Evaluation of the statistical significance

Network	I&A	I&B	I&C	I&D	I&E	I&F	I&G	I&H
DP	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$
YT	$< 2.2^{-16}$	$< 2.2^{-16}$	4.1×10^{-13}	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$
LJ	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	-	$< 2.2^{-16}$	$< 2.2^{-16}$	3.2×10^{-13}
OR	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	$< 2.2^{-16}$	-	$< 2.2^{-16}$	$< 2.2^{-16}$	2.5×10^{-9}

Table 1. p-values of the differences between OPT-R-RDFS and any other baseline. A-I are MDC, QDC, GrCon, LCTC, DCPC, HK, LLSA, MRW, and OPT-R-RDFS respectively. Given a method X, the numbers in the column of I&X are the p-values of the differences between I and X. For example, 2.5×10^{-9} is the p-value of the difference between OPT-R-RDFS and MRW in Orkut.

Using the well-established Wilcoxon signed rank test [1] which is also evaluated in [2], we evaluate the p-value of each difference between OPT-R-RDFS (the proposed method) and any other baseline in each network based on the results in Table 3 of the original paper.

DCPC is an index-based method. As the index size of LJ and that of OR are too large to load into memory, only the p-values of the differences between OPT-R-RDFS and DCPC on DP and YT are reported.

Table 1 shows the p-values. Most of the p-values are smaller than 2.2×10^{-16}. The largest p-value is 2.5×10^{-9}. As all the p-values are much smaller than 0.01, each difference between OPT-R-RDFS and any other baseline in each network is statistically significant (A result is statistically significant if its p-value is lower than 5% [3]).

References

1. Janez Demsar: Statistical Comparisons of Classifiers over Multiple Data Sets. Journal of Machine Learning Research 7: 1-30 (2006)
2. Nicola Barbieri, Francesco Bonchi, Edoardo Galimberti, Francesco Gullo: Efficient and effective community search. Data Min. Knowl. Discov. 29(5): 1406-1433 (2015)
3. Regina Nuzzo: Scientific method: Statistical errors. Nature 506(7487): 150-152 (2014)