Domain Parking Automatic Detection

**Domain parking** refers to the registration of an internet domain name without that domain being associated with any services such as e-mail or a website. This may have been done with a view to reserving the domain name for future development, and to protect against the possibility of cybersquatting.

The parking landing page is usually a single-page website that people see when they enter the domain.

Domain parking can be classified as monetized and non-monetized. In the former, advertisements are shown to visitors and the domain is "monetized". In the latter, an "Under Construction" or a "Coming Soon" message may or may not be put up on the domain by the registrar or reseller.

The main target of this project is to automatically detect domain parking pages.

**Phases**
1. Build a dataset of domain parking pages.
2. Learn how domain parking pages can be detect and find features that can be relevant when trying to classify a web page as parking page or not.
3. Develop an algorithm for classifying a URL as a parking page
4. Test the algorithm

**Deliverables**
1. All data from the research
2. Code for the classifier which detects if a given web page is a parking web page or not.
   a. Input: suspected web page url
   b. Output: whether the web page is domain parking page or not
3. A report:
   a. Summary of all phases of the research
   b. Statistical evidence to support the decisions in the algorithm
   c. Test results
   d. A final recommendation of the feasibility and usefulness of the technique

Alon Freund  
alonf@il.ibm.com

Aviv Ron  
rona@il.ibm.com