Towards Privacy Preserving Keyword Search via MapReduce

林恩洋 9822020
Alex Leontiev 9822058
Motivation

• Cloud computing is very popular
• It is dangerous to upload plain data (cloud is untrusted)
• Solution: encrypt these data before outsourcing
• But, then we can find the data we want only if we download all of them. Is there any better choice?

• Our Basic Goal: Make the cloud be able to do **keyword search** but learn nothing about the data and search
• We want to create model of such cloud by using **Hadoop** software and **MapReduce** framework
Scenario

- In our scenario data owner uploads images (encrypted) and their descriptions (tags) to cloud. Descriptions are encrypted in such way, so search can be performed. Like Google Images

- Trapdoor - something that only owner can generate. Server can perform search only if it has trapdoor
Possible Attempts

• Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data
  – Critics: Newest paper, so no known critics
• Privacy Preserving Keyword Searches on Remote Encrypted Data
  – Critics: it's computationally expensive, index is big and difficult to handle keyword addition. No trapdoor.
• Secure Indexes by E.J. Goh
  – Critics: not very secure - statistically info is leaked, also they mention that there will be false positive matches, which is bad for mobile users.
• Searchable Public Key Encryption
  – Critics: Computationally expensive
• Searchable Symmetric Key Encryption
  – Critics: Only fixed length words
Current State

• What I have done thus far:
  – Studied papers
  – Established the simulation model of the system
  – Tested and compare SPKE and SSKE (the last two in the previous page)

• What I will finish by next semester?
  – Finish the cloud program
  – Finish the user interface programming – mobile computing
  – Develop program to extract description from images itself - image analysis