Towards Privacy Preserving Keyword Search via MapReduce

Oleksii Leontiev (歐立思)
9822058
Motivation

- Cloud computing is very popular
- Commonly user just uploads plain data to cloud server
- May be too dangerous for some data (cloud is untrusted)

- Solution: encrypt these data before outsourcing
- But, how to process them then?

Our Basic Goal: encrypt data in such way, so cloud will be able to do **keyword search** (like Google), but will learn nothing more

We want to create model of such cloud using **Hadoop** software and **MapReduce** framework
Scenario

- **Trapdoor** - something that only owner can generate. Server can perform search only if it has trapdoor.
- In our scenario data owner uploads images and their descriptions (tags) to cloud. Descriptions are encrypted in such way, so search can be performed. Like Google Images.

Cloud (=Hadoop) is "honest but curious"
System overview

Basic
1. “Setup & Encrypt” – by Owner
2. “Generate trapdoor” – by Owner
3. “Search” – by Cloud

Optional
1. “Refresh” – by Owner
2. “Retrieve Images” – by Cloud, initiated by User
3. ....
What we did (in Fall):
- Learned papers
- Developed detailed design for system with basic functionality
- Found two theoretical implementations for this design
- Developed prototype

Our plan (for Spring):
Divide project in two independent parts:
- Develop program to extract description from images itself - image analysis
- Develop user (and owner) client for our application - mobile computing