

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**

**ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ**

**Παχυλάκης Μιχαήλ  
Μεταπτυχιακός Φοιτητής**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης  
Επόπτης Μεταπτ. Εργασίας: Καθηγητής, Ε. Μαρκάτος**

**Παρασκευή, 17/01/2020, 13:00**

**Αίθουσα B106, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης**

**“Ανάλυση μετρήσεων του διαδικτυακού οικοσυστήματος διαφημίσεων  
πονταρίσματος κεφαλίδας”**

#### **ΠΕΡΙΛΗΨΗ**

Τα τελευταία χρόνια, το ποντάρισμα κεφαλίδας (Header Bidding) έχει κερδίσει έδαφος μεταξύ των διαδικτυακών εκδοτών, αλλάζοντας τα δεδομένα του διαδικτυακού διαφημιστικού οικοσυστήματος. Σε αντίθεση με το παραδοσιακό πρότυπο καταρράκτη, το ποντάρισμα κεφαλίδας στοχεύει να δώσει τον έλεγχο των διαφημίσεων πίσω στους εκδότες, αυξάνοντας τη διαφάνεια, τη δικαιοσύνη και τον ανταγωνισμό μεταξύ των διαφημιστών, αποτέλεσμα του οποίου θα είναι οι μεγαλύτερες τιμές για κάθε διαφήμιση. Αν και πολλά υποσχόμενο, λίγα είναι γνωστά για το πώς λειτουργεί αυτό το πρωτόκολλο: Ποιες είναι οι δυνατές υλοποιήσεις του, ποιοι είναι οι μεγάλοι παίκτες, και ποιο είναι το κόστος του στο δίκτυο αλλά και στην εμπειρία πλοήγησης του χρήστη; Σε αυτή την εργασία. Παρουσιάζουμε το HBDetector μια πρωτότυπη μεθοδολογία που ανιχνεύει το ποντάρισμα κεφαλίδας σε μια ιστοσελίδα, σε πραγματικό χρόνο, βασιζόμενο σε διάφορα σήματα. Τέτοια σήματα περιλαμβάνουν τα γεγονότα DOM τα οποία ενεργοποιούνται σε μια ιστοσελίδα από τις βιβλιοθήκες πονταρίσματος

κεφαλίδας και τα διαδικτυακά αιτήματα που στέλνονται στους διαφημιστικούς συντάιρους.

Χρησιμοποιώντας το HBDetector προσπελάσαμε τις κορυφαίες 35,000 ιστοσελίδες με βάση την κατάταξη Alexa, όπου καταφέραμε να μαζέψουμε και να αναλύσουμε δεδομένα από 800,000 δημοπρασίες διαδικτυακών διαφημίσεων. Βασιζόμενοι στα δεδομένα που μαζέψαμε μπορέσαμε να αναγνωρίσουμε τρεις διαφορετικούς τύπους πονταρίσματος κεφαλίδας τους οποίους χρησιμοποιούν οι εκδότες. Επιπλέον, βρήκαμε ότι (i) 14.28% από τις κορυφαίες ιστοσελίδες χρησιμοποιούν το πρωτόκολλο πονταρίσματος κεφαλίδας. (ii) Οι εκδότες προτιμούν να συνεργάζονται με λίγους Συντάιρους Ζήτησης (Demand Partners) οι οποίοι κυριαρχούν ήδη και στην αγορά του πρωτοκόλλου καταρράκτη. (iii) Η καθυστέρηση από το πρωτόκολλο κεφαλίδας μπορεί να είναι σημαντικά πιο υψηλή (έως και 3 φορές πάνω) σε σχέση με αυτή του πρωτοκόλλου καταρράκτη.

Σε αυτή την εργασία, παρουσιάσαμε το σχεδιασμό και την υλοποίηση του HBDetector, και εκτελέσαμε την πρώτη σε βάθος ανάλυση του διαδικτυακού οικοσυστήματος πονταρίσματος κεφαλίδας. Παραθέτουμε μια λεπτομερή ανάλυση σχετικά με το πώς αυτό το νέο πρωτόκολλο λειτουργεί, πως μπορεί να ανιχνευτεί και ρίξαμε φως σχετικά με τις λεπτομέρειες του στο διαδίκτυο.

**Michalis Pachylakis**

**M.Sc. Thesis**

**Computer Science Department**

**University of Crete**

**Master's Thesis Supervisor: Professor, E. Markatos**

**Friday, 17/01/2020, 13:00**

**Room B106, Computer Science Dept., University of Crete**

**“A measurement study of the online header bidding ad-ecosystem”**

## **ABSTRACT**

In recent years, Header Bidding has gained popularity among web publishers, challenging the status quo in the ad ecosystem. Contrary to the traditional waterfall standard, Header Bidding aims to give back to publishers control of their ad inventory, increase transparency, fairness and competition among advertisers, resulting in higher ad-slot

prices. Although promising, little is known about how this ad protocol works: What are Header Bidding's possible implementations, who are the major players, and what is its network and UX overhead?

In this thesis, we present HBDetector, a novel methodology to detect Header Bidding auctions on a website in real time based on different signals. Those signals include the DOM events that are being triggered in a webpage by the Header Bidding libraries and web requests that are being sent to the advertising partners. Using HBDetector we crawled the 35,000 top Alexa websites, where we managed to collect and analyze a dataset of 800,000 auctions. Based on the data collected we were able to identify three different facets of Header Bidding currently used by the publishers. Furthermore, we find that: (i) 14.18% of the top websites utilize Header Bidding. (ii) Publishers prefer to collaborate with a few Demand Partners who also dominate the waterfall market. (iii) Header Bidding latency can be significantly higher (up to 3x in median cases) than waterfall.

In this thesis, we present the design and implementation of HBDetector, and conduct the first in depth analysis of the Header Bidding advertising ecosystem. We provide a detailed analysis on how this new advertising standard works, how it can be detected and we shed light on its mechanics on the web.