

Supplementary Material

Anonymous Author(s)

No Institute Given

1 Proofs

Theorem 1. *The following statements hold:*

1. MG is connected.
2. $P(L) = P(L_{\text{MG}})$
3. Consider for each peak $v \in P(L) \setminus \{\max(L)\}$ of L the set $N'_{\text{MG}}(v) := \{u \in N_{\text{MG}}(v) \mid \exists v' \in N_{\text{MG}}(u) : h(v') > h(v)\}$. Then:

$$u \in N'_{\text{MG}}(v) \Rightarrow \exists u' \in C_L(v) : h(u') \geq h(u).$$

4. It holds for $v \in P(L) \setminus \{\max(L)\}$ that:

$$\text{prom}_L(v) = \min_{u \in N'_{\text{MG}}(v)} (h(v) - h(u)).$$

Proof. 1) Let $u \in V_{\text{MG}} \setminus \{\max(L)\}$. We show that u is connected to $\max(L)$. As every $u' \in K(L)$ is connected to some $\hat{u} \in P(L)$ in G_{MG} , it is sufficient to consider the case that $u \in P(L)$. For each dominator $u_1 \in D_L(u)$ of u it holds that $h(u_1) > h(u)$ and u_1 is connected to u in G_{MG} . Analogously, if $u_1 \neq \max(L)$, there exists u_2 with $h(u_2) > h(u_1) > h(u)$ such that u_2 is connected to u_1 and thus to u . Repeating this argument as long as the last element has a dominator leads to chains (u, u_1, \dots, u_m) where all elements are connected and $h(u_{i-1}) < h(u_i)$. As V_{MG} is finite, there must be an $m \in \mathbb{N}$ such that $u_m = \max(L)$ holds.

2) “ \subseteq .” Let $v \in P(L)$ and $u \in N_{\text{MG}}(v)$. Then there exists a mountain path $p = (v_0, \dots, u, \dots, v_m)$ in L such that $v_0 = v$ or $v_m = v$ and such that $c(p) = u$. If $v_0 = v$ we have

$$h(v) \underset{v \in P(L)}{\geq} h(v_1) \geq h(c(p)) = h(u).$$

Analogously, if $v = v_m$, we have

$$h(v) \underset{v \in P(L)}{\geq} h(v_{m-1}) \geq h(c(p)) = h(u).$$

Since $u \in N_{\text{MG}}(v)$ was chosen arbitrary, it follows that $v \in P(L_{\text{MG}})$. “ \supseteq .” Let $v \notin P(L)$. We show $v \notin P(L_{\text{MG}})$. If $v \notin V_{\text{MG}}$, we are finished. Otherwise $v \in K(L)$ and hence it must exist $u \in P(L)$ such that $v \in K(u)$ and thus $\{u, v\} \in E_{\text{MG}}$. With an analog argumentation as above, we can see that $h(u) > h(v)$ and thus $v \notin P(L_{\text{MG}})$.

3) Let $u \in N'_{\text{MG}}(v)$. It either holds that $u \in K(v)$ or that there exists $\tilde{v} \in P(L)$ such that $u \in K(\tilde{v})$ and there is an ascending path $p = (\tilde{v}, \dots, u, \dots, v) \in \uparrow_L(v)$ with $c(p) = u$. In the first case, we are finished by choosing $u' = u$. Hence, we assume the existence of $\tilde{p} = (\tilde{v} = v_0, \dots, u = v_i, \dots, v = v_m) \in \uparrow_L(\tilde{v})$. Let now $v' \in N_{\text{MG}}(u)$ with $h(v') > h(v)$. Hence, there exists $v'' \in P(L)$ and a path $p'' = (u_0 = v'', \dots, u_j = u, \dots, u_l = v') \in \text{MG}$ with $c(p'') = u$ and 1. $p'' \in \uparrow_L(v'')$ or 2. $(u_l, \dots, u_0) \in \uparrow_L(v')$. Let $p''' := (v = v_m, \dots, v_i = u = u_j, u_{j+1}, \dots, u_l = v')$. If p''' is not a path, there can only be pairs of doubled vertices of the form (v_k, u_s) as p' and p'' are paths. Let s be maximal such that this is the case. Then, $\hat{p} := (v = v_m, v_{m-1}, \dots, v_k, u_{s+1}, \dots, u_l = v')$ is a path. If $\hat{p} \notin \uparrow_L(v)$, there must exist $n \in \{s+1, \dots, l-1\}$ with $u_n \in P(L)$ and $h(u_n) > h(v)$. Choose n minimal that this is the case and set $\hat{v} = u_n$. Then set $p := (v = v_m, \dots, v_k, u_{s+1}, \dots, u_n = \hat{v})$. If $\hat{p} \in \uparrow_L(v)$ instead simply set $p := \hat{p}$. In both cases, we get $p \in \uparrow_L(v)$. Let $u' := c(p)$. As either $u' \in p'$ or $u' \in p''$ and $c(p'') = c(p') = u$, it follows $h(u') \geq h(u)$.

4) “ \geq ”: Let $u \in K(v)$. It holds that

$$\text{prom}_L(v) = h(v) - h(u) \geq \min_{u' \in N'_{\text{MG}}(v)} (h(v) - h(u'))$$

“ \leq ”: Let $u \in N'_{\text{MG}}(v)$. Choose $u' \in C(v)$ with $h(u') \geq h(u)$. Hence:

$$\text{prom}_L(v) = \min_{\tilde{u} \in C(v)} h(v) - h(\tilde{u}) \leq h(v) - h(u') \leq h(v) - h(u).$$

Since u was chosen arbitrary from $N'_{\text{MG}}(v)$, it directly follows that $\text{prom}_L(v) \leq \min_{u \in N'_{\text{MG}}(v)} (h(v) - h(u))$.

Theorem 2. *If for each peak $v \in P(L) \setminus \{\max(L)\}$ the line parent is unique, then $\text{LP}(L)$ is a tree.*

Proof. Let $n := |V_{\text{MG}}|$. As shown for example by Deo [1], it is sufficient to show that $\text{LP}(L)$ is connected and has $n - 1$ edges.

Connected: We show, that each $u \in P(L)$ is connected to $\max(L)$. Let $\text{LP} : P(L) \setminus \{\max(L)\} \rightarrow P(L)$ map every other peak to its line parent. Since for each $u \in P(L)$ it holds that $h(\text{LP}(u)) > h(u)$ and since $P(L)$ is finite, there is $m \in \mathbb{N}$ with $\text{LP}^m(u) = \max(L)$. Let $v_0 := u$ and for $i \in \{1, \dots, n\}$ let $v_i := \text{LP}^i(u)$. It follows that $(u = v_0, v_1, \dots, v_m = \max(L))$ is a path from u to v .

Edges: It holds that $E_{\text{LP}} = \{\{u, \text{LP}(u)\} \mid u \in P(L) \setminus \{\max(L)\}\}$. Also, $\{u, \text{LP}(u)\} = \{w, \text{LP}(w)\} \Rightarrow u = w$ since $u < \text{LP}(u)$ and $v < \text{LP}(v)$. Thus, $|E_{\text{LP}}| = |\{\{u, \text{LP}(u)\} \mid u \in P(L) \setminus \{\max(L)\}\}| = |\{u \mid u \in P(L) \setminus \{\max(L)\}\}| = n - 1$.

Theorem 3 (Connectivity of relative neighborhood graph). *Let (M, d) be a finite metric space. Then $\text{RNG}(d)$ is connected.*

Proof. Assume that $\text{RNG}(d)$ is not connected. As M is finite we can choose an unconnected pair $(m_1, m_2) \in M \times M$ such that $d(m_1, m_2)$ is minimal. Since m_1, m_2 are not connected they are especially not adjacent in $\text{RNG}(d)$. Thus, we

find $m_3 \in M$ such that $d(m_1, m_3) < d(m_1, m_2)$ and $d(m_2, m_3) < d(m_1, m_2)$. As m_1, m_2 are chosen as the pair with the minimal distance among the unconnected pairs, the pairs (m_1, m_3) and (m_2, m_3) are both connected. Hence, we would have a connection from m_1 to m_2 . \nmid

Theorem 4 (RNG as Edge-Reduction). *Let $G = (V, E, w)$ be a connected, undirected and weighted graph and $d_{\text{SP}} : V \times V \rightarrow \mathbb{R}_{\geq 0}$ be the shortest path metric on G . Then it holds that $E_{\text{RNG}(d_{\text{SP}})} \subseteq E$.*

Proof. Let $u, v \in V$ such that $\{u, v\} \notin E$. We will show that $\{u, v\} \notin E_{\text{RNG}(d_{\text{SP}})}$. As u, v are not adjacent there must exist a shortest path $p = (u = v_0, \dots, v_n = v)$ with $n \geq 2$. As p is a shortest path we have $d(v_{i-1}, v_i) = w(v_{i-1}, v_i)$ for $i \in \{1, \dots, n\}$. We thus have $d(u, v) = \sum_{i=1}^n d(v_{i-1}, v_i) > d(v_0, v_1) = d(u, v_1)$. As $(v_1, \dots, v_n = v)$ is a shortest path between v_1 and v we also have $d(v_1, v) = \sum_{i=2}^n d(v_{i-1}, v_i) < \sum_{i=1}^n d(v_{i-1}, v_i) = d(u, v)$. It hence follows that $\{u, v\} \notin E_{\text{RNG}(d_{\text{SP}})}$.

2 Creation of Datasets

2.1 Twitter Data

The networks **Twitter>10K** and **Twitter>100K** are generated as follows. First, we count the number of followers for all Twitter users in the original network¹. Then, we extract the users with at least 10,000 followers and 100,000 followers, respectively. These users are the vertices. Two users are connected, if they have a common follower. For two users u_1, u_2 we weight their edge via

$$w(\{u_1, u_2\}) := 1 - \frac{|F(u_1) \cap F(u_2)|}{|F(u_1) \cup F(u_2)|},$$

where $F(u)$ is the set of followers of u . The height of u is given via $h(u) := |F(u)|$.

2.2 Co-Author Networks

We generate co-author networks for three conferences which are strongly related to Knowledge Discovery, namely the **ECML/PKDD**², the **KDD**³ and the **PAKDD**⁴. We use the dump from 2021-04-01 of the Semantic Scholar Research Corpus⁵ for this. To weaken the influence of old publications and social connections that not exist anymore, we only use publication with a “year” attribute between 2000 and 2020.

To identify the authors for a specific venue we identify authors with publications at the specific venue. To discard authors which are only weakly connected

¹ <https://snap.stanford.edu/data/twitter-2010.html>

² <https://2023.ecmlpkdd.org/>

³ <https://kdd.org/kdd2023/>

⁴ <https://pakdd2023.org/>

⁵ <https://api.semanticscholar.org/corpus>

to the community, we only choose authors which are co-author of at least 2 publications at the specific venue. We generate an edge between two authors if they have at least one joint publication. Here, the edges between a_1, a_2 are weighted via

$$w(\{u_1, u_2\}) := 1 - \frac{|P(a_1) \cap P(a_2)|}{|P(a_1) \cup P(a_2)|},$$

where $P(a)$ is the set of publications co-authored by a between 2000 and 2020. As the height function we choose the h-index. For this, we again only consider publications and citations which are dated between the years 2000 and 2020.

2.3 Connectivity and Vanishing Distances

When generating the above networks, two problems occur:

Connectivity. The networks will probably not be connected. To solve this issue, we restrict them to the biggest connected component with respect to the amount of vertices.

Vanishing Edges. It may happen that an edge has weight 0 which results in the shortest path distance not being a metric. This happens for example if two Twitter users are followed by exactly the same persons. This problem can be solved by replacing zero weights by minimal weights: We first choose the minimal positive edge weight $w_0 > 0$. Then, if an edge has weight zero, we replace its weight with $\frac{w_0}{2}$. In this way, the zero edges are still the edges with the minimal weight but the shortest path distance is a metric. The resulting networks after these two steps are the ones used in the paper.

3 Implementation Details

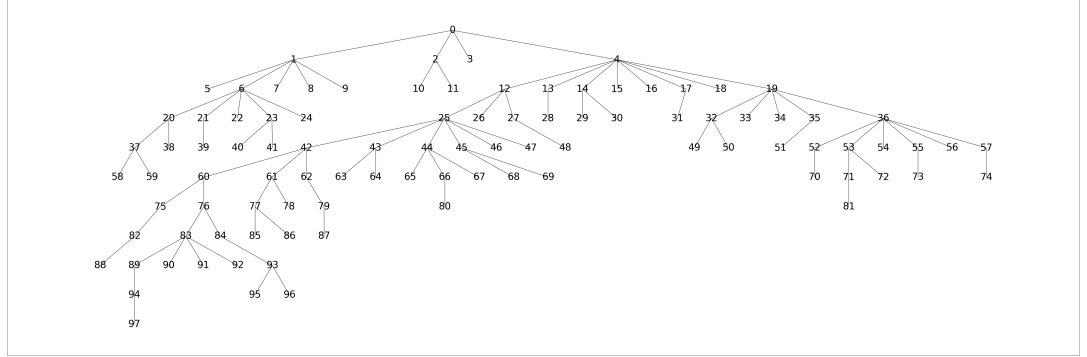
All our experiments were implemented in Python 3.7 and carried out on a Xeon-Gold 5122. To work with graph data, we used the network library networkx 2.6.3. The shortest path lengths between all pairs of vertices were computed via `all_pairs_dijkstra_path_length`, except for the twitter datasets. Here, for performance reason, we switched to the `shortest_paths` function of the Python API of igraph 0.9.7.

4 Trees

Here, we display the labeled trees for the co-author networks. In the brackets, we display the height and prominence of each author.

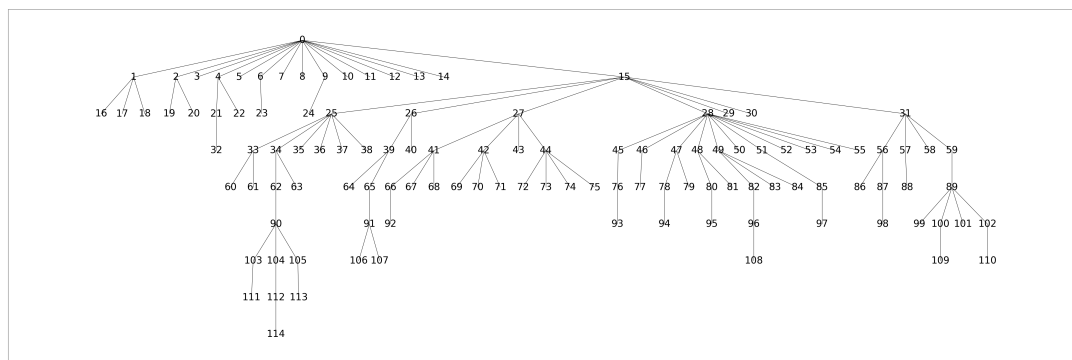
References

1. Deo, N.: Graph Theory with Applications to Engineering and Computer Science. Prentice-Hall series in automatic computation (1974)



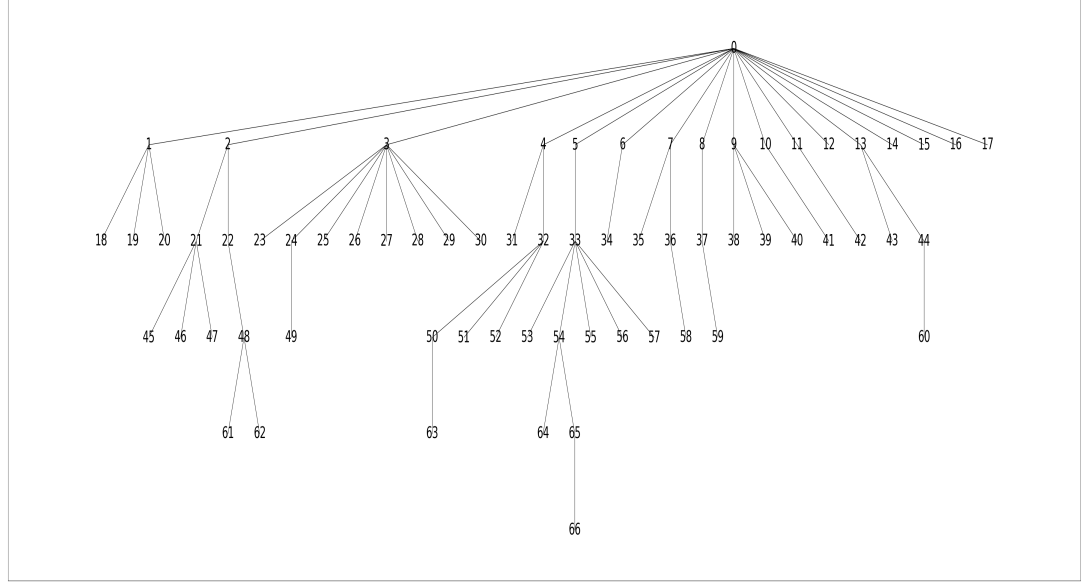
id	name	id	name	id	name	id	name
0	Yoshua Bengio (154, 154)	25	Zhi-Hua Zhou (90, 36)	50	Klemens Böhm (29, 2)	75	Shie Mannor (56, 17)
1	Alexander J. Smola (100, 36)	26	Congfu Xu (14, 2)	51	Kathryn B. Laskey (30, 15)	76	Andreas Hotho (51, 2)
2	Jürgen Schmidhuber (77, 20)	27	S. Muthukrishnan (67, 21)	52	Ahmed AbdelWahab (22, 3)	77	Richard Nock (30, 5)
3	Marc Toussaint (37, 2)	28	Volker Tresp (42, 22)	53	Gennady L. Andrienko (52, 7)	78	Aapo Hyvärinen (49, 30)
4	Jiawei Han (139, 85)	29	Nikunj C. Oza (22, 2)	54	Tomás Pevný (29, 24)	79	Marcos André Gonçalves (44, 19)
5	Christian Wolf (31, 6)	30	Peter J. Stuckey (43, 4)	55	Eyke Hüllermeier (53, 15)	80	Jakub Mareček (13, 9)
6	Zoubin Ghahramani (94, 41)	31	Marko Jereminov (11, 1)	56	Pavlos Protopapas (24, 4)	81	Irina Rish (26, 14)
7	Bernd Bischl (27, 2)	32	Quan Z. Sheng (44, 2)	57	Céline Robardet (24, 2)	82	Bernard De Baets (54, 16)
8	Devis Tuia (47, 19)	33	Eli Upfal (38, 11)	58	Ambuj K. Singh (35, 22)	83	Saso Dzeroski (48, 7)
9	Qian Zhang (79, 23)	34	Iordanis Koutsopoulos (29, 9)	59	Pieter Libin (9, 5)	84	Osmar R. Zaiane (47, 10)
10	Marco Zaffalon (26, 7)	35	Fabio Crestani (33, 1)	60	Steven C. H. Hoi (59, 4)	85	Marc Boullé (17, 4)
11	Ross T. Whitaker (49, 27)	36	Luc De Raedt (55, 8)	61	Masashi Sugiyama (59, 4)	86	Boris Chidlovskii (15, 2)
12	Yong Yu (95, 18)	37	Gianluca Bontempi (49, 20)	62	Chris H. Q. Ding (64, 11)	87	Wagner Meira (40, 15)
13	Matej Oresic (75, 55)	38	Tristan Snowsill (13, 2)	63	Elizabeth S. Burnside (32, 16)	88	Tapio Pahikkala (27, 5)
14	Bhavani M. Thuraisingham (48, 2)	39	Holger Fröning (12, 9)	64	Milind Tambe (71, 37)	89	Louis Wehenkel (39, 1)
15	Jinbo Xu (40, 27)	40	Battista Biggio (35, 6)	65	Sarah Nogueira (4, 2)	90	Fabio Mercorio (15, 1)
16	Rui Chen (48, 16)	41	Jesús Manuel de la Cruz (20, 5)	66	Matthias Weidlich (37, 6)	91	Kaihua Liu (14, 8)
17	Lawrence T. Pileggi (42, 32)	42	Feiping Nie (70, 16)	67	George Tsatsaronis (32, 8)	92	Hongyu Guo (15, 4)
18	Nassir Navab (70, 61)	43	Raymond J. Mooney (72, 23)	68	Sridhar Mahadevan (37, 11)	93	Jerzy Stefanowski (33, 2)
19	Aristides Gionis (56, 9)	44	Ioannis P. Vlahavas (47, 7)	69	Markus Schedl (38, 18)	94	Patrick Gallinari (36, 14)
20	Rémi Munos (59, 23)	45	Samuel Kaski (45, 17)	70	Lena A. Jäger (8, 6)	95	Mark Last (30, 1)
21	Franz Pernkopf (24, 8)	46	Wolfgang Rhode (78, 69)	71	Francesco Calabrese (32, 13)	96	Jesús Cid-Sueiro (15, 1)
22	Kazumi Saito (26, 3)	47	Alex Alves Freitas (54, 15)	72	Elke A. Rundensteiner (43, 20)	97	Nicolas Vayatis (23, 1)
23	C. Lee Giles (70, 33)	48	Barry O'Sullivan (29, 3)	73	Mohamed Nadif (19, 11)		
24	Yuji Matsumoto (42, 19)	49	Sen Wang (35, 10)	74	Henrik Grosskreutz (17, 1)		

Fig. 1. ECML/PKDD



id	name	id	name	id	name	id	name
0	Jiawei Han (139, 139)	29	Andrew McCallum (89, 6)	58	Ruslan Salakhutdinov (86, 47)	87	Ming Zhou (69, 19)
1	Eamonn Keogh (79, 30)	30	David Lobell (82, 47)	59	Feiping Nie (70, 10)	88	Tianyang Zhang (7, 2)
2	Tat-Seng Chua (80, 6)	31	Fei Wang (90, 7)	60	Stephen P. Boyd (88, 21)	89	Masashi Sugiyama (59, 4)
3	Louise E. Moser (23, 14)	32	Sangram Ganguly (55, 22)	61	Kota Tsubouchi (10, 4)	90	Ravi Kumar (66, 9)
4	Jason I. Hong (60, 8)	33	Wei Zhang (102, 6)	62	George Karypis (71, 11)	91	Achla Marathe (23, 5)
5	Dinggang Shen (85, 14)	34	Eric E. Schadt (102, 20)	63	Stefan Savage (70, 5)	92	Salvatore Scellato (30, 8)
6	Martin Ester (51, 7)	35	Max Welling (64, 7)	64	Benjamin J. Bachman (6, 2)	93	Klaus Ackermann (5, 2)
7	Vitaly Shmatikov (57, 22)	36	Tong Zhang (72, 15)	65	Feng Chen (46, 2)	94	Prantik Bhattacharyya (8, 2)
8	Alex Deng (12, 3)	37	Céline Robardet (24, 4)	66	Cecilia Mascolo (64, 23)	95	Hiroki Sugiura (8, 5)
9	Thomas S. Huang (115, 45)	38	Thore Graepel (44, 13)	67	Virgilio A. F. Almeida (43, 6)	96	Luc De Raedt (55, 6)
10	David L. Small (17, 4)	39	Lawrence A. Donehower (52, 8)	68	Liang-Hao Huang (11, 5)	97	Kungpeng Liu (8, 1)
11	Nick Koudas (55, 10)	40	Sam Shah (12, 3)	69	Ioana Giurgiu (9, 1)	98	Chuan Qin (32, 3)
12	Jie Wang (106, 33)	41	Wei Chen (111, 51)	70	Ming Lei (4, 2)	99	Yasuhiro Fujiwara (39, 7)
13	Cheng-Wei Wu (15, 6)	42	Zheng Liu (88, 9)	71	Yi Li (80, 38)	100	Hiroshi Nakagawa (22, 3)
14	Paul M. Thompson (137, 89)	43	Jayant Kalagnanam (27, 2)	72	Xiaoqiang Zhu (17, 1)	101	Bernhard Pfahringer (45, 5)
15	Michael I. Jordan (134, 59)	44	Yong Yu (95, 18)	73	Kewei Chen (53, 8)	102	Pierre-Alain Fouque (40, 17)
16	Bilson J. L. Campana (9, 2)	45	Yehuda Koren (49, 5)	74	Lawrence Carin (69, 2)	103	Chih-Jen Lin (56, 18)
17	Hiroshi Mamitsuka (28, 5)	46	Robert Chen (66, 3)	75	Ruiyuan Li (20, 3)	104	Duncan Watts (50, 32)
18	Evangelos E. Milios (37, 4)	47	Hans-Peter Kriegel (71, 10)	76	Joe Walsh (9, 2)	105	Vahab S. Mirrokni (50, 7)
19	Guoliang Li (62, 4)	48	Hiroshi Sawada (41, 18)	77	Fanjin Zhang (4, 2)	106	Graham Katz (18, 10)
20	Evangelos Kanoulas (27, 8)	49	Eric Horvitz (80, 8)	78	Souvik Ghosh (12, 4)	107	Jaime Arredondo (13, 4)
21	Dimitrios Gunopoulos (58, 6)	50	Matthias Grossglauser (29, 8)	79	Michael Mitzenmacher (66, 7)	108	Anthony J. Bagnall (29, 15)
22	Ed H. Chi (51, 13)	51	João Gama (46, 3)	80	Ryo Asaka (22, 19)	109	Yukihiko Tagami (7, 3)
23	Yong Ge (34, 6)	52	Prashant J. Shenoy (67, 19)	81	Ken-ichi Kawarabayashi (37, 14)	110	Christine Lagouët (9, 7)
24	Murali Ramanathan (41, 2)	53	Diane J. Cook (54, 7)	82	Pedro Domingos (76, 15)	111	Wei Di (19, 4)
25	Bernhard Schölkopf (119, 32)	54	Andrew J. Connolly (69, 25)	83	Yina Tang (6, 2)	112	Yundong Tian (26, 8)
26	Inderjit Dhillon (71, 26)	55	Jennifer Widom (69, 8)	84	Rodica Pop-Busui (47, 44)	113	Adam D. Smith (47, 5)
27	Anil K. Jain (130, 53)	56	Xiaoming Li (84, 19)	85	Zimu Zhou (34, 7)	114	Estevam R. Hruschka (19, 2)
28	Christos Faloutsos (107, 20)	57	Yunfei Lu (9, 4)	86	Wei Wang (75, 14)		

Fig. 2. KDD



id	name	id	name	id	name	id	name
0	Jiawei Han (139, 139)	17	Chun Chen (58, 7)	34	Felix Cheung (8, 4)	51	Lior Rokach (52, 15)
1	Christos Faloutsos (107, 52)	18	Ramayya Krishnan (39, 5)	35	Manuel Campos (17, 9)	52	Gang Li (40, 6)
2	Yong Yu (95, 24)	19	Alex Beutel (31, 15)	36	Sadok Ben Yahia (25, 3)	53	Svetha Venkatesh (46, 11)
3	Zhi-Hua Zhou (90, 35)	20	Nan-Chen Hsieh (8, 1)	37	Niloy Ganguly (33, 3)	54	Manabu Okumura (27, 13)
4	Wei Zhang (102, 35)	21	Xiaoming Li (84, 6)	38	Shen Furao (18, 7)	55	Houkuan Huang (20, 2)
5	Eamonn J. Keogh (79, 30)	22	Qian Zhang (79, 6)	39	Elaheh ShafieiBavani (8, 1)	56	Si Quang Le (12, 3)
6	Minghua Zhang (19, 7)	23	Meng Chen (48, 13)	40	Xin Zuo (5, 3)	57	Tharshan Vaithianathan (13, 4)
7	Francesco Bonchi (47, 7)	24	Yusuke Suzuki (37, 26)	41	Samir Mustapha (13, 4)	58	Maguelonne Teisseire (24, 2)
8	Cyrus Shahabi (57, 11)	25	Hao Chen (34, 5)	42	Chris H. Q. Ding (64, 4)	59	Tanya Goyal (12, 3)
9	Chen Wang (71, 6)	26	Hiroyuki Toda (29, 10)	43	Gianluigi Greco (26, 10)	60	Wayne Wobcke (14, 5)
10	Boualem Benatallah (54, 8)	27	Mineichi Kudo (17, 8)	44	Nada Lavrac (41, 8)	61	Md. Rafiul Hassan (14, 5)
11	George Karypis (71, 11)	28	Alexandros Kalousis (26, 13)	45	Chunmei Dong (7, 2)	62	Viktor K. Prasanna (53, 5)
12	Osmar R. Zaiane (47, 10)	29	Ulf Johansson (18, 6)	46	Claudia Eckert (29, 17)	63	Yun Xiong (15, 1)
13	Irwin King (58, 2)	30	Tony F. Chan (65, 47)	47	Hsin-Min Wang (29, 3)	64	Sanparith Marukatat (13, 5)
14	Xueping Zhang (5, 1)	31	Peng Wang (99, 17)	48	Albert Y. Zomaya (66, 18)	65	Kazumi Saito (26, 11)
15	Luc De Raedt (55, 25)	32	Yiyu Yao (63, 25)	49	Kenichi Takahashi (11, 1)	66	Junichiro Mizusaki (15, 6)
16	Tamás D. Gedeon (32, 8)	33	Geoffrey I. Webb (48, 1)	50	Yaqin Wang (25, 11)		

Fig. 3. PAKDD