

**CPS196.03 Information Management and Mining - Spring  
2009**

**Notes for Internet Advertising and Click-Fraud Detection**

---

**Question 1**

Suppose we have four advertisers who advertise on Google Adwords for “Toyota Camry”. The four advertisers and their details are:

1. MarcJacobsen.com: This Web site is a Durham-based dealer of Toyota cars. They pay 20 cents per click, and have a per-day budget of 100 dollars. Google computes a quality score of 0.8 for this site for a “Toyota Camry” query.
2. BuyFordFusion.com: This Web site is a dealer of Ford cars. Their (negative) advertisement on “Toyota Camry” is based on a recent study that found Ford Fusion to have a better safety score than Toyota Camry. They pay 50 cents per click, and have a per-day budget of 500 dollars. Google computes a quality score of 0.1 for this site.
3. RaleighMazdaAndToyota.com: This Web site is a Raleigh-based dealer of Mazda and Toyota cars. They pay 30 cents per click, and have a per-day budget of 200 dollars. Google computes a quality score of 0.6 for this site.
4. Cars.com: This Web site is an online marketplace for cars. They pay 60 cents per click, and have a per-day budget of 1000 dollars. Google computes a quality score of 0.2 for this site.

Use the above information to answer the following three questions.

- 1.1 What would the rank of each advertisement be for a “Toyota Camry” search query? Give ranks in 1-4, where 1 is the best rank, and 4 is the worst rank. Assume that none of the advertisers have used up their budget for the day.

1.2 If the user who did the above search clicks on the advertisement for MarcJacobsen.com, then how much would MarcJacobsen.com have to pay Google? Again, assume that none of the advertisers have used up their budget for the day.

1.3 Illustrate clearly the two types of click fraud that MarcJacobsen.com can commit in the above scenario. Use the numbers given in the question as needed.

**Question 2** The following questions are based on invalid clicks and Google's efforts to detect them.

1. What is the definition of an invalid click?
2. What is the "Fundamental Problem of the Pay-Per-Click Advertising Model" (also referred to as the fundamental problem of invalid clicks)?
3. List the techniques that Google uses to detect invalid clicks. What additional techniques does Tuzhilin recommend, and how feasible are they?
4. How does Google evaluate how well its techniques for detecting invalid clicks are performing?
5. What are Tuzhilin's recommendations regarding the evaluation of techniques for detecting invalid clicks are performing? (Hint: estimating False Positives, False negatives, etc. on Page 26.) What would it take to implement these recommendations?