# Pengembangan dan Analisis Metode Penilaian Risiko Privasi pada Aplikasi Mobile Berbasis Android Menggunakan Multiple Application Attributes = Development and Analysis Privacy Risk Assessment Method for Android-Based Mobile Applications Using Multiple Application Attributes

R. Ahmad Imanullah Zakariya, author



---

Abstrak

Pengembangan aplikasi yang tidak dilengkapi dengan informasi detail mengenai aspek keamanan aplikasi menyebabkan pengguna mengalami kesulitan untuk menilai dan memahami risiko keamanan privasi yang mereka hadapi, sehingga banyak informasi sensitif yang terungkap tanpa sepengetahuan pengguna. Penelitian ini mengembangkan desain penilaian risiko privasi melalui pendekatan analisis statik dengan memanfaatkan permission dan beberapa atribut aplikasi (multiple application attributes), serta menggunakan majority voting ensemble learning dengan menerapkan teknik pemilihan fitur Random Forest Feature Importance untuk mendeteksi keamanan aplikasi. Nilai risiko diperoleh dari sebuah matriks risiko yang dibentuk dari dua aspek penilaian, yaitu frekuensi terjadinya risiko (likelihood) dan tingkat keparahannya (severity). Penilaian likelihood dilakukan dengan mengkombinasikan prediksi ensemble learning dan atribut aplikasi, sementara penilaian severity berdasarkan pada karakteristik dan jumlah permission. Untuk mengevaluasi model pembelajaran dan desain penilaian risiko privasi digunakan dataset CIC-AndMal2017 yang terdiri dari 2126 file APK. Jumlah data yang digunakan untuk membentuk model memiliki proporsi 80% data training dan 20% data testing, serta metode klasifikasi data yang digunakan adalah binary class (malicious dan benign). Penelitian ini menerapkan bahasa pemrograman Python dan menggunakan parameter default pada proses pembentukan model pembelajaran. Hasil percobaan menunjukkan bahwa model ensemble learning yang dibentuk dari algoritma Decision Tree, K-Nearest Neighbor, dan Random Forest memiliki performa model yang lebih baik dibandingkan single classification model, dengan accuracy sebesar 95.2%, precision 93.2%, dan F1-Score sebesar 92.4%. Penerapan teknik pemilihan fitur mampu meningkatkan efisiensi waktu selama pembelajaran model dengan total waktu sebesar 263 ms. Serta, hasil penilaian risiko mampu memberikan informasi yang komprehensif dan logis mengenai keamanan privasi aplikasi kepada pengguna. Hal ini menunjukkan bahwa desain penilaian risiko yang dibuat dapat menilai aplikasi secara efektif dan objektif.

.....Lack of detailed information about the application's security aspects leads to the user's inability to assess and understand the risk of privacy breaches and leads to the disclosure of a great deal of sensitive information without the user's knowledge. This study proposes a privacy risk assessment development through employing static analysis with permission and multiple application attributes and using majority voting ensemble learning with the Random Forest Feature Importance technique to detect app security. The risk score is obtained from a risk matrix based on two assessment aspects, namely the frequency of risk (likelihood) and its severity. The likelihood assessment is performed by combining ensemble learning predictions and information on multiple application attributes, while the severity assessment is performed by utilizing the number and characteristics of permissions. The dataset CIC-AndMal2017, which consists of 2126 APK files, was used to evaluate learning models and privacy risk assessment design. The amount of

data used to build models consists of 80% data training and 20% data testing, while the data classification method used is binary class (malicious and benign). This study employs Python programming and implements default parameters in building a learning model. The experimental results show that ensemble learning model built from Decision Tree, K-Nearest Neighbor, and Random Forest algorithms provides better model performance than single classification models with accuracy of 95.2%, precision of 93.2%, and F1-Score of 92.4%. By applying feature selection technique, it could improve the efficiency of time used to learn the model with a total time of 263 milliseconds. Moreover, the results of the risk assessment provide comprehensive and rational information about the security of application privacy to users. This shows that the risk assessment design can assess the applications effectively and objectively.