of small, portable, and wireless computing and communication devices.
- Being able to use a computing device even when being on the move.
- Portability is one aspect of mobile computing
  - portable vs. mobile.

Distributed IS duties

- Data persistency.
- Data exchange between heterogonous applications.
- Data distribution on distant sites.
- Data permanent consistency management.
- Platforms interoperability.
- Applications portability.
- Concurrent access management.
- Legacy systems integration.
- Openness.
- Security.

Pervasive IS duties

- Distributed IS duties.
- Scalability.
- Invisibility.
- Context-awareness.
- Intelligence (« smartness »).
- Pro-action « all the time everywhere ».

Scalability

- Scaling.
- Manage increasing volumes of
  - Users.
  - Applications.
  - Connected devices.
- Develop applications whose their “heart” is independent from the volume, the users, and the devices.
- Use adaptation techniques to be able to give answers for each case.

Invisibility

- Requires minimal human intervention.
- Self adaptation to changes.
- Self-learning.
- Example:
  - Dynamic reconfiguration of network characteristics of a device.
  - Space resources access according to geographic zone encapsulating that space.
  - Out of space= limits definition of space.

Context-awareness

- Perception of the environment to interact more 'naturally' with the user.
- Sensors of the physical environment.
- Self-descriptive devices.
- Persons description.
- Applications meta-data.

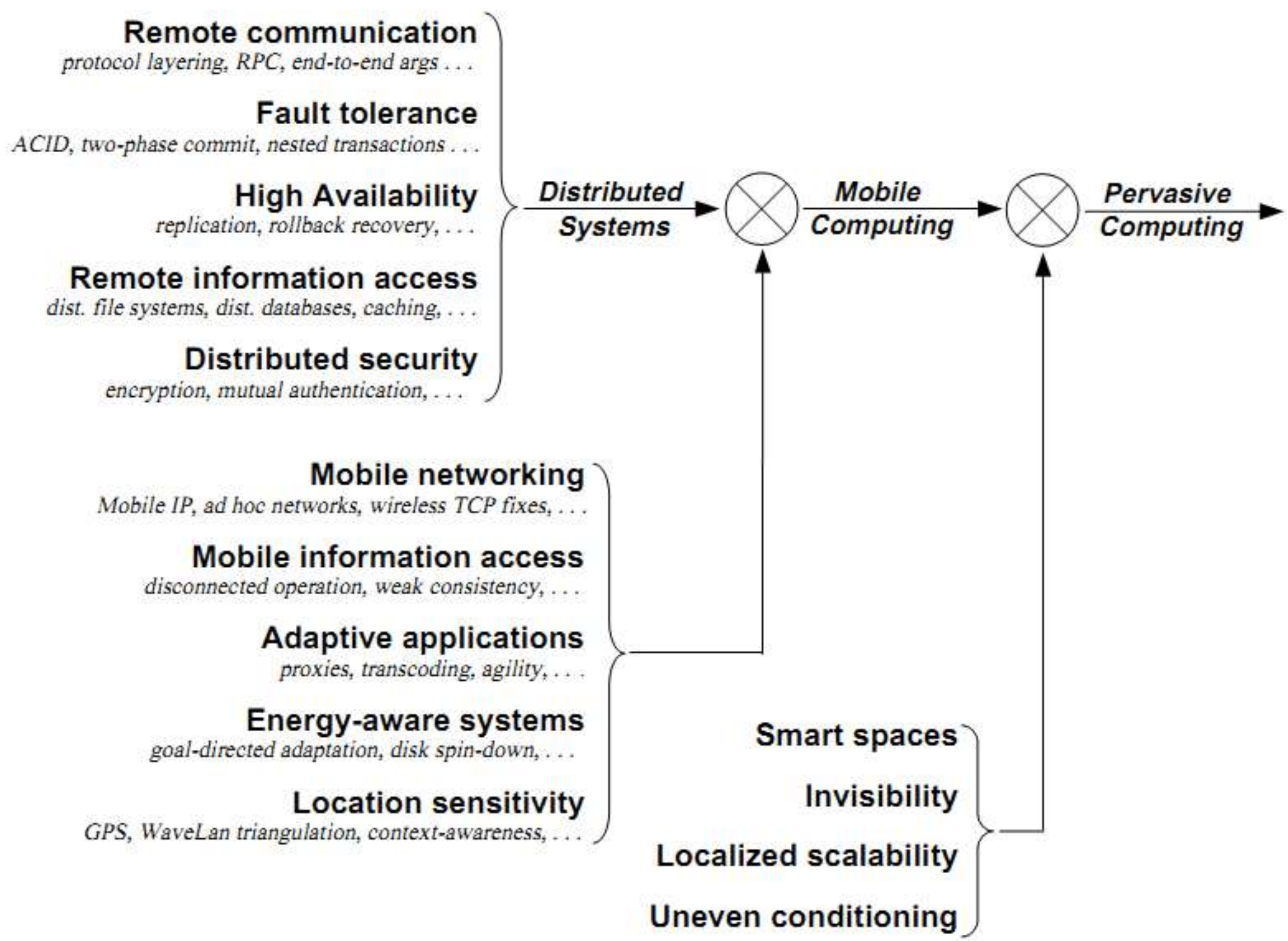
« Smartness »

- Smart = intelligent, quick-witted, malignant, resourceful.
- Perceive the execution context is not sufficient.
- Must effectively use context information.
- Example : smart home.

Pro-action

- Suggest, propose corrective actions to the user depending on the context present or predicted.
  - For example, move 100 meters to a more efficient network and thus accomplish a task in a correct time.
- Implies to know:
  - Predict an event, a situation, etc.
  - Assess a current or a possible situation.
  - Compare two situations and judge the best.
  - That "it's worth" to break invisibility.

Pervasive computing

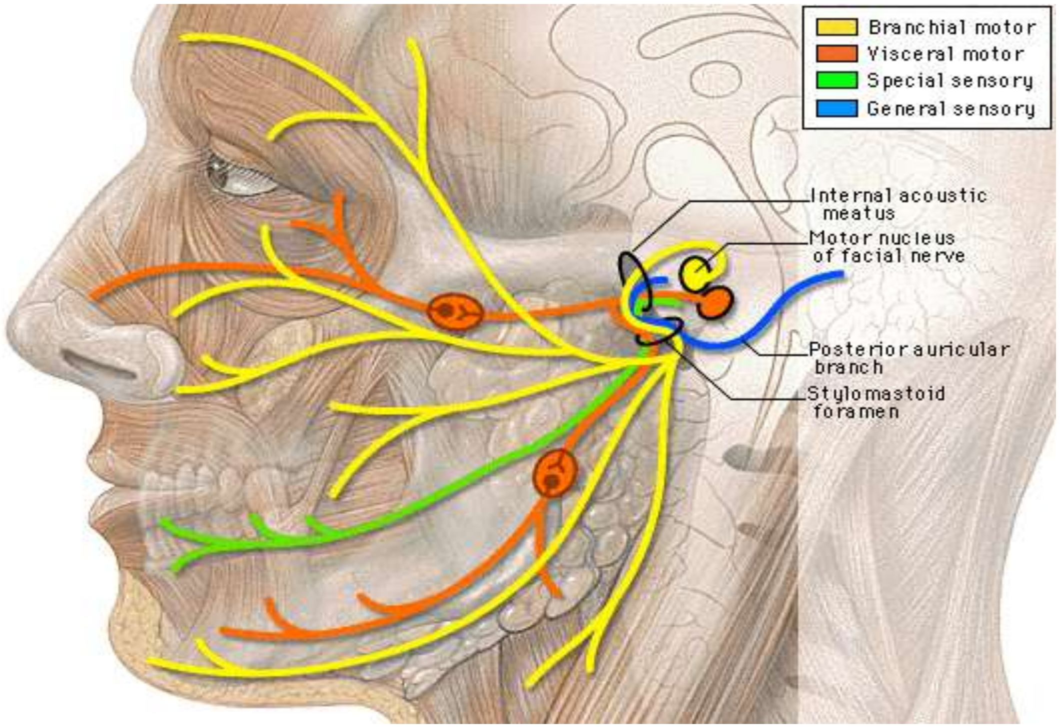


Facial paralysis use case

Motivation scenario: Medical case

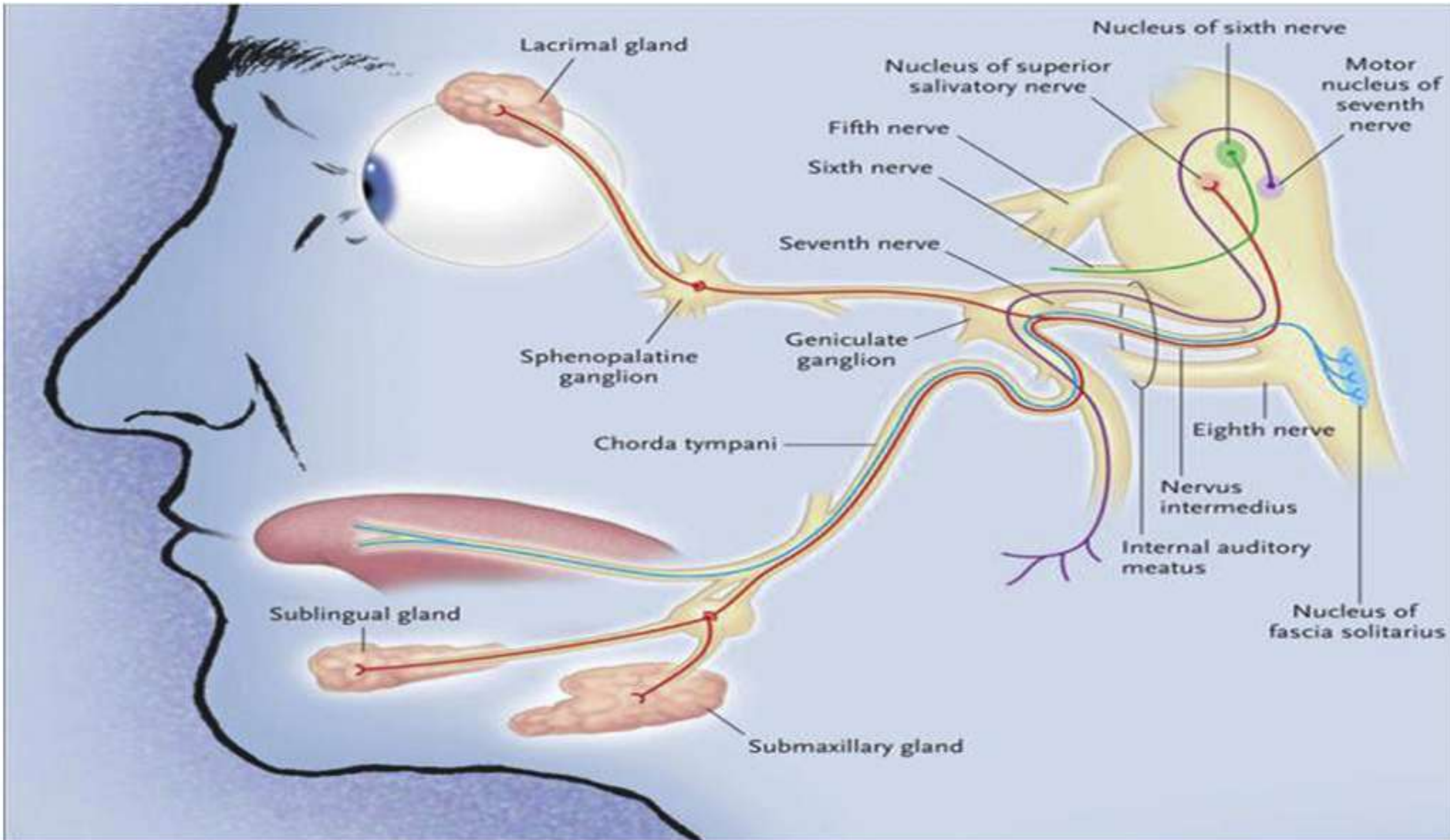
- Let us consider a health care organization which is interested:
  - In analyzing mobility data in different areas such as in facial nerve.
  - To decide upon bell's patients recovery and about the disease behavior.
- It is interested in analyzing:
  - The recovery process of a patient in time intervals.
  - The demographical profiles of patients coming from different geographical zones, belonging to different age intervals, having different family antecedents, etc.
- This knowledge will enable physicians:
  - To understand the disease behavior.
  - To apply more effective strategies according to patients recovery.

Facial nerve components modeling



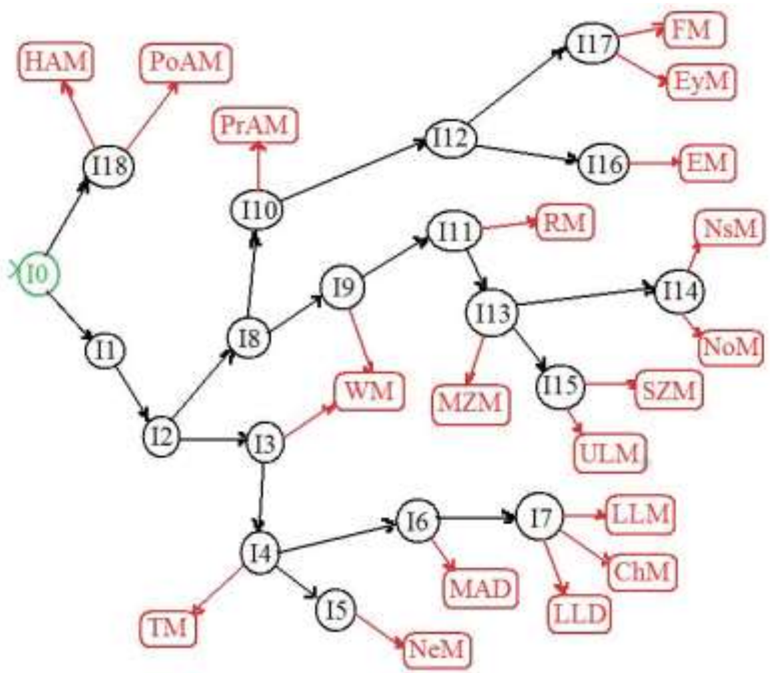
Modeling the facial nerve stream as a moving object circulating into the facial nerve “network”.

Anatomy of facial nerve

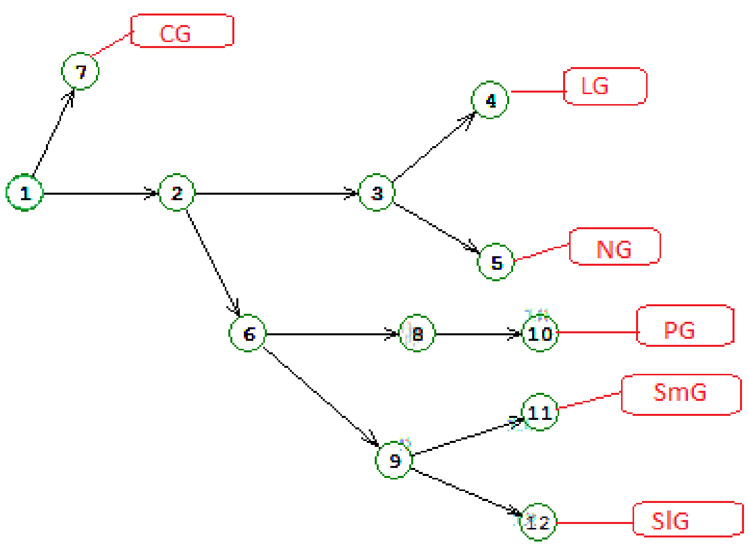


Facial nerve graph

Muscles graph



Glands Graph



Mobility data analysis

Analyzing mobility data (1/3)

- Modern communication and computing devices are **pervasive** and carried by various objects: people, vehicles, etc.
- The consequence is that object and its activity in a space may be sensed:
  - Not necessarily on purpose.
    - We just collect data.
  - As a side effect of the services provided to mobile users.
    - Calls made and/or received, SMS, emails, shopping, diagnosis, etc.
- Wireless mobile devices network is an infrastructure able:
  - To gather mobility data..
  - To analyze them and gain insights about objects movements.
    - Trajectories (stop, moves, activities, etc.).

Analyzing mobility data (2/3)

- Usage of location aware devices allows access to large spatiotemporal datasets.
- The space-time nature of this kind of data:
  - Results in the generation of huge amounts of spatiotemporal data.
  - Imposes new challenges regarding the analytical tools to be used for transforming raw data to knowledge.
- Necessity to investigate the extension of traditional analytical techniques to be applicable on mobility data.

Analyzing mobility data (3/3)

- The analysis of mobility data raises opportunities for discovering behavioral patterns to be exploited in applications like:
  - Mobile marketing.
  - Mobile information collections.
  - Mobile hospitals.
  - Mobile physicians.
  - Stream nerve detection.
  - Heart disease supervision.
  - Robots.
  - Traffic management, etc.

- OLAP and DM techniques can be employed in order to convert this vast amount of raw data into useful knowledge.

Expectations from mobility data analysis (1/3)

- To propose innovative analytical techniques aiming to extract useful patterns from spatiotemporal data.
  - Extraction, Transformation, and Loading.
  - Mobile devices, MOD, TDW.
- To identify the difference between two types of spatiotemporal data: mobility and immobility data.
  - Stream nerve vs muscles.
- To focus on data warehousing and mining techniques that can be applied on MODs.
  - Patients, physicians, devices, medical staffs, hospitals, cities, etc.

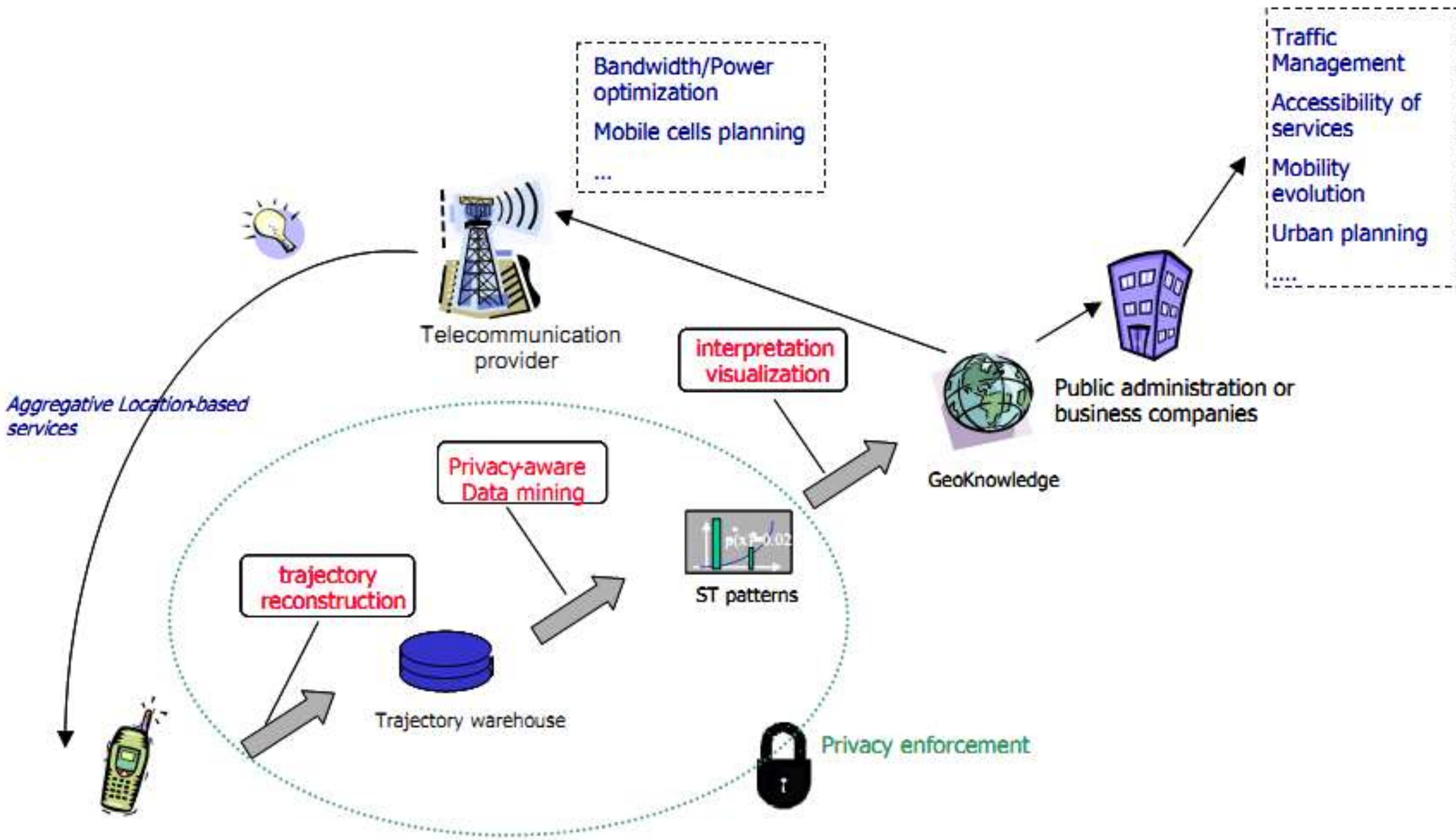
Expectations from mobility data analysis (2/3)

- How traditional data cube model is adapted to TDWs in order to transform raw location data into valuable information.
- How ETL procedure feeds a TDW with aggregated trajectory data.
- How to aggregate cube measures for OLAP purposes.

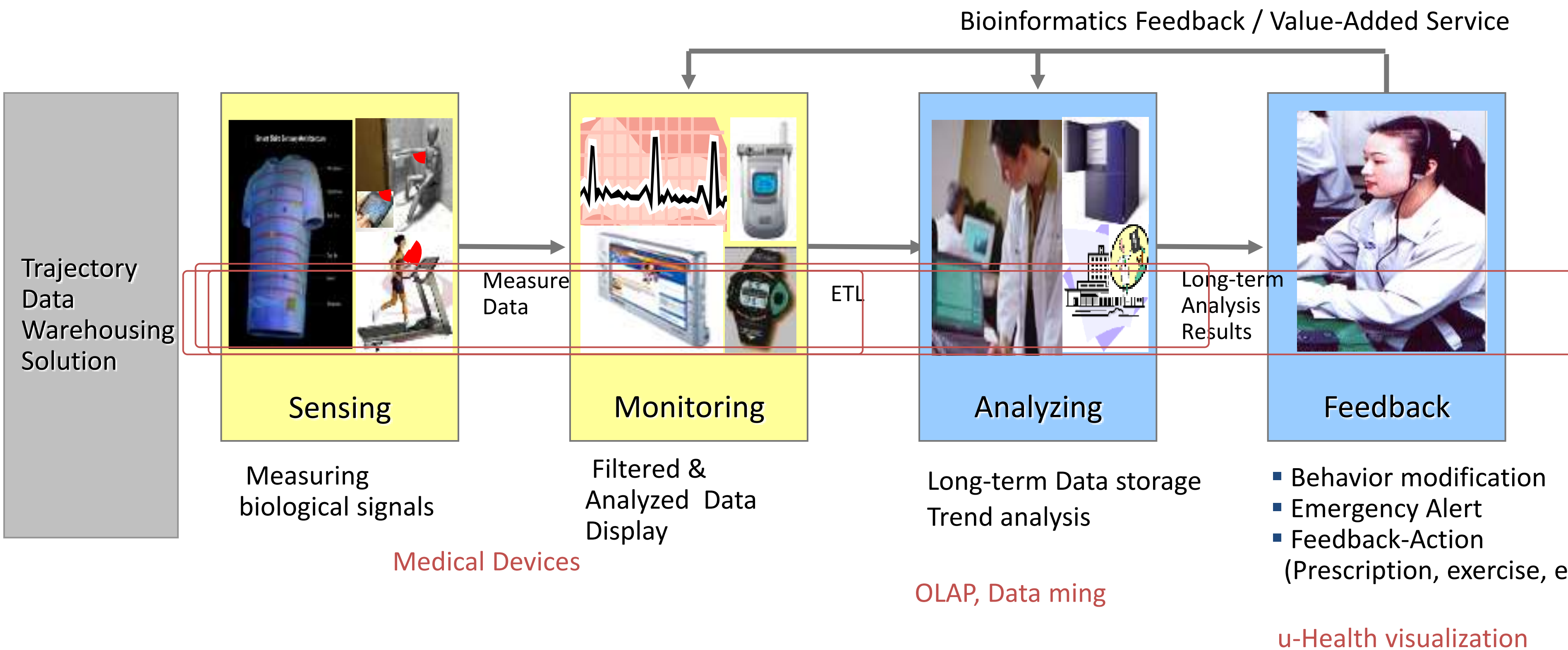
Expectations from mobility data analysis (3/3)

- To study a new approach in designing trajectory data cubes.
- To give answers to ad hoc OLAP queries related to various applications.
- To propose a new OLAP data model that include a flexible fact table that:
  - Can answer queries considering semantic definitions of trajectories.
  - Provides the option to choose the appropriate semantic for aggregation queries over trajectory data.

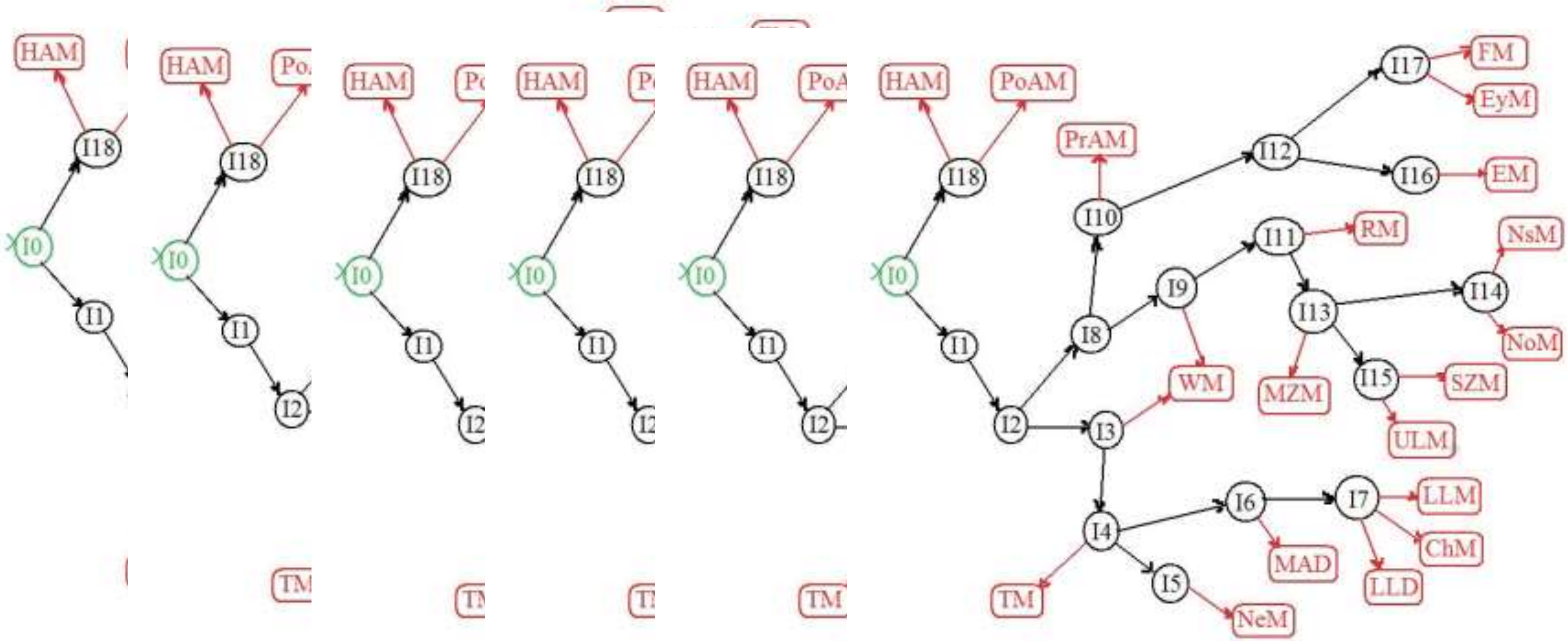
Moving object data management, warehousing and mining



Health care decisional process



Facial nerve TDW



Concepts on spatiotemporal data

Basic concepts on spatiotemporal data (1/3)

- Generation of dissimilar, dynamic, and geographically distributed spatiotemporal data has exploded, thanks to advances in:
  - Mobile devices and remote sensors.
  - Networks.
  - Location sensing devices.
- 2 types of spatiotemporal data: mobile and static.

Basic concepts on spatiotemporal data (2/3)

- The rate at which geospatial data are generated exceeds the ability to organize and analyze them to extract patterns in a timely manner.
- CS and Geo-informatics collaborate to provide innovative and effective solutions.

Basic concepts on spatiotemporal data (3/3)

- A typical category of mobility data is the time-stamped location data:
  - Collected by location-aware devices.
  - Allowing access to large datasets consisting of time-stamped geographical locations.
- The traditional database technology has been extended into MODs that handle:
  - Modeling.
  - Indexing.
  - Query processing issues for trajectories.
- The challenge after storing the data is:
  - The implementation of analytics to extract useful knowledge.

Immobile entities (1/2)

- Sensing technologies feel, record and study phenomena.
  - Human body, diseases, etc.
- At least data collection occurs every one small time unit depending on applications.
- Patients data collection is huge and rapidly increasing.
- Medical staffs record information to describe and study patients bodies activities.
- Analysts find a valuable “data treasure”, to process and analyze to discover knowledge from this data.

Immobile entities (2/2)

- Human body phenomena are instantly recorded by a number of organizations worldwide.
- A system collecting and analyzing most accurate data among different sources is needed.
- Some sources provide data about the same phenomena though with differences in their details.
- The need for a generic architecture will be able to integrate the remote sources in a proper way by refining and homogenizing raw data.

Data warehousing and mining components

Data warehousing and mining components

- Tools for data exploration and inspection.
- Algorithms for generating historic profiles of activities related to specific spaces and time periods.
- Techniques providing the association of data with other geophysical parameters of interest:
  - Patient morphology, disease and recovery evolution, etc.
  - Visualization components using geographic and other thematic-oriented maps for:
    - Presentation of data to users such as medical staffs.
    - Supporting sophisticated user interaction.

Trajectory data warehousing and mining users profiles

- Physicians are interested in:
  - Constructing and visualizing patients' profiles of certain body regions during specific time periods of a disease evolution.
  - Discovering regions of similar behavior.
  - Disease activity, thus querying the system for properties of general interest.

Facial paralysis warehousing and mining

- DWMS architecture provides users a wealth of information about patients recovery from Bell’s palsy recovery.
- Collected data can be stored in a local database and/or a data warehouse (for simple querying and analysis for decision making, respectively).
- Data within the database is dynamic and detailed; while that within the data warehouse is static and summarized.

DWMS querying functionality

- Retrieval of spatial information given a temporal instance.
  - When we are dealing with records including:
    - Position (segments).
    - Time of facial nerve partial or total recovery together with attributes like intensity, segments, paths, muscles, etc.
- Retrieval of spatial information given a temporal interval.
  - Evolution of spatial objects (stream nerve) over time.
  - Recording the duration of partial and total recoveries and how certain parameters of the phenomenon vary throughout the time interval of its duration.

Examples of typical queries

- Find the number of recoveries realized during the past four months, which reside more closely to a given location.
- Find all sequelae of patients residing in a certain region.
- Find the number of recoveries occurred in a specified time interval.

Maintaining summary for data analysis (1/3)

- Two popular techniques for analyzing data and interpreting their meaning are:
  - OLAP analysis.
  - Data mining.
- Summarized health care data can study the phenomenon from a higher level and search for hidden, previously unknown knowledge.
- View part of the historical recovery profile:
  - Example. the number of cases that leads to a surgery in the past twenty years, over a specified region.
- View the same information over a country, continent, the world.
  - More detailed view, formally a drill-down operation.
  - Worldwide. More summarized view, formally a roll-up operation.

Maintaining summary for data analysis (2/3)

- Slice and dice, for selecting parts of a data cube by imposing conditions on a single or multiple cube dimensions.
- Pivot, which provides the user with alternative presentations of the cube.
- Integrating data analysis and mining techniques into an DWMS aims to the discovery of interesting, implicit and previously unknown knowledge.

Examples of useful patterns found through Knowledge Discovery & Delivery (KDD) process

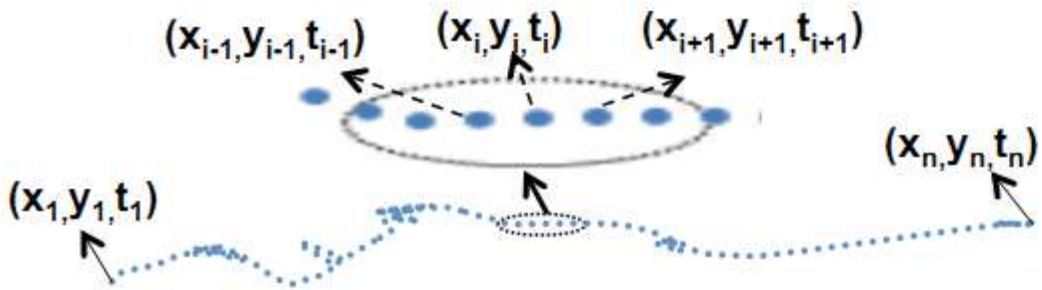
- Clustering of information.
  - Cases occurred closely in space and/or time.
  - Cases related to kids, adults, etc.
- Classification of phenomena with respect to recovered areas, and detecting phenomena semantics by using pattern finding techniques, etc.
  - Characterizing the main recovery aspects in recovery sequences.
  - Measuring the similarity of sequelae sequences, according to a similarity measure specified by the domain expert, etc.

Mobile entities

The case of trajectory data

- Moving objects are geometries (i.e. points, lines, areas) changing over time.
  - Pills, stunts, stream nerve, etc.
- Trajectory data describes the movement of these objects.
- Movement implies two dimensions: the spatial and the temporal (recovery).
- Movement can be described as continuous change of position in the geographical space and through comparison between two different temporal instants.

Spatio-temporal trajectory

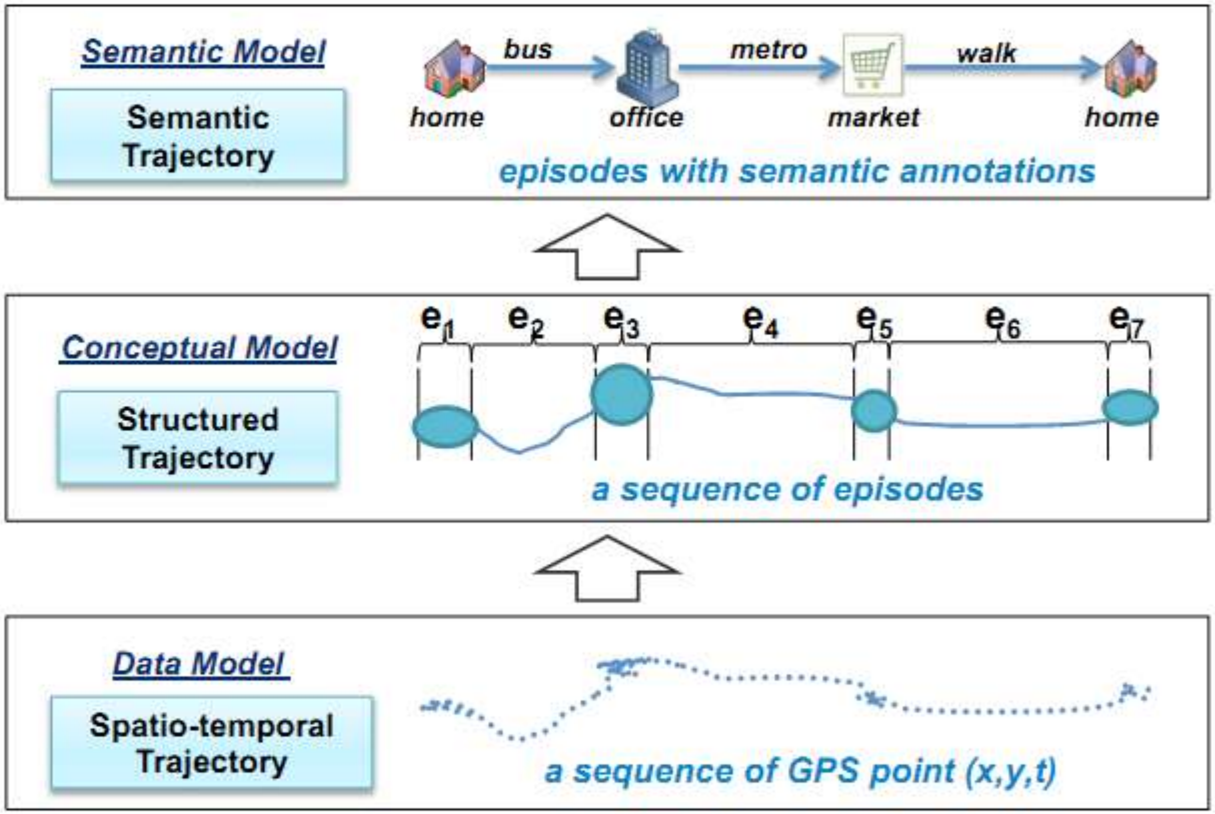


- A sequence of spatiotemporal points
  - In a road network.
  - In a human body network.
  - In a part a body network: The face for example.

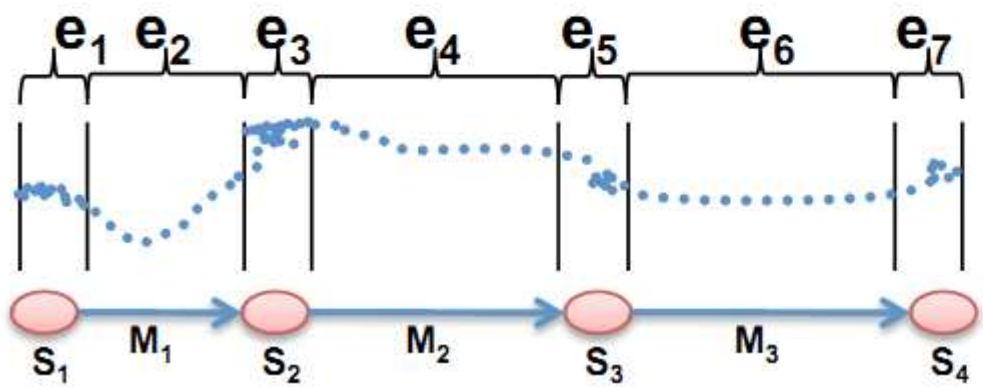
A semantic trajectory example

- Semantic enrichment integrate structured trajectories with semantic knowledge from the two semantic viewpoints, i.e.:
  - Geographic view.
  - Application domain view.

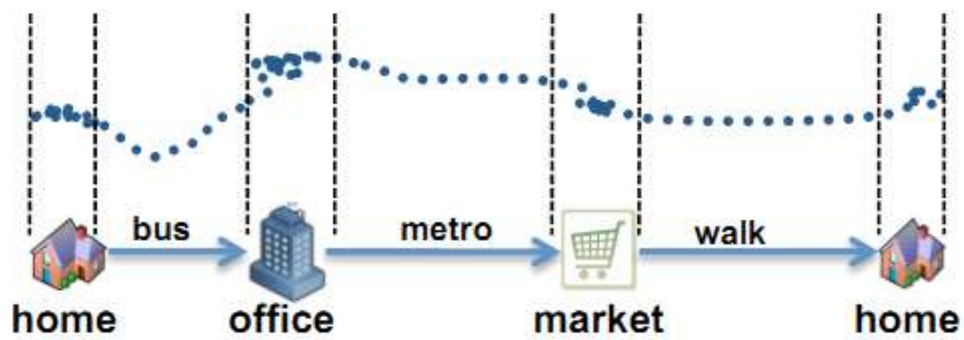
A hybrid semantic trajectory model



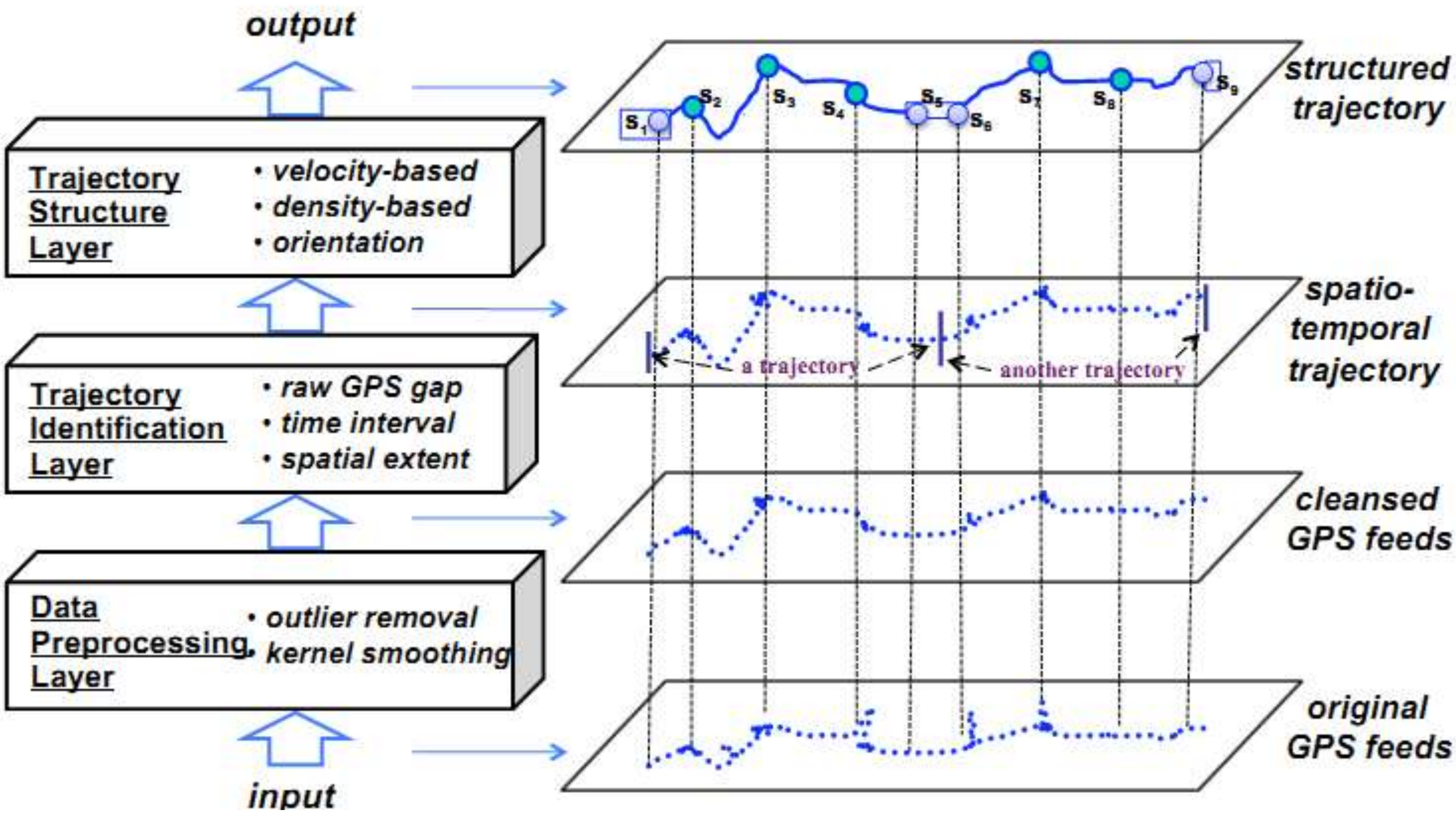
Structured trajectory (a sequence of episodes)



Semantic trajectory (a sequence of semantic episodes)



Offline trajectory computing framework



Formalization

Formalization (1/3)

- A trajectory T is a continuous mapping from the temporal $I \subseteq R$ to the spatial domain (R^2, the 2D plane).
- $I \subseteq R \rightarrow R^2$: $t \rightarrow a(t) = (a_x(t), a_y(t))$.
- $T = \{(a_x(t), a_y(t), t) \mid t \in I\} \subset R^2 x R$.
- Where $(a_x(t), a_y(t), t)$ are the sample points contained in the available dataset.

Formalization (2/3)

- From an application point of view:
  - A trajectory is the recording of an object‘s motion.
  - Example. The recording of the positions of an object at specific timestamps.
- The actual trajectory consists of a curve.
- Real-world requirements imply that the trajectory has to be built upon a set of sample points.
  - The time-stamped positions of the object.

Formalization (3/3)

- Trajectories of moving points are often defined as sequences of (x, y, t) triples.
- $T=\{(x_1, y_1, t_1), (x_2, y_2, t_2), \ldots, (x_n, y_n, t_n)\}$, where $x_i, y_i, t_i \in R$, and $t_1 < t_2 \ldots < t_n$.
- The main objective is to include appropriate techniques for the representation, querying, indexing and modeling of moving object's trajectories.

Innovation

The need for innovation in decision support techniques (1/2)

- Traditional decision support techniques developed as a set of applications and technologies for:
  - Gathering.
  - Storing.
  - Analyzing.
  - Providing access to data.
- Example:
  - Data warehousing.
  - Online analytical processing.
  - Data mining.
  - Visualization.

The need for innovation in decision support techniques (2/2)

- These techniques are embedded in decision support systems.
  - To support business and organizational decision-making activities.
- Such systems help decision makers to identify, analyze and solve problems as well as make decisions, by combining:
  - Raw data.
  - Documents.
  - Personal knowledge.
  - Business models, etc.
- Decision support techniques were developed to satisfy the changeable and complicated needs of current business and technological environment.

Decision support techniques extensions (1/2)

- The extension of traditional techniques to deliver new analytics, suitable for mobility data.
- To serve emerging applications (e.g. mobile health care) that need to convert raw location data into useful knowledge.
- A TDW can help towards computing aggregations on trajectory data and thus studying them in a higher level of abstraction.

Decision support techniques extensions (2/2)

- Data mining techniques are used to discover unknown, useful patterns.
- The vast amount of available mobility data requires the extension of traditional mining techniques so as to be suitable for this new kind of data.
- Discovering spatiotemporal associations, clusters, predicting actions, etc., lead to mobility patterns:
  - That could help to construct summary and useful abstractions of large volumes of raw location data and gain insights on movement behaviors.

Efficient trajectory data warehousing

Efficient trajectory data warehousing (1/2)

- Data warehousing is a technology for integrating all sorts of transactional data, dispersed within organizations.
- A DW is defined as a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management of decision making process.
- In a DW, data are organized and manipulated in accordance with the concepts and operators provided by a multidimensional data model, which views data in the form of a data cube.

Efficient trajectory data warehousing (2/2)

- A data cube allows data to be modeled and viewed in multiple dimensions:
  - Each dimension represents some business perspective.
  - It is typically implemented by adopting a star, snowflake, or constellations schema model.
- A DW consists of a fact table surrounded by a set of dimensional tables related with it, which contains keys and measures.

Dimensions and measures

- Dimensions represent the analysis axes, while measures are the variables being analyzed over the different dimensions.
- Each dimension is organized as a hierarchies of dimension levels; each level corresponding to a different granularity for the dimension.
- The members of a certain dimension level (months) can be aggregated to constitute the members of the next higher level (years).
- The measures are also aggregated following this hierarchy by means of an aggregation function.

OLAP Operations (1/2)

- DWs are optimized for OLAP operations.
- Typical OLAP operations include:
  - The aggregation or de-aggregation of information (called roll-up and drill-down, respectively) along a dimension.
  - The selection of specific parts of a cube (slicing and dicing).
  - The reorientation of the multidimensional view of the data on the screen (pivoting).

OLAP Operations (2/2)

- Data warehousing and OLAP techniques can be employed in order to convert vast amount of raw data into useful knowledge.
- Conventional techniques were not designed for analyzing trajectory data.
- There is the need for extending data warehousing technology so as to handle mobility data.
- Such a warehouse could analyze measures: he number of patients in specific spatial areas, the average intensities of facial nerve, the maximum and average speed of a stream nerve.
- This analysis could be done through appropriate dimensions that will allow to explore aggregated data under different granularities.

Motivation issues (1/2)

- Transform raw trajectories to valuable information used for decision making in ubiquitous applications.
- What makes extracting valuable information from such spatiotemporal data a hard task:
  - The high volume of raw data produced by sensing and positioning technologies.
  - The complex nature of data stored in trajectory databases.
  - The specialized query processing demands.
- Extend traditional aggregation techniques to produce summarized trajectory information and provide OLAP style analysis.

Motivation Issues (2/2)

- Extending traditional (i.e., non-spatial), spatial or spatiotemporal models to incorporate semantics driven by the nature of trajectories introduce specific requirements:
  - Support high level OLAP analysis.
  - Facilitate knowledge discovery from TDWs.
- The basic analysis constituents in a TDW (i.e. facts) are the trajectories themselves.
- We categorize the identified requirements into modeling, analysis and management requirements.
  - The first considers logical and conceptual level challenges introduced by TDWs.
  - The second goes over OLAP analysis requirements.
  - The third focuses on more technical aspects.

Data cube modeling

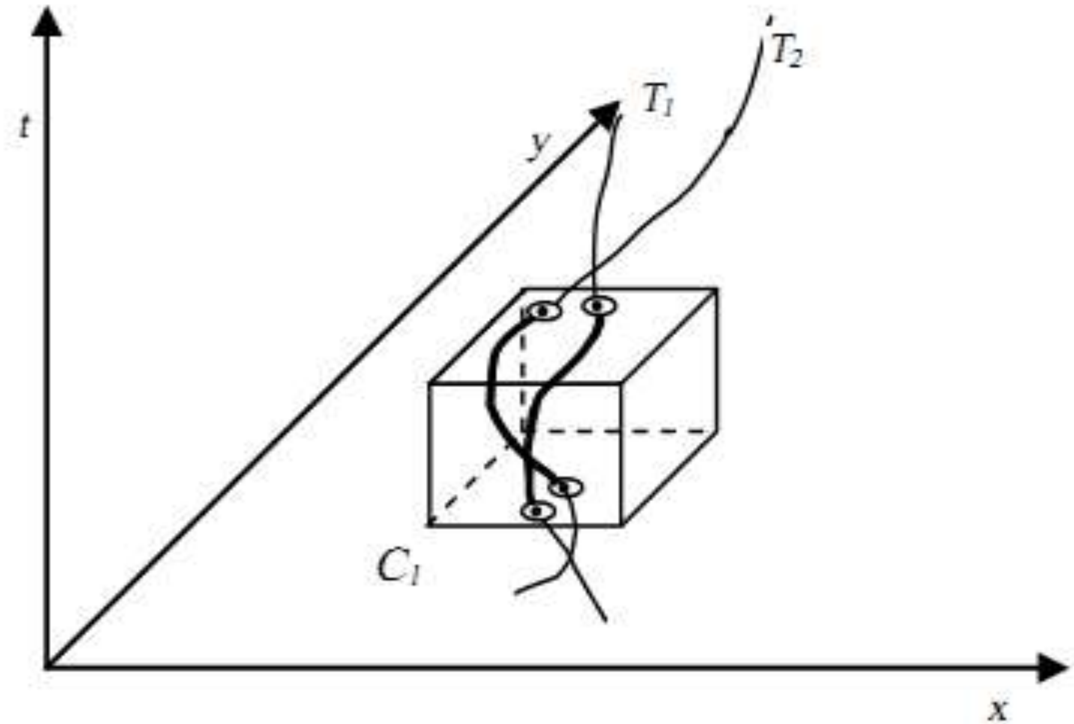
Data cube modeling issues

- Investigate the prerequisites and the constraints imposed:
  - When describing the design of a TDW from a user perspective (i.e. conceptual model).
  - When describing the final result as a system in a platform-independent tool (i.e. logical model).

Thematic, spatial, temporal measures

- From a modeling point of view:
  - A trajectory is a spatial object whose location varies in time.
- Trajectories have thematic properties:
  - Usually are space and time dependent.
- Characteristics of trajectories are described to be analyzed:
  - Numeric characteristics, such as the average speed of the trajectory, its direction, its duration.
  - Spatial characteristics, such as the geometric shape of the trajectory.
  - Temporal characteristics, such as the timing of the movement.
  - Spatiotemporal characteristics; such as a representative trajectory or a cluster of trajectories.

The portions of trajectories that lie within a cell



Thematic, spatial, temporal measures

- Depending on the application and user requirements, several numeric measures could be considered.
- The number of trajectories found in the cell (or started/ended their path in the cell; or crossed/entered/left the cell, and so on).
- The {average/min/max} distance covered by trajectories in the cell.
- The {average/min/max} time required to cover this distance.

Thematic, spatial, temporal dimensions (1/4)

- Regarding the supported dimensions, as starting point a TDW should support:
  - The classic spatial dimensions (e.g. coordinate, roadway, district, cell, city, province, country).
  - Temporal dimensions (e.g. second, minute, hour, day, month, year).
  - Hierarchies, describing the underlying spatiotemporal framework wherein trajectories are moving.

Thematic, spatial, temporal dimensions (2/4)

- It is important to allow space-time related dimensions to interact with thematic dimensions describing other sorts of information regarding trajectories like:
  - Technographic (mobile device used).
  - Demographic data (age and gender of users).
- This allow an analyst :
  - To query TDW about the number of objects crossed an area of interest: get quantitative information.
  - To identify the objects in question: get qualitative information.

Thematic, spatial, temporal dimensions (3/4)

- A rich TDW schema should include the following dimensions:
  - Temporal (time).
  - Geographical (location).
  - Demographics (gender, age, occupation, marital status, home and work postal code, etc.).
  - Technographics (mobile device, sensors, subscriptions in special services, etc.).
- Technographics and demographics dimensions:
  - Enhance the warehouse with semantic information.
  - Allow the grouping of trajectories according to demographical characteristics or based on the technological characteristics of their devices.

Thematic, spatial, temporal dimensions (4/4)

- Trajectory is a set of sampled locations in time, for which:
  - The in-between positions are calculated through some kind of interpolation.
  - The lowest level information is that of spatial coordinates.
- This implies a huge discretization of the spatial dimension: cell positions could be used instead of point positions.

OLAP

OLAP Requirements

- In traditional DWs, data analysis is performed interactively by applying a set of OLAP operators.
- In spatial data warehousing, particular OLAP operators have been defined to tackle the specificities of the domain.
- We expect algebra of OLAP operators to be defined for trajectory data analysis.
- Such an algebra should include:
  - Traditional operators, such as roll-up, drill-down and selection properly tailored to trajectories.
  - Additional operators which account of the spatiotemporal data type.

Roll-up

- Roll-up operation allows to navigate from a detailed to a more general level of abstraction:
  - Either by climbing up the concept hierarchy (e.g. from the level of ‘city‘ to the level of ‘country‘).
  - Or by some dimension reduction (e.g. by ignoring the ‘time‘ dimension and performing aggregation only over the ‘location‘ dimension).

Drill-down

- Drill-down operation is the reverse of roll-up.
- It allows to navigate from less detailed to more detailed information by:
  - Either stepping down a concept hierarchy for a dimension (‘country‘ to ‘city‘).
  - Or by introducing additional dimensions (e.g. by considering not only the ‘location‘ dimension but the ‘time‘ dimension also).

Slice, Dice

- Slice operation performs a selection over one dimension (‘city=Tunis‘), whereas dice operation involves selections over two or more dimensions (‘city=Swarthmore and year=2006‘).
- The conditions can involve not only numeric values but also more complex criteria, like spatial and/or temporal query windows.
- To support these operations, the selection criteria can be transformed into a query against the TDW and processed by adequate query processing methods.
- OLAP operations should be also supported by a TDW since they provide meaningful information.

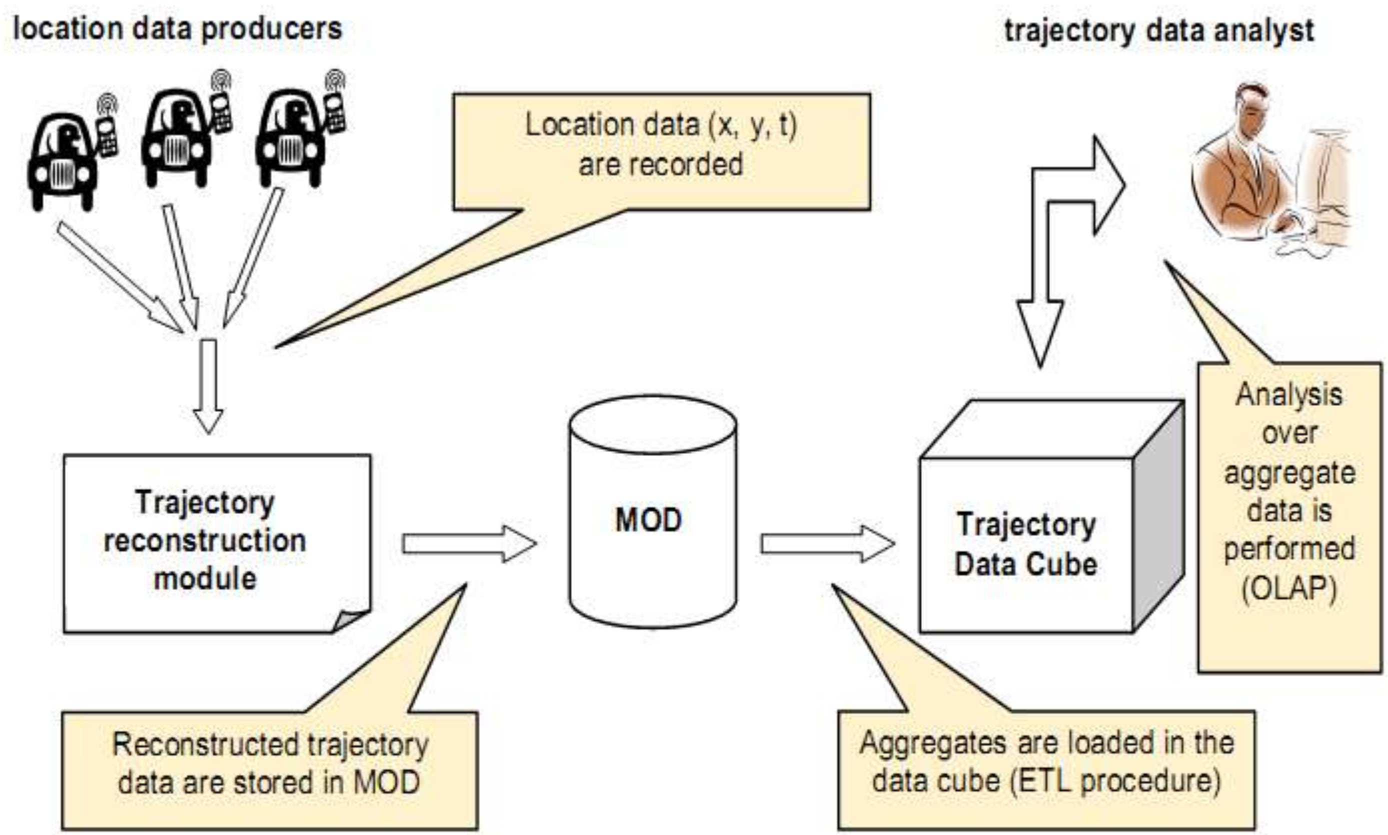
Other operations

- Other operations dedicated to trajectories might be defined. Examples include:
  - Operators that dynamically modify the spatiotemporal granularity of measures representing trajectories.
  - Medoid etc. operators which apply advanced aggregation methods, such as clustering of trajectories to extract representatives from a set of trajectories.
  - Operators to propagate/ aggregate uncertainty and imprecision present in the data of the TDW.

Trajectory Data Warehousing

Trajectory Data Warehousing

The framework



Key questions

- How to store and query trajectory data?
  - Database technology is extended: Moving Object Databases.
- How to reconstruct a trajectory from raw logs?
  - Position devices provide us information just about location points and not about trajectories.
- How to analyze trajectory data
  - Data Warehousing technology is adapted to handle trajectory data.
- Are there any spatiotemporal patterns in my data?
  - New data mining algorithms are needed so as to discover such patterns.
  - Mobile objects health care mining is a very interesting topic in this area.

Data mining (1/2)

- Clustering, the discovery of groups of similar trajectories, together with a summary of each group.
- Which are the main paths of recovery during time intervals.
- Which are the main paths of non recovery during time intervals.

Data mining (2/2)

- Frequent patterns, the discovery of frequently recovered sub paths.
- Classification, the discovery of behavior rules, aimed at explaining the behavior of current patients.
- Clustering moving object trajectories, for example, requires finding out both:
  – A proper spatial granularity level (points, segments, lines, etc.).
  – A significant temporal sub domain (e.g., recovery “rush time “ might be informative for defining a clustering structure over recovery data.

References

- Jalel Akaichi. Mobility data modeling and analysis, Master sc. Course, Jalel Akaichi.
- J. Akaichi. Pervasive systems and cloud computing, Master sc. Course.
- Frederique laforest. Systèmes d'information pervasifs, Master sc. Course.
- Gerasimos D. Marketos. Data Warehousing & Mining Techniques for Moving Object Databases. PhD Thesis. 2009.
- Akaichi, J., Oueslati, W. (2011) 'Facial nerve stream modeling for facial paralysis patients states evolution', IEEE IRI, pp. 429-433.
- Oueslati, W., Akaichi, J. (2011) 'A Trajectory UML profile For Modeling Trajectory Data: A Mobile Hospital Use Case'. International Journal of Advanced Research in Computer Science.
- Oueslati, W., Akaichi, J. (2010) 'Pre-learning from mobile professionals' Trajectory'. International Conference on Computer Systems and Applications (AICCSA), IEEE/ACS, pp. 1-7.
- Oueslati, W., Akaichi, J. (2010) 'Mobile Information Collectors' Trajectory Data Warehouse Design', International Journal of Managing Information Technology (IJMIT) Vol.2, No.3, pp. 1-20.
- Oueslati, W., Akaichi, J. (2010) 'A Survey on Data Warehouse Evolution', International Journal of Database Management Systems (IJDMS), Vol.2, No.4, pp. 11-24.
- Akaichi, J., Elloumi, L. (2009) 'Bell's palsy Disease Recovery Surveillance: Facial Nerve Trajectory Data Modelling', Data Integration in the Life Science Workshops (DILS 2009), Manchester.
- Khayati, M., Akaichi, J. (2009) 'Pervasive Assistance System for Mobile Physicians on the Road', International Journal of Healthcare Technology Management (IJHTM), Volume 10, No. 1, pp. 82-100.
- Akaichi, J. (2009) 'Mobile Long Life Learner Pervasive Assistance System', Learning Technology, publication of IEEE Computer Society's, Technical Committee on Learning Technology (TCLT), Volume 11, Issue 3.

Thank you