WEBKDD2000-WebMiningforE-Commerce

MyraSpiliopoulou
Otto-von-GuerickeUniversitaet.
Universitaetsplatz2,D-39016
Magdeburg
myra@iti.cs.uni-magdeburg.de

JaideepSrivastava
UniversityofMinnesota&Yodlee,Inc
3600BridgeParkway,Suite200
RedwoodCity,CA94065
srivasta@cs.umn.edu

RonKohavi
BlueMartiniSoftware
2600CampusDrive,SanMateo,
California94403
ronnyk@bluemartini.com

BrijM.Masand
RedwoodInvestmentSystems
76SummerStreet
Boston,MA02110-1225USA
brij@redwood.com

ABSTRACT

Inthispaper,weprovideasummaryoftheWEBKDD2   000
workshop,whosethemewas'WebMiningforE-Commerc   e'.
ThisworkshopwasheldinconjunctionwiththeACM   SIGKDD
InternationalConferenceonKnowledgeDiscoveryin   Databases
(KDD-2000).

Keywords

Webmining,e-commerce,personalization,clickstrea   manalysis.

1. THEME

The ease and speed with which business transactions   can be
carriedoutovertheWebhasbeenakeydrivingfor   ceintherapid
growth of electronic commerce. In addition, custome   r
interactions, including personalized content, e-mai   l campaigns,
andonlinefeedbackprovidenewchannelsofcommuni   cationthat
werenotpreviouslyavailableorwhereveryineffic   ient. TheWeb
is revolutionizing the way businesses interact with   each other
(B2B) and with each customer (B2C). It has introduc   ed entirely
newwaysofdoingcommerce,includinge.g.auctions   andreverse
auctions. It also made it imperative for organizati   ons and
companiestooptimizetheirelectronicbusiness.

Knowledge about the customer is fundamental for the
establishmentofviablee-commercesolutions.Webm   iningfore-
commerceistheapplicationofwebminingtechnique   stoacquire
thisknowledgefore-commerce.Typicalconcernsin   e-commerce
includeimproved cross-sells, up-sells, personalize   d ads, targeted
assortments,improvedconversionrates,andmeasure   mentsofthe
effectivenessofactions.

The WEBKDD 2000 workshop is the second workshop hel   d in
conjunctionwiththeACMSIGKDDInternationalConfe   renceon
KnowledgeDiscoveryinDatabases(KDD)anddedicate   dtothe
challengesofwebmining.WEBKDD'99focusedonthe   aspectsof
web mining related to user profiling; the long vers   ion of its
proceedings has appeared as volume 1836 of the Lect   ures in
Artificial Intelligence series (LNAI) by Springer V   erlag. In
response to call for papers,   WEBKDD 2000 received 31
contributions. Each was reviewed by at least three   program
committee members. Seven submissions were selected   for
presentationaslongpapers,andsixasshortpaper   sreportingon
goodideasataratherpreliminaryphase.

The URL   http://robotics.stanford.edu/~ronnyk/WEBKDD2000
contains the final versions of the workshop papers   and the slide
presentations.

2. WORKSHOP

The KDD community responded very enthusiastically t   o the
WEBKDD2000workshop,andwereceivedfarmorerequ   estsfor
attendance (approximately 110) than there was space   . About 85
people attended the workshop, which brought togethe   r e-
commercepractitioners,toolvendorsanddataminin   gresearchers.
Thepaperpresentationwasdividedintothreesessi   ons.

The first session, titled   *Web personalization and recommender
systems*, focused on how web mining can address one of the
fundamental issues of B2C e-commerce, namely person   alized
customer experience. As often described by Jeff Bez   os, CEO of
Amazon.com, and mentioned by Joseph Pine in his The
Experience Economy [1], customer experience is the   key to
building customer loyalty to an on-line store, sinc   e leaving the
storeisexactlyoneclickaway.Inthissessionwe   hadthreelong
andtwoshortpapers,whichpresentedtheleadinge   dgeideasin
this important area. Mobasher, Dai, Luo, Nakagawa,   Sun, and
Wiltshire's paper, titled   *Discovery of Aggregate Usage Profiles
forWebPersonalization* ,describedhowusagedatafromweblogs
canbeanalyzed/minedtobbuilduserprofiles,andh   owthesecould
be use to enhance the user's browsing experience. V   ucetic and
Obradovic's paper, titled   *A Regression-Based Approach for
Scaling-Up Personalized Recommender Systems in E-co   mmerce*,
presented an approach to applying regression techni   ques to
understand user preferences for recommender systems   . This
approach is interesting since statistical technique   shave not been
appliedsufficientlytothisproblem.Sarwar,Karyp   is,Konstanand
Riedl's paper, titled   *Application of Dimensionality Reduction in
Recommender Systems – A Case Study*   , presented a novel
applicationofdimensionalityreductiontechniques   fromscientific
computingtotherecommendationsystemproblem.The   twoshort
papers in this session, namely Lin, Alvarez and Rui   z's
*Collaborative Recommendation via Adaptive Associati   on Rule
Mining*andChangandYuan's   *ASynthesizedLearningApproach
for Web-Based CRM* , presented early results in alternative
approachestotherecommendationsystemproblem.In   summary,
the wealth of interest in applying various techniqu   es to the
recommendation system problem shows the centrality   of this
problemtoWebpersonalization.

The second session, titled *Mining frameworks and case studies*, presented experience reports from four groups on us ing web mining in various e-commerce applications. Any fiel d of enquiry must have its 'proof of the pudding is in eating it ' component, and the case studies presented here provide exactly tha t flavor. Ansari, Kohavi, Mason and Zheng's paper, titled *Integrating E-Commerce and Data Mining: Architecture and Challeng es*, provided a comprehensive overview of the issues in applying data mining techniques to E-commerce. The authors bring a lot of real world perspective from their experience at BlueMar tini Software, especially from the viewpoint of an E-commerce solu tion provider. Theusinger and Huber's *Analyzing the footsteps of your customers, A case study by ASK|net and SAS Institut e GmbH* presented the experience of applying data mining in E-commerce from a solutions perspective, where SAS's tools wer e used to solve a problem for ASK|net. Sanford Gayle, in his paper *The Marriage of Market Basket Analysis to Predictive Mo deling: The Essential Challenge in Exploiting Web-Log Files for Prediction*, presented an approach to using association and corr elation analysis to extract predictive models from web logs . Coenen, Swinnen, Vanhoof and Wets' paper *A Framework for Self Adaptive Websites: Tactical versus Strategic Challe nge* examined the various issues in building such sites. The esse ntial tension seems to be between making a website personalized to individual users - maybe even dynamically change it based on t he particular user's behavior – and the information overload it c an cause.

The third session, titled *Navigation analysis*, focused on how clickstream data can be analyzed to extract valuabl e e-commerce knowledge from it. Being able to analyze clickstrea m data provides an unprecedented opportunity to understand in detail the process leading up to a buy/not buy decision vs. ju st recording the final outcome - as is the case with point-of-saled ata. Clickstream data is over 95%+ of all data collected in most lar ge-scale e-commerce environments, and contains a wealth of kno wledge embedded in it. Berendt's paper, titled *Web Usage Mining, Site Semantics, and the Support of Navigation,* provided a general overview of the issues in clickstream analysis, and how the mined

knowledge can be used for supporting site navigatio n. Kato, Nakayama and Yamane's paper *Navigation Analysis Tool based on the Correlation between Content Distribution and Access Patterns,* presented an approach whereby mined patterns from ite content can be correlated with mined patterns from site usage, and a tool based on this approach. Investigating applic ations of such a tool would be an interesting line of inquiry. Gaul and Schmidt-Thieme's *Mining Web Navigation Path Fragments* presented some novel algorithms for extracting navigational p ath fragments. Finally, Tan and Kumar's *Modeling of Web Robot Navigational Patterns* addressed the challenging and commercially importa nt problem of separating the site visits of web robots from humans. This is crucial for at least two applications: (1) as competitive pressures increase, commerce sites would like to bl ock robots that collect sensitive information, and (2) accurate mod eling of human users' e-commerce behavior requires that web robot accesses be filtered out. While the papers in this session pres ent some of the leading ideas, the research in this area is just be ginning and we have barely scratched the surface.

3. CONCLUSION

WEBKDD 2000 turned out to be a very successful work shop by all measures. More than 110 people showed interest in the workshop and over 85 attended it. The quality of pa pers was excellent, the discussion was lively, and a number of interesting directions of research were identified. This is a s trong endorsement of the level of interest in this rapidl y emerging field of inquiry.

4. REFERENCES

[1] B. Joseph Pine, James H. Gilmore, B. Joseph Pin e II. The Experience Economy. Harvard Business School Pr; ISB N: 0875848192, http://www.amazon.com/exec/obidos/ASIN/0875848192/ref =sc_b_1/103-2009916-9046229