

PatentMiner: Topic-dirven Patent Analysis and Mining

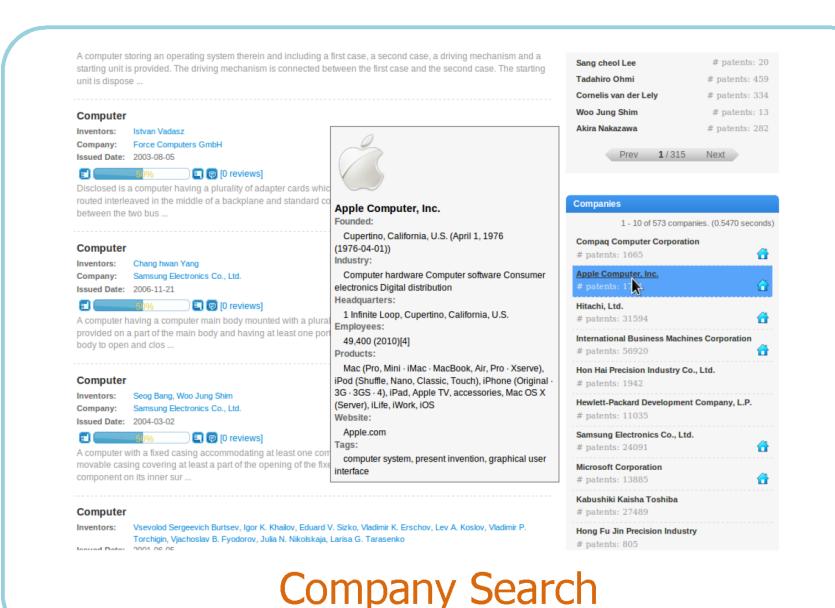
Jie Tang, BoWang, Yang Yang, Po Hu, Yanting Zhao, Xinyu Yan, Bo Gao, Minlie Huang, Peng Xu, Weichang Li, and Adam K. Usadi **Department of Computer Science, Tsinghua University, China ExxonMobil Research and Engineering Company, USA**

Introduction

PatentMiner is a free online service used for analyzing and mining patent networking data. By now, we collected 8,000,000+ patents, 400,000+ companies, 2,000,000+ inventors. PatentMiner provides in-depth topic-level analysis

functions.

System Overview



Major functions in PatentMiner: □ Patent Search

Find prolific inventors, top company, and best patent

Company Analysis

Extract company basic information, and discover topic evolution

Topic Cataloging

Model patents and companies with mixture topic distribution

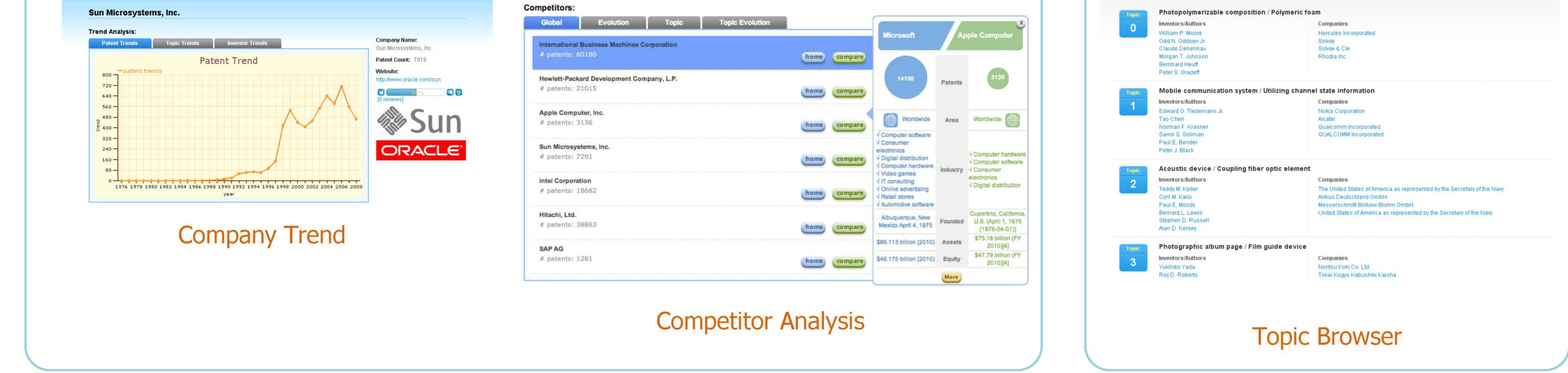
Competitive Analysis

Identify your competitors and analyze their technology evolution

	ary of " data mining" :						
Result Cat	egories (U.S. Class)					Clo	se
	to focus, right click to return.						
			ł	All results			
	Mining or in s	situ disintegration of hard i	material		Hydraulic and	d earth engineering	
Hard material disintegrating machines Processes		Processes	Cutter tooth or Automatic control; tooth head signaling or indicating		Earth trea	Earth treatment or control	
				In situ conversion of solid to fluid			11 16 13 13 14 19 11 11 12 20 10 24
Patent Fou	nd for "data mining"				Inventors	# query related / a	
Order by	relevance date authority	1	Found 3	211 patents, used time 0.012	seconds William R. Ker	nnedy # patents: 38/4	16
					John M. Kenn	edy # patents: 37/4	13
	e data mining framework				Gerhard Merte	en # patents: 23/6	i4
Inventors:	Raman S. Iyer, Ioan Bogdan Criv Oveson Rong J. Guan Zhaohlu				Valter Weirich	# patents: 21/4	19
Company:	Oveson, Rong J. Guan, ZhaoHui Tang, Pyungchul Kim, Irina G. Gorbach Company: Microsoft Corporation			Rakesh Agraw	val # patents: 21/17	/1	
Issued Date: 2008-06-03					John C. Stank	cus # patents: 19/5	51
	disclosure pertains to extensible	* *			Nunipert Deck	ker # patents: 17/2	22
mining system is disclosed that supports plug-in or integration of non-native mining algorithms, perhaps provided by third parties, such			Raymond L. V	Wright # patents: 17/2	20		
					Maurice K. Lei	Begue # patents: 16/3	10
Patient da	ta mining				James J. Fallo	on # patents: 14/2	24
Inventors: Company: Issued Date:	R. Bharat Rao, Sathyakama San Stefan Niculescu, Arun Kumar G Siemens Medical Solutions USA, 2000 11, 10	oel, Thomas R. Warrick	s, Radu 🧯	1 - 69% - 9 6 P) reviews] Pre	ev 1/451 Next	
		g framework for mining hi	ah-quality strue	ctured clinical information	The		
The present invention provides a data mining framework for mining high-quality structured clinical information. The data mining framework includes a data miner that mines medical information from a computerized patient record (CPR)						# query related / a	10
based on do	main-sp				International i # patents: 1	Business Machines Corporation 52/60180	9
Clustering	module for data mining					Eisenhutte Westfalia	
Inventors:	Marcos M. Campos		6) <u>73%</u> () () ()) reviews] # patents: 1		
Company: Issued Date:	Oracle International Corporation				The United St the Secretary	ates of America as represented by of the Navy	
		roaram product for parfor	nina clusterina	based data mining that in	di natantas d		2
A system, software module, and computer program product for performing clustering based data mining that improved performance in model building, good integration with the various databases throughout the enterprise, flexible							

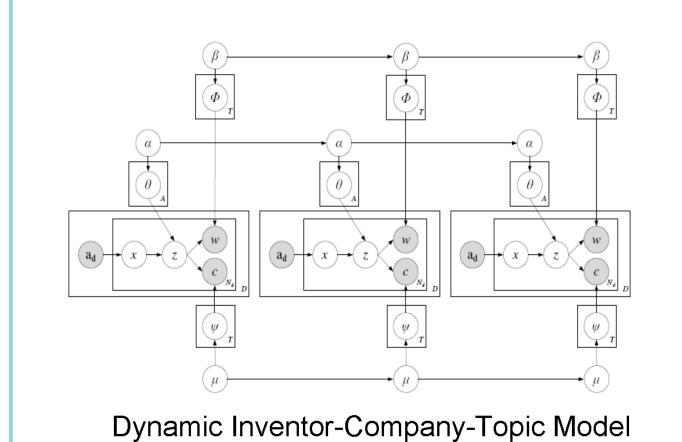
ate	ntMiner				
	Photopolymerizable composition / Polymeric foam				
Topic	Photopolymerizable composition	1 / Polymeric roam			
0	Inventors/Authors	Companies			
	William P. Moore	Hercules Incorporated			
	Odd N. Oddsen Jr.	Solvay			
	Claude Dehennau	Solvay & Cie			
	Morgan T. Johnson	Rhodia Inc.			
	-				
	Bernhard Heuft				

entMine	r			
Sun Microsyster	ns Inc			
Trend Analysis:	ns, mc.			
Patent Trends	Topic Trends	Inventor Trends	Company Name: Sun Microsystems, Inc.	
	Pat	Patent Count: 7019		
800 - patent tre	nds		Website: http://www.oracle.com/sun	



Technique Issue

Modeling Patent Network

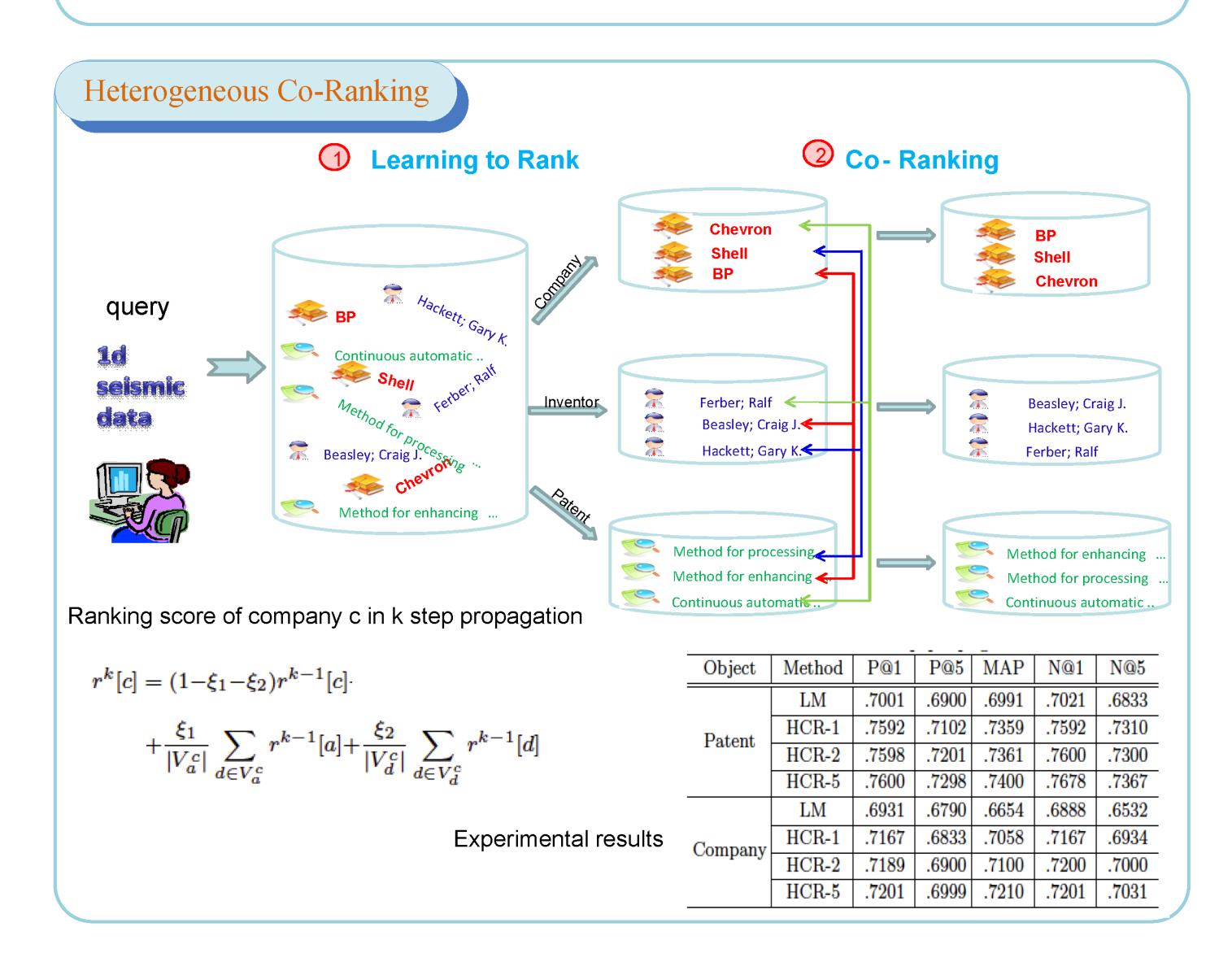


Inventor-topic smoothing Company-topic smoothing $\Omega_1 = \sum_z (\theta_{az}^t - \theta_{az}^{t-1})^2$ $\Omega_2 = \sum_{z} (\psi_{cz}^t - \psi_{cz}^{t-1})^2$ Topic smoothing $\Omega_3 = \sum_{z} (P(z)^t - P(z)^{t-1})^2$ Objective function $\mathcal{O}(\mathbf{D}) = -\mathcal{L}(\mathbf{D}) + \gamma_1 \Omega_1 + \gamma_2 \Omega_2 + \gamma_3 \Omega_3$ $\mathcal{L}(\mathbf{D}) = P(\mathbf{x}, \mathbf{z}, \mathbf{w}, \mathbf{c} | \Theta, \Phi, \Psi, \mathbf{a}) =$ $\prod_{d=1}^{M} \prod_{i=1}^{N_d} \frac{1}{A_d} \times \prod_{z=1}^{K} \left(\prod_{x=1}^{A} \theta_{xz}^{m_{xz}} \prod_{j=1}^{W} \phi_{zw_j}^{n_{zw_j}} \prod_{c=1}^{C} \psi_{zc}^{n_{zc}} \right)$

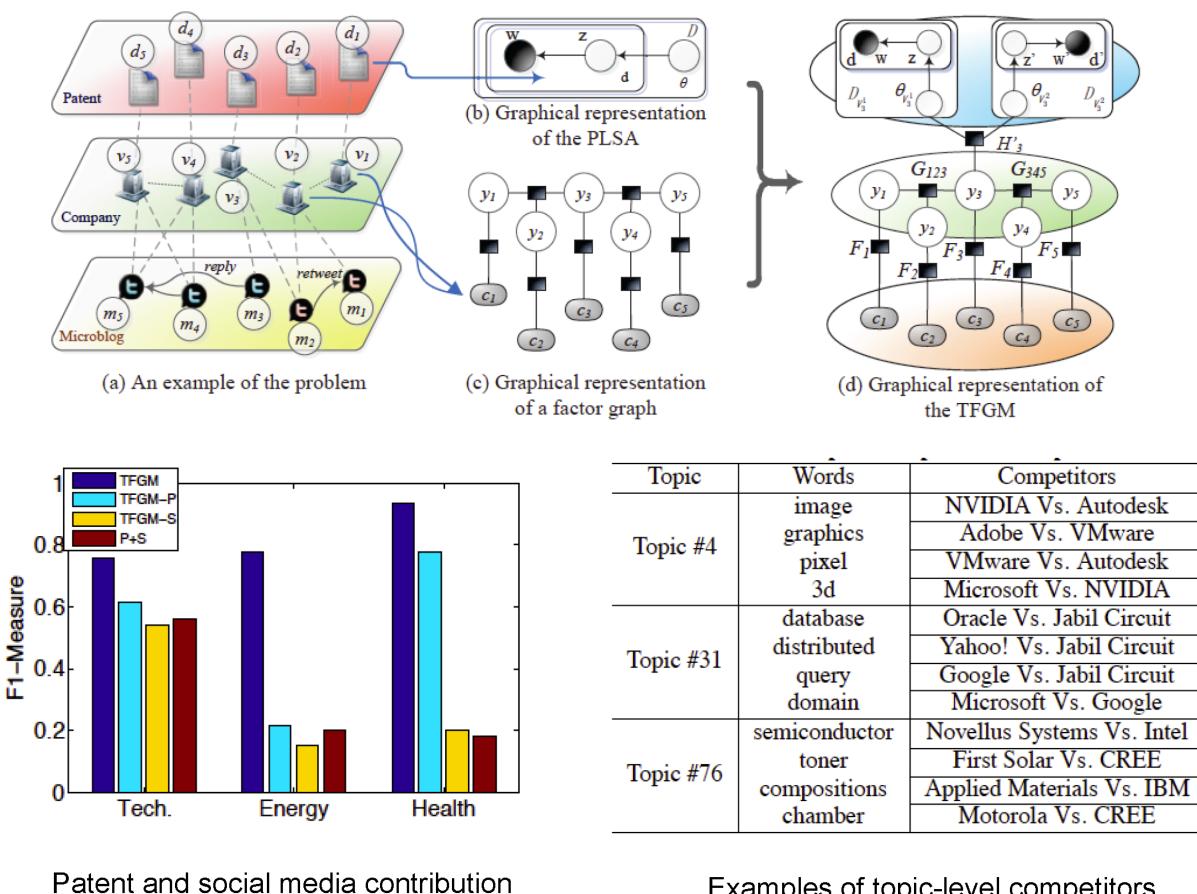
Competitor Analysis	
Method 1: topic comparison based on DICT	
Global Competitor Analysis	<u> </u>
$S(c, c') = \left(\sum_{i=1}^{n} p(z_i c)p(c' z_i)\right)^2 + \eta(\mathcal{D}_c - \mathcal{D}_{c'})^2$	Global
Topic-level Competitor Analysis $S(c, c', z) = (p(c z) - p(c' z))^2$	Topic

-						
	Methods	P@1	P@5	MAP	N@1	N@5
	WBS	.2009	.1087	.2904	.2009	.2841
Global	TopCom+TBD	.1731	.0846	.3078	.1731	.2871
	TopCom+PBC	.2098	.1161	.2920	.2098	.3085
Topic	LM+LDA	.1536	.1221	.2643	.1536	.2524
	TopCom+DBC	.1369	.1270	.2388	.1469	.2446
	TopCom+HBC	.1620	.1366	.2781	.1620	.2874

Experimental results



Method 2: combing patent records and social medias



Examples of topic-level competitors