1. Furnas G W, Landauer T K, Gomez L M, et al. The vocabulary problem in human-system communication[J]. Communications of the ACM, 1987,30(11):964-971.
2. Haiduc S, Bavota G, Marcus A, et al. Automatic query reformulations for text retrieval in software engineering[C]. In proceedings of the 2013 International Conference on Software Engineering, 2013:842-851.
3. YANG Jinqiu, LIN Tan. Swordnet: Inferring semantically related words from software context[J]. Empirical Software Engineering, 2014,19(6):1856-1886.
4. Tian Y, Lo D, Lawall J. Sewordsim: software-specific word similarity database[C]. In proceedings of the 36th International Conference on Software Engineering, 2014:568-571.
5. Sando. Sando code search tool[EB/OL]. [2016-01-05]. http://sandosearch.weebly.com/.
6. Aragon Consulting Group, Inc. Krugle code search[EB/OL]. [2016-01-05]. http://www.krugle.com/.
7. Linstead E, Bajracharya S, Ngo T, et al. Sourcerer: mining and searching internet-scale software repositories[J]. Data Mining and Knowledge Discovery, 2009,18(2):300-336.
8. Merriam-Webster, Incorporated. Merriam-webster english dictionary and thesaurus[EB/OL]. [2016-01-05]. <http://www.merriam-webster.com/>.
9. Princeton University. WordNet[EB/OL]. [2016-01-05]. http://wordnet.princeton.edu/.
10. Sridhara G, Hill E, Pollock L, et al. Identifying word relations in software: A comparative study of semantic similarity tools. In proceedings of the 16th IEEE International Conference on Program Comprehension, 2008:123-132.
11. Shepherd D, Pollock L, Vijay-Shanker K. Towards supporting on-demand virtual remodularization using program graphs[C]. In proceedings of the 5th international conference on Aspect-oriented software development, 2006:3-14.
12. Shepherd D, FRY Z P, HILL E, et al. Using natural language program analysis to locate and understand action-oriented concerns[C]. In proceedings of the 6th International Conference on Aspect-oriented Software Development, 2007:212-224.
13. Hill E. Integrating natural language and program structure information to improve software search and exploration[Dissertation]. University of Delaware, 2010.
14. Hill E, Pollock L, Vijay-Shanker K. Improving source code search with natural language phrasal representations of method signatures[C]. In proceedings of the 26th IEEE/ACM International Conference on Automated Software Engineering, 2011:524-527.
15. YANG Jinqiu, LIN Tan. Inferring semantically related words from software context[C]. In proceedings of the 9th IEEE Working Conference on Mining Software Repositories, 2012:161-170.
16. Howard M J, Gupta Samir, Pollock L, et al. Automatically mining software-based, semantically-similar words from comment-code mappings[C]. In proceedings of the 10th Working Conference on Mining Software Repositories, 2013:377-386.
17. Wang S, Lo D, Jiang L. Inferring semantically related software terms and their taxonomy by leveraging collaborative tagging[C]. In proceedings of the 28th IEEE International Conference on Software Maintenance, 2012:604-607.
18. Tian Y, Lo D, Lawall J. Automated construction of a software-specific word similarity database[C]. In proceedings of the 2014 Software Evolution Week - IEEE Conference on Software Maintenance, Reengineering, and Reverse Engineering, 2014:44-53.
19. Ye X, Shen H, Ma X, et al. From word embeddings to document similarities for

improved information retrieval in software engineering[C]. In proceedings of the 38th International Conference on Software Engineering, 2016:404-415.

1. Word embedding[EB/OL]. [2016-10-27]. https://en.wikipedia.org/wiki/Word\_embedding.
2. Turian J, Ratinov L, Bengio Y. Word representations: a simple and general method for semi-supervised learning[C]. In proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, 2010:384-394.
3. Hinton G E. Learning distributed representations of concepts[C]. In proceedings of the 8th annual conference of the cognitive science society, 1986:1-12.
4. Clark A, Fox C, Lappin S. The Handbook of Computational Linguistics and Natural Language Processing[M]. Wiley-Blackwell, 2010:74-104.
5. 王晓龙，关毅. 自然语言处理[M]. 北京：清华大学出版社，2005：47-65.
6. Bengio Y, Ducharme R, Vincent P, et al. A neural probabilistic language model[J]. Journal of Machine Learning Research. 2003, 3:1137-1155.
7. Mikolov T, Kombrink S, Deoras A, et al. Rnnlm-recurrent neural network language modeling toolkit[C] In proceedings of the 2011 IEEE Workshop on Automatic Speech Recognition & Understanding, 2011:196-201.
8. Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[J]. arXiv preprint arXiv:1301.3781, 2013.
9. Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C]. In proceddings of the 2013 Neural Information Processing Systems. 2013:3111-3119.
10. Morin F, Bengio Y. Hierarchical Probabilistic Neural Network Language Model[C]. In proceedings of the international workshop on artificial intelligence and statistics, 2005:246-252.
11. Information retrieval[EB/OL]. [2016-10-17]. https://en.wikipedia.org/wiki/Word\_embedding.
12. Standard Boolean model[EB/OL]. [2016-10-01]. https://en.wikipedia.org/wiki/Standard\_Boolean\_model.
13. Vector space model[EB/OL]. [2016-10-17]. https://en.wikipedia.org/wiki/Vector\_space\_model.
14. Extended Boolean model [EB/OL]. [2016-10-19]. https://en.wikipedia.org/wiki/Extended\_Boolean\_model.
15. MSDN Blogs[EB/OL]. [2016-01-05]. https://blogs.msdn.microsoft.com/.
16. Bing Search[EB/OL]. [2016-01-05]. http://global.bing.com/.
17. Stack Overflow[EB/OL]. [2016-01-05]. http://stackoverflow.com/.
18. The Porter Stemming Algorithm[EB/OL]. [2016-01-05]. https://tartarus.org/martin/PorterStemmer/.
19. word2vec[EB/OL]. [2016-01-05]. https://code.google.com/archive/p/word2vec/.
20. List of English Stop Words[EB/OL]. [2016-01-05]. http://xpo6.com/list-of-english-stop-words/.
21. Maaten L V D. Accelerating t-sne using tree-based algorithms[J]. The Journal of Machine Learning Research, 2014,15(1):3221-3245.
22. Ko A J, Myers B A, Coblenz M J, et al. An Exploratory Study of How Developers Seek, Relate, and Collect Relevant Information during Software Maintenance Tasks[J]. IEEE Transactions on Software Engineering, 2006:32(12):971-987.
23. Lv F, Zhang H, Lou G, et al. Codehow: Effective code search based on API understanding and extended boolean model[C]. In proceedings of the 30th International Conference on Automated Software Engineering, 2015: 260-270.
24. Elastic Stack and Product Documentation[EB/OL]. [2016-10-05]. https://www.elastic.co/guide/index.html.
25. Elasticsearch 权威指南[EB/OL]. [2016-10-05]. <http://www.learnes.net/>.
26. JDT[EB/OL]. [2016-10-05]. <https://projects.eclipse.org/projects/eclipse.jdt>.
27. Shepherd D. Action-oriented concerns[EB/OL]. [2016-01-05]. https://www.eecis.udel.edu/~gibson/context/action\_oriented\_concerns.txt.
28. Shepherd D, Damevski K, Ropski B, et al. Sando: an extensible local code search framework[C]. In proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering, 2012:1-2.
29. GitHub[EB/OL]. [2016-01-05]. https://github.com/.
30. Bajracharya S K, Ossher J, Lopes C V. Leveraging usage similarity for effective retrieval of examples in code repositories[C]. In proceedings of the 8th ACM SIGSOFT international symposium on Foundations of software engineering, 2010:157-166.
31. Mihalcea R, Corley C, Strapparava C. Corpus-based and knowledge-based measures of text semantic similarity[C]. In proceedings of the 21st national conference on Artificial intelligence, 2006:775-780.
32. Kim J, Lee S, Hwang S, et al. Towards an intelligent code search engine[C]. In proceedings of the 24th AAAI Conference on Artificial Intelligence, 2010:1358-1363.