Perekhodko *Applied Network Science* (2024) 9:34
https://doi.org/10.1007/s41109-024-00638-y

Applied Network Science

RESEARCH

Open Access

Local 2-connected bow-tie structure of the Web and of social networks

Eugenia Perekhodko[1*]

*Correspondence:
e.perekhodko@northeastern.edu

[1] Northeastern Network Science Institute, Huntington 177, Boston 02115, USA

Abstract

The explosive growth of the Web and of social networks motivates the need for analyzing the macroscopic structure of their underlying graphs. Although the characterization of the structure of a graph with respect to its pairwise connectivity has been known for over 15 years, just one subsequent study analyzed the world inside the giant strongly connected component, where it has been shown that the largest strongly connected component has its own *microscopic bow-tie structure* defined with respect to pairwise 2-connectivity among its vertices. In this paper, we introduce the *local microscopic bow-tie structure* of the largest strongly connected component, demonstrating its *self-similarity* property. Our experiments, conducted on the several Web graphs and social networks demonstrate clear structural differences between considered Web and social networks.

Keywords: Network analysis, Network structure, 2-connectivity, Bow-tie, Locality, Centrality measures, Web graphs, social networks, Web mining, Social network analysis

Introduction

The explosive growth of the World Wide Web during the past two decades has created a vast infrastructure that serves daily on-line needs of people all over the world. In parallel with the growth of the World Wide Web, in the past decade several social networks emerged and were adopted quickly by a large portion of the world population. The most popular social networks nowadays consist of hundreds of millions up to a billion of active users and they naturally capture important social activities. Both the Web and social networks can be represented by directed graphs. In Web graphs, the vertices represent static HTML pages and the edges the hyperlinks among the pages. In social networks, the vertices correspond to users of a social network and the edges represent the who-follows-whom information. The analysis of the structure of the underlying graphs in both Web and social networks is undoubtedly important in many scenarios. For instance, the structure of the Web can be beneficial for improving the browsing experience (Carrière and Kazman 1997). Moreover, the knowledge of the macroscopic structure of the Web has been used for designing efficient algorithms for computing PageRank values (Kamvar et al. 2003).

The existing literature provides a good understanding of the connectivity structure in Web graphs and social networks. The bow-tie structure, that was observed for the first time more than 15 years ago, has been applied on different graphs revealing the macroscopic structure of several graphs, and also helped to study the evolution of the Web and several social networks. All these studies observed a giant strongly connected component containing a substantial portion of the vertices, which is considered the core of the directed graph for both in Web graphs and social networks, and which is continuously growing as the graphs evolve. Although the characterization of the structure of a graph with respect to its pairwise connectivity has been known for over 15 years, just one subsequent study further analyzed the structure inside the giant strongly connected component of Web graphs and of social networks. In this scenario, it is necessary to analyze further the core of Web graphs and social networks in order to understand the structure of the core of the Web and of social networks. In this paper, we applied new theoretical notions and algorithms in order to conduct a novel analysis of the core of the Web and of social networks. We considered a natural extension of the methodology used in the pioneering study by Broder et al, by analyzing pairwise 2-connectivity of Web and social networks on the local level. Our findings shed light on the structural properties of the core of the Web and of social networks.

Background and definitions

Let $G = (V, E)$ be a directed graph (digraph), with m edges and n vertices. Digraph G is *strongly connected* if there is a directed path from each vertex to every other vertex. The *strongly connected components* of G are its maximal strongly connected subgraphs. Two vertices $u, v \in V$ are *strongly connected* if they belong to the same strongly connected component of G. The *size* of a strongly connected component is given by its number of vertices. An edge (resp., a vertex) of G is a *strong bridge* (resp., a *strong articulation point*) if its removal increases the number of strongly connected components. Two vertices $u, v \in V$ are said to be *2-edge-connected* (resp., *2-vertex-connected*), and we denote this relation by $u \longleftrightarrow_{2e} v$ (resp., $u \longleftrightarrow_{2v} v$), if there are two edge-disjoint (resp., two internally vertex-disjoint) directed paths from u to v and two edge-disjoint (resp., two internally vertex-disjoint) directed paths from v to u (note that a path from u to v and a path from v to u need not be edge-disjoint or vertex-disjoint). A *2-edge-connected component* (resp., *2-vertex-connected component*) of a digraph $G = (V, E)$ is defined as a maximal subset $B \subseteq V$ such that $v \longleftrightarrow_{2e} w$ (resp., $v \longleftrightarrow_{2v} w$) for all $v, w \in B$Georgiadis et al. (2018). A visualisation of mentioned definitions could be seen in Fig. 1.

Related works

Previous studies provided a good understanding of so called *macroscopic* connectivity structure of the underlying graphs for both Web and social networks. For the Web graph the first such understanding was given by Broder et al. (2000) more than two decades ago. In this study Broder et al. observed an existence of a *macroscopic* bow-tie structure

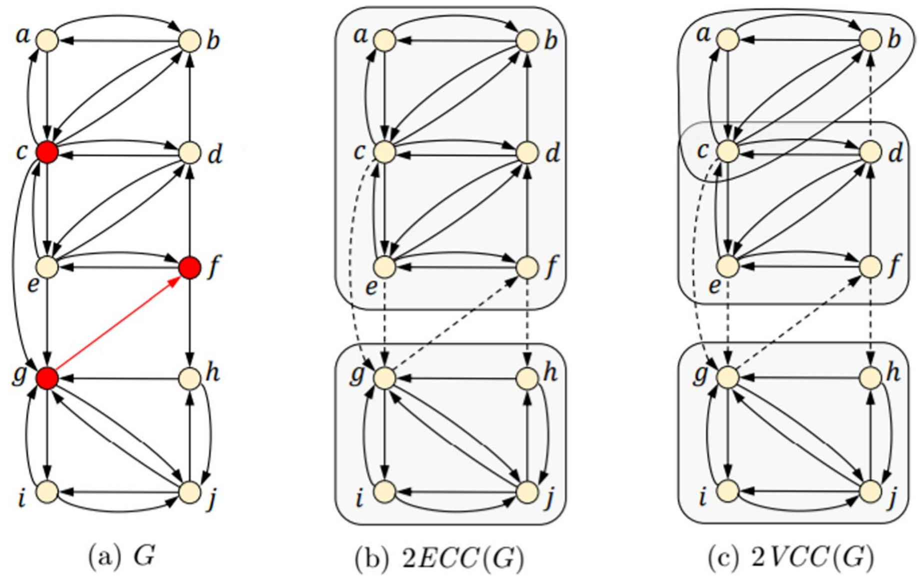

Fig. 1 **a** A strongly connected digraph *G*, with strong articulation points and strong bridges shown in red (better viewed in color); **b** The 2-edge-connected components of *G*; **c** The 2-vertex-connected components of *G* Georgiadis et al. (2018)

that was defined as follows. *Macroscopic* bow-tie structure is formed around the largest strongly connected component (in short *LSCC*). The rest of the vertices were partitioned into six distinct subsets according to their relation to the *LSCC*. Precisely, the vertices that have a directed path to a vertex in the *LSCC* form the *IN* subset; the vertices that have a directed path from a vertex in *LSCC* form the *OUT* subset; the vertices that have a directed path from a vertex that belongs to the *IN* subset form a subset called *IN-TENDRILS*; the vertices that have a directed path to a vertex that belongs to the *OUT* subset form a subset called *OUT-TENDRILS*; the vertices that have both a path from a vertex in the *IN* subset and a path to a vertex in the *OUT* subset form the subset called *TUBES*. The rest of the vertices which do not have a path neither from the *LSCC* or to the *LSCC* and also do not have a path neither from *IN* subset or to *OUT* subset form *DISCONNECTED* subset. See Fig. 2 for an illustration of the mentioned relations. At that time Broder et al. reported the existence of largest strongly connected component in the Web that contained about 28% of all the vertices. Although later studies showed that several structural properties are crawler-sensitive (Serrano et al. 2007), such as the degree distribution or the sizes of the different components of the graph, they *all* agreed on the existence of a giant strongly connected component.

Subsequent studies proposed different structures of the Web graphs, such as the “daisy” structure (Donato et al. 2005) and the “teapot” structure (Zhu et al. 2008), both defined with respect to the largest strongly connected component. The “daisy” structure of the Web discovered by Donato et. al. revealed that the Web has a dense core or the largest strongly connected component and fragmented *IN* and *OUT* components hanging from the largest strongly connected component. In their study Zhu et. al. analyzed the Chinese Web on the different aggregation levels and reported that the Web on the page

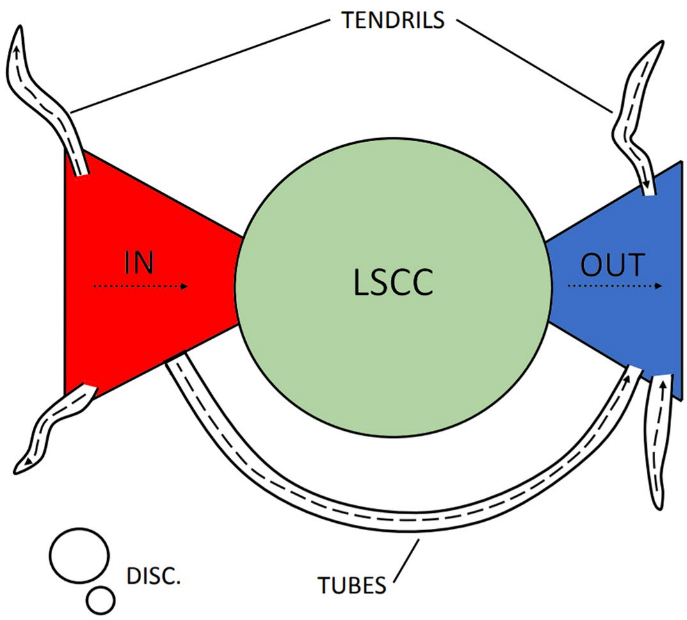

Fig. 2 A visualization of the *global* bow-tie structure of the Web (Broder et al. 2000)

level has the "teapot" structure, where the largest strongly connected component has the biggest size, *IN* component has a medium size and *OUT* component has a small size. Their findings also revealed that with the increment of aggregation levels, the structure of the Chinese Web becomes increasingly close to the "daisy" structure.

In a much recent work Meusel et al. (2014) revisited aforementioned study of Broder et al. (2000) by collecting and analyzing a 2012 crawl of the Web containing 3.5 billion vertices and 128 billion edges. Their study revealed that the largest strongly connected component of the Web consisted of more than 50% of the total number of vertices. Their findings demonstrated that the Web became much more connected as it evolved, which was witnessed by the fact that the fraction of the vertices contained in the core of the Web (i.e., its *LSCC*) substantially increased.

Gabielkov et al. (2014) conducted analysis of the *macroscopic* structure of the underlying graph of Twitter social network, consisting of 505 millions of vertices and 23 billions of edges. They reported existence of the bow-tie structure of the underlying social network formed around the largest strongly connected component containing as well about half of the total number of vertices. In their study Gabielkov et al. (2014) characterized the largest strongly connected component as *"the core of the regular Twitter activity"*.

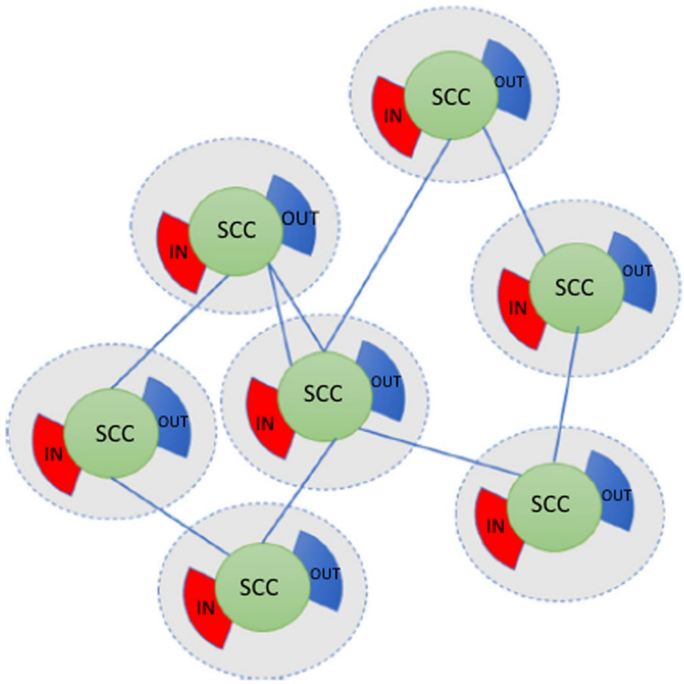

Fig. 3 A visualization of the *local* bow-tie structure of the Google Web network (Fujita et al. 2019). This study showed that each community has its own bow-tie structure that consists of *strongly connected component*, *IN*, and *OUT* parts (better viewed in color)

All aforementioned studies gave us an understanding of the *macroscopic* or *global* connectivity structure of the Web and of social networks, which is centered around the *core* of the networks - their largest strongly connected component. Fujita et al. (2019) proposed a *local* bow-tie structure of the Web, where they discovered that each community in the Web graph has its own bow-tie structure, which is formed in the same way as a global bow-tie structure, but not just around the largest strongly connected component, but around each strongly connected component. Their findings demonstrated a *self-similarity* property of the Web, i.e that the structure of the Web as a whole repeats itself in its communities (see Fig. 3).

The first study that analyzed the *core* of the Web and of social networks (i.e. their *LSCC*) was made by Italiano et al. (2017), where it was shown that the largest strongly connected component has its own $microscopic\ bow-tie\ structure$. In their study they extended the notions of pairwise *strong* connectivity used in the pioneering study by Broder et al. by analyzing the pairwise $2-connectivity$ inside the *LSCC*. Their experiments were conducted on eight Web graphs and seven social networks and revealed an inner bow-tie structure of the *LSCC* that was remarkably consistent across different Web graphs and across different social networks. The *microscopic* bow-tie

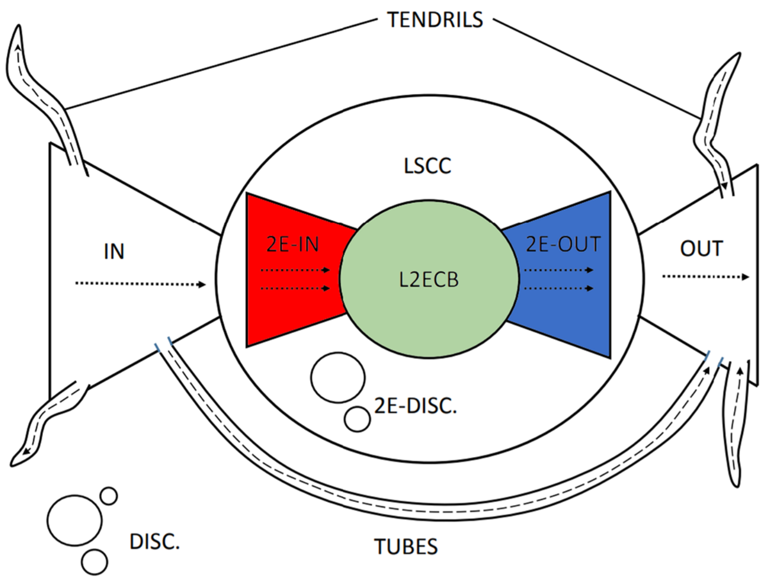

Fig. 4 A visualization of the *microscopic* bow-tie structure of the *LSCC* Italiano et al. (2017). As we can see the structure consists of the largest 2 − *edge* − *connected block*, *IN*, and *OUT* components (better viewed in color). To simplify the picture we depicted the *OTHER* component of the structure as *DISCONNECTED* set

structure with respect to the pairwise 2-edge connectivity between the nodes of the *LSCC* was defined as follows. The core of the structure was defined its largest 2-edge-connected block (in short, *L2ECB*), where all the vertices have two edge-disjoint paths to each other. The sets *2E-IN* and *2E-OUT* (in short, *IN* and *OUT*, respectively) were defined as the sets which contain the vertices that have two edge-disjoint paths to *L2ECB*, and from *L2ECB*, respectively. The vertices that do not have two edge-disjoint paths to *L2ECB* and also do not have two edge-disjoint paths from *L2ECB*, form the set *2E-OTHER* (in short, *OTHER*), which also could be considered as an "analogy" of the *DISCONNECTED* set of the macroscopic bow-tie structure. The sets *L2ECB*, *IN*, *OUT* and *OTHER* were considered to be *the components* of proposed microscopic bow-tie structure. The visualization of *microscopic* bow-tie structure of the *LSCC* could be seen in Fig. 4. Italiano et al. (2017) demonstrated that the *L2ECB* component occupies in total about 40% of vertices in Web graphs and around 60% in social networks, while the component *2E-IN* occupies about 43% and 8%, and the component *2E-OUT* occupies about 6% and 10% of total amount of vertices, respectively. Their findings showed that the structure of the largest strongly connected component is remarkably consistent across different Web graphs and across social networks. Although the analysis in this study demonstrated noticeable differences in the structures of the *LSCC* of Web and of social networks, it undoubtedly needed further investigation of their structural properties.

In this paper we make a step further in the structural analysis made by Italiano et al. (2017) and extend the notion introduced by Fujita et al. of local bow-tie structure with respect to *strong connectivity* to the notion of local bow-tie structure with respect to pairwise $2-$ *connectivity* among the vertices inside the giant strongly connected component. We demonstrate existence of the *local microscopic bow* $-$ *tie structure* of the largest strongly connected component, showing in this way *self* $-$ *similarity* property of the largest strongly connected component itself. Next, we analyze structural differences of the Web and of social networks by considering the distributions of average *PageRank*, *Outdegree* and *Indegree* values over the components of their local microscopic bow-tie structures. Our findings demonstrate differences in the structure of the *LSCC* for Web and for social networks.

The structure of this paper is organized as follows. In Sect. 2 we describe the datasets on which we performed our analysis. Section 3 contains methodology behind it. In Sect. 4 we present distributions of centrality measures among components of the microscopic *local* bow-tie structure for considered Web and social networks. We also present distributions of 2-connectivity characteristics among the components of the local structure. Based on *local microscopic bow* $-$ *tie structure* we demonstrate *self-similarity* property of the largest strongly connected component of considered Web and of social networks and show structural differences between two types of networks.

Data

We conducted our analysis on the collection of 3 Web graphs and 3 social networks. This amount of networks were processed due to computational challenges to process bigger graphs. Among Web graphs we considered the following datasets. *Stanford* and *Berkstan* are the Web graphs of the Stanford and UC-Berkley University domains. The *Google* graph represents the Web pages under the google.com domain, as they were released as part of the Google programming contest in 2002. The datasets *Stanford*, *Berkstan*,

Table 1 Characteristics of the Web and of social networks that we considered in ascending order of their number of edges

Graph	n	m	δ_{avg}	n_{LSCC}	m_{LSCC}	Amount of local bow-ties
Stanford	281K	2.31M	8.3	150K (53%)	1.57M (68%)	1.2K
Google	875K	5.10M	5.8	434K (50%)	3.41M (67%)	4.6K
Berkstan	685K	7.60M	11.1	334K (49%)	4.52M (60%)	2.9K
Epinions	75.80K	508K	6.7	32.30K (42%)	443K (87%)	70
Academia	200K	1.39M	7.0	147K (74%)	1.33M (88%)	222
Google+	211K	1.50M	7.1	86.70K (41%)	1.01M (67%)	563

By n and m we refer to the number of vertices and edges, respectively; δ_{avg} is the average degree and *Google* were taken from the SNAP data repository (Leskovec and Sosič 2014). We also considered the following social networks. *Epinions* (Richardson et al. 2003) represents a user oriented product review website. *Academia* (Fire et al. 2013) is the underlying graph of the network of researchers registered on academia.edu. *Google+* (Fire et al. 2013) contains a small subgraph of Google's famous social network. Table 1 summarizes the characteristics of the datasets that we considered with their largest strongly connected components.

Methodology

Our methodology behind our microscopic *local* structural network analysis is the following. We consider microscopic bow-tie structure introduced in Italiano et al. (2017) and extend it to the *local level*, that is, we consider bow-tie structures, each of which is formed not around the *largest 2-edge-connected block* (*L2ECB*), but around of the *each* of 2-edge-connected blocks (*2ECBs*) of the largest strongly connected component of the digraph *G*. In other words local microscopic bow-tie structure consists of *2-edge-connected blocks*, *IN*, *OUT* and *OTHER* components inside the largest strongly connected component. Note, that *IN*, *OUT* and *OTHER* components are defined in the same way as for global microscopic bow-tie structure. The visualization of proposed structure could

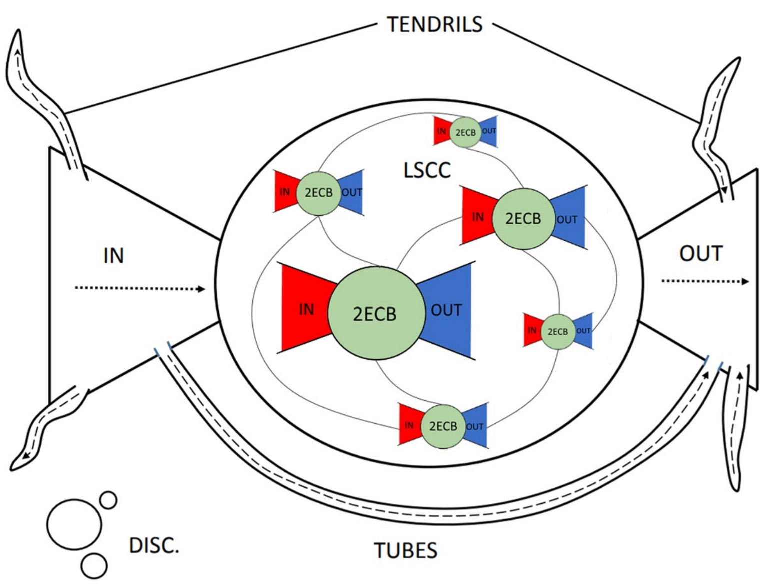

Fig. 5 A visualization of the *local microscopic bow-tie structure* of the largest strongly connected component proposed in this study. As we can see microscopic bow-tie structure consists of *2-edge-connected blocks*, *IN*, and *OUT* parts (better viewed in color). To simplify the picture we did not depict *OTHER* part of the structures. We would like to highlight that proposed structure is valid for any strongly connected component of any directed graph

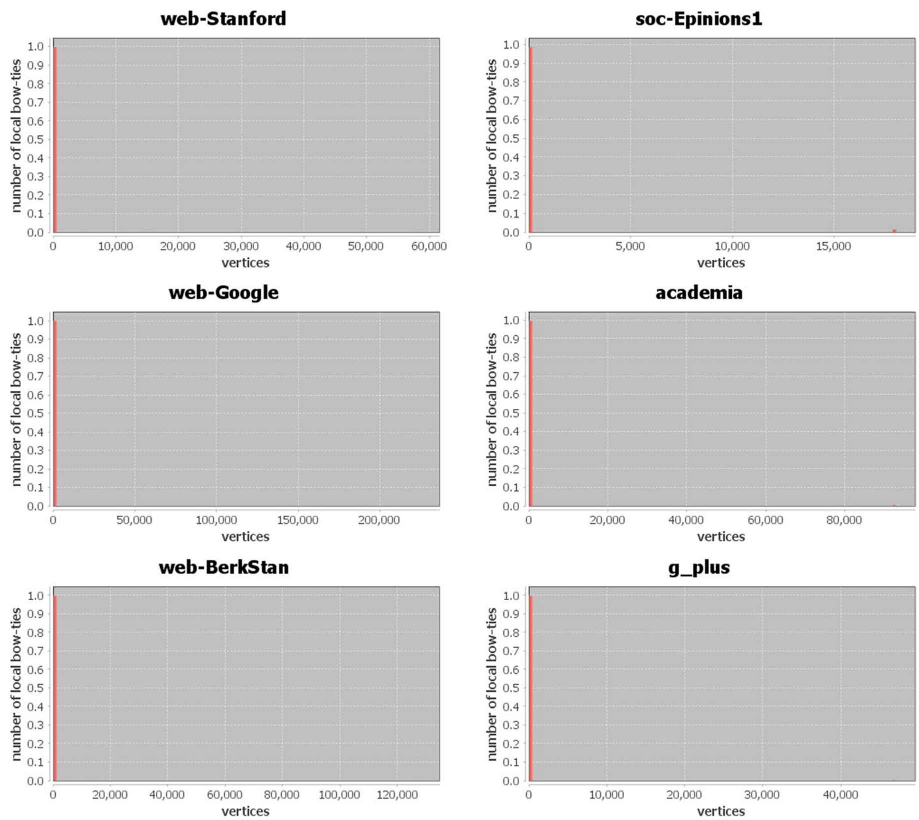

Fig. 6 Distribution of the *vertices* among 2*ECBs* for the Web and social networks that we considered. As we can see most of the local bow-ties occupy considerably small amount of the vertices in Web and social networks

be seen in Fig. 5. It is worth mentioning that methodology for extracting local microscopic bow-tie structure could be applied to *any* strongly connected component of *any* directed graph G.

Microscopic *local* bow-tie structure

In this section we present local microscopic bow-tie structure of the largest strongly connected component for considered Web and social networks. We also analyze components of the microscopic *local* bow-tie structure for considered Web and social networks. We compute several *centrality measures* for *2ECB*, *OUT*, *IN* and *OTHER* components of local structure and show how they differ for Web and social networks. More specifically, we compute average *Pagerank*, average *Outdegree* and average *Indegree* for *2ECB*, *OUT*, *IN* and *OTHER* components of the local structure. We also compute the portion of the vertices of each set of *OUT*, *IN* and *OTHER* that are not 2-edge-connected

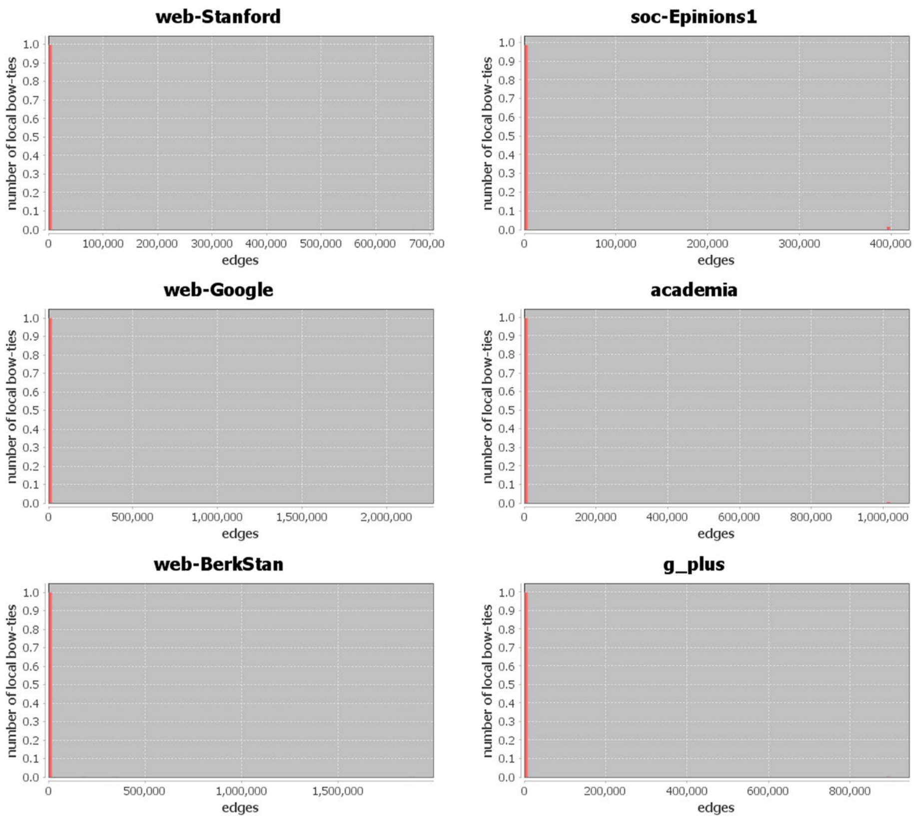

Fig. 7 Distribution of the *edges* among 2*ECBs* for the Web and social networks that we considered. As we can see most of the local bow-ties also occupy a considerably small amount of the edges in Web and social networks

with any other vertex, i.e. they form *singleton 2-edge-connected blocks*, as well as the portion of the *pairs* that are *2-edge-connected*.

In the following four subsections we analyze each set of the *2ECB*, *OUT*, *IN* and *OTHER* components of local bow-tie structure for the Web and social networks that we considered. For each set of the *2ECB*, *OUT*, *IN* and *OTHER* components of local bow-ties we present distributions of the *vertices*, *edges*, average *PageRank*, *Outdegree* and *Indegree* centralities, as well as distributions of $2 - edge - connected$ *pairs* and *singleton* $2 - edge - connected$ *blocks* among all of the components of the local structure. In this way we can see the differences in the distributions of the mentioned metrics among the components of the local structures of the Web and of social networks. We ask the interested reader to proceed with the following subsections.

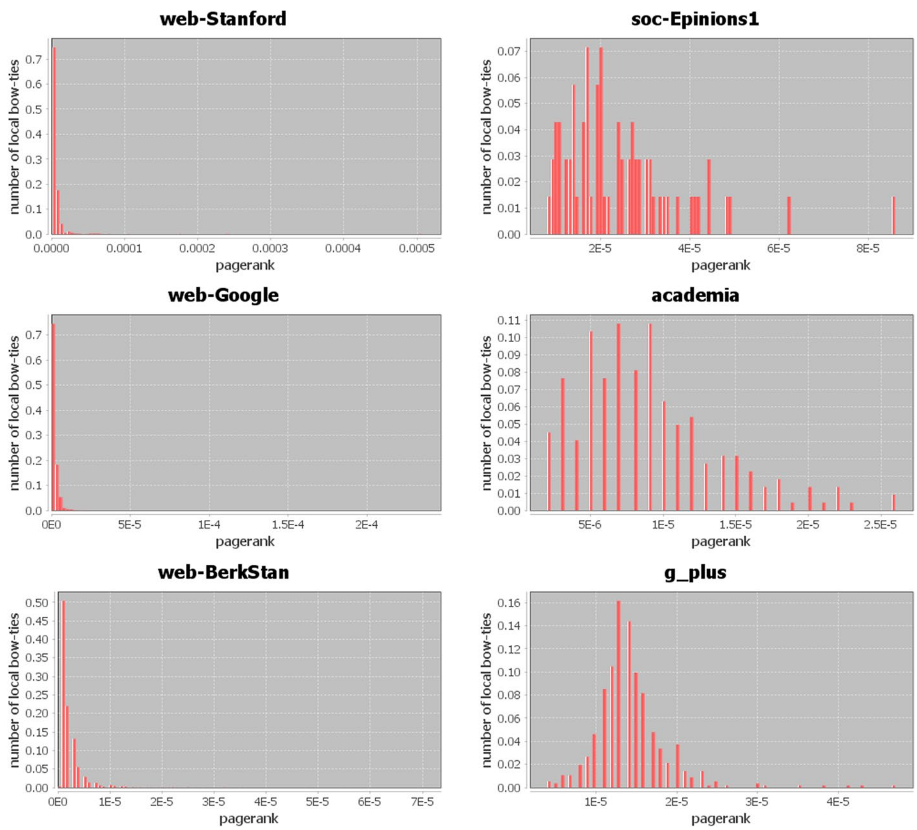

Fig. 8 Distribution of the *PageRank* values among 2*ECBs* for the Web and social networks that we considered. We can observe uniform distributions separately among Web and social networks, which differ significantly between them

Analysis of *2-edge-connected blocks* of the local bow-tie structure

In the following in order to show differences in the structure of the Web and of social networks we placed the histograms for the Web graphs on a right side of the figures and the histograms for the social networks on the left side of the figures. In Figs. 6 and 7 we can see the distributions of the *vertices* and *edges* among all of the 2*ECBs* of local bow-tie structures. We can observe that for both Web and social networks almost 100% the 2*ECBs* occupy considerably small amount of both vertices and edges in our structures. In these plots there is no difference in the core structure between the Web and the social networks.

In Fig. 8 we can see the distributions of the average *PageRank* values among 2*ECBs* of the local bow-tie structures for considered Web and social networks. As we can see the distributions have clear and consistent patterns for both Web and social networks,

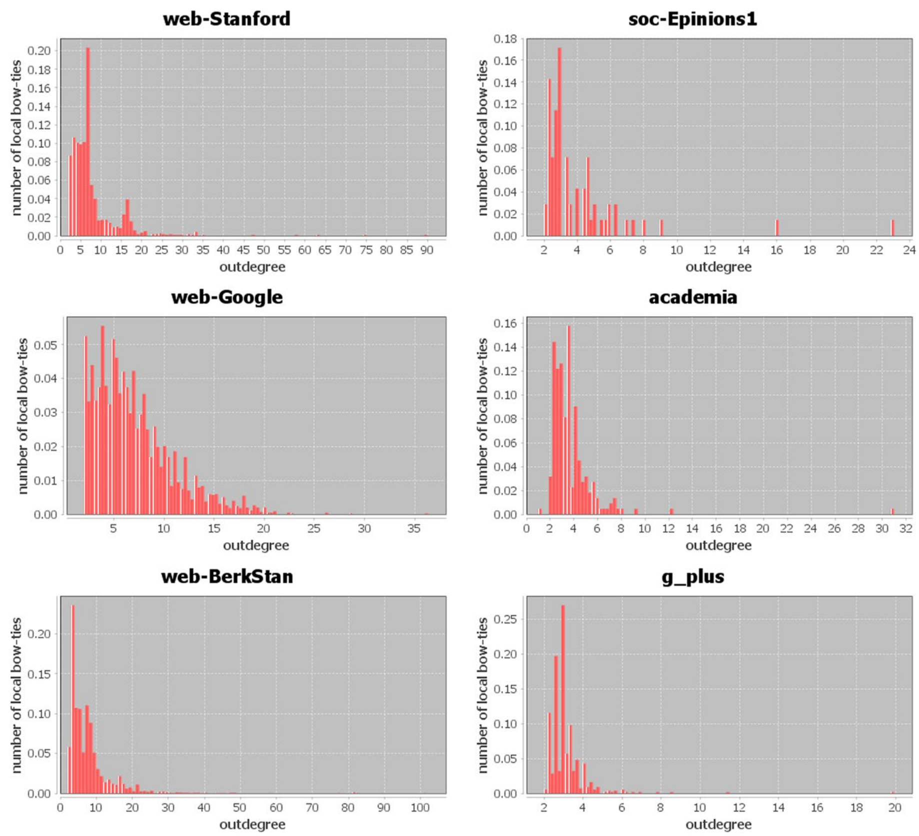

Fig. 9 Distribution of the *Outdegree* values among 2*ECBs* for the Web and social networks that we considered

which differ significantly between them. More precisely, for more than 50% of the local bow-ties the average *PageRank* value is close to 0 and only a few 2*ECBs* have higher values of the average *PageRank*. The distribution of the average *PageRank* values among 2*ECBs* of the local bow-tie structures for considered social networks is much different. More specifically, only a few blocks have the maximum average *PageRank* value, and the rest of the blocks are distributed in decreasing or increasing order.

In Figs. 9 and 10 we can see the distributions of the average *Outdegree* and *Indegree* values among 2*ECBs* of the local bow-tie structures for our Web and social networks. As we can observe the average *Oudegree* has a decreasing order for the Web and for social networks. The pick of the average *Outdegree* for the Web networks lies between 5 and 10 values, while for the social networks, it lies between 2 and 4 values. Similar observation could be made for the average *Indegree* of the Web and of social networks.

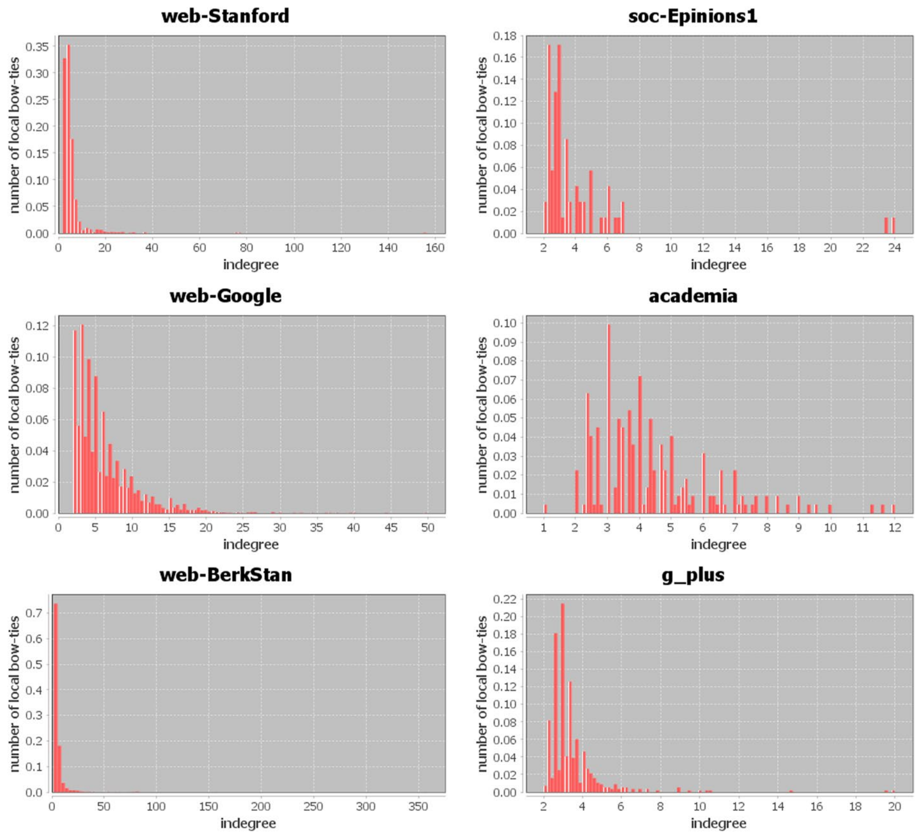

Fig. 10 Distribution of the *Indegree* values among 2*ECBs* for the Web and social networks that we considered

Analysis of *OUT* components of the local bow-tie structure

As we can see in Fig. 11 all our Web graphs have very similar distribution of the vertices among *OUT* components of our local structures. The picture is slightly different for the social networks, where for the *Epinions* social network the majority of the *OUT* components occupy about one third of the vertices in the largest strongly connected component of the network. The similar observation could be made for the edges among *OUT* components, shown in Figure 12.

In Fig. 13 we show the distribution of the *PageRank* values among *OUT* components of the networks that we considered. As we can see the distributions are quite uniform for all our Web graphs. The distribution is more skewed comparing to the distribution of the average *PageRank* for the 2*ECBs*. The distribution is different for our social networks, where we can observe that for the less amount of *OUT* components average *PageRank* values are close to 0.

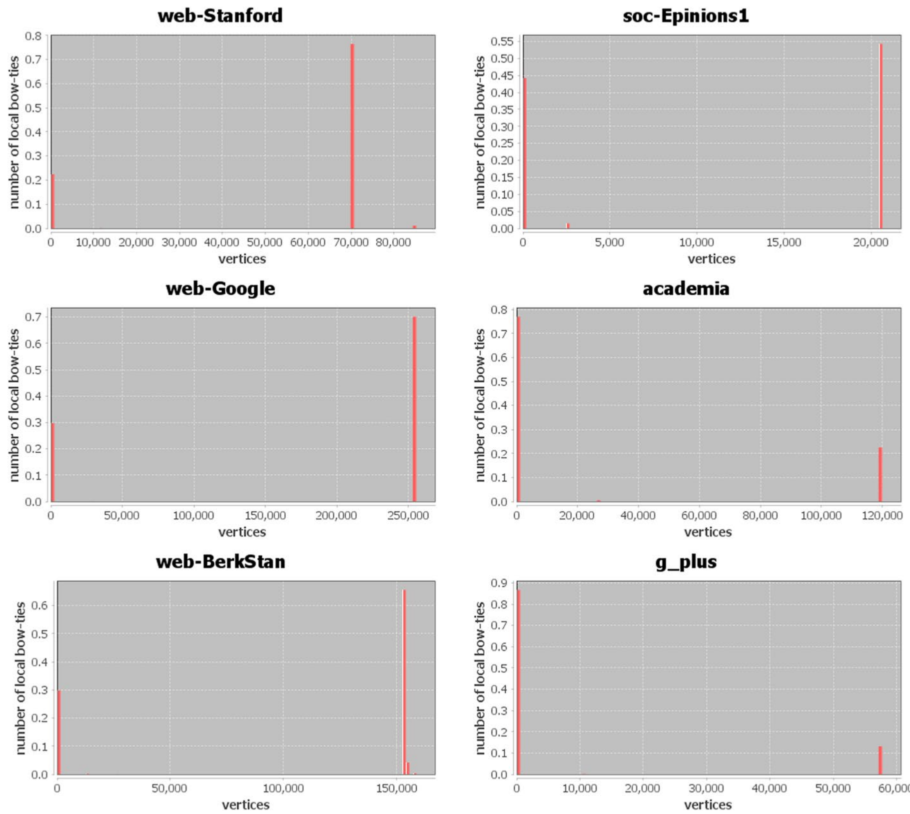

Fig. 11 Distribution of the *vertices* among *OUT* components for the Web and social networks that we considered

In Fig. 14 we present the distribution of *Outdegree* values for the Web and social networks that we considered. As we can see the Web graphs again have a quite uniform distribution among our components. For the social networks we can observe different pattern of distribution, which remains consistent among our networks. Similar observation could be actually made for the distribution of *Indegree* values for the Web and social networks that we considered (Fig. 15).

In Fig. 16 we show the distribution of $2 - edge - connected\ pairs$ of our Web and social networks. As we can see for the Web graphs about 30% of 2-edge-connected pairs have the value close to the 0 and the rest 70% have in average the value of 2e—10. Different observation could be made for our social networks. As we can see about 80% of $2 - edge - connected\ pairs$ have the value close to 0 for *academia* and *Google+* social networks, while for *Epinions* social network this percentage drops to 45%.

Similar differences between distributions of the *singletons* appear in Fig. 17.

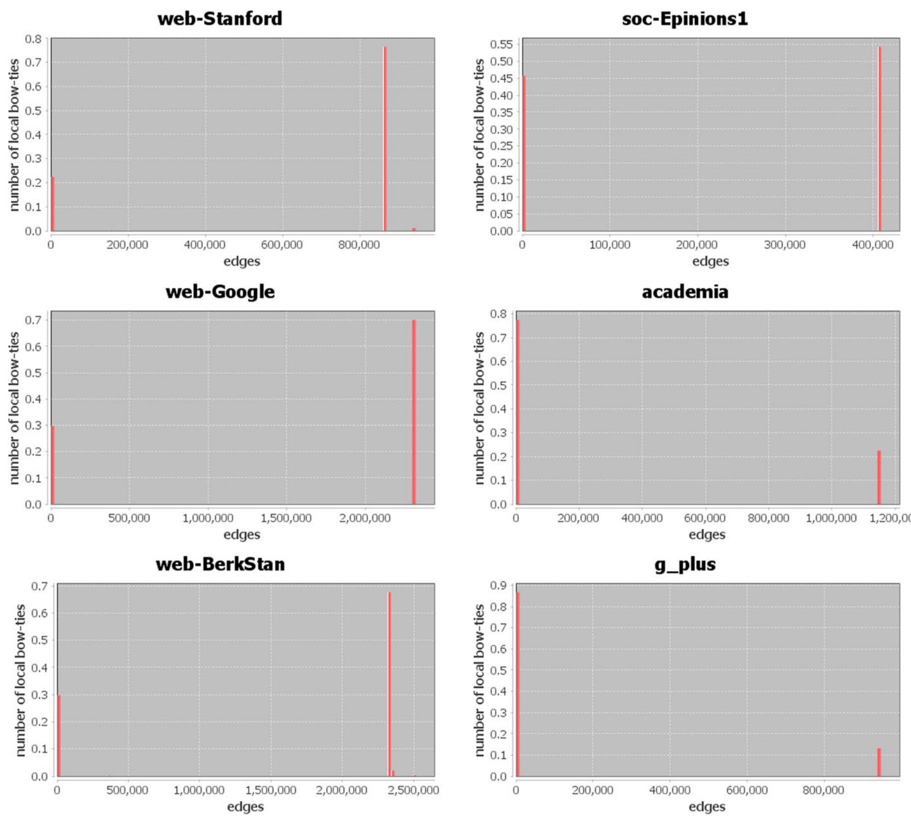

Fig. 12 Distribution of the *edges* among *OUT* components for the Web and social networks that we considered

Analysis of *IN* components of the local bow-tie structure

In Fig. 18 we show distribution of the vertices among *IN* components in our local bow-tie structures. As we can see for our considered Web graphs about 80% of our *IN* components occupy very low amount of vertices. For social networks corresponding value is 60, 30 and 50% for *Epinions*, *Academia* and *Google+* networks, respectively. Similar picture could be observed for the distributions of edges in our *IN* components as shown in Fig. 19.

In Fig. 20 we can observe distributions of the average *PageRank* values among *IN* components of our local bow-tie structures. As it could be seen in Fig. 20 both our Web and social networks have quite distinguishable distributions.

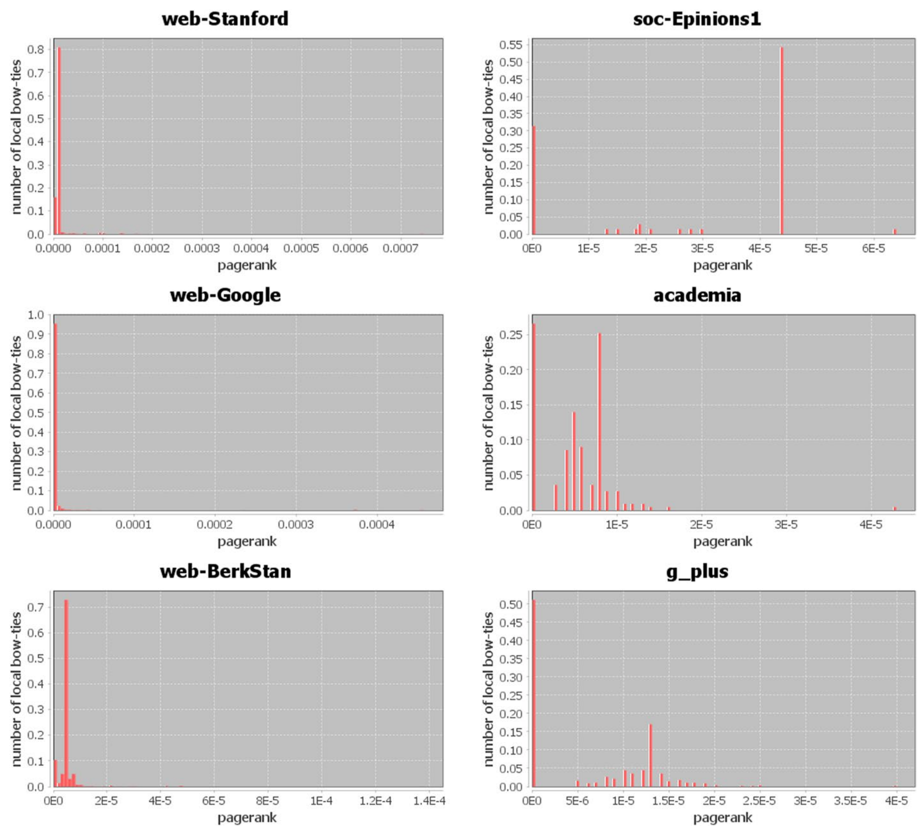

Fig. 13 Distribution of the *PageRank* values among *OUT* components for the Web and social networks that we considered

In Fig. 21 we can see distributions of average *Outdegree* values for our Web and social networks. We can observe a clear pattern for the social networks. About 30% of *IN* components have very low average outdegree. For the social networks the pattern of distribution is quite different, about 50% of all *IN* components have outdegree values equal to 15 on average. Similar distinguished distributions could be observed for the *Indegree* average values among *IN* components, as shown in Fig. 22.

In Fig. 23 we can observe the distribution of the $2-edge-connected\ pairs$ among *IN* components of our local bow-tie structures. As we can see for the Web graphs about 80% of *IN* components have the amount of $2-edge-connected\ pairs$ close to zero, while for the social networks corresponding value represented by about 50% of the components on average. Similar observation could be made for the distribution of the singletons among the same components. As we could observe also with the previous metrics we can see consistent pictures of the distributions among Web and social networks (Fig. 24).

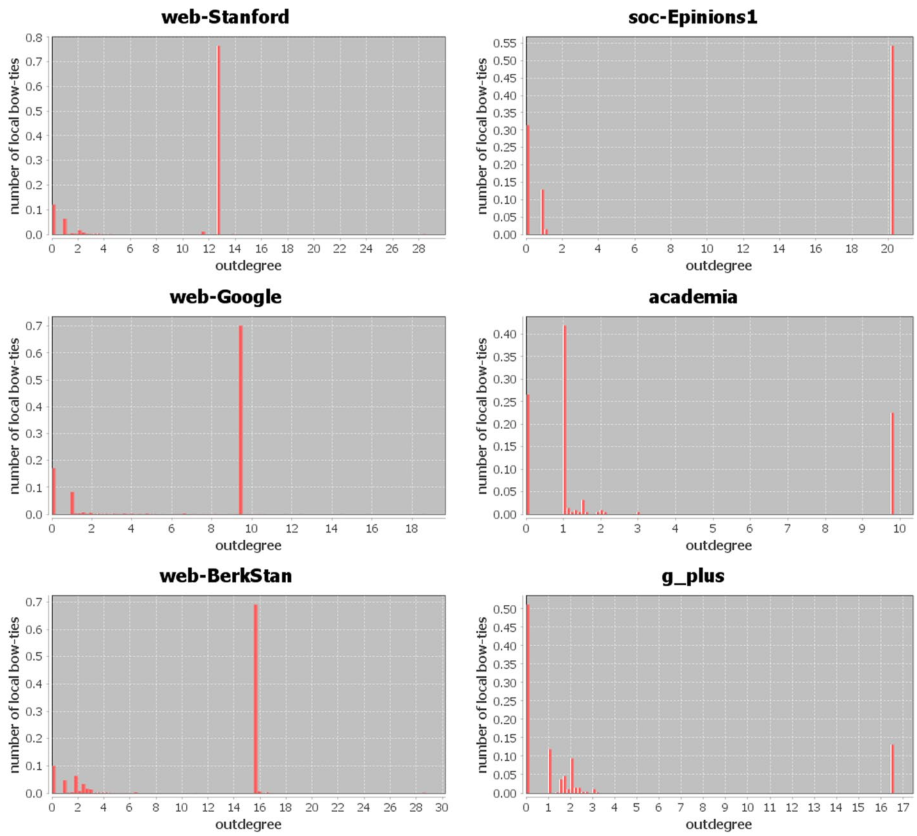

Fig. 14 Distribution of the *Outdegree* values among *OUT* components for the Web and social networks that we considered

Analysis of *OTHER* components of the local bow-tie structure

In Figs. 25 and 26 we show the distributions of the vertices and edges over *OTHER* components for our Web and social networks. As in the cases of previous components we see clear difference between the same distributions for Web and for social networks. At the same time we could observe a surprisingly consistent picture among the distributions of the same type of the network.

In Fig. 27 we can see the distributions of the average *Pagerank* values among *OTHER* components in our local bow-tie structures. As we can observe about 70% of the components have the value of average *Pagerank* equal to 1e−6. Different picture could be seen for the social networks where the majority of the components have 0.5e−6, 4e−6 and 9e−6 average *Pagerank* values.

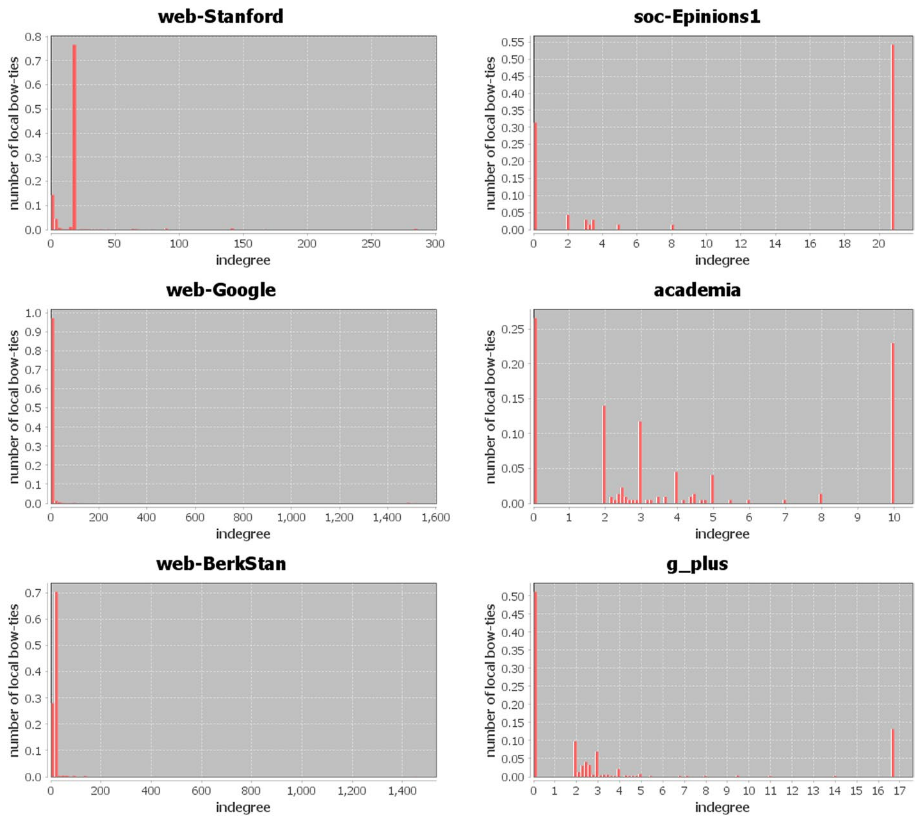

Fig. 15 Distribution of the *Indegree* values among *OUT* components for the Web and social networks that we considered

In Fig. 28 we show the distributions of the average *Outdegree* values among *OTHER* components of our structures. We can observe very distinguishable distributions for the Web and for social networks that we considered. As we can see about 70% of our components have the average values 8.5, 5.5 and 11.5 for *Stanford*, *Google* and *Berkstan* Web graphs, respectively. That is different for the social networks that we considered, where about 50% of the components have the average *Outdegree* values have the value of 1.5 on average.

As we can see in Fig. 29 about 70% of our components have the average *Indegree* values about 3 for social networks. For Web graphs the corresponding value drops to 50–60% of the amount of all components.

As we can see in Figs. 30 and 31 the distributions of $2-edge-connected\ pairs$ and *singletons* among *OTHER* components are quite similar for considered Web and social networks.

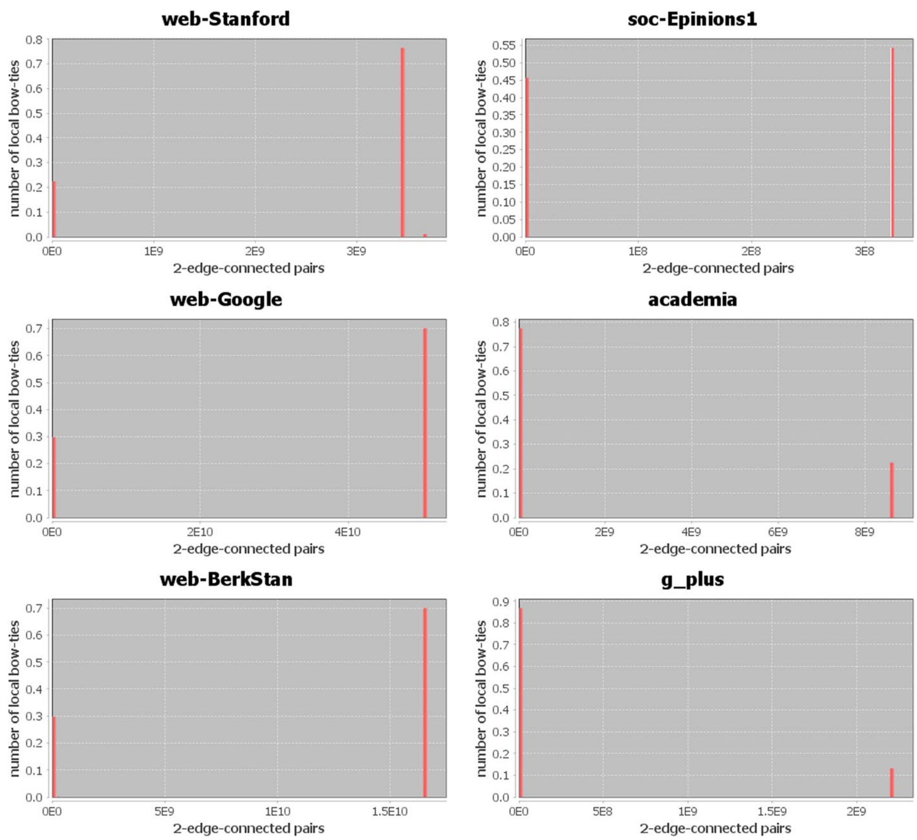

Fig. 16 Distribution of the $2-edge-connected\ pairs$ values among *OUT* components for the Web and social networks that we considered

Conclusion

In this paper we extended the notion of local bow-tie structure with respect to strong connectivity to the local bow-tie structure with respect to pairwise 2-connectivity among the vertices inside the largest strongly connected component. We have conducted our experiments on several Web and social networks and unveiled $local\ microscopic\ bow-tie\ structure$ inside the largest strongly connected component with the microscopic bow-tie formed around each 2-connected block. Revealed local microscopic bow-tie structure demonstrates $self-similarity$ property of the largest strongly connected component. In order to investigate structural differences between Web and social networks that we considered, we analyzed distributions of the *vertices*, *edges*, average *PageRank*, *Outdegree*, and *Indegree* centralities, as well as

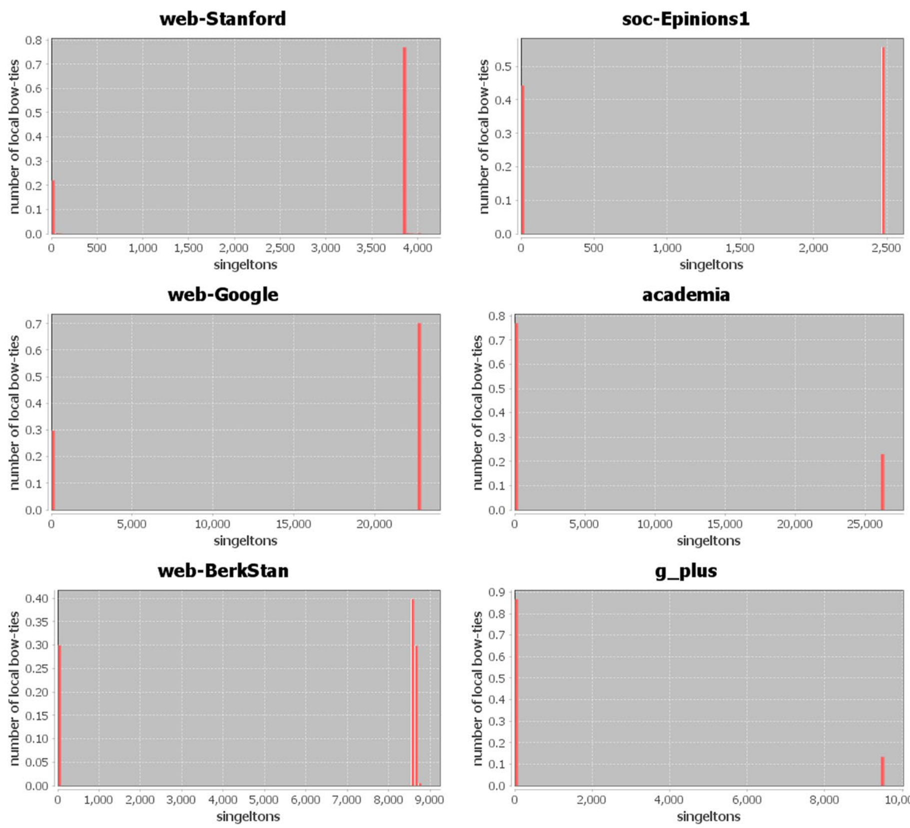

Fig. 17 Distribution of the *singletons* values among *OUT* components for the Web and social networks that we considered

2 – *edge* – *connected pairs* and *singletons* among all the components of their microscopic local bow-tie structures. Our results show how average *PageRank*, *Outdegree*, and *Indegree* centralities change in the structure of the Web and of social networks and demonstrate the dependency of the centrality measures on the 2-connected structure of the Web and of social networks. Our results also show the difference in the structures of Web and of social networks, proving the structure being one of the fundamental characteristics of the network.

Our study is the first to apply the notions of 2-connectivity to analyze microscopic structure of Web graphs and of social networks on the local level. We believe that carrying out this analysis on more types of networks such as biological networks will shed more light on their structural properties and will help to see structural differences between different types of networks.

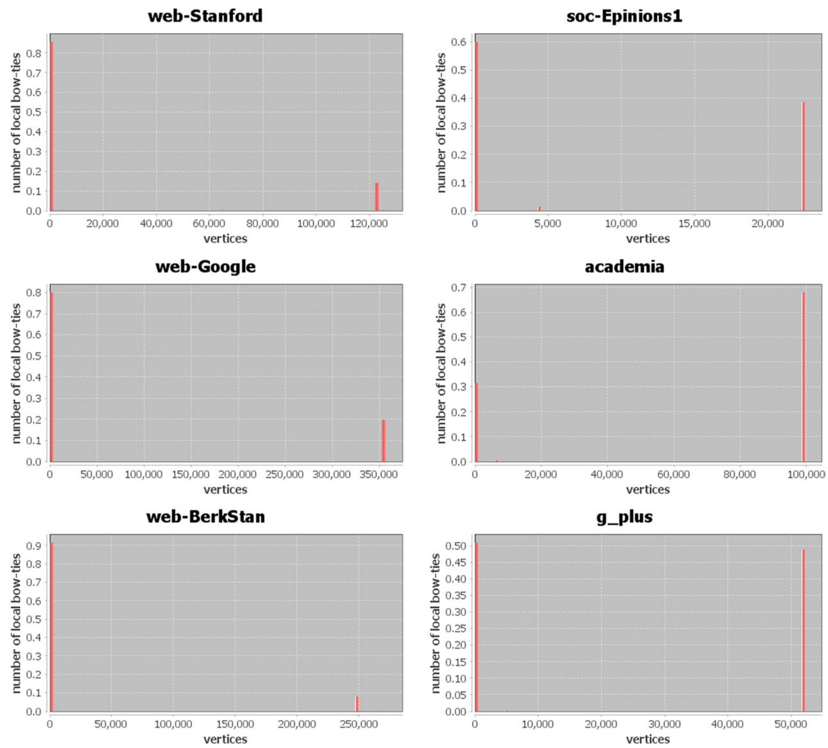

Fig. 18 Distribution of the *vertices* among *IN* components for the Web and social networks that we considered

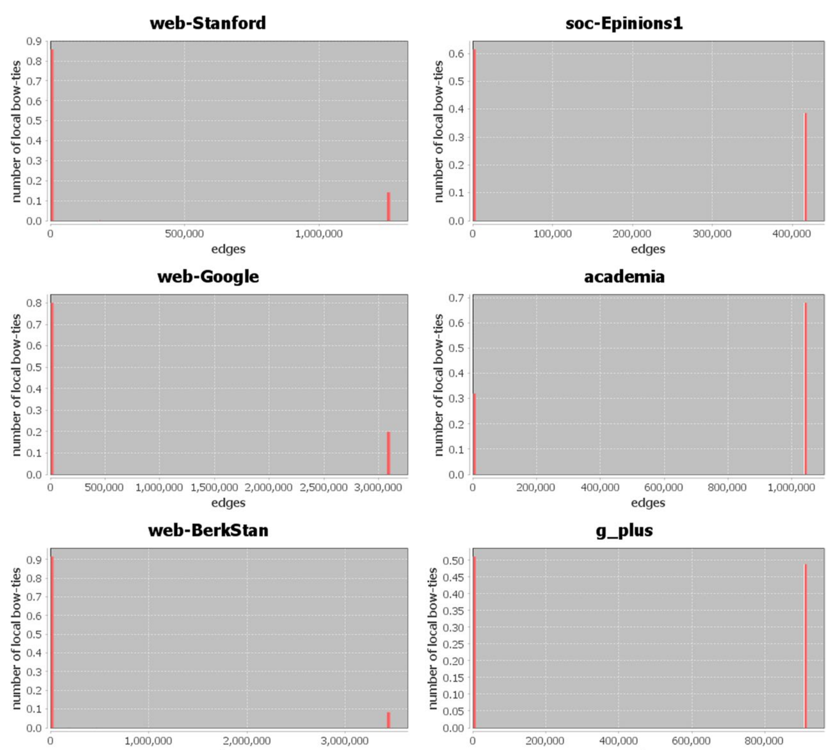

Fig. 19 Distribution of the *edges* among *IN* components for the Web and social networks that we considered

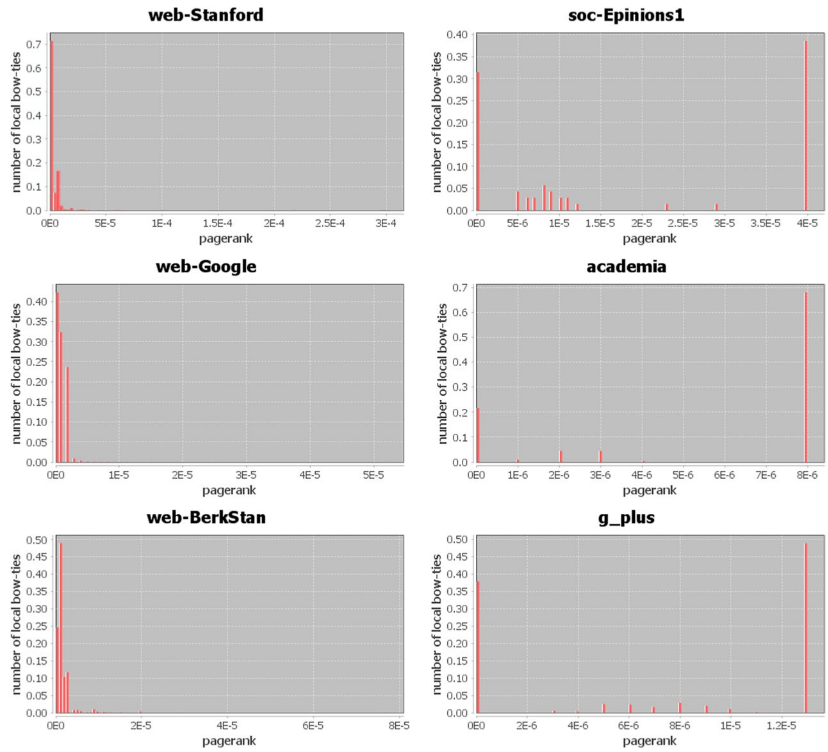

Fig. 20 Distribution of the *PageRank* values among *IN* components for the Web and social networks that we considered

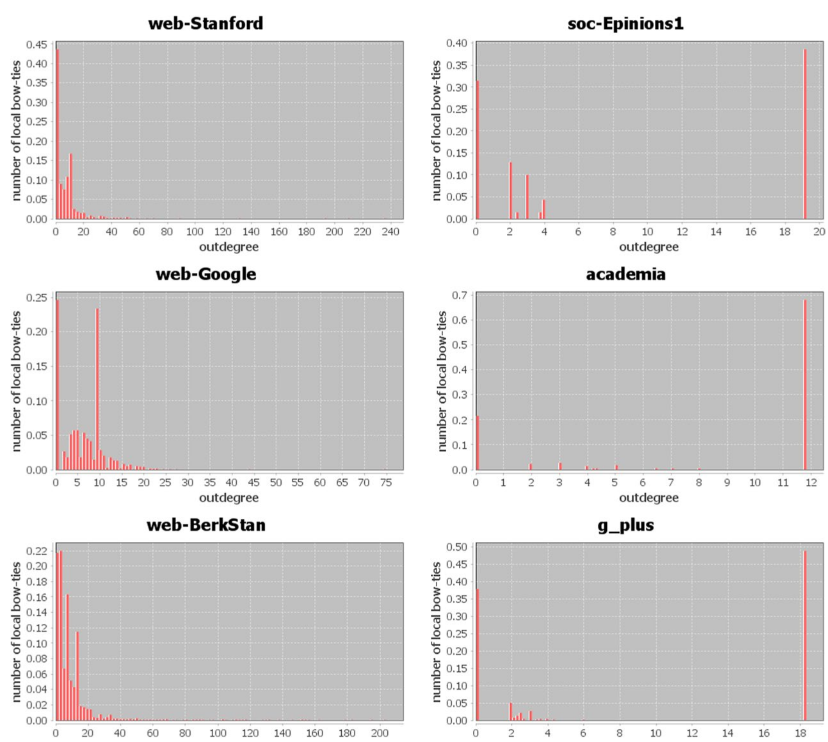

Fig. 21 Distribution of the *Outdegree* values among *IN* components for the Web and social networks that we considered

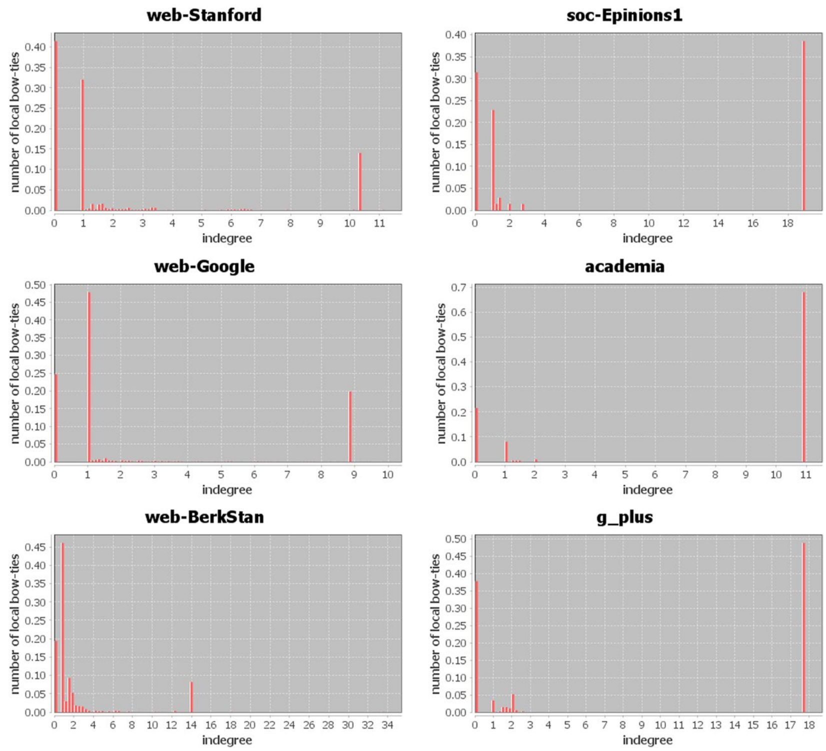

Fig. 22 Distribution of the *Indegree* values among *IN* components for the Web and social networks that we considered

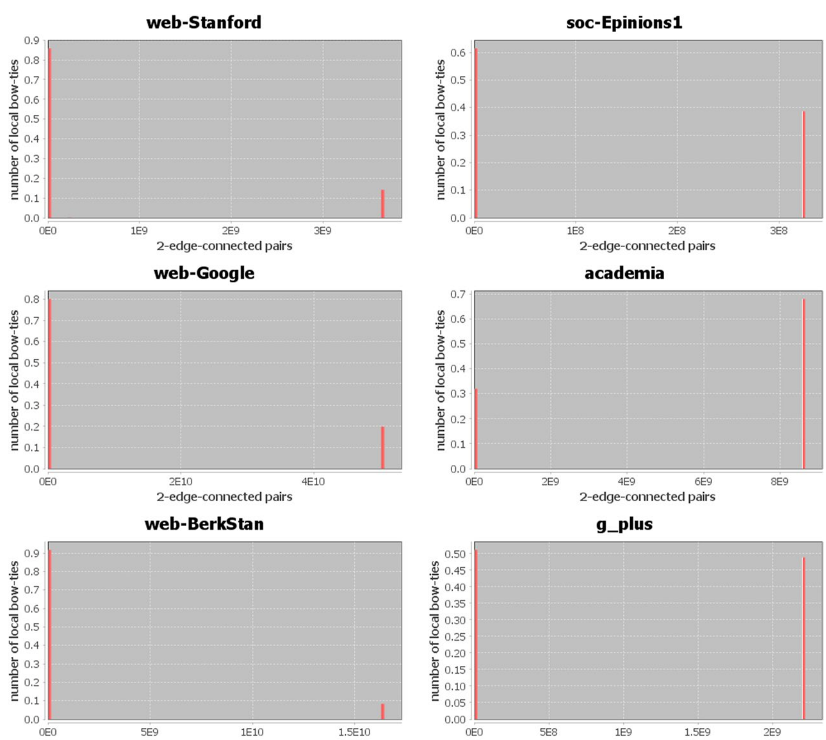

Fig. 23 Distribution of the $2-edge-connected\ pairs$ among *IN* components for the Web and social networks that we considered

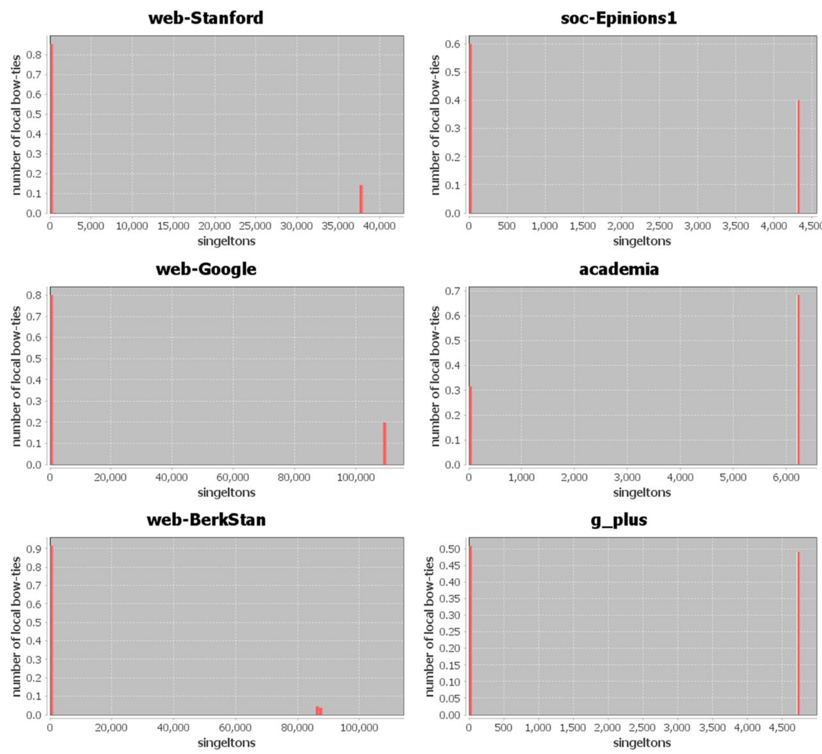

Fig. 24 Distribution of the *singletons* among *IN* components for the Web and social networks that we considered

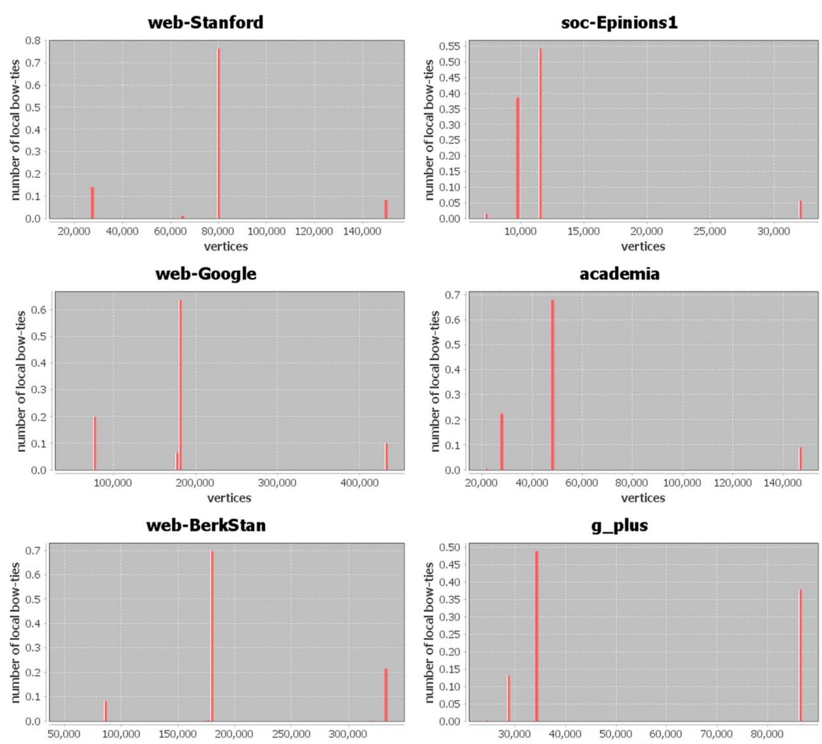

Fig. 25 Distribution of the *vertices* among *OTHER* components for the Web and social networks that we considered

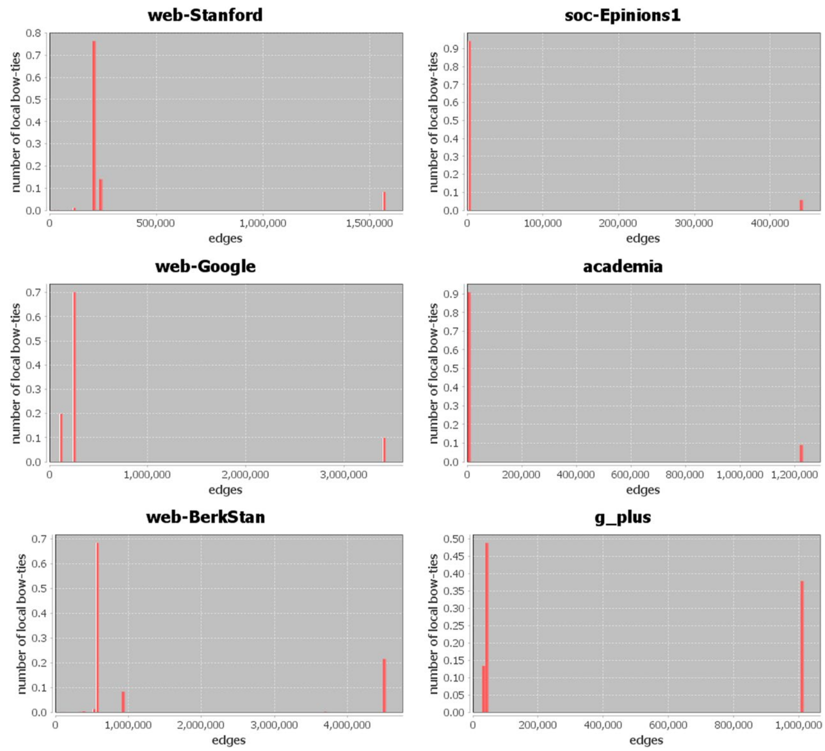

Fig. 26 Distribution of the *edges* among *OTHER* components for the Web and social networks that we considered

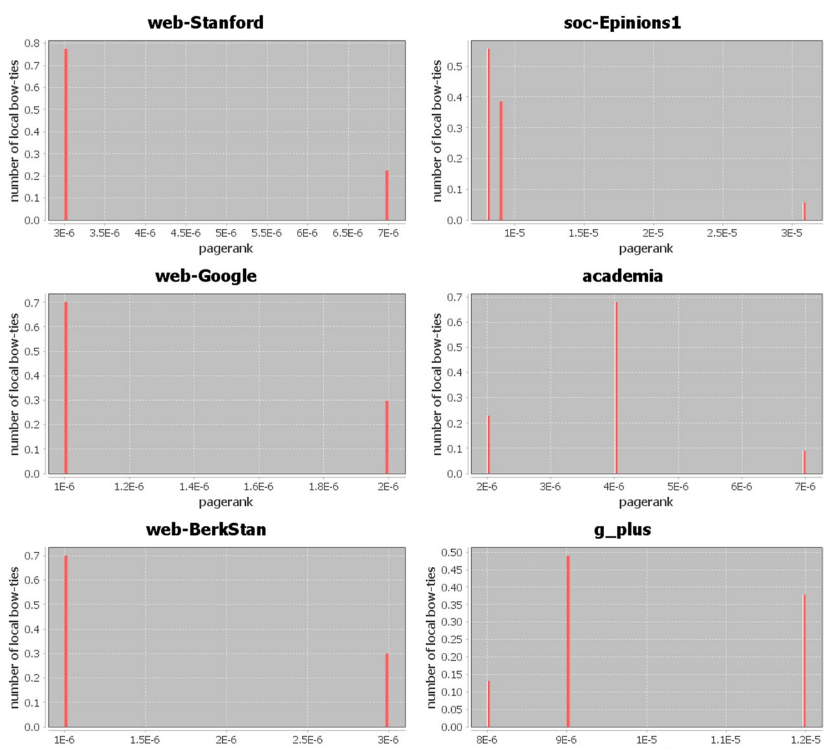

Fig. 27 Distribution of the *PageRank* values among *OTHER* components for the Web and social networks that we considered

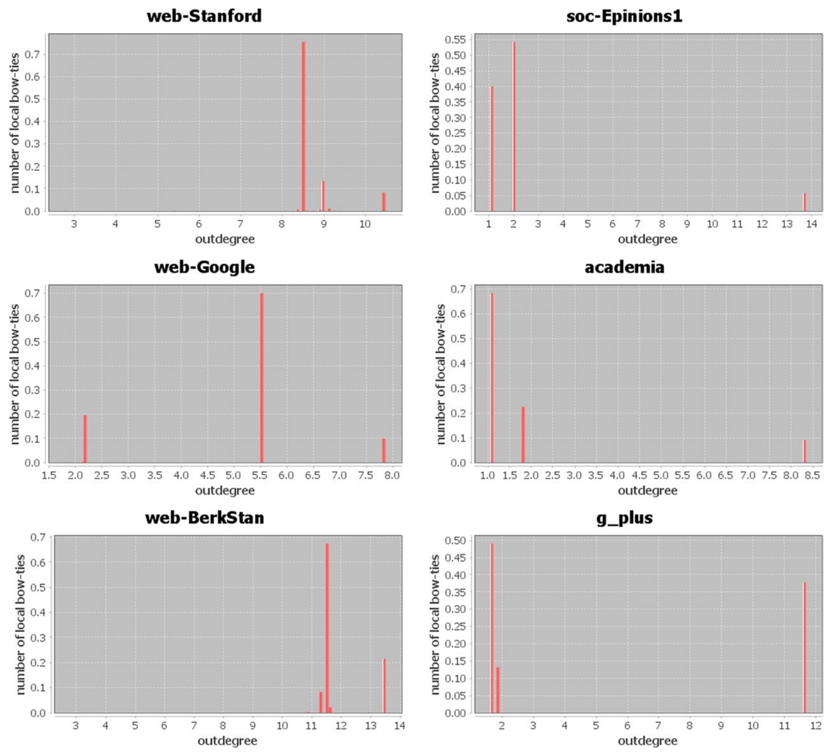

Fig. 28 Distribution of the *Outdegree* values among *OTHER* components for the Web and social networks that we considered

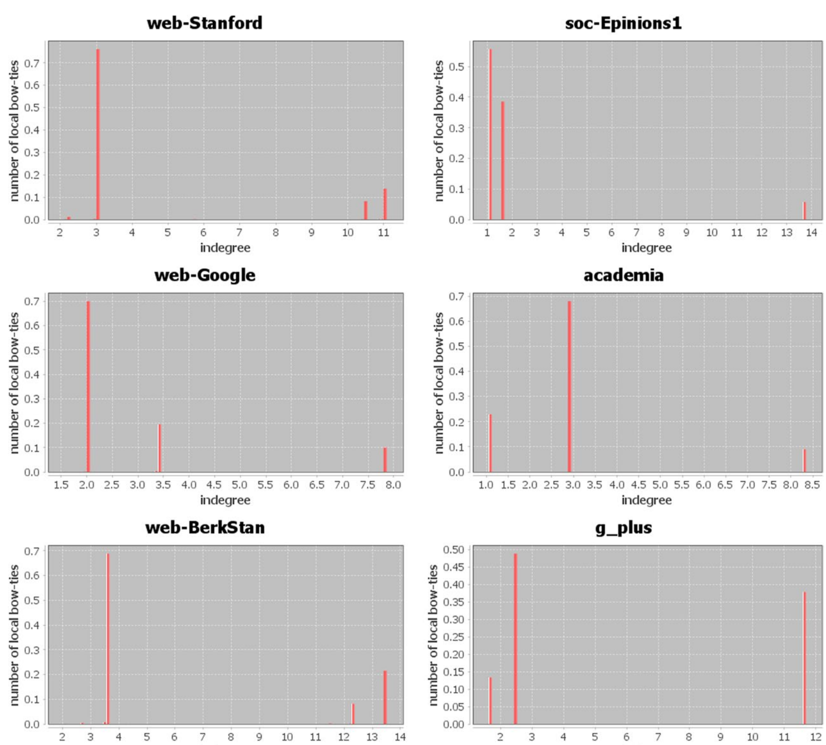

Fig. 29 Distribution of the *Indegree* values among *OTHER* components for the Web and social networks that we considered

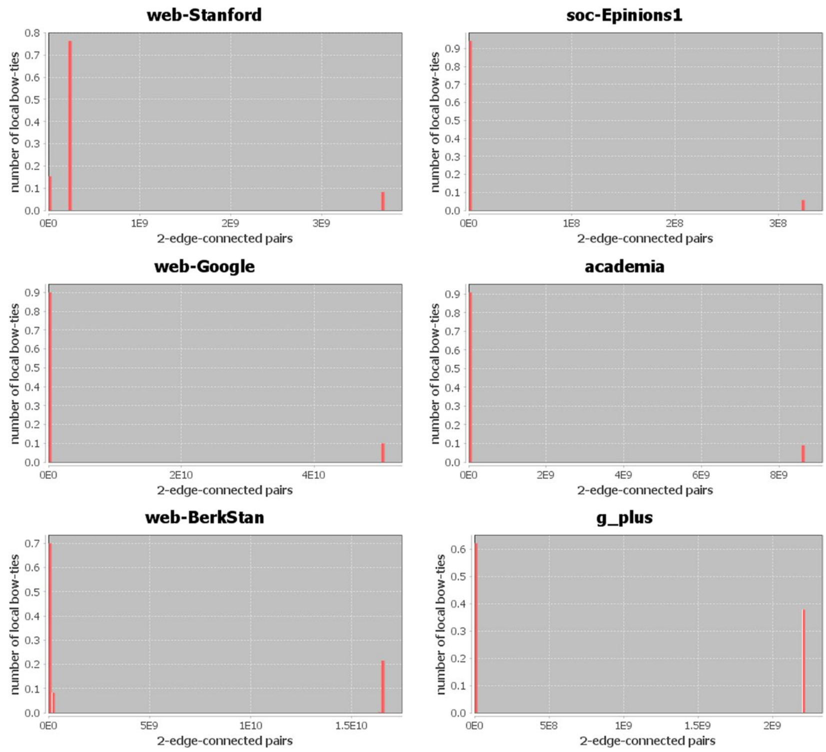

Fig. 30 Distribution of the 2 — *edge* — *connected pairs* among *OTHER* components for the Web and social networks that we considered

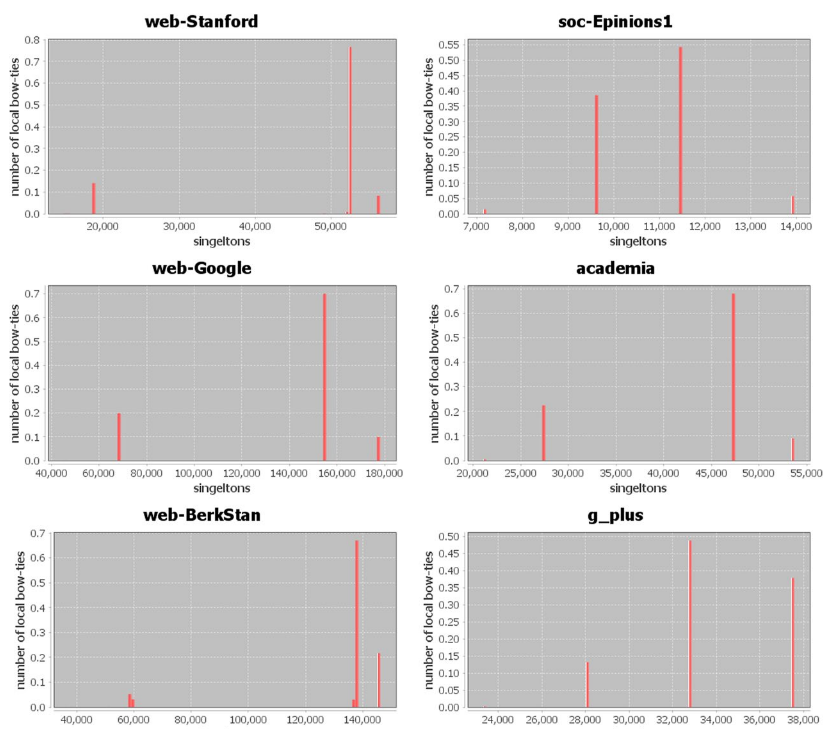

Fig. 31 Distribution of the *singletons* among *OTHER* components for the Web and social networks that we considered

Finally, in the global bow-tie structure with respect to the strongly connected components of a digraph, there exist three additional sets of vertices referred as *OUT-TENDRILS*, *IN-TENDRILS*, and *TUBES*. These are the vertices that have a path to a vertex in *OUT*, a path from a vertex in *IN*, and both a path from a vertex in *OUT* and to a vertex in *IN*, respectively. Although these sets can be naturally extended to fit the structure of 2-connectivity, it is not yet known whether these sets can be computed efficiently (i.e., in linear time), which is a crucial bottleneck when dealing with large scale graphs. We leave as a last open problem whether these sets can be computed in linear time.

Analysis conducted in our study is valid for any strongly connected component of any directed graph.

Abbreviations

G	Directed graph
V	Set of the vertices of a directed graph *G*
E	Set of the edges of a directed graph *G*
SCC	Strongly connected component of a directed graph *G*
2*ECB*(*G*)	Two-edge-connected blocks of a directed graph *G*
2*VCB*(*G*)	Two-vertex-connected blocks of a directed graph *G*
LSCC	Largest strongly connected component of a directed graph *G*
OUT	*OUT* component of a macroscopic bow-tie structure of a directed graph *G*
IN	*IN* component of a macroscopic bow-tie structure of a directed graph *G*
IN−*TENDRILS*	*IN-TENDRILS* component of a macroscopic bow-tie structure of a directed graph *G*
OUT−*TENDRILS*	*OUT{-}TENDRILS* component of a macroscopic bow-tie structure of a directed graph *G*
TUBES	*TUBES* component of a macroscopic bow-tie structure of a directed graph *G*
DISCONNECTED	*DISCONNECTED* component of a macroscopic bow-tie structure of a directed graph *G*
DISC.	*DISCONNECTED* component of a macroscopic bow-tie structure of a directed graph *G*
L2ECB	Largest two-edge-connected block of the largest strongly connected component of a directed graph *G*
2*ECBs*	Two-edge-connected blocks of the largest strongly connected component of a directed graph *G*
2*ECB*	Two-edge-connected block of the largest strongly connected component of a directed graph *G*
2*E*−*IN*	Component *IN* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
2*E*−*OUT*	Component *OUT* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
2*E*−*OTHER*	Component *OTHER* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
2*E*−*DISC*.	Component *OTHER* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
IN	Component *IN* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
OUT	Component *OUT* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
OTHER	Component *OTHER* of a microscopic bow-tie structure of the largest strongly connected component of a directed graph *G*
SNAP	Stanford Network Analysis Project

Acknowledgements

I dedicate this work to God, Saint Nicholas, Saint Sergius, Saint Luka Voyno- Yasenetsky, Saint Eugenia, Saint Ksenia and my parents.

Author contributions

Major contributor—EP.

Funding

Open access funding provided by Northeastern University Library This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Availability of data and materials

The data that support the findings will be available in repository 'StructuralNetworkAnalysis' at https://github.com/EugeniyaP/StructuralNetworkAnalysis.git from the date of publication.

Declarations

Competing interests

The authors declare that they have no Conflict of interest

Received: 15 December 2023 Accepted: 22 June 2024
Published online: 09 July 2024

References

Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, Wiener J (2000) Graph structure in the web. Comput Netw 33(1):309–320

Carrière SJ, Kazman R (1997) Webquery: searching and visualizing the web through connectivity. Comput Netw ISDN Syst 29(8):1257–1267

Donato D, Leonardi S, Millozzi S, Tsaparas P (2005) Mining the inner structure of the web graph. In: WebDB, pp 145–150

Fire M, Tenenboim-Chekina L, Puzis R, Lesser O, Rokach L, Elovici Y (2013) Computationally efficient link prediction in a variety of social networks. ACM Trans Intell Syst Technol 5(1):10

Fujita Y, Kichikawa Y, Fujiwara Y, Souma W, Iyetomi H (2019) Local bow-tie structure of the web. Appl Netw Sci 4(1):15

Gabielkov M, Rao A, Legout A (2014) Studying social networks at scale: macroscopic anatomy of the twitter social graph. In: ACM SIGMETRICS performance evaluation review, vol 42, pp 277–288. ACM

Georgiadis L, Italiano GF, Karanasiou A, Parotsidis N, Paudel N (2018) Computing 2-connected components and maximal 2-connected subgraphs in directed graphs: an experimental study. In: 2018 Proceedings of the twentieth workshop on algorithm engineering and experiments (ALENEX), pp 169–183. SIAM

Italiano GF, Parotsidis N, Perekhodko E (2017) What's inside a bow-tie: analyzing the core of the web and of social networks. In: Proceedings of the 2017 international conference on information system and data mining, pp 39–43

Kamvar S, Haveliwala T, Manning C, Golub G (2003) Exploiting the block structure of the web for computing PageRank. Stanford University Technical Report

Leskovec J, Sosič R (2014) SNAP: a general purpose network analysis and graph mining library in C++. http://snap.stanford.edu/snap

Meusel R, Vigna S, Lehmberg O, Bizer C (2014) Graph structure in the web—revisited: a trick of the heavy tail. In: Proceedings of the companion publication of the 23rd international conference on World Wide Web Companion, pp 427–432. International World Wide Web Conferences Steering Committee

Richardson M, Agrawal R, Domingos P (2003) Trust management for the semantic web. In: The Semantic Web-ISWC 2003, pp 351–368. Springer

Serrano M, Maguitman A, Boguñá M, Fortunato S, Vespignani A (2007) Decoding the structure of the www: a comparative analysis of web crawls. ACM Trans Web (TWEB) 1(2):10

Zhu JJ, Meng T, Xie Z, Li G, Li X (2008) A teapot graph and its hierarchical structure of the Chinese web. In: Proceedings of the 17th international conference on world wide web, pp 1133–1134. ACM

Publisher's Note