Computing Strong Articulation Points and Strong Bridges in Large Scale Graphs

Donatella Firmani, ***Luigi Laura***.

Dept. of Computer Science and Systems, *Sapienza Univ. of Rome*

Federico Santaroni, ***Giuseppe F. Italiano***.

Dept. of Computer Science, Systems and Production, *Univ. of Rome Tor Vergata*

Alessio Orlandi.

Dept. of Computer Science, *Univ. of Pisa*

11° Symposium on Experimental Algorithms (SEA 2012)

Connectivity in *undirected* graphs

Let $G = (V,E)$ be an ***undirected* graph**, with m edges and n vertices.

- G is **connected** if there is a *path* from each vertex to every other vertex in G.

  The **connected components** of G are its maximal *connected* subgraphs.

- A vertex $v \in V$ is an **articulation point** if its removal increases the number of *connected components* of G.

  A graph G is **2-vertex connected** if it has ***no*** *articulation points*.

  → An *articulation point* is a **vertex cut** for G.

Connectivity in *directed* graphs

Let $G = (V,E)$ be a ***directed*** **graph**, with m edges and n vertices.

- G is ***strongly*** **connected** if there is a *directed path* from each vertex to every other vertex in G.

  The ***strongly*** **connected components** **(*SCCs*)** of G are its maximal *strongly connected* subgraphs.

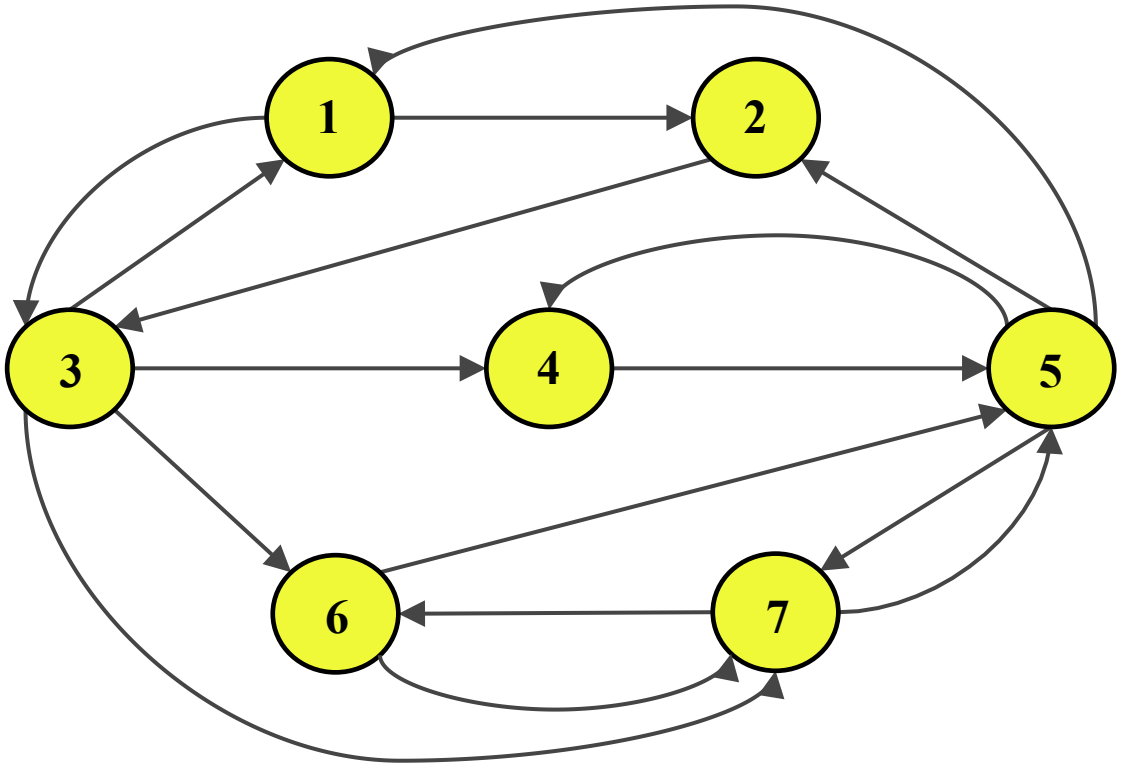
- A vertex $v \in V$ is a ***strong*** **articulation point (SAP)** if its removal increases the number of *strongly connected components* of G.

  A digraph G is **2-vertex connected** if it has ***no*** *strong articulation points*.

  → A *strong articulation point* is a **vertex cut** for G.

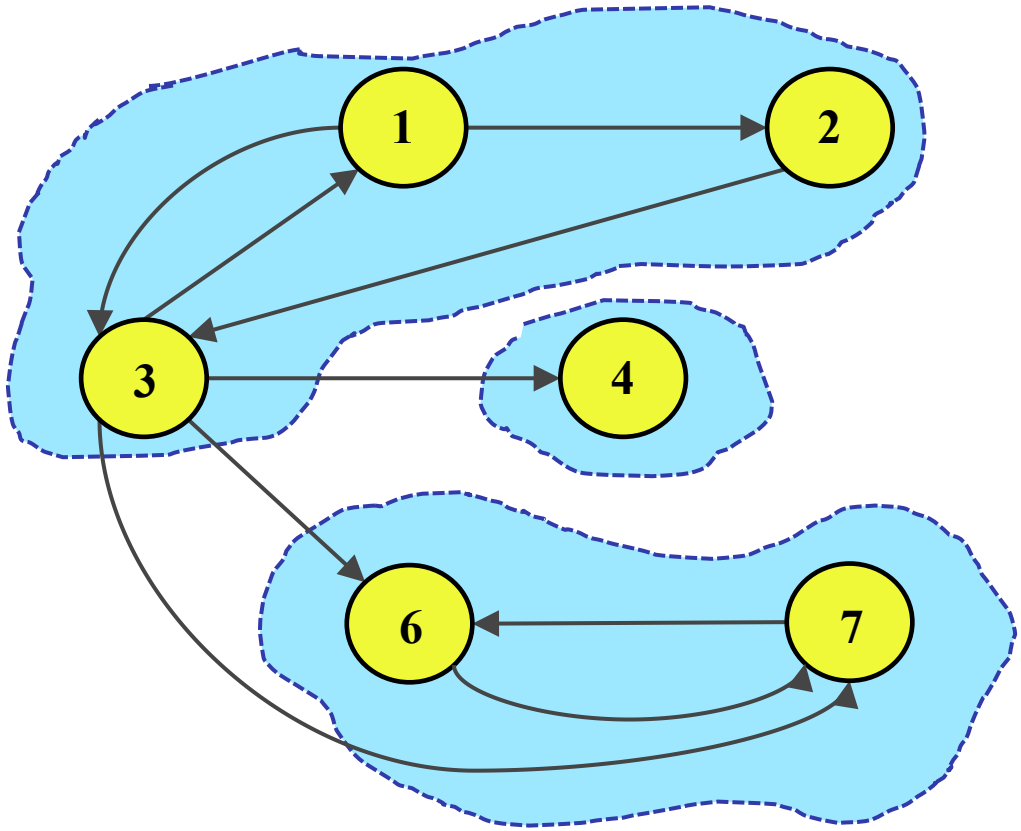
Connectivity in *directed* graphs

A vertex $v \in V$ is a **strong articulation point** if its removal increases the number of *strongly connected components* of G.



Connectivity in *directed* graphs

A vertex $v \in V$ is a **strong articulation point** if its removal increases the number of *strongly connected components* of G.



Vertex **5** is a ***strong articulation point***

Connectivity in *undirected* graphs

An edge $(u,v) \in E$ is a **bridge** if its removal increases the number of *connected components* of G.

A graph G is **2-edge connected** if it has ***no*** *bridges*.

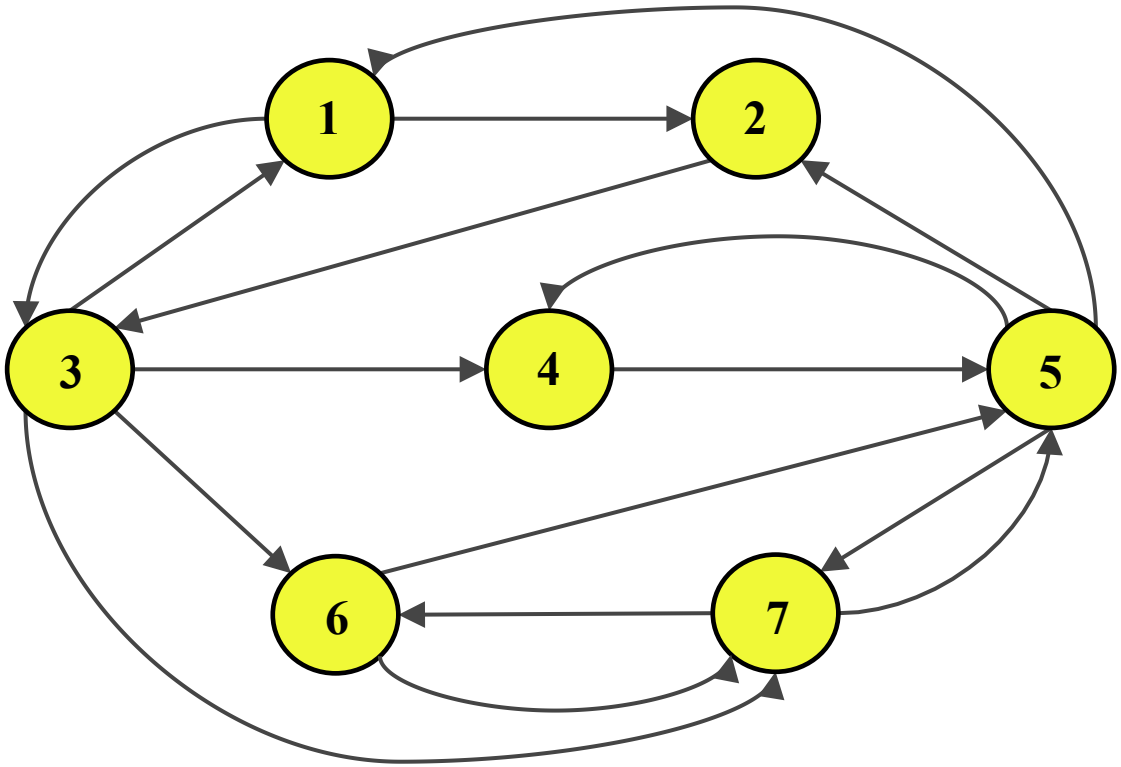
→ A bridge is an **edge cut** for G.

Connectivity in *directed* graphs

An edge $(u,v) \in E$ is a **strong bridge** if its removal increases the number of *strongly connected components* of G.

A digraph G is **2-edge connected** if it has ***no*** *strong bridges*.

→ A strong bridge is an **edge cut** for G.

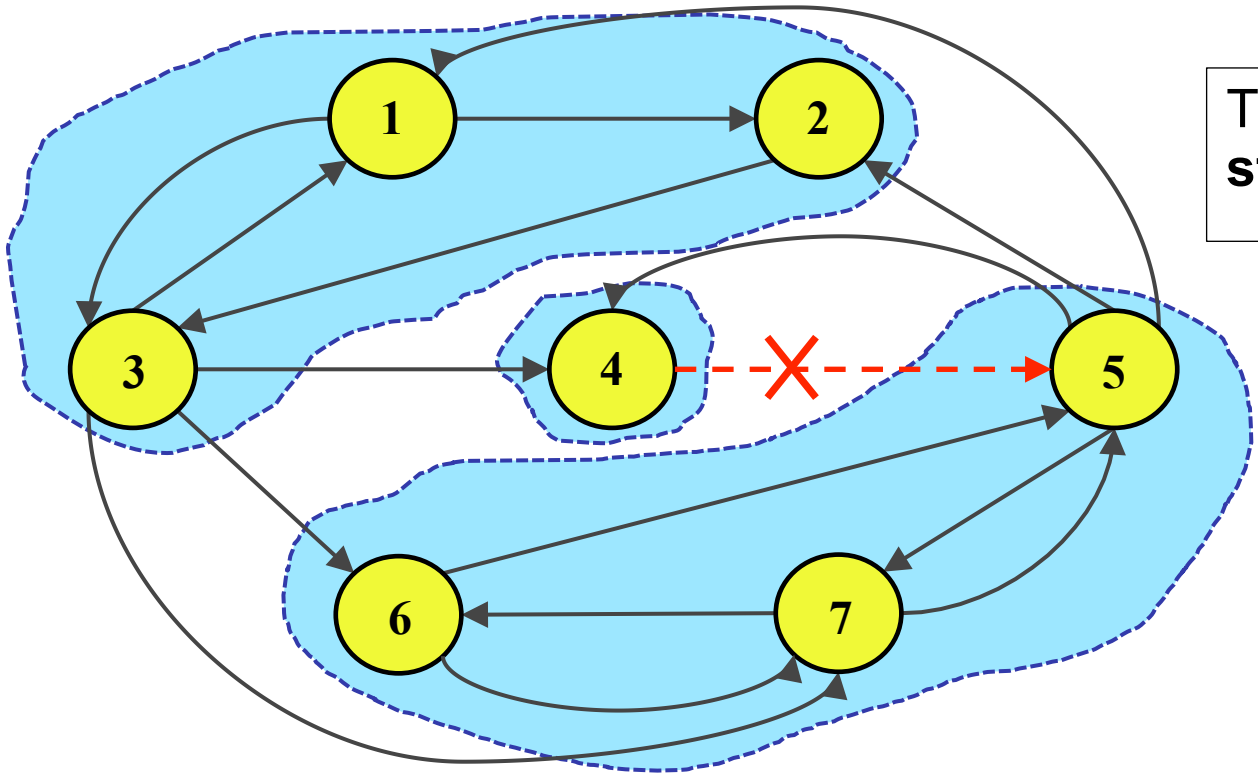


Connectivity in *directed* graphs

An edge $(u,v) \in E$ is a **strong bridge** if its removal increases the number of *strongly connected components* of G.

A digraph G is **2-edge connected** if it has ***no*** *strong bridges*.

→ A strong bridge is an **edge cut** for G.



The edge **(4,5)** is a **strong bridge**.

Theoretical bounds

Directed graphs

- *Testing* whether *G* is **2-vertex connected**
  I.e., find *one* **vertex-cut** (*if any*)

  O(n+m)
  [Georgiadis 10]

- *Testing* whether *G* is **2-edge connected**
  I.e., find *one* **edge-cut** (*if any*)

  O(n+m)
  [Tarjan 76] +
  [Gabow & Tarjan 83]

- *Computing **all*** connectivity cuts
  (**strong articulation points**,
  strong bridges) in *G*

  O(n+m)
  [Italiano, Laura and S. 2010]

Same problem for **undirected graphs** was solved in *linear time* by **Tarjan** in **1973** !

Some applications

- Constraint programming - filtering for tree constraint: computing all the *strong articulation points* was an open problem posed by Beldiceanu et al [2005]

- Reliability in directed networks

- Connectivity and flow of information in social networks [Mislove et al. 2007]

- Speed up computation of matrix determinants
  [Bini & Pan 1994] [Maybee et al. 1989]

…

Naive algorithms

Consider a directed **strongly connected** graph G.

Algorithm ***Naive(SAP)***

- For each vertex *v* check whether it is a **strong articulation point** by computing the *SCCs* of *G/{v}*.

Total time: ***O(n(n+m))***

Algorithm ***Naive(SB)***

- Choose a vertex *v* and compute an out-branching $T^+(v)$ and an in-branching $T^-(v)$;
- Consider the graph $G' = T^+(v) \bigcup T^-(v)$:
  - for each edge *e* in ***G'*** check whether it is a **strong bridge** by computing the *SCCs* of *G/{e}*.

Observation: *G'* has the **same *strong bridges*** of *G* and *G' has* ***2n-2*** *edges*. ***[Italiano et al]***

Total time: ***O(n(n+m))***

Dominators

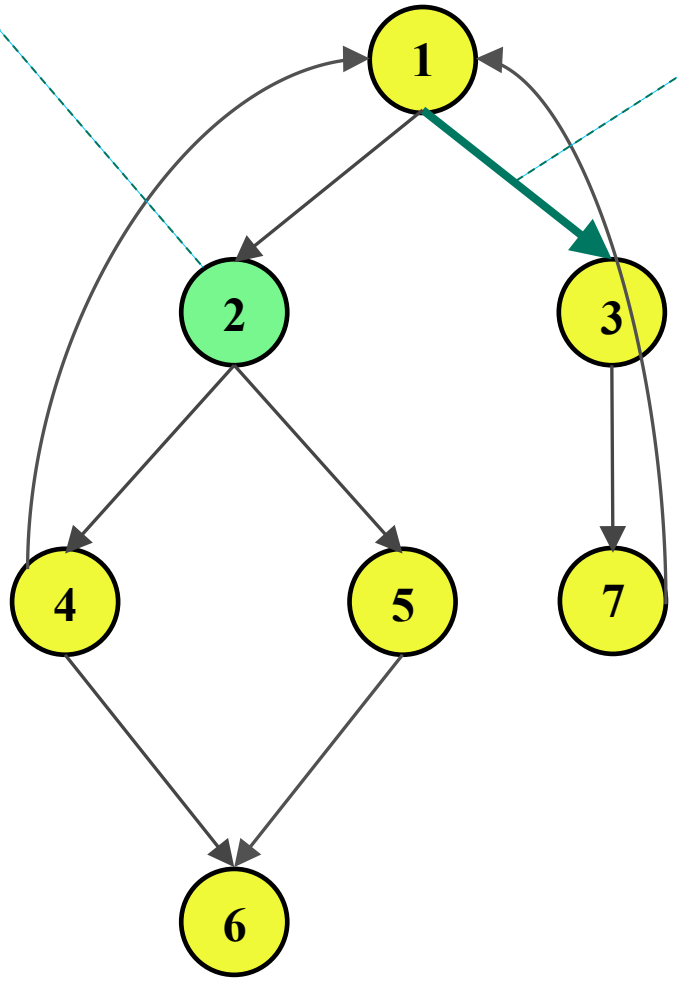
We need few more definitions…

Vertex dominator
*(or simply **dominator**)*

Vertex v *dominates vertex* u *if every path from* s *to* u *includes v.*

Trivial dominators
For a vertex v*, its trivial dominators are* s *and* v.

Vertex dominators of G(s) *can be computed in **linear time** **O(n+m)***



Edge dominator

Edge e *dominates vertex* u *if every path from* s *to* u *includes* e.

Edge dominators of G(s) *can be computed in **linear time** **O(n+m)***

Flowgraph G(s)=(V,E,s)

A digraph with a distinguished vertex s from which every other vertex is reachable

Linear time algorithms

Consider a **strongly connected graph** $G = (V, E)$, its **reversal graph** $G^R = (V, E^R)$, where $E^R = \{(v,u) \mid (u,v) \in E\}$, and a vertex $s \in V$.

Note that *A vertex v is a **SAP** in $G \Leftrightarrow$ it is a **SAP** in G^R.*
*An edge (u,v) is a **SB** in $G \Leftrightarrow (v,u)$ is a **SB** in G^R.*

Theorem *[Italiano et al]*

A vertex $v \neq s$ is a **strong articulation point** in G
$\Leftrightarrow v \in D(s) \bigcup D^R(s)$, *where $D(s)$* is the set of **non-trivial dominators** of $G(s)$ and $D^R(s)$ is the set of **non-trivial dominators** of $G^R(s)$.

An edge (u,v) is a **strong bridge** in G
$\Leftrightarrow (u,v) \in DE(s)$ **or** $(v,u) \in DE^R(s)$, where $DE(s)$ is the set of **edge-dominators** of $G(s)$ and $DE^R(s)$ is the set of **edge-dominators** of $G^R(s)$.

Linear time algorithms

Consider a **strongly connected** graph $G=(V,E)$

Algorithm *ILS(SAP)*

- Choose a vertex **s** and compute $D(s)$;
- *Reverse* the graph and compute $D^R(s)$;
- *Check if **s** is SAP* by removing it from G and by computing the SCCs;
- *Return* $D(s) \bigcup D^R(s)$, and **s** *if* the previous test succeeded.

Algorithm *ILS(SB)*

- Choose a vertex s and compute $DE(s)$;
- *Reverse* the graph and compute $DE^R(s)$;
- *Return* $\{(u,v) \mid (u,v) \in DE(s) \vee (v,u) \in DE^R(s)\}$.

Observation: They can be applyed to general digraphs.

Total time: O(n+m)

Experimental analysis

Experimental setup:

- CPU Intel Xeon X5650 (6 cores) @ 2.67GHz 12MB cache
- 32GB of DDR3 RAM @ 1GHz
- Linux Red Hat 4.1.2-46 (Kernel 2.6.18)
- Java Virtual Machine 1.6.0_16 (64-Bit)
- WebGraph library 3.0.1
- Implementations written in Java to exploit features offered by WebGraph (designed to deal with large graphs)

Experimental analysis

Naive [*O(n(n+m))*] *VS linear time* algorithms

Experiments over ***small instances***.

Graph	n	m	$\#sap$	$\#sb$	ILS(SAP)	Naive(SAP)	ILS(SB)	Naive(SB)
c.elegans	0.3 K	2.3 K	36	45	0.097	0.231	0.067	0.313
political_blogs1	0.6 K	2.7 K	43	84	0.049	0.214	0.064	0.296
political_blogs2	1.0 K	8.2 K	80	154	0.050	0.822	0.068	1.479
political_blogs3	1.4 K	19 K	115	216	0.101	2.049	0.099	4.238
p2p-Gnutella04	11 K	40 K	1344	1674	0.150	29.22	0.180	42.73
wiki-Vote	7 K	103 K	143	152	0.180	12.57	0.180	24.77
enron	69 K	276 K	796	4967	0.580	371.39	0.690	700.37

ILS(SAP) and ILS(SB) algorithms are faster than naive algorithms for small instances, too.

Experimental analysis

Linear time algorithm evaluation

Dataset

Real world graphs including:

- Web graphs
- Social networks
- Communication graphs
- Peer2peer networks
- Product co-purchase graphs

The size of the graphs is up to **22 million** vertices and more than one **half billion** edges.

Graph	n	m
p2p-Gnutella04	11 K	40 K
wiki-Vote	7 K	103 K
enron	69 K	276 K
email-EuAll	265 K	420 K
soc-Epinions1	76 K	509 K
soc-Slashdot0811	77 K	905 K
soc-Slashdot0902	82 K	948 K
amazon0302	262 K	1.2 M
web-NotreDame	325 K	1.4 M
uk-2007-05@100K	100 K	3 M
cnr-2000	325 K	3.2 M
amazon0312	400 K	3.2 M
amazon-2008	735 K	5.1 M
wiki-Talk	2.3 M	5.0 M
web-Google	875 K	5.1 M
web-BerkStan	685 K	7.6 M
in-2004	1.3 M	16.9 M
eu-2005	862 K	19.2 M
uk-2007-05@1M	1 M	41.2 M
soc-LiveJournal1	4.8 M	68.9 M
ljournal-2008	5.3 M	79.0 M
indochina-2004	7.4 M	194 M
uk-2002	18.5 M	298 M
arabic-2005	22.7 M	640 M

Experimental analysis

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Running times (secs)

Experimental analysis

…able to process massive graphs (billion edges) in 10-15 min

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Running times (secs)

Experimental analysis

SAPs appear often (15-25% co-purchase, 11-18% social)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

SAPs appear often (15-25% co-purchase, 11-18% social)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

SAPs appear often (15-25% co-purchase, 11-18% social)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

The vast majority of SAPs are in big SCC (less for Web graphs)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

The vast majority of SAPs are in big SCC (less for Web graphs)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

SBs are less frequent (except for email graphs)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

SBs are less frequent (except for email graphs)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

SBs are also in big SCC (less for some social graphs)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

SBs are also in big SCC (less for some social graphs)

Graph	n	m	$\#sap$	$\#sb$	n_{scc}	m_{scc}	$\#sap_{scc}$	$\#sb_{scc}$	ILS(SAP)	ILS(SB)
p2p-Gnutella04	11 K	40 K	1.3 K	1.6 K	4.3 K	18.7 K	1.3 K	1.6 K	0.15	0.18
wiki-Vote	7 K	103 K	143	152	1.3 K	39.4 K	143	152	0.18	0.18
enron	69 K	276 K	796	4.9 K	8.2 K	147 K	781	4.8 K	0.58	0.69
email-EuAll	265 K	420 K	962	46.0 K	34 K	151 K	960	46.0 K	1.32	1.80
soc-Epinions1	76 K	509 K	8.4 K	23.5 K	32 K	443 K	8.1 K	20.9 K	1.07	1.20
soc-Slashdot0811	77 K	905 K	14.0 K	417	70 K	888 K	13.0 K	3	1.76	2.19
soc-Slashdot0902	82 K	948 K	14.2 K	501	71 K	912 K	14.1 K	69	1.79	2.35
amazon0302	262 K	1.2 M	71.9 K	75.7 K	241 K	11.3 M	69.6 K	73.3 K	3.55	4.77
web-NotreDame	325 K	1.4 M	13.8 K	61.8 K	54 K	304 K	9.6 K	31.9 K	2.48	3.27
uk-2007-05@100K	100 K	3 M	9.4 K	47.0 K	53 K	1.6 M	2.8 K	16.8 K	3.00	3.46
cnr-2000	325 K	3.2 M	32.5 K	104 K	112 K	1.6 M	14.6 K	44.1 K	4.28	5.08
amazon0312	400 K	3.2 M	69.5 K	83.2 K	380 K	3.0 M	69.0 K	82.6 K	11.37	12.40
amazon-2008	735 K	5.1 M	103 K	159 K	627 K	4.7 M	102 K	156 K	25.81	21.89
wiki-Talk	2.3 M	5.0 M	14.8 K	86.7 K	111 K	14.7 M	14.8 K	85.5 K	19.02	18.58
web-Google	875 K	5.1 M	102 K	267 K	434 K	3.4 M	89.8 K	211 K	13.59	15.48
web-BerkStan	685 K	7.6 M	108 K	297 K	334 K	4.5 M	53.6 K	164 K	9.91	12.15
in-2004	1.3 M	16.9 M	82.0 K	421 K	480 K	7.8 M	33.5 K	216 K	32.39	39.02
eu-2005	862 K	19.2 M	104 K	160 K	752 K	17.9 M	99.3 K	146 K	23.95	27.67
uk-2007-05@1M	1 M	41.2 M	147 K	415 K	593 K	22.0 M	82.5 K	259 K	20.90	24.51
soc-LiveJournal1	4.8 M	68.9 M	654 K	1.3 M	3.8 M	65.8 M	649 K	1.3 M	260.03	273.60
ljournal-2008	5.3 M	79.0 M	734 K	1.3 M	4.1 M	74.9 M	727 K	1.3 M	275.53	299.57
indochina-2004	7.4 M	194 M	774 K	2.2 M	3.8 M	98.8 M	503 K	1.4 M	155.83	192.06
uk-2002	18.5 M	298 M	2.3 M	6.1 M	12.0 M	232 M	1.8 M	4.8 M	404.92	478.13
arabic-2005	22.7 M	640 M	2.7 M	6.7 M	15.1 M	473 M	2.2 M	5.2 M	681.47	837.89

Experimental analysis

Average *indegree*, *outdegree* and *Page Rank* of SAPs higher than average !

The only exception are product co-purchase graphs

Graph	δ^-_{avg}		δ^+_{avg}		PR_{avg}	
	V	sap	V	sap	V	sap
p2p-Gnutella04	3.68	4.87	3.68	9.60	$9.19 \cdot 10^{-5}$	$1.12 \cdot 10^{-4}$
enron	3.99	62.79	3.99	103.48	$1.46 \cdot 10^{-5}$	$1.63 \cdot 10^{-4}$
email-EuAll	1.58	280.43	1.58	103.06	$3.80 \cdot 10^{-6}$	$4.29 \cdot 10^{-4}$
wiki-Talk	2.1	69.35	2.10	290.50	$4.18 \cdot 10^{-7}$	$2.86 \cdot 10^{-6}$
amazon0302	4.71	4.65	4.71	4.89	$3.82 \cdot 10^{-6}$	$3.42 \cdot 10^{-6}$
amazon0312	7.99	6.81	7.99	8.65	$2.50 \cdot 10^{-6}$	$2.06 \cdot 10^{-6}$
amazon-2008	7.02	8.21	7.02	8.97	$1.36 \cdot 10^{-6}$	$2.05 \cdot 10^{-6}$
wiki-Vote	12.5	79.29	12.50	68.91	$1.21 \cdot 10^{-4}$	$6.83 \cdot 10^{-4}$
soc-Epinions1	6.71	34.29	6.71	32.47	$1.32 \cdot 10^{-5}$	$5.83 \cdot 10^{-5}$
soc-Slashdot0811	11.7	38.73	11.70	39.90	$1.21 \cdot 10^{-5}$	$3.06 \cdot 10^{-5}$
soc-Slashdot0902	11.54	39.08	11.54	40.48	$1.16 \cdot 10^{-5}$	$2.92 \cdot 10^{-5}$
soc-LiveJournal1	14.23	35.89	14.23	35.47	$2.07 \cdot 10^{-7}$	$5.69 \cdot 10^{-7}$
ljournal-2008	14.73	38.46	14.73	37.56	$1.88 \cdot 10^{-7}$	$5.10 \cdot 10^{-7}$
web-NotreDame	4.6	14.13	4.60	13.63	$3.11 \cdot 10^{-6}$	$1.25 \cdot 10^{-5}$
uk-2007-05@100K	30.51	139.95	30.51	46.40	$1.01 \cdot 10^{-5}$	$4.32 \cdot 10^{-5}$
cnr-2000	9.88	27.59	9.88	16.79	$3.12 \cdot 10^{-6}$	$9.08 \cdot 10^{-6}$
web-BerkStan	11.09	21.08	11.09	10.87	$1.46 \cdot 10^{-6}$	$3.36 \cdot 10^{-6}$
web-Google	5.57	17.80	5.57	9.24	$1.09 \cdot 10^{-6}$	$3.30 \cdot 10^{-6}$
in-2004	41.25	222.67	41.25	45.78	$1.01 \cdot 10^{-6}$	$5.06 \cdot 10^{-6}$
eu-2005	22.3	49.35	22.30	25.20	$1.16 \cdot 10^{-6}$	$2.99 \cdot 10^{-6}$
uk-2007-05@1M	12.23	41.71	12.23	19.61	$7.32 \cdot 10^{-7}$	$2.09 \cdot 10^{-6}$
indochina-2004	26.18	62.23	26.18	27.60	$1.37 \cdot 10^{-7}$	$4.09 \cdot 10^{-7}$
uk-2002	16.1	43.93	16.10	20.82	$5.48 \cdot 10^{-8}$	$1.43 \cdot 10^{-7}$
arabic-2005	28.14	82.68	28.14	34.96	$4.44 \cdot 10^{-8}$	$1.31 \cdot 10^{-7}$

➔ **Further indication of the importance of SAPs in real world graphs!**

Experimental analysis

Current work

*A **core** of a* directed *network is the minimal set of vertices whose removal breaks the graph into many small SCCs.*

Following an approximation used for Web graphs, a **core** of a directed social network can be *approximated by the 10% of highest degree*.

Mislove et al [2007]

How many of the highest degree vertices are SAP?

Very few of them!

Can we exploit SAPs to do better?!

Graph	$\#sap_{scc}$	$\#sap_{scc}$ in $\delta^-(10\%)$	$\#sap_{scc}$ in $\delta^+(10\%)$
p2p-Gnutella04	1.3 K	132	69
enron	781	24	8
email-EuAll	960	3	5
wiki-Talk	14.8 K	53	152
amazon0302	69.6 K	2497	4801
amazon0312	69.0 K	998	3628
amazon-2008	102 K	374	1825
wiki-Vote	143	3	11
soc-Epinions1	8.1 K	89	90
soc-Slashdot0811	13.0 K	1	3
soc-Slashdot0902	14.1 K	2	5
soc-LiveJournal1	649 K	3057	3113
ljournal-2008	727 K	3729	3666
web-NotreDame	9.6 K	797	478
uk-2007-05@100K	2.8 K	24	153
cnr-2000	14.6 K	1587	741
web-BerkStan	53.6 K	1640	4179
web-Google	89.8 K	5150	4037
in-2004	33.5 K	2641	2907
eu-2005	99.3 K	10441	12928
uk-2007-05@1M	82.5 K	5825	8360
indochina-2004	503 K	30252	38724
uk-2002	1.8 M	152773	183997
arabic-2005	2.2 M	139907	176214