

# **Practical Detection of Click Spams Using Classification-Based Algorithms**

**Mahdieh Fallah**

**Sajjad Zarifzadeh**

Department of Electrical and Computer Engineering, Yazd University, Yazd, Iran

## **ABSTRACT**

Most of today's internet services utilize user feedback (clicks) to improve the quality of their services. For example, search engines use click information as a key factor in document ranking. As a result, some websites cheat to get a higher rank by fraudulently absorbing clicks to their pages. This phenomenon, known as "Click Spam", is initiated by programs called "Click Bot". Thus, the problem of distinguishing bot-generated traffic from the user traffic is critical for the viability of Internet services, like search engines. In this paper, we propose a novel classification-based system to effectively identify fraudulent clicks in a practical manner. We first model user sessions as a set of features. Then, we classify user sessions with a one-class classification algorithm that works based on the well-known K-Nearest Neighbor algorithm. Finally, we analyze our methods with the real log of a Persian search engine. Experimental results show that the proposed algorithm can detect fraudulent clicks with a precision of up to 96% which outperform the previous works by more than 5%.

**Keywords:** Click Spam, Bot Detection, Anomaly, Machine Learning.