3rd ACM International Workshop on **Services Integration in Pervasive Environments**
July 6-10, 2008, Sorrento, Italy

User-centric Inference Based on History of Context Data in Pervasive Environments

Nikos Kalatzis
National Technical University of Athens
Greece

Summary of presentation

- Concepts & Motivation
- Research objectives
- Context Management in Daidalos PSP
- Context prediction mechanism
- Evaluation and experiments
- Conclusions

Concepts & Motivation

- **Pervasive computing**
  - envisions a world where people are surrounded by intelligent & intuitive interfaces embedded in everyday objects around them. These IFs recognize and respond to the situation and behavior of an individual in a personalized manner, while considering the surrounding environment's conditions.
- **Context awareness (CA) and services**
  - CA provides pervasive computing environments with the ability to adapt the services they provide, by implicitly sensing and automatically deriving the users' needs from the context that surrounds them.
- **History of Context (HoC)**
  - the overall set of past context values; requirement for proactiveness

Pervasive computing systems need to be strongly proactive in order to minimize human-machine interaction.

Research Objectives

- We propose an innovative approach for modeling, storing and exploiting history-of-context (HoC), which supports the accurate prediction of future context information and the estimation of currently unknown/unavailable context values.
- The designed context prediction technique is based on the fact that repeated patterns are usually detected in human behavior.
- The main design objectives to achieve this include: type-independent context prediction, efficient context summarization, and minimisation of the required processing, time & storage resources.

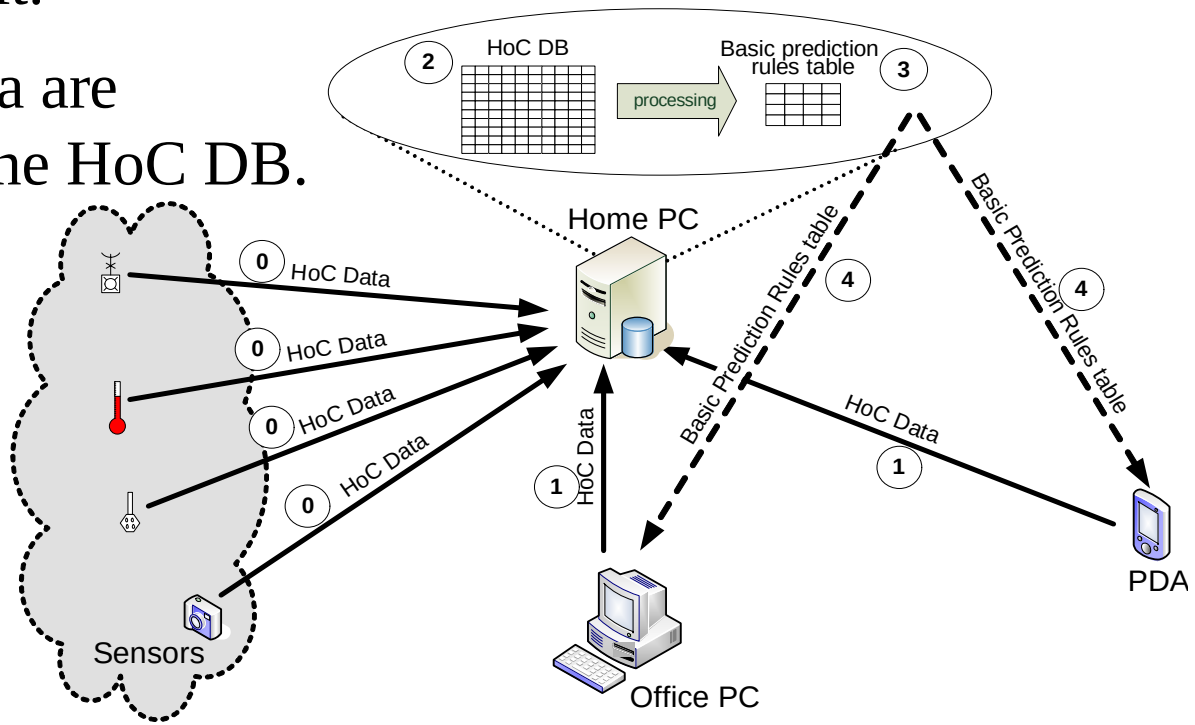
Context Management in the Daidalos PSP

- **Pervasive Service Platform (PSP):** Establishes the context awareness related functionality, coupled with features such as personalization, service discovery & composition, security & privacy.
- **Context Management system (CM):** Part of the PSP that provides a uniform way to access context information and offers to context sources interfaces enabling them to feed context data in the system.
- **Context Inference Engine (CIE):** Part of the CM that enhances the proactiveness of the system by enabling context prediction/estimation based on the history of context collected. The proposed HoC handling technique is implemented by this component.

Automatic initiation or configuration of services/devices is supported, depending on user habits & current context data.

History of Context lifecycle

- **Step 0 –** Various sensors synchronously disseminate context information to a resource rich system (Home PC).
- **Step 1 –** Other CM nodes also collect context information and asynchronously forward it.
- **Step 2 –** Past context data are processed and stored in the HoC DB.
- **Step 3 –** New Basic Prediction Rules (BPRs) are created.
- **Step 4 –** These BPRs are propagated to the CM nodes.



Context Data Acquisition, Abstraction & Caching

- **History of Context** (HoC): the overall set of past context values.
- Raw context data collected from various sensors are abstracted based on the Daidalos Context Model (DCM) and Context MAnagement oNTOlogy (COMANTO)
  - Abstracted raw data are represented by tags such as: “Home”, “Office” for location , “HomePc”, “Pda” for device, “Free”, “Busy” for activity
- Context types selected for monitoring and prediction in a human centric approach are: **Time**, **PersonID**, **Location**, **Device**, **Activity**.

Entity ID	Start Time	End Time	Location	Device	Activity
...person#471	18022007|080536	18022007|082253	Home	HomePc	Browsing
...person#471	18022007|082315	18022007|082817	Home	HomePc	Emailing
...person#471	18022007|083527	18022007|085401	KatehakiStr	CarPC	Driving
...person#471	18022007|085811	18022007|090935	Office	OfficePC	Typing
...person#471	18022007|091148	18022007|091824	Office	OfficePhone	Talking
...person#471	18022007|092012	18022007|100450	Office	OfficePC	Typing
...person#471	18022007|100618	18022007|100804	Office	PDA	Walking
...person#471	18022007|100932	18022007|101523	Office	Printer	Printing
...person#471	18022007|101702	18022007|104211	Office	OfficePc	Browsing

Example of HoC cached data

HoC Database

- A DB entry is a combination of context values (entity, location, device, activity).
- The HoC DB is structured based on 30-minute time frames (TFs).
  - 48 TFs per day each assigned with none, one, or multiple entries.
- Each entry is associated with a daily calculated time depended score:

a) This score is increased by 1 if the exact same context situation is observed.

b) In order to expresses the aging rate of entry occurrences score is decreased by: $S = S_0 \lambda^{\kappa-1}$

S_0: current score

λ : daily attenuation factor $0 < \lambda < 1$

k : number of days (since the context combination observed or the last calculation of score)

- Score values range: $0 \leq S \leq \frac{1}{1-\lambda}$
- Entries with low score are removed

TF Index	entityID-location-device-activity	Score
0	…person#471-Home-NONE-Sleeping	77.74255339
	…person#471-Home-TV-WatchingTV	13.32932862
	…person#471-Pub-PDA-Drinking	4.866635658
	…person#471-Restaurant-PDA-Eating	2.491089311
1	…person#471-Home-NONE-Sleeping	79.46278259
	…person#471-Home-TV-WatchingTV	15.57818485
	…person#471-Pub-PDA-Drinking	2.002311926

Context Prediction (I)

- Context prediction is based on **Basic Prediction Rules** (BPRs).
  - A rule expresses a group of context data occurring simultaneously.
  - Rules are extracted from the HoC Database.
- Given the time and the current context values, the missing context is estimated. Estimated value is accompanied by probability of occurrence value.

Example of BPR table segment of the HoC DB

TF Index	location-device-activity	Probability
10	Home-NONE-Sleeping	0.937
11	Home-NONE-Sleeping	0.892
12	Home-NONE-Sleeping	0.821
13	Home-NONE-Sleeping	0.748
14	Home-NONE-Sleeping	0.693
15	Home-NONE-Showering	0.658
16	Home-NONE-GettingDressed	0.642
17	Home-HomePC-Browsing	0.687
18	KatehakiStr-CarPC-Driving	0.849
19	Office-OfficePC-Typing	0.876
20	Office-OfficePhone-Talking	0.669

Context Prediction (II)

- The BPRs are generated based on **Rule Generation Models** (RGM). Examples of RGMs:
  - ***Max score model***: the BPRs reflect the entry demonstrating the maximum score in each time frame.
  - ***Minimum score threshold model***: the BPRs generated reflect the entries demonstrating score above a predefined threshold.
  - ***Minimum score difference threshold model***: the BPRs are extracted based on the minimum score difference of entries in the same time frame.
- Probability of occurrence is related with the calculated scores:

  S_{ij} :score of an entry in a timeframe

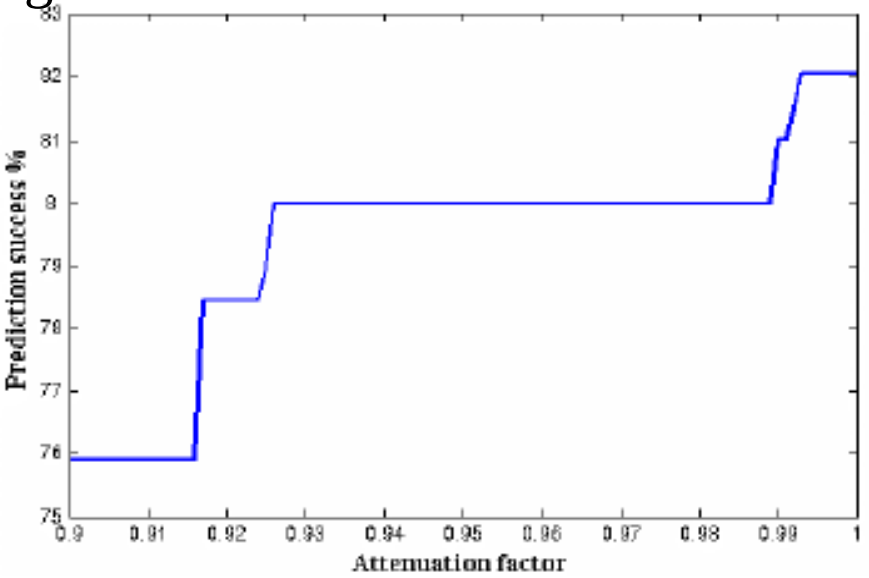
  J_i : the set of all context data combinations in TF_i.

$$P_{ij} \approx \frac{S_{ij}}{\sum_{j \in J_i} S_{ij}}$$

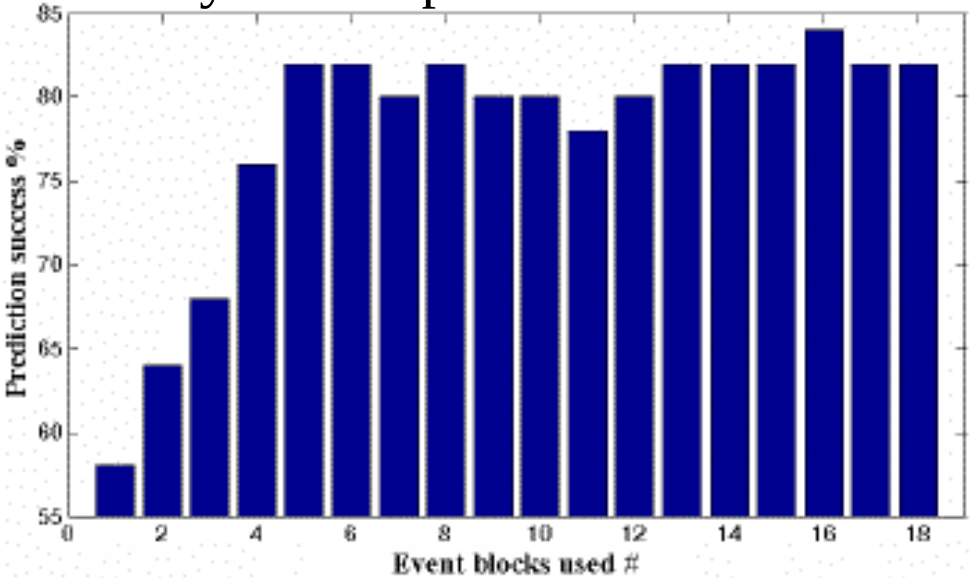
- In case the BPR table contains more than one entry in each time frame, user or application decides on which one to use.

Experiments and Evaluation (I)

The proposed methodology has been validated based on the **Max Score Model** against a real location-based dataset collected by mobile phone users.



User location prediction success ratio over varying attenuation factors. Max prediction success (82%) for λ>0.993.

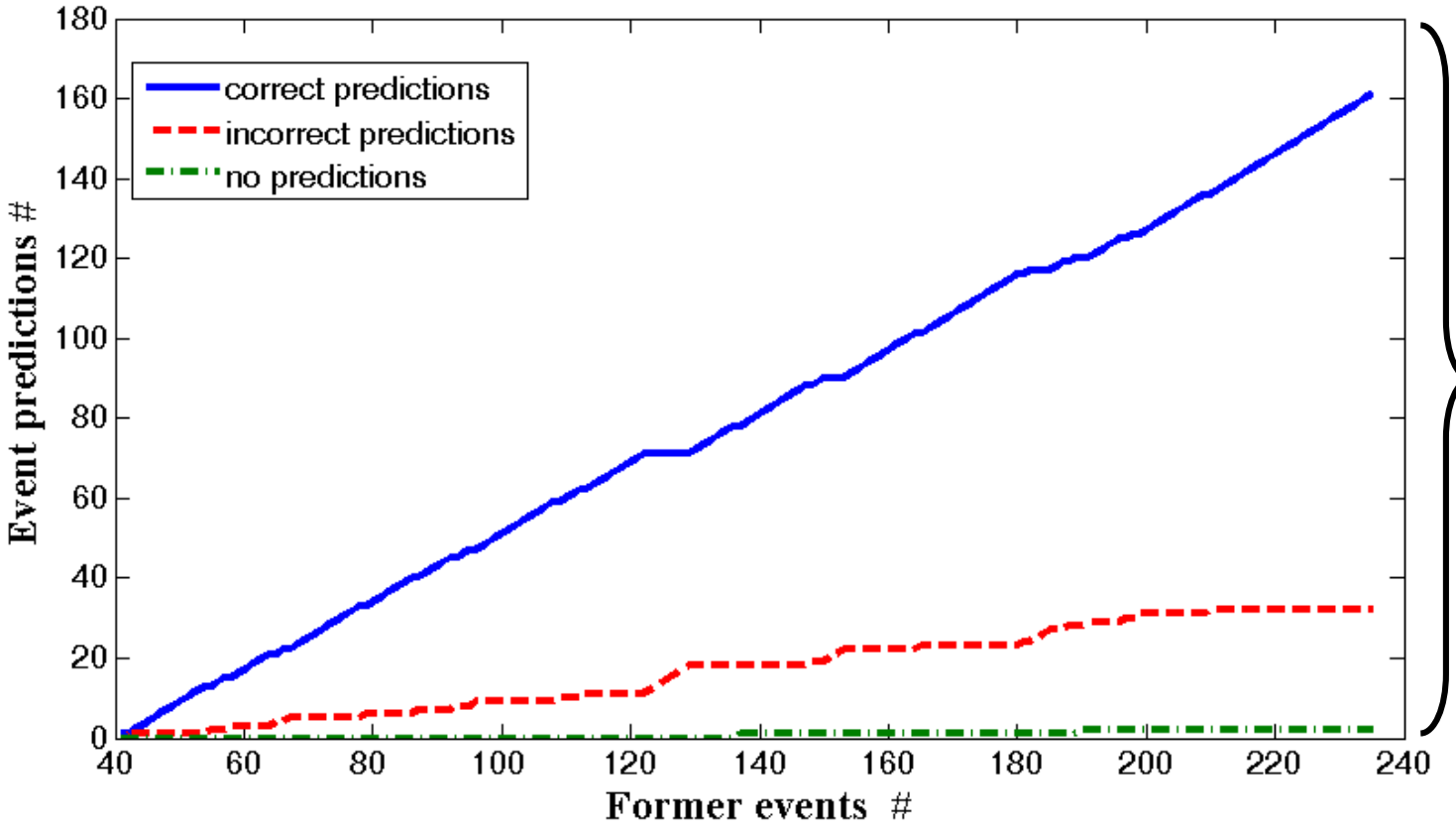


User location prediction success ratio over varying volume of context data used to create the BPRs.

Average prediction success > 80% → observed when ~28% (~3 weeks) of the available past context data are used for training.

Experiments and Evaluation (II)

- Best results achieved when the system is continuously trained with the new context data observed.
- Average prediction success → **83.4%**



Number of event predictions (correct/incorrect/ impossible) over the number of former events available for learning

Major Outcomes, Results, Future Plans

Main advantages of the described approach:

- Type-independent context prediction
  - Methodology is applicable to any type of context data
  - Prediction success ratio is ~**84%** and is estimated to further increase
- Efficient context summarization
  - Volume of processed historical context data stored is less than **6%** of the volume of the HoC data originally collected
- Minimal required processing, time and storage resources
  - Proposed mechanism ideal for resource-limited devices, such as PDAs and mobile phones

On going research & future plans:

- Extend approach to be based on both periodical and sequential patterns during the prediction rule generation process + day of the week information
- Comparison of performance of the four RGMs + automatic identification of most appropriate RGM considering system capabilities and application requirements

Thank you for your attention!

Email: **nikosk@cn.ntua.gr**