

6. CONCLUSIONS

This paper introduces the crawl process of web crawler to obtain recruitment information taking www.lagou.com for example, provide more comprehensive analysis on many aspects which job seekers care much for “big data” position, shows the operation process including data scrubbing, data analysis, data visualization, offer some reference for the majority of interested applicants, reduce the asymmetry of information, make search for jobs more efficiently and more closely match, this paper laid a certain foundation for future research work, the future work should be improved according to the following deviations.

Possible deviations:

1. The limitation of sample size may cause errors to the analysis results because of only 183 samples;

2. This paper crawl the recruitment information of www.langou.com only for the Internet industry, actually, other industries also have demands for big data talents, which may lead to incomplete information.

ACKNOWLEDGMENT

This paper is sponsored by Project No. 2018dfy006 of the key projects of School of Dongfang, Zhejiang University of Finance and Economics.

REFERENCES

[1] Ali Research Institute. (2015). *Internet+ from IT to DT*. Mechanical Industry Press, Beijing.

[2] Wang, T. (2018). Analysis and implementation of recruitment information for software technical personnel based on Python. *Fujian Computer*, 11: 118-119. <https://doi.org/10.16707/j.cnki.fjpc.2018.11.058>

[3] Liu, G.P., Liu, N., Duan, H.Y. (2018). Talent recruitment data collection based on focused web crawler technology. *Computer Programming Skills & Maintenance*, 5: 69-70.

[4] Kumar, M., Bhatia, R., Rattan, D. (2017). A survey of Web crawlers for information retrieval. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(6): 1-8. <https://doi.org/10.1002/widm.1218>

[5] Wang, B.Y. (2017). Research on Python-based web crawler technology. *Digital Technology and Application*, 5: 76.

[6] Madhusudan, P.A., Lambhate, P.D. (2017). Deep web crawling efficiently using dynamic focused web crawler. *International Research Journal of Engineering and*

Technology, 4(6): 3303-3306.

[7] OH, H.J., Won, D.H., Kim, C., Park, S.H., Kim, Y. (2018). Design and implementation of crawling algorithm to collect deep web information for web archiving. *Data Technologies and Applications*, 52(2): 266-277. <https://doi.org/10.1108/DTA-07-2017-0053>

[8] Kim, T.J., Kim, H.J. (2017). Machine learning-based topical web crawler: An ensemble approach incorporating meta-features. *Journal of Engineering and Applied Sciences*, 12(18): 4651-4656. <https://doi.org/10.36478/jeasci.2017.4651.4656>

[9] Guo, L.R. (2017). Programming of web crawler based on Python. *Electronic Technology and Software Engineering*, 23: 248-249.

[10] Xing, L., Wu, M.N. (2018). Design and application of recruitment theme crawler. *Computer Knowledge and Technology*, 14(25): 73-75.

[11] Wu, Y.C. (2019). Discussion on web data grabbing and analyzing with Python crawler technology. *Computer Era*, 8: 94-96. <https://doi.org/10.16644/j.cnki.cn33-1094/tp.2019.08.027>

[12] Chang, F.J., Li, Z.H., Wen, J., Chang, F.J. (2019). The design and implementation of recruitment data crawler using Python. *Software Guide*, 7: 16-17. <https://doi.org/10.11907/rjdk.191156>.

[13] Huang, Q. (2019). Data visualization method and system implementation based on Python. *China Computer & Communication*, 14: 137-140.

[14] Li, P. (2019). Research on Python-based WebCrawler and anti-reptile technology. *Computer & Digital Engineering*, 47(6): 1415-1420. <https://doi.org/10.3969/j.issn.1672-9722.2019.06.028>.

[15] Jia, N.Y. (2019). Job data analysis based on python crawler-take www.lagou.com as an example. *Information Technology and Informatization*, 4: 64-66. <https://doi.org/10.3969/j.issn.1672-9528.2019.04.018>

[16] Li, J.H. (2018). Data analysis based on Python. *Electronic Technology & Software Engineering*, 17: 167.

[17] Wes MCKinney. Xu Jingyi. (2019). *Python for Data Analytics: Data Wrangling with Pandas, NUMPy, and IPython*, China Machine Press, Beijing.

[18] Zhu, Y.Z., Jing, J. (2019). Chinese word segmentation technology based on Python language. *Communications Technology*, 52(7): 1612-1619. <https://doi.org/10.3969/j.issn.1002-0802.2019.07.012>.

[19] Yan, M., Zheng, C.X. (2018). Word segmentation and word cloud production in Python environment. *Modern Computer*, 34: 86-89. <https://doi.org/10.3969/j.issn.1007-1423.2018.34.021>