Final Project Proposal

Computational Advertising

Group 04

Jang, Jinil (jinil@nyu.edu)
Shin, Soobok (ss6321@nyu.edu)
Heo, Jinyoung (jinyoung@nyu.edu)

Summary of Project

Online advertising is increasing significantly over time. Since the amount of time that people use the internet is getting higher, the online market industry is also growing fast. Thus, suggesting proper and right advertisement is considered highly important factor. To implement Computational advertising system, we can use key words, geographic information, behavioral patterns and so on. With such data, we can publish the advertisements base on these information dynamically. In this project, we will develop online advertising using behavioral patterns and some other information that we have.

One Major Feature We Will Add to Our System

Our program provide advertisement depends on the query, result of the search engine and the click log of the people. Each advertisement have keywords to represent, so this keywords will match with the query. Also, the result of the search engine will use to examine which advertisement is related. Moreover, we'll use click log. With click log, we will get the information such as if some people search specific query tend to interesting certain kind of advertisement or people who likes A advertisement also interested in B.

How The Feature Should Be Tested

First of all, we are going to prepare some virtual advertisements which will be suggested beside of search result page. Without user interaction, system can be judged whether it suggests ads based on given sort of ranking formula correctly or not. Second, keywords of each ad, which are registered advertising system, are going to be adjusted to suggest improper ad first. Suggesting unrelated advertisements in high position could happen in real world pure key words matching system. Then, we will simulate users' interaction with the ads. If we implement system in right way, advertisements suggested by the system is getting better result.

Preliminary Design Description

- Query representation: when the user inputs queries on the search engine, the server compute a query representation based on each advertisements' keywords based on below formula:

$$P(q|A) = \frac{\sum_{a \in A} \#(q, a)}{\sum_{q'} \sum_{a \in A} \#(q', a)} \quad (q : query,\ A : advertise)$$

- Finding relevance advertisement: sponsor provides the set of keywords representing their advertisement. Note that the number of keywords is unlimited. However, we will calculate the relevance as accuracy percentage based on total keywords of each advertisement. Thus if the number of accurate keywords are same, we will show upper the advertisement with less keywords set.

$$Relevance = \frac{\sum_{k \in K} \#(k=q)}{\sum_{k \in K} \#(k)}$$

- Advertise ranking : the server logged the query and click event. That is, when user search some query, and click an advertisement link, the server log the information. Based on this log, we will approach HMM model / Viterbi algorithm.

$$b_j(k) = \prod_{i=1}^{|v_t|} P(w_t^i | q_j)$$

$b_j(k)$ means the edge weight(probability) from a 'query' to 'click one ad link'. With this edge weight, relevance and advertising cost, we compute the rank of advertisements.

$$Rank = (\sum_{q \in Q} b(q)) \times \frac{\sum_{k \in K} \#(k=q)}{\sum_{k \in K} \#(k)} \times (Ad\ cost)$$