NLP Techniques for Android Security

Olga Gadyatskaya

University of Luxembourg
olga.gadyatskaya@uni.lu

Thesis Details

Natural Language Processing (NLP) was recently applied to Android security with many promising results [3, 1, 2]. It allows to automatically infer information from text, such as app reviews, or even app code and the manifest file, and to use this information in security analysis. For example, the WHYPER framework automatically identifies risky Android applications that ask for some sensitive permissions but do not justify these permissions usage in the textual description of the app on Google Play [3]. The CHABADA system automatically verifies that textual app description matches actual app behavior [1]. These systems leverage NLP and machine learning techniques in order to enhance security analysis and to assist the analyst and the end-user in taking security decisions.

In this thesis you will design your own NLP system for enhancing Android security analysis that will match textual description of the app in question with its behaviour.

References

1. Gorla, A., Tavecchia, I., Gross, F., Zeller, A.: Checking app behavior against app descriptions. In: Proc. of ICSE. pp. 1025–1035. ACM (2014)
2. Ou, Z., Rastogi, V., Zhang, X., Chen, Y., Zhu, T., Chen, Z.: AutoCog: Measuring the description-to-permission fidelity in Android applications. In: Proc. of CCS. ACM (2014)
3. Pandita, P., Xiao, X., Yang, W., Enck, W., Xie, T.: Whyper: Towards automating risk assessment of mobile applications. In: Proc. of USENIX Security (2013)