Discovering temporal hidden contexts in web sessions for user trail prediction





Julia Kiseleva(Eindhoven University of Technology),Hoang Thanh Lam(Eindhoven University of Technology),Mykola Pechenizkiy(Eindhoven University of Technology),Toon Calders(Université Libre de Bruxelles)

User Trail Prediction

- User web log: $S = \{s_1, s_2, ..., s_n\}$
- Alphabet of user actions: $A = \{a_1, a_2, ..., a_k\}$



- Can we find contexts helping to build more accurate local models (performing better than a global model)?
- Can we use the same problem formulation to discover some meaningful contexts?

User Web Log

Alphabet of actions: {Search, Refine Search, Click on Banner, Product View, Payment}

Session 1	Search	Refine Search	Click on Banner	Product View	Payment
Session 3	Product View	Payment			
Session 3	Search	Refine Search	Refine Search	Click on Banner	
Session 4	Search	Refine Search	Click on Banner	Product View	Payment
Session 5	Product View	Click on Banner			

User Navigation Graph



- Approaches to create local models:
 - o Horizontal partitions (context is geographical position)

Session 1	Search	Refine Search	Click on Banner	Product View	Payment	Users
Session 3	Product View	Payment				from Europe
Session 3	Search	Refine Search	Refine Search	Click on Banner		
Session 4	Search	Refine Search	Click on Banner	Product View	Payment	Users from
Session 5	Product View	Click on Banner				South America

- Where user location is context
- Continents are contextual categories C:
- C1= "Europe", C2 = "South America"

Contextual Local Models

Users from Europe:





Users from South America:



- Approaches to create local models:
 - o Horizontal partition
 - o Vertical partition :
 - Two types of behavior:
 - Ready to buy (Product View, Payment)

Session 1	Search	Refine Search	Click on Banner	Product View	Payment
Session 2	Product View	Payment			
Session 3	Search	Refine Search	Refine Search	Click on Banner	
Session 4	Search	Refine Search	Click on Banner	Product View	Payment
Session 5	Product View	Click on Banner			

- Approaches to create local models:
 - o Horizontal partition
 - o Vertical partition:
 - Two types of behavior:
 - Ready to by (Product View, Payment)
 - o Just browsing (Search, Refine Search, Click on Banner)

Session 1	Search	Refine Search	Click on Banner	Product View	Payment
Session 2	Product View	Payment			
Session 3	Search	Refine Search	Refine Search	Click on Banner	
Session 4	Search	Refine Search	Click on Banner	Product View	Payment
Session 5	Product View	Click on Banner			

- Approaches to create local models:
 - o Horizontal partition
 - o Vertical partition:
 - Two types of behavior:
 - Ready to buy (Product View, Payment)
 - o Just browsing (Search, Refine Search, Click on Banner)

Session 1	Search	Refine Search	Click on Banner	Product View	Payment
Session 2	Product View	Payment			
Session 3	Search	Refine Search	Refine Search	Click on Banner	
Session 4	Search	Refine Search	Click on Banner	Product View	Payment
Session 5	Product View	Click on Banner			

Vertical Partitioning



Context-Awareness Example



Context-Awareness: (G,H)

Context-Awareness



Optimization problem

$$F_{c_i}(T_{validation_i}, M_i) = \sum_{a_j \in T_{validation_i}} F(a_j, M_i)$$

$$EF(E) = \sum_{i=1}^{h} F_{c_i}(T_{validation_i}, M_i)$$

$$EF^* = \underset{E}{\operatorname{arg\,max}} EF(T_{validation}, M))$$



ababababcdcdababcdcdcd

The number of true predictions = 0





ababababcdcdababcdcdcd

The number of true predictions = 12





ababababcdcdababcdcdcd

The number of true predictions = 20





ababababcdcdababcdcdcd

The number of true predictions = 20



Stop as long as no additional prediction benefit of merging



Mastersportal.eu -Homepage

The European Study Choice Platform Masters	Contact Advertising Register
Start your search for Masters Search the largest database in Europe for Masters Any discipline Any subdiscipline Any subdiscipline Find & company 2 0 6 2 Masters	sters here! unting SEARCH sters in Europe!
ITI Home	spotnight
Banner Senglish Click Advertise here! Abord UNIVERSITY Advertise here!	Universities in the spotlight International Hellenic University INTERNATIONAL HELLENIC UNIVERSITY The International Hellenic University (IHU) is the first Greek state University where the teaching and student administration of its More information

Mastersportal.eu - Search

Search	Try e.g. Business adn	ninistration		SEARCH	
I Home Search Results					
Refine your search Discipline	United Kingdom (2334) Germany (509) France (486) Show all countries	<u>Netherlands</u> (378) <u>Spain</u> (239) <u>Ireland</u> (164)	<u>Sweden</u> (163) <u>Switzerland</u> (153) <u>Italy</u> (142)	<u>Austria</u> (115)	Search our DistanceLearningPortal
Business & Economics 🔹					
Any subdiscipline 🔹	You have searched to	r a Business & Econo	omics Master's study.		
Language of tuition	Sort by: Relevance		- 10	1 ² Click Progr	on am is
Any language	M.Sc. Health Econ Barcelona Graduate School Start date: Sep 2013	omics and Policy of Economics Barcelona, Tuition Fee:	Spain € 2,000 per year + Dur	Searc	h Result
Location Any country	barcelona gse	There is no do indicators of a a more intens decades.	bt that the health status of th hation's wellbeing. In terms of regulatory activity and reform	ne population is one of the main of policy, it is hard to find a sector m, and more so in the last two	rwith
Filte	Tick on			Detailed Description	

Dataset



Period	May 2012
#sessions	350.618
#requests	1.775.711

Results





In theory:

- Ideal local Markov models will be at most as good as an ideal global Markov model

In practice:

- We observe that local models become more accurate



Possible explanation:

- 1. Users' behavior doesn't follow Markov property
- 2. "Ambiguous" transition probabilities. Local models can get better estimates

Further study of this phenomenon is required

Explaining Results





Resulted Clusters

Id	Summary	Cluster
1	Intensive Search	Basic Search, Refine Search, Empty Search Result
2	Explore information related to program	Program impression in search result, Banner click, Program click , Click on university link
3	Start of browsing	University Spotlight impression, Quick search
4	Explore country information	File view, Click on country link
5	Explore search result	Program Impression in related programs, University impression on nearby universities
6	Explore program	Program in landing page, Submit inquiry
7	Outlier	Submit question, X-node

Site Map

- Page is represented as set of possible actions
 - Example: Homepage is (Quick Search, University Spot light impression, Question Submit)
 - o Calculate Jaccard similarity between Page and Cluster

Site map



Conclusion

- We formulate the problem of temporal context discovery as an optimization problem.
- A hierarchical clustering method is proposed to determine the optimal number of hidden contexts and mine temporal contexts.
- We show a real-world use case in which the contexts as we defined them do exist and are useful for prediction.

Future Works

- Considering multidimensional contextual features (i.e. sessions and activities live in vector space).
 - o Strategies for vertical and horizontal partitioning
- Introducing a mechanism for detecting/predicting context switching within a web-session.
- Testing our method on more datasets.

Thank You!

- Context identification and integration it into prediction models
- Accurately predicting users' desired actions and understanding behavioral patterns of users in various webapplications
- Personalization and adaptation to diverse customer need and preferences
- Accounting for the practical needs within the considered application areas.









Conclusion

- We formulate the problem of temporal context discovery as an optimization problem.
- A hierarchical clustering method is proposed to determine the optimal number of hidden contexts and mine temporal contexts.
- We show a real-world use case in which the contexts as we defined them do exist and are useful for prediction.

