# Cloud Assisted Security Services: F-Secure's View

Alexey Kirichenko

2017-04-21

F-Secure

# CloSer and F-Secure's Strategy

**Data Analysis and Cloud Support for Security Services**

- The recently updated company strategy emphasizes importance of data analytics and cloud technologies for F-Secure's future.
- We have several types of high-quality end-point sensors, large volumes of incoming object data, technologies for extracting static and behavioral features of objects. While the use of Machine Learning and Data Mining techniques is gradually increasing, we are in the beginning of the road.
- Most of our services rely on the cloud support, in particular, object reputation information, threat intelligence, end-point sensor data analysis and correlation.

# App Reputation

**Automated Classification of Android Apps**

- First versions of the APK classifier were built in CloSe.
- To stay competitive with very modest resource allocations, we have to explore ways of achieving highly-trustworthy automated decision making for incoming APK's.
- Key directions:
  - Identification of new meaningful static features; static analysis of embedded code.
  - Analysis methods based on meta-information (prevalence, signing keys, etc.)
  - Optimal ways to utilize classifiers within the overall sample processing stack.

# APK Handling

1. Hash-based ORSP query
2. Local scan with rule-based engines
3. Hash-based query to the Security Cloud, FRS check; if found, the APK is sent to the Security Cloud and analyzed with detection rules and the APK classifier.
4. Otherwise, if allowed, the APK is sent to the Security Cloud and analyzed with detection rules and the APK classifier.

- Near-real-time processing
- Privacy concerns?

# Web Content Reputation

**Browsing Protection and Digital Parenting**

- As of today, our classifiers perform at satisfactory levels only for a few categories.
- Web resources with little text or mostly “non-English” cannot be reliably classified.
- Key directions:
  - Topic modeling-based approaches
  - Image labeling and classification methods
  - Machine Learning-based approaches for detecting phishing sites
  - Optimal ways to utilize classifiers within the overall sample processing stack

# Web Content Analysis

- Many detection and categorization rules.
- A classifier: an SVM for each (source, category) combination, “confidence” values as the output, ELM to ensemble the SVMs’ output.
- Good results for the “Adult” category, decent for a few more.
- Topic modelling (LDA-based) engine, several heuristic choices.
- ~ 40 topics at the moment, rules for mapping those to categories.
- Compared with the APK case, the training sets size and quality are no more than modest.

# Support for Advanced Attack Detection

**F-Secure's Rapid Detection Service (RDS)**

- The detections are now primarily rule- and human expertise-based. We want to complement those with machine intelligence, both for misuse and anomaly detection.
- End-point (or user) profiling may help reduce false alarm rates for Machine Learning-based detection algorithms.
- Protecting RDS mechanisms and confidentiality of locally collected data in cases of partial end-point compromise.