Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών *Ασφάλεια Υπολογιστών και Δικτύων*

Μεταπτυχιακή Διατριβή

Συμπεριφορική διαφήμιση και δημιουργία προφίλ στο διαδίκτυο

Χρήστος Μπαρμπούδης

Επιβλέπων Καθηγητής
Κωνσταντίνος Λιμνιώτης

Μάιος 2018

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών *Ασφάλεια Υπολογιστών και Δικτύων*

Μεταπτυχιακή Διατριβή

Συμπεριφορική διαφήμιση και δημιουργία προφίλ στο διαδίκτυο

Μπαρμπούδης Χρήστος

Επιβλέπων Καθηγητής
Κωνσταντίνος Λιμνιώτης

Η παρούσα μεταπτυχιακή διατριβή υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων για απόκτηση μεταπτυχιακού τίτλου σπουδών στην Ασφάλεια Υπολογιστών και Δικτύων από τη Σχολή Θετικών και Εφαρμοσμένων Επιστημών του Ανοικτού Πανεπιστημίου Κύπρου.

Μάιος 2018

Περίληψη

Αντικείμενο της παρούσας διατριβής αποτελεί η συμπεριφορική διαφήμιση και η δημιουργία προφίλ χρηστών στο διαδίκτυο. Η συμπεριφορική διαφήμιση βασίζεται στα δεδομένα που αντλούνται από την διαδικτυακή περιήγηση του χρήστη και επιτρέπει την προβολή διαφημίσεων, οι οποίες αντανακλούν τα προσωπικά ενδιαφέροντα του. Για να επιτευχθεί αυτό, έχουν ανακαλυφθεί διάφοροι τρόποι παρακολούθησης του χρήστη κατά την διάρκεια της περιήγησης τους στο διαδίκτυο, κάποιους από τους οποίους οι χρήστες δεν γνωρίζουν την ύπαρξη τους.

Ως μέσο αντιμετώπισης αυτού, παρατηρείται μια αύξηση από εφαρμογές στα κινητά τηλέφωνα με σκοπό τον αποκλεισμό διαφημίσεων (Ad Blocking apps),οι πάροχοι των οποίων ισχυρίζονται ότι παρέχουν ασφαλή και ιδιωτική περιήγηση.

Αντικείμενο της παρούσας διατριβής είναι η μελέτη τεχνικών για την συμπεριφορική διαφήμιση και τη δημιουργία προφίλ χρηστών, με έμφαση στο κατά πόσον οι εφαρμογές αποκλεισμού διαφημίσεων πράγματι προστατεύουν τα δεδομένα των χρηστών τους. Για το εν λόγω ερευνητικό ερώτημα, μελετήθηκαν πέντε δημοφιλείς τέτοιες εφαρμογές μέσω δυναμικής ανάλυσης τους σε περιβάλλον Android.

Τα αποτελέσματα της διατριβής καταδεικνύουν ότι παρόλο που οι εφαρμογές αυτές αποκλείουν τις διαφημίσεις, εν τέλει δεν προστατεύουν τα προσωπικά δεδομένα των χρηστών και σε ορισμένες περιπτώσεις επιτρέπουν σε τρίτες οντότητες να συλλέγουν πληροφορίες των χρηστών εν αγνοία τους.

Λέξεις κλειδιά: μέθοδοι παρακολούθησης, αποτυπώματα, εφαρμογές Ad blocking, ιδιωτικότητα, real time bidding

Summary

This thesis studies the behavioral advertising and the creation of the user's profiles on the internet. Behavioral advertising is based on the data derived from the user's web browsing and allows for the display of the advertisements that reflects his or her personal interests. To achieve this, has been discovered various ways of tracking the user's during their internet browsing, some of which user's are unaware of their existence.

As a means of addressing this, there is an increase in ad blocking apps, whose providers claim to provide safe and private browsing.

The purpose of this thesis is to study techniques for behavioral advertising and the creation of user's profiles, with an emphasis on whether the ad blocking applications actually protect their user's data. For this research question five popular such application has been studied through their dynamic analysis in Android.

The results of the thesis show that although these applications block adds they do not protect user's personal data and in some cases allow third parties entities to gather user information without their knowledge.

Key words: tracking mechanisms, fingerprints, ad blocking apps, privacy, real time bidding

Ευχαριστίες

Θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα καθηγητή μου Κωνσταντίνο Λιμνιώτη για την πολύτιμη βοήθεια και καθοδήγηση σε όλη την διάρκεια υλοποίησης της υπάρχουσας εργασίας.

Επίσης θα ήθελα να ευχαριστήσω την οικογένεια μου για όλη την στήριξη και υπομονή που έδειξε καθ’ όλη την διάρκεια των μεταπτυχιακών μου σπουδών.

Περιεχόμενα

Κεφάλαιο 1
Εισαγωγή

Τα τελευταία χρόνια, έχει παρατηρηθεί το φαινόμενο της παρακολούθησης χρηστών στο διαδίκτυο και απόκτησης όσο το δυνατόν περισσότερων πληροφοριών που αφορούν τους χρήστες. Αυτές οι πληροφορίες συγκεντρώνονται από τις εταιρείες με στόχο την βελτίωση της απόδοσης των ιστότοπων καθώς επίσης, όπως αναφέρουν οι ίδιες εταιρείες, για ευκολότερη πλοήγηση των χρηστών. Όμως το κύριο χαρακτηριστικό της συλλογής δεδομένων και παρακολούθησης των χρηστών δεν αφορά την βελτίωση. Σε πολλές περιπτώσεις παρατηρούμε διαρροή των πληροφοριών προς τρίτες εταιρείες (third party trackers) για σκοπούς συμπεριφορικής διαφήμισης. Οι εταιρείες παρακολουθούν όλο και περισσότερο την διαδικτυακή συμπεριφορά των χρηστών και χρησιμοποιούν τις πληροφορίες που συλλέγονται για να προβάλουν στους χρήστες στοχευμένες διαφημίσεις, με βάση την προτίμηση τους. Το φαινόμενο αυτό ονομάζεται **συμπεριφορική διαφήμιση** (**Behavioral Advertising)** και τα τελευταία χρόνια αποτελεί την κύρια στρατηγική πολλών εταιρειών παγκοσμίως. Παρά το γεγονός ότι οι διαφημιστικές εταιρείες μπορούν να επωφεληθούν αρκετά από την συμπεριφορική διαφήμιση, η πρακτική ανταλλαγής προσωπικών δεδομένων μεταξύ των εταιρειών/οργανισμών και η συνεχής παρακολούθηση χρηστών εγείρει ανησυχίες σχετικά με την προστασία της ιδιωτικής ζωής και των προσωπικών δεδομένων των χρηστών.

Σε ένα απλό παράδειγμα της συμπεριφορικής διαφήμισης, ένα διαφημιστικό δίκτυο (δηλ. μια εταιρεία που προβάλλει διαφημίσεις σε χιλιάδες ιστότοπους) παρακολουθεί τις επισκέψεις ιστότοπου ενός καταναλωτή. Εάν ένας καταναλωτής επισκέπτεται αρκετούς ιστότοπους σχετικά με τα αυτοκίνητα, το δίκτυο υποθέτει ότι ο καταναλωτής ενδιαφέρεται για τα αυτοκίνητα. Το δίκτυο μπορεί στη συνέχεια να εμφανίζει διαφημίσεις για αυτοκίνητα μόνο σε άτομα που υποτίθεται ότι ενδιαφέρονται για αυτοκίνητα. Κατά συνέπεια, όταν δύο άτομα επισκέπτονται ταυτόχρονα τον ίδιο ιστότοπο, μπορεί ο ένας εξ αυτών να δει διαφημίσεις αυτοκινήτου ενώ ο άλλος χρήστης (ο οποίος δεν επισκέφθηκε ιστότοπους για αυτοκίνητα αλλά για έπιπλα) μπορεί να δει διαφημίσεις επίπλων [01]. Χαρακτηριστικό είναι πως αυτή η επιλογή διαφημίσεων ανάλογα με τις προτιμήσεις των χρηστών εφαρμόζεται μέσω μιας "**άτυπης" δημοπρασίας** που πραγματοποιείται σε πραγματικό χρόνο στο background, μεταξύ των διαφημιστών, γνωστή και ως **Real Time Bidding** (RTB). Η υποβολή προσφορών σε πραγματικό χρόνο (RTB) είναι ένα αναδυόμενο και ελπιδοφόρο επιχειρησιακό μοντέλο για την ηλεκτρονική υπολογιστική διαφήμιση στην εποχή των μεγάλων δεδομένων. Με βάση την ανάλυση των τεράστιων δεδομένων που παράγονται από χρήστες του Διαδικτύου, η διαφήμιση RTB έχει τη δυνατότητα να εντοπίζει σε πραγματικό χρόνο τα χαρακτηριστικά και τα ενδιαφέροντα του χρήστη-στόχου, προσφέροντας αυτόματα τις καλύτερες αντιστοιχήσεις διαφημίσεων. Η RTB διαφήμιση γνώρισε εκρηκτική ανάπτυξη από τη γέννησή της τα τελευταία χρόνια. Στις διεθνείς αγορές, αναφέρεται ότι 88% των διαφημιζομένων της Βόρειας Αμερικής έχουν μεταβεί στο RTB και το μέγεθος της αγοράς RTB αναμένεται να αυξηθεί στα 8,49 δισ. Δολ. το 2018, αντιπροσωπεύοντας το 29% των προϋπολογισμών διαφημιστικής προβολής [02].

Στην αντίπερα όχθη οι χρήστες, για να προστατευτούν από τον καταιγισμό διαφημίσεων, χρησιμοποιούν προγράμματα για αποκλεισμούς διαφημίσεων - τους λεγόμενους «Ad Blockers». Η χρήση των Ad Blockers αυξάνεται συνεχώς, ειδικότερα στα κινητά τηλέφωνα, όπου έχουμε την εμφάνιση εφαρμογών περιήγησης με εγκατεστημένους Ad Blockers, που παρέχουν αποκλεισμό διαφημίσεων. Επιπρόσθετα οι εταιρείες που παρέχουν τις συγκεκριμένες εφαρμογές τις χαρακτηρίζουν απολύτως ασφαλείς για τον χρήστη και σύμφωνες με τις προϋποθέσεις νομιμότητας αναφορικά με προστασία τον προσωπικών δεδομένων. Ωστόσο ελάχιστες, μέχρι στιγμής, έρευνες που έχουν πραγματοποιηθεί πάνω στους Ad Blockers, έδειξαν ότι εγείρουν βασικά θέματα κατά της ιδιωτικότητας των χρηστών. Στο πλαίσιο αυτό η διατριβή μελετά πέντε πιο γνωστές εφαρμογές Ad Blocking για κινητά τηλέφωνα (Free Ad Blocker browser, CM browser, Ad Blocker browser, Brave browser,

Dolphin – Best Web browser)και τα ζητήματα παραβίασης ιδιωτικότητας που ενδέχεται να προκαλούν.

Με βάση όλα τα παραπάνω, η συγκεκριμένη διατριβή έρχεται για να μελετήσει τους μεθόδους παρακολούθησης των χρηστών, να διατυπώσει την λειτουργία του Real Time Bidding, και να εξετάσει μέσα από ένα περιβάλλον δοκιμών, την συμπεριφορά των Ad Blocking apps στα κινητά τηλέφωνα και κατά πόσο αποτελούν απειλή για τα προσωπικά δεδομένα των χρηστών. Για τους ανωτέρω ερευνητικούς σκοπούς, αναπτύχθηκε κατάλληλο περιβάλλον δοκιμών, στο οποίο χρησιμοποιήθηκε το δημοφιλές λειτουργικό σύστημα Android με σκοπό να αναλυθούν γνωστές Ad Blocking εφαρμογές που γνωρίζουν τεράστια απήχηση από τους χρήστες.

1.1 Δομή της διατριβής

Όπως προαναφέρθηκε αντικείμενο της παρούσας διατριβής είναι αφενός η μελέτη των μηχανισμών παρακολούθησης και της άτυπης δημοπρασίας που πραγματοποιείται μεταξύ των διαφημιστών (RTB) και αφετέρου η ανάλυση των Ad Blocking apps ως προς το κατά πόσο εγείρουν θέματα κατά της ιδιωτικότητας. Ειδικότερα η δομή της διατριβής είναι ως εξής:

Στο κεφάλαιο 2 γίνεται μια μελέτη των τεχνολογιών που χρησιμοποιούνται για παρακολούθηση χρηστών, καθώς και για προβολή στοχευμένων μηνυμάτων.

Στο κεφάλαιο 3 γίνεται μια εισαγωγή στα δίκτυα διαφημίσεων και στην άτυπη δημοπρασία (Real Time Bidding) που πραγματοποιείται μεταξύ των διαφημιστών σε πραγματικό χρόνο. Συγκεκριμένα παρουσιάζεται η εξέλιξη των δικτύων διαφημίσεων με την πάροδο του χρόνου, και αναλύεται η βασική λειτουργία του Real Time Bidding.

Στο κεφάλαιο 4 γίνεται η ανάλυση πέντε πολύ γνωστών εφαρμογών Ad Blocking ως προς τα δεδομένα και τις προσβάσεις που αποκτούν στο κινητό τηλέφωνο του χρήστη, και στο κατά

πόσο υπάρχει κίνδυνος με βάση την ιδιωτικότητα από αυτές τις εφαρμογές σε τρίτες οντότητες.

Στο κεφάλαιο 5 γίνεται μία σύνοψη των ερευνητικών αποτελεσμάτων της διατριβής και καταγραφή συμπερασμάτων.

Κεφάλαιο 2

Μηχανισμοί παρακολούθησης χρηστών στο διαδίκτυο

Η ιδιωτικότητα φαίνεται να είναι η αχίλλειος πτέρνα του σημερινού ιστού. Οι υπηρεσίες διαδικτύου κάνουν συνεχείς προσπάθειες για να λάβουν όσο το δυνατόν περισσότερες πληροφορίες σχετικά με τα πράγματα που ψάχνουμε, τους ιστότοπους που επισκεπτόμαστε, τους ανθρώπους με τους οποίους επικοινωνούμε και τα προϊόντα που αγοράζουμε. Η παρακολούθηση η αλλιώς Web Tracking, γίνεται συνήθως για εμπορικούς σκοπούς. Μέσω του Web Tracking υπάρχει η δυνατότητα δημιουργίας διαδικτυακού προφίλ του χρήστη, από τις πληροφορίες που έχουν συλλεχθεί κατά την περιήγηση στο διαδίκτυο. Οι μηχανισμοί Web Tracking μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες: τους μηχανισμούς που λαμβάνουν χώρα στην πλευρά του πελάτη (**client side mechanisms**) και στους μηχανισμούς δημιουργίας αποτυπωμάτων (**Fingerprinting**) από την πλευρά των διαφημιστών.

Οι μηχανισμοί client side λαμβάνουν χώρα στον υπολογιστή του χρήστη μέσω των προγραμμάτων περιήγησης, που έχουν αναλάβει την αποθήκευση δεδομένων ή αρχείων στον υπολογιστή του χρήστη. Τα δεδομένα αυτά παρακολουθούν τον χρήστη και αργότερα μοιράζονται στους διακομιστές με στόχο την δημιουργία διαδικτυακού προφίλ. Οι μηχανισμοί client side μπορούν να χωριστούν σε τρεις κατηγορίες [03] :

1) Μηχανισμοί παρακολούθησης βάσει πλοήγησης (Session only tracking mechanisms)
2) Μηχανισμοί βάσει αποθήκευσης (Storage based tracking mechanisms)
3) Μηχανισμοί αποθήκευσης βάσει της κρυφής μνήμης (Cashe based tracking mechanisms)

2.1 Μηχανισμοί παρακολούθησης βάσει πλοήγησης

Οι ιστορικά πρώτοι γνωστοί μηχανισμοί παρακολούθησης βασίστηκαν αποκλειστικά στις διαδικτυακές συνεδρίες (sessions). Μια συνεδρία είναι μια συζήτηση μεταξύ του διακομιστή και ενός χρήστη και αποτελείται από μια σειρά συνεχών αιτημάτων και απαντήσεων. Οι μέθοδοι αυτοί ήταν σχετικά απλές και δεν δημιουργούσαν σημαντικές απειλές για τους χρήστες [03].

Αναγνωριστικά συνεδρίας αποθηκευμένα σε κρυφά πεδία

Πριν από το 1994, ο μόνος τρόπος με τον οποίο ένας ιστότοπος θα μπορούσε να λαμβάνει δεδομένα σχετικά με το χρήστη από την άλλη άκρη ήταν η τοποθέτηση ενός αναγνωριστικού συνεδρίας (session id) στη διεύθυνση URL ή ως τιμή σε μια φόρμα [04]. Το αναγνωριστικό αυτό μπορεί να είναι οποιαδήποτε συμβολοσειρά που να είναι σε θέση να παρακολουθεί μοναδικά τον χρήστη κατά την διάρκεια μιας μόνο περιήγησης. Μπορεί να αποτελείται για παράδειγμα από χρονική σήμανση και έναν τυχαίο αριθμό. Σε αντίθεση με άλλους μηχανισμούς παρακολούθησης το αναγνωριστικό αυτό δεν αποθηκεύεται και εξαφανίζεται όταν ο χρήστης τερματίσει την σελίδα. Αν και αυτή η τεχνική συνεχίζει να λειτουργεί μέχρι και σήμερα, δεν αποτελεί μία σοβαρή απειλή κατά τις ιδιωτικότητας των χρηστών [03].

Αυθεντικός Έλεγχος ταυτότητας ιστού

Μία άλλη πιθανότητα για εντοπισμό του χρήστη είναι η ζήτηση ή απαίτηση από τον χρήστη να εγγραφεί στην ιστοσελίδα που έχει επισκεφτεί. Στην συνέχεια οι πόροι που παρέχονται από

τον ιστότοπο είναι διαθέσιμοι μόνο για τους χρήστες που είναι εγγεγραμμένοι. Αυτό καθιστά την αναγνώριση του χρήστη πολύ εύκολη και ακριβή. Αυτή η μέθοδος είναι ανεξάρτητη από το πρόγραμμα περιήγησης, το λειτουργικό σύστημα ή τον υπολογιστή που χρησιμοποιείται, καθώς επίσης και τον τόπο όπου ο χρήστης είναι συνδεδεμένος στο διαδίκτυο. Υπάρχουν ωστόσο δύο σημαντικά θέματα σε αυτήν την μέθοδο : 1) Ο χρήστης πρέπει να αποδεικνύει την ταυτότητα του κάθε φορά που χρησιμοποιεί τον ιστότοπο, 2) ο έλεγχος ταυτότητας (authentication) είναι έγκυρος μόνο εντός της τρέχουσας σύνδεσης, αφού ο χρήστης έχει συνδεθεί. Η συγκεκριμένη μέθοδος δεν αποτελεί παραβίαση της ιδιωτικότητας διότι ο χρήστης είναι ενήμερος ότι έχει συνδεθεί και ότι όλα αυτά που κάνει στον ιστότοπο ενδέχεται να καταγραφούν [03].

Window.name Document object model (DOM) property

Το Document object model είναι μία διεπαφή ανάμεσα σε πλατφόρμες και έχει ως στόχο την πρόσβαση και την αλληλεπίδραση των Web εγγραφών (πχ HTML, XHTML,XML). Το Document object model είναι κοινό για όλα τα προγράμματα περιήγησης ιστού και περιλαμβάνει το παράθυρο ιδιότητας window.name. Το παράθυρο είναι προσβάσιμο μέσω κώδικα JavaScript και μπορεί να αποθηκεύσει αρκετά megabyte δεδομένων (2-32 MB). Κάθε καρτέλα του προγράμματος περιηγητή (tab) έχει την δική της ιδιότητα window.name, η οποία είναι κενή κατά την δημιουργία. Η ιδιότητα αυτή δεν επηρεάζεται από επαναφορτώσεις σελίδων και είναι προσβάσιμη από άλλα domains, κάτι που σημαίνει ότι μπορεί να χρησιμοποιηθεί για την ανταλλαγή πληροφοριών μεταξύ των domains, το οποίο θέτει άμεση ασφάλεια και απειλές για την προστασία των προσωπικών δεδομένων [05]. Ωστόσο ο συγκεκριμένος μηχανισμός παρακολούθησης είναι πιο ασφαλής από ότι τα cookies - τα οποία περιγράφονται στην συνέχεια - λόγω της μη συμμετοχής του διακομιστή (web server), γεγονός που τον καθιστούν λιγότερο ευάλωτο σε επιθέσεις cookie sniffing.

2.2 Μηχανισμοί βάσει αποθήκευσης

Η επόμενη ομάδα μηχανισμών παρακολούθησης εξαρτάται από την αποθήκευση δεδομένων στους υπολογιστές των χρηστών. Αυτές οι μέθοδοι φαίνεται να είναι οι πιο συχνά χρησιμοποιούμενες. Γενικά, είναι πολύ πιο προηγμένες από τις μεθόδους που βασίζονται σε συνεδρίες, και οι ικανότητές τους είναι επίσης υψηλότερες. Κάθε ένας από αυτούς τους μηχανισμούς αποτέλεσε, με την εμφάνισή του, τη μεγαλύτερη απειλή για την προστασία της ιδιωτικότητας των χρηστών. Οι περιηγητές άρχισαν να εφαρμόζουν την εκκαθάριση αυτών των αποθηκευτικών χώρων κατόπιν αιτήματος του χρήστη ή με την πάροδο συγκεκριμένου χρονικού διαστήματος.

COOKIES

Ένας από τους πιο γνωστούς τρόπους για αναγνώριση του χρήστη είναι η χρήση των λεγόμενων cookies (πολλές φορές αποκαλούνται HTTP cookies). Τα cookies είναι μικρά κομμάτια δεδομένων (έως 4KB) που δημιουργούνται από έναν διακομιστή (web server) και αποθηκεύονται στον υπολογιστή του χρήστη ή στο πρόγραμμα περιήγησης (web browser), παρέχοντας πληροφορίες σχετικά με τη δραστηριότητα περιήγησης του χρήστη. Παρόλο που αντιπροσωπεύουν απλή δέσμη αριθμών και γραμμάτων, η φύση τους είναι να μεταφέρουν ουσιαστικές πληροφορίες για τον χρήστη, η οποία θεωρείται από τους υπερασπιστές των δεδομένων ως προσωπική εισβολή στην ιδιωτική ζωή [06].

Πρέπει εξάλλου να σημειωθεί ότι υπάρχει ειδική νομοθεσία που ρυθμίζει το ζήτημα της νόμιμης εγκατάστασης και περαιτέρω χρήσης των cookies: στην Ευρώπη, είναι σε ισχύ η Οδηγία 2002/58/ΕΚ (e-Privacy Directive), η οποία έχει ενσωματωθεί σε όλα τα Κράτη-Μέλη της Ευρωπαϊκής Ένωσης: βάσει αυτής, δεν επιτρέπεται η εγκατάσταση cookies στον τερματικό εξοπλισμό ενός χρήστη, εφόσον το cookie αυτό δεν είναι απολύτως απαραίτητο για την παροχή της υπηρεσίας την οποία αιτείται ο χρήστης, παρά μόνο με την ειδική συγκατάθεσή του: προφανώς, η κοινή πρακτική καταδεικνύει ότι η ως άνω απαίτηση καταστρατηγείται στην πράξη. Στα προσεχή χρόνια, η Οδηγία 2002/58/ΕΚ επίκειται να αντικατασταθεί από Κανονισμό της Ευρωπαϊκής Ένωσης, σχέδιο του οποίου είναι αυτή τη στιγμή υπό διαμόρφωση.

Υπάρχουν δύο τύποι HTTPcookies που μπορούν να χρησιμοποιηθούν [03]:

- Session Cookies: Αυτά είναι cookies τα οποία λήγουν μόλις ο χρήστης κλείσει τον περιηγητή ιστού του. Εάν ένα cookie δεν περιέχει ημερομηνία λήξης, θεωρείται ένα

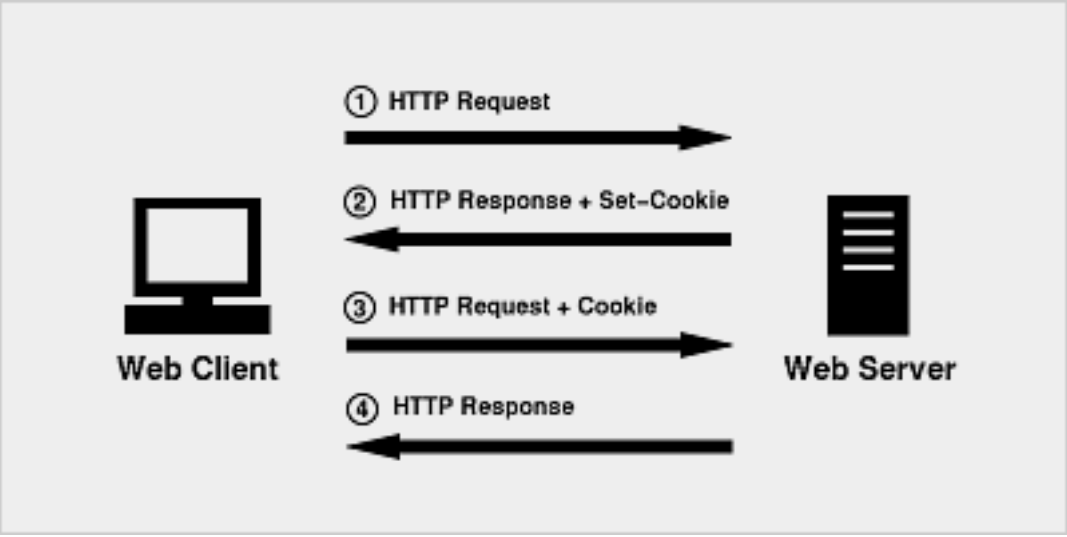
session cookie. Τα session cookies αποθηκεύονται στη μνήμη και δεν γράφονται ποτέ στο δίσκο του συστήματος. Όταν το πρόγραμμα περιήγησης κλείσει το cookie χάνεται.

- Persistent Cookies: Εάν το cookie περιέχει ημερομηνία λήξης, θεωρείται persistent cookie. Την ημερομηνία που καθορίζεται στη λήξη, το cookie θα αφαιρεθεί από το δίσκο. Υπάρχουν πολλά διαφορετικά πεδία που μπορεί να περιέχει ένα cookie, χωρισμένα με άνω-κάτω τελείες όπως αναφέρεται στο παρακάτω παράδειγμα:

Expires="Wdy, DD−Mon−YYYYHH: MM: SSGMT"

Η συγκεκριμένη εντολή καθορίζει πότε πρέπει να διαγραφεί ένα cookie [07]

Ο τρόπος εγκατάστασης των cookies στον υπολογιστή είναι σχετικά απλός.
Όταν ο χρήστης επισκέπτεται για πρώτη φορά έναν ιστότοπο, ένα αρχείο cookie με ένα μοναδικό αναγνωριστικό χρήστη, αποθηκεύεται στον υπολογιστή του [08]. Αυτό το μοναδικό αναγνωριστικό αποτελείται από έναν κωδικό που αναφέρεται στο συγκεκριμένο cookie. Έπειτα ο ιστότοπος μπορεί να ανακτήσει το αναγνωριστικό αυτό κάθε φορά που ο χρήστης επισκέπτεται τον συγκεκριμένο ιστότοπο. Με αυτόν τον τρόπο ο ιστότοπος διατηρεί ένα ιστορικό επισκέψεων του χρήστη, εκτός εάν ο χρήστης διαγράψει το cookie από τον υπολογιστή του [03]. Στην εικόνα 1 παρουσιάζεται ένα παράδειγμα από τον τρόπο λειτουργίας των cookies.



Εικόνα 1 : Ανταλλαγή cookie HTTP μεταξύ πελάτη &διακομιστή[9]

Κατά την σύνδεση ενός χρήστη ο απομακρυσμένος διακομιστής ιστού(web server) λέει στο πρόγραμμα περιήγησης του χρήστη να αποθηκεύσει το αρχείο δεδομένων που ονομάζεται Set-Cookie. Το παρακάτω είναι ένα παράδειγμα πως μπορεί να μοιάζει μια συναλλαγή HTTP cookie:

HTTP response from web server

{. . .}

Set-cookie: first.lastname

Εάν γίνει δεκτό από το πρόγραμμα περιήγησης, το cookie αποθηκεύεται και αποστέλλεται στο διακομιστή μέσα σε ένα πεδίο κεφαλίδας HTTP που ονομάζεται cookie:

HTTP GET from the client

{. . .}

Cookie: first.lastname

Στην παραπάνω συναλλαγή, ο διακομιστής ιστού είπε στον χρήστη να δημιουργήσει το cookie "first.lastname". Την επόμενη φορά που ο χρήστης ζητά ένα αντικείμενο από αυτόν τον διακομιστή ιστού, αποστέλλει το cookie μαζί με το αίτημα. Αυτό υποδεικνύει τον τρόπο με τον οποίο ένας διακομιστής ιστού μπορεί να είναι σε θέση να ανακαλέσει ορισμένες πληροφορίες, όπως η είσοδος χρηστών [07]. Αυτό γίνεται κάθε φορά που ο πελάτης υποβάλλει ένα νέο αίτημα στον ίδιο ιστότοπο (domain) [05]. Ωστόσο η ακρίβεια τους έγκειται στο κατά πόσον ο χρήστης επιτρέπει την χρήση cookies, δεν καθαρίζει την προσωρινή μνήμη των cookies από το πρόγραμμα περιήγησης, και χρησιμοποιεί πάντα το ίδιο πρόγραμμα περιήγησης (web browser) για να επισκεφτεί το συγκεκριμένο ιστότοπο (domain).

First Party Cookies / Third Party Cookies

Τα cookies μπορούν να δημιουργηθούν είτε από τον πάροχο της ιστοσελίδας (first party cookies), είτε από μία άλλη ανεξάρτητη τρίτη οντότητα (third party cookies). Τα first party cookies χρησιμοποιούνται συνήθως για λειτουργικούς ή στατιστικούς σκοπούς από τον πάροχο και αποθηκεύουν τις πληροφορίες που ο χρήστης παρείχε στον ιστότοπο, για παράδειγμα, το όνομα ή τη διεύθυνση, και τη διατηρούν στη μνήμη του διακομιστή, ώστε να μην χρειάζεται να ξαναπαίρνονται οι ίδιες πληροφορίες. Αυτός ο τύπος cookie δεν θεωρείται απαραίτητα ενοχλητικός, καθώς επιταχύνει τη δραστηριότητα περιήγησης του χρήστη. Τα third party cookies όμως, τοποθετούνται σε έναν ιστότοπο από μία τρίτη οντότητα, για

παράδειγμα από τους διαφημιζόμενους, προκειμένου να συλλέγουν τις προτιμήσεις, του χρήστη και να κάνουν στην συνέχεια, στοχευμένες διαφημίσεις [10]. Ορισμένοι περιηγητές ιστού εφαρμόζουν περιορισμούς cookie που εμποδίζουν τους ιστότοπους τρίτων μερών να ορίσουν cookies στο πρόγραμμα περιήγησης [11]. Αυτός ο περιορισμός όμως μπορεί εύκολα να παρακαμφθεί μεταφέροντας τον χρήστη μέσω JavaScript στον ιστότοπο τρίτου μέρους που θα ορίσει ή θα διαβάσει τα cookies. Έτσι ο χρήστης αποπροσανατολίζεται σε σχέση με τον ιστότοπο που επισκέφτηκε αρχικά. Με αυτόν τον τρόπο, το περιεχόμενο τρίτων εμφανίζεται ότι προέρχεται από ιστότοπους πρώτων μερών. Για το σκοπό αυτό, αντί να ανακατευθύνει τον χρήστη, ο ιστότοπος μπορεί να χρησιμοποιήσει αναδυόμενα παράθυρα στα οποία το περιεχόμενο τρίτου μέρους εμφανίζεται ως προερχόμενο από το πρώτο μέρος [03].

Flash Cookies

Τα Flash Cookies γνωστά και ως Local Shared Objects (LSO) είναι ένα αρχείο κειμένου που αποστέλλεται από ένα διακομιστή(web server) σε έναν υπολογιστή χρήστη όταν αυτός ζητά περιεχόμενο που υποστηρίζεται από το Adobe Flash, ένα δημοφιλές plug-in πρόγραμμα περιήγησης που χρησιμοποιείται για κινούμενο διαδραστικό, διαδικτυακό περιεχόμενο (animated interactive web content). Χρησιμοποιούνται κυρίως για να αποθηκεύουν πληροφορίες που αφορούν το Flash, όπως το σημείο στο οποίο σταμάτησε να παίζει το βίντεο του χρήστη ή η διαφήμιση με κινούμενα σχέδια που έπαψε να περιστρέφεται. Τα Flash Cookies αποθηκεύονται ως αρχεία τύπου .sol στο σύστημα του χρήστη [03].

Τα Flash Cookies προσφέρουν χαρακτηριστικά που τα καθιστούν πιο μόνιμα από τα τυπικά HTTP Cookies. Μπορούν να περιέχουν μέχρι 100KB πληροφοριών (τα HTTP Cookies αποθηκεύουν μόνο 4KB) καθώς επίσης δεν έχουν ημερομηνία λήξης από προεπιλογή. Ένα πολύ σημαντικό πλεονέκτημα των Flash Cookies ήταν ότι δεν ελέγχονταν κατά το παρελθόν από το πρόγραμμα περιήγησης. Έτσι, η διαγραφή των HTTP Cookies, η εκκαθάριση του ιστορικού, η διαγραφή της προσωρινής μνήμης ή η επιλογή της διαγραφής ιδιωτικών δεδομένων στο πρόγραμμα περιήγησης δεν επηρέαζε τα Flash Cookies [12]. Αυτό έκανε τα Flash Cookies μια ελκυστική επιλογή ως μέθοδο παρακολούθησης των χρηστών. Τα τελευταία χρόνια η εταιρεία Adobe έχει κάνει αλλαγές ως προς την ιδιωτικότητα που επιτρέπουν την εκκαθάριση αυτών των τιμών, υποστηρίζοντας το API ClearSiteData από την έκδοση 10.3. Πλέον τα LSOs μπορούν να διαγραφούν με παρόμοιο τρόπο όπως τα HTTP Cookies [13]. Εναλλακτικά, οι χρήστες θα μπορούσαν να καταργήσουν LSOs από το σύστημα αρχείων

άμεσα, αλλά δεδομένου ότι οι LSO αποθηκεύονται σε κρυφό φάκελο, η επιλογή αυτή δεν ήταν πολύ συνηθισμένη.

Εκτός από τα Local Shared Objects (LSO), το Flash χρησιμοποιεί επίσης και τα Remote Shared Objects(RSO). Τα δύο αυτά συνεργαζόμενα αντικείμενα της Flash, μπορούν να έχουν πρόσβαση στα αποθηκευμένα περιεχόμενα ενός τοπικά μόνιμου αντικειμένου, σε ένα αρχείο με κατάληξη .sor [11].

Zombie Cookies

Ενώ τα cookies όπως είδαμε είναι ένας αποτελεσματικός μηχανισμός για την παρακολούθηση των χρηστών, περισσότεροι άνθρωποι πλέον αντιλαμβάνονται τις επιπτώσεις της ιδιωτικής ζωής και εκκαθαρίζουν τακτικά τα cookies τους (HTTP και Flash Cookies) [05]. Λόγω αυτού σχεδιάστηκαν τα Zombie Cookies. Τα Zombie Cookies, γνωστά και ως Overcookies ή Super Cookies, θεωρούνται ως ένα πιο μόνιμο είδος Cookie. Μεγάλο τους πλεονέκτημα αποτελεί η δυνατότητα χρήσης πολλών χώρων αποθήκευσης :

• HTTP cookies,
• Flash cookies,
• Silverlight Isolated Storage,
• Web storage,
• Web history,
• Browser cache,
• Window.name DOM property.

Κάθε φορά που ένας από αυτούς τους χώρους διαγράφεται από το σύστημα του χρήστη, το cookie Zombie χρησιμοποιεί κώδικα JavaScript από την πλευρά του χρήστη για να τους αναδημιουργήσει ξανά μέσα από τα υπόλοιπα δεδομένα αποθήκευσης [14]. Επίσης έχουν τη δυνατότητα να προσδιορίσουν ένα χρήστη, ακόμη και όταν όλα τα άλλα είδη cookies (συμπεριλαμβανομένων των Flash cookies) έχουν διαγραφεί. Τα zombie cookies μπορούν να εξαπλωθούν σε διαφορετικά προγράμματα περιήγησης στον ίδιο υπολογιστή, παρέχουν έναν εξαιρετικά ανθεκτικό μηχανισμό παρακολούθησης και έχει βρεθεί ότι χρησιμοποιούνται από πολλές δημοφιλείς τοποθεσίες για να παρακάμψουν τις εσκεμμένες ενέργειες των χρηστών [15].

Microsoft Silverlight Cookie

Η Microsoft εισήγαγε την προσθήκη Silverlight το 2007 ως ανταγωνιστή του Flash της Adobe. Η μόνιμη αποθήκευση των cookies Silverlight μπορεί να επιτευχθεί χρησιμοποιώντας την απομονωμένη αποθήκευση. Μπορεί να περιέχει 100KB ανά ιστότοπο και αποθηκεύεται στο προφίλ του χρήστη, ακριβώς όπως τα Flash Cookies. Η απομονωμένη αποθήκευση είναι απενεργοποιημένη σε ιδιωτική λειτουργία. Η Microsoft δεν δημοσιεύει στατιστικά στοιχεία ως προς τη διείσδυση του Silverlight, αλλά η διείσδυση εκτιμάται ότι υπερβαίνει το 50%, κυρίως λόγω των αυτόματων εγκαταστάσεων του Windows Update. Η Microsoft ανακοίνωσε το πέρας της τελευταίας έκδοσης του Silverlight το 2021 και οι μελλοντικές εξελίξεις πιθανότατα θα επικεντρωθούν στο Html 5. Αλλά μέχρι να καταργηθεί, το Silverlight παρέχει τον ίδιο τύπο αποθήκευσης όπως το Flash [16]. Αυτή η αποθήκευση μπορεί να «καθαριστεί» μόνο με μη αυτόματο τρόπο (διαγράφοντας αρχεία από κρυφό φάκελο στο σύστημα αρχείων ή χρησιμοποιώντας επιλογές αποθήκευσης στην εφαρμογή Silverlight) [05].

HTML5 Web Storage

Η HTML Web Storage αφορά έναν μηχανισμό παρακολούθησης βασισμένο στην αποθήκευση δεδομένων στην πλευρά των χρηστών. Προσφέρει μόνιμη αποθήκευση δεδομένων, όπως τα HTTP και Flash Cookies, αλλά με περισσότερη χωρητικότητα και χωρίς την απαίτηση αποθήκευσης πληροφορίας σε κεφαλίδες HTTP. Η αρχική ιδέα της HTML 5 ήταν να παρέχει μηχανισμούς παγκόσμιας αποθήκευσης (Global Storage), και δυνατότητες αποθήκευσης δεδομένων σε ιστοσελίδες, κάτι που όμως δεν προχώρησε από κανένα πρόγραμμα περιήγησης λόγω παραβίασης της πολιτικής SOP (Same Origin Policy) [03]. Υπάρχουν δύο τύποι αποθήκευσης: 1) Αποθήκευση συνεδρίας (Session Storage), 2) Τοπική Αποθήκευση (Local Storage).

Η Τοπική αποθήκευση(Local Storage) έχει σχεδιαστεί για αποθήκευση δεδομένων που εκτείνεται σε πολλαπλά παράθυρα και διαρκεί πέρα από την τρέχουσα περίοδο λειτουργίας. Τα δεδομένα αυτά μπορούν να φτάσουν σε χωρητικότητα μέχρι και 5MB, κάτι που δίνει μεγάλο πλεονέκτημα σε σχέση με τα HTTP Cookie (4KB) και Flash Cookie (100KB). Παράλληλα, και σε αντίθεση με τα cookies στα οποία πρόσβαση μπορεί να έχει τόσο η πλευρά του διακομιστή (server side) όσο και αυτή του χρήστη (client side), η HTML5 web storage εμπίπτει αποκλειστικά στην επίβλεψη του client-side scripting. Η τεχνολογία client-side scripting αναφέρεται στα διαδικτυακά προγράμματα που εκτελούνται από τον διακομιστή στην πλευρά

του χρήστη και όχι στην πλευρά του διακομιστή [17]. Ένα ακόμη χαρακτηριστικό της τοπικής αποθήκευσης είναι ότι τα δεδομένα δεν διαγράφονται από προεπιλογή ή από συγκεκριμένη ημερομηνία λήξης. Ο μόνος τρόπος διαγραφής είναι είτε από απόφαση της ιστοσελίδας, είτε χειρωνακτικά (manually) από τον ίδιο τον χρήστη [18].

Η αποθήκευση συνεδρίας (Session Storage)είναι παρόμοια με την τοπική αποθήκευση: διατηρεί την πολιτική ίδιας προέλευσης και τα αποθηκευμένα αντικείμενα μπορεί να έχουν μέγεθος μέχρι και 5 MB. Ωστόσο, τα αντικείμενα είναι διαθέσιμα μόνο στο τρέχον παράθυρο του προγράμματος περιήγησης και διαγράφονται όταν κλείνει το παράθυρο [03].

Αρκετοί έχουν επισημάνει τους κινδύνους ιδιωτικής ζωής που παρουσιάζονται από το HTML 5. Άλλοι υποστήριξαν ότι το HTML 5 έχει μεγάλες δυνατότητες να επιτρέψει τη χρήση περισσότερων μοντέλων που προστατεύουν την προστασία της ιδιωτικής ζωής. Ωστόσο, δεν έχει πραγματοποιηθεί ενδελεχής έρευνα σχετικά με τις πρακτικές απορρήτου HTML5 [17]. Παρακάτω παρουσιάζεται ένας πίνακας (πίνακας 1) με βασικές διαφορές ανάμεσα σε HTTP Flash Cookies και HTML Web Storage.

	HTTP Cookies	Flash Cookies	HTML5 Storage
Χωρητικότητα	4KB	100KB(Από προεπιλογή)	5MB(Από προεπιλογή)
Λήξη	Συνεδρία(Από προεπιλογή)	Μόνιμα(Από προεπιλογή)	Μόνιμα(Από προεπιλογή)
Αποθήκευση	Αρχείο SQL(Firefox)	Έξω από το πρόγραμμα περιήγησης	Αρχείο SQL(Firefox)
Πρόσβαση	Μόνο από πρόγραμμα περιήγησης	Διάφορα προγράμματα περιήγησης σε ίδιο υπολογιστή	Μόνο από πρόγραμμα περιήγησης

Πίνακας 1 : Βασικά χαρακτηριστικά HTTP Cookies, Flash Cookies, HTML5 Cookies [17]

HTML5 Indexed DB

Το Indexed DB είναι μια νέα ιδέα με βάση το HTML5 για την αποθήκευση των δεδομένων μέσα στο πρόγραμμα περιήγησης του χρήστη(web storage). Προήλθε από τις προδιαγραφές του W3C για την υλοποίηση του χώρου αποθήκευσης στον ιστό το 2009 και στο παρελθόν ήταν γνωστό και ως Web Simple DB [19]. Το Indexed DB ουσιαστικά είναι μια βάση δεδομένων από την πλευρά του χρήστη που ενσωματώνεται στο πρόγραμμα περιήγησης. Τα αρχεία και τα

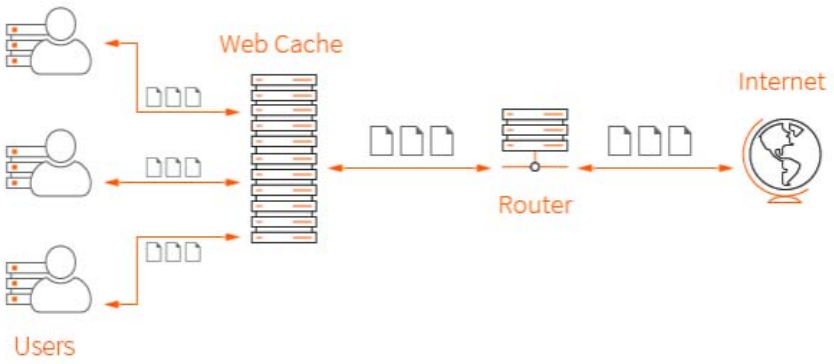
δεδομένα που αποθηκεύονται από το πρόγραμμα περιήγησης διατηρούνται στο σκληρό δίσκο του υπολογιστή του χρήστη. Η βάση δεδομένων Indexed DB, αποθηκεύει τα δεδομένα, ακόμα και όταν τερματίζεται το πρόγραμμα περιήγησης. Πρόκειται για μια αποθήκευση δεδομένων από την πλευρά του χρήστη / πελάτη, κάτι που σημαίνει ότι τα δεδομένα μπορούν να ανακτηθούν ακόμα και αν το πρόγραμμα περιήγησης είναι εκτός σύνδεσης. Επομένως, τα αρχεία βρίσκονται στο σύστημα αρχείων χρήστη, και μπορούν να ανακτηθούν έως ότου αντικατασταθούν από άλλα αρχεία. Ο αποθηκευτικός χώρος με βάση το πρόγραμμα περιήγησης, όπως το Indexed DB, μπορεί να χρησιμοποιηθεί σε πολλαπλά προγράμματα περιήγησης και είναι συμβατό με πολλαπλές πλατφόρμες. Το Indexed DB επεκτείνει την τοπική αποθήκευση (local storage) του HTML 5 παρέχοντας εφαρμογές web με αποθήκευση εκτός σύνδεσης (offline storage). Αυτό μπορεί να χρησιμοποιηθεί από τα καταστήματα ηλεκτρονικού εμπορίου για την αποθήκευση των προτιμήσεων των πελατών [20]. Συνοψίζοντας, το Indexed DB λειτουργεί υπό τις ίδιες συνθήκες όπως το HTML 5 Local Storage, οπότε έχει το ίδιο αντίκτυπο στην ιδιωτική ζωή των χρηστών [03].

2.3 Μηχανισμοί παρακολούθησης με βάση την κρυφή μνήμη cache

Μια άλλη ομάδα μεθόδων παρακολούθησης χρησιμοποιεί επίσης αποθηκευτικό χώρο που βασίζεται σε πελάτες / χρήστες. Αλλά σε αντίθεση με την προηγούμενη ομάδα που χρησιμοποίησε αποθήκευση με στόχο την διατήρηση δεδομένων, αυτή η ομάδα εκμεταλλεύεται τις πιθανότητες αναγνώρισης μέσω των προγραμμάτων περιήγησης και προσδιορίζει τους ιστότοπους που επισκέφτηκε ένας χρήστης με την χρήση της κρυφής μνήμης (web cashe) [03].

Η προσωρινή αποθήκευση στο Web είναι ένα βασικό χαρακτηριστικό γνώρισμα του πρωτοκόλλου HTTP που αποσκοπεί στην ελαχιστοποίηση της κυκλοφορίας του δικτύου ενώ βελτιώνει την αντιληπτή ανταπόκριση του συστήματος στο σύνολό του. Οι κρυφές μνήμες (cache) βρίσκονται σε κάθε επίπεδο της διαδρομής ενός περιεχομένου από τον αρχικό διακομιστή στο πρόγραμμα περιήγησης. Η προσωρινή αποθήκευση στο Web λειτουργεί με προσωρινή αποθήκευση των απαντήσεων HTTP για αιτήματα σύμφωνα με ορισμένους κανόνες. Μεταγενέστερα αιτήματα για αποθηκευμένο περιεχόμενο μπορούν στη συνέχεια να

ικανοποιηθούν από μια μνήμη cache που βρίσκεται πιο κοντά στον χρήστη αντί να στέλνουν το αίτημα μέχρι το διακομιστή ιστού [21]. Στην εικόνα 2 φαίνεται ένα παράδειγμα της χρήσης κρυφής μνήμης στο διαδίκτυο.



Εικόνα 2 : Παράδειγμα λειτουργίας Web Cache(Πηγή https://www.maxcdn.com/one/visual-glossary/web-cache/)

Υπάρχουν διάφοροι τρόποι με τους οποίους μπορεί να γίνει αξιοποίηση της προσωρινής μνήμης του περιηγητή οι οποίοι αναλύονται στην συνέχεια.

Ενσωμάτωση αναγνωριστικών σε αποθηκευμένα έγγραφα

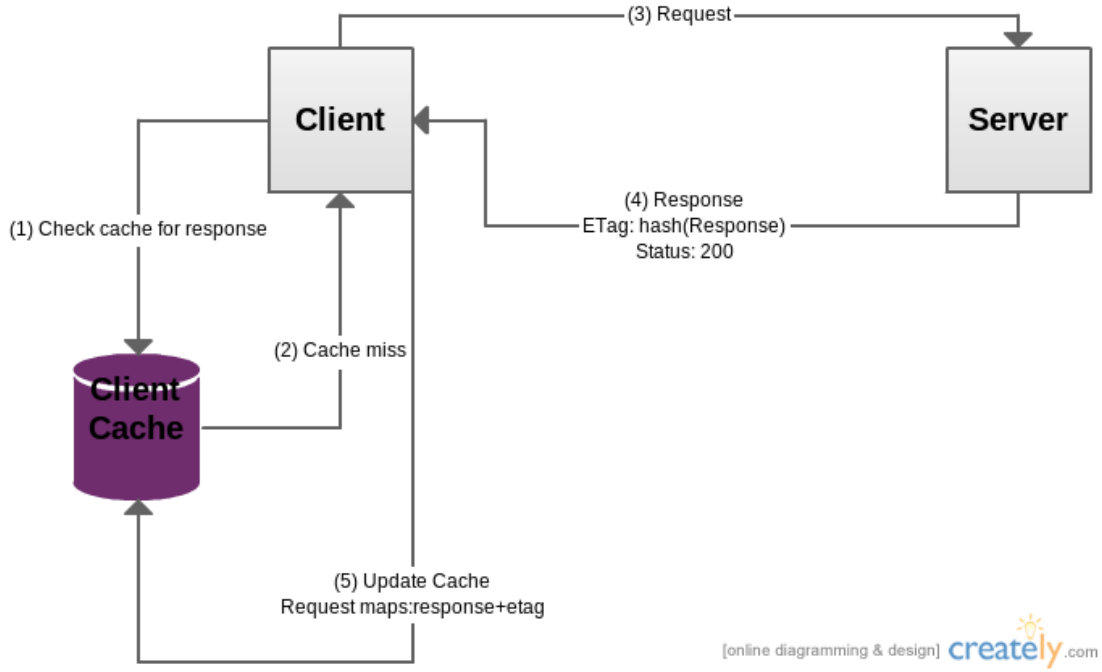
Τα προγράμματα περιήγησης Web υποστηρίζουν την προσωρινή αποθήκευση του περιεχομένου που έχει καθοριστεί μέσω της κρυφής μνήμης του περιηγητή (Browser Cache). Οι ίδιοι οι περιηγητές ιστού διατηρούν μια μικρή μνήμη cache. Συνήθως, ο περιηγητής ορίζει μια πολιτική που υπαγορεύει τα πιο σημαντικά στοιχεία για τη μνήμη cache. Αυτό μπορεί να είναι περιεχόμενο για συγκεκριμένο χρήστη ή περιεχόμενο που θεωρείται ‘’δαπανηρό’’ για λήψη (download) και ενδέχεται να ζητηθεί ξανά [21].

Ένας απλός τρόπος προσέγγισης της μνήμης cache, για σκοπούς παρακολούθησης, είναι η απευθείας ενσωμάτωση ενός αναγνωριστικού παρακολούθησης. Αυτό το αναγνωριστικό μπορεί να εγκατασταθεί όταν ζητείται ένα αρχείο από τον διακομιστή για πρώτη φορά. Το αναγνωριστικό μπορεί να ζει μόνιμα στην προσωρινή μνήμη του προγράμματος περιήγησης, όταν στέλνεται με τις κεφαλίδες ελέγχου της προσωρινής μνήμης HTTP [11]. Στη συνέχεια μπορεί να καλείται από ιστοσελίδες και να χρησιμοποιείται από υπηρεσίες με σκοπό την παρακολούθηση του χρήστη [22].

Ετικέτες Οντότητας

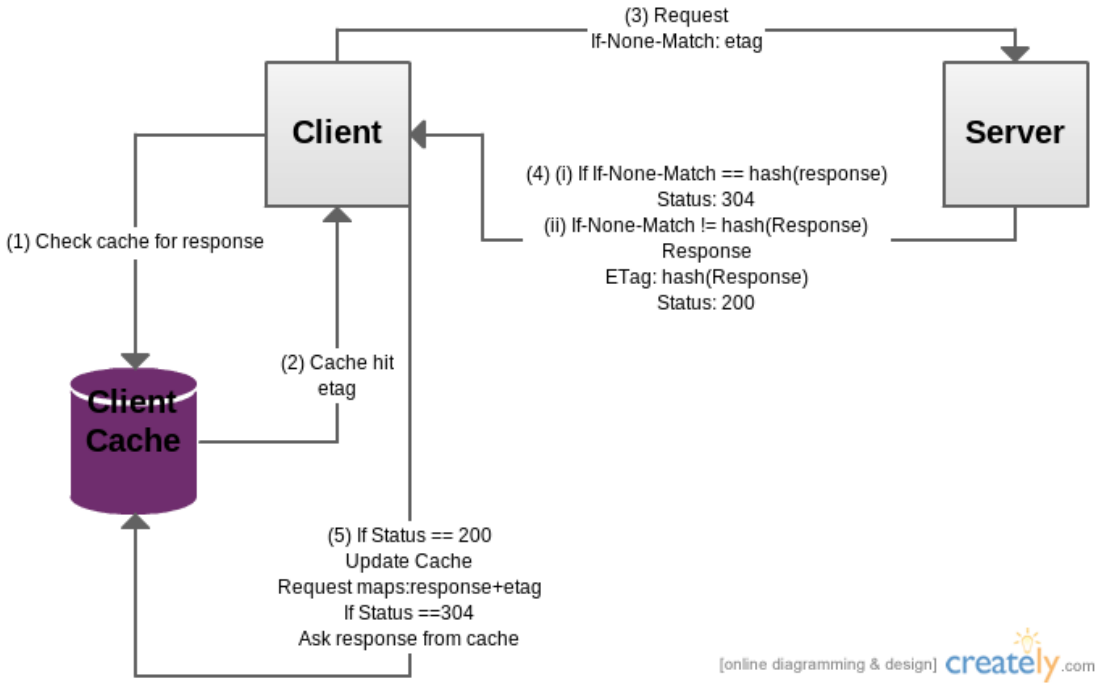
Οι ετικέτες οντοτήτων είναι ένας τρόπος ενσωμάτωσης της προσωρινής αποθήκευσης στο πρωτόκολλο HTTP. Χρησιμοποιούνται ως υπογραφές που επικυρώνουν αν η προσωρινή μνήμη του χρήστη ενημερώνεται, και για αυτόν τον λόγο αποτελούν μηχανισμό παρακολούθησης που βασίζονται στην μνήμη cache.

Μια ετικέτα οντότητας (ETag) είναι μια κεφαλίδα HTTP που χρησιμοποιείται για επικύρωση της Web cache του χρήστη. Όταν ένας διακομιστής επιστρέφει μια απάντηση, μπορεί να επισυνάψει ένα ETag header που έχει μια τιμή που αντιπροσωπεύει την κατάσταση του αντικειμένου που επιστράφηκε ως απάντηση στο αίτημα του χρήστη. Όταν ο χρήστης υποβάλλει μεταγενέστερα αιτήματα για την ίδια απάντηση στο ίδιο διακομιστή, μπορεί να στείλει πίσω το ETag στην αίτησή του χρησιμοποιώντας την κεφαλίδα If-None-Match. Αν η τρέχουσα κατάσταση ταιριάζει με αυτήν που αποστέλλεται από τον χρήστη, τότε ο διακομιστής στέλνει την ετικέτα ETag με κωδικό 304 (Not modified - Μη τροποποιημένο), διαφορετικά θα στείλει καινούργιο ETag [23]. Στις παρακάτω εικόνες (3,4) μπορούμε να δούμε δύο παραδείγματα της χρήσης των ETags.



Εικόνα 3: Παράδειγμα χρήσης των ETags, Σενάριο 1ο (Πηγή http://dweebslair.blogspot.gr/2014/03/client-side-cache-controlling-content.html)

Στο πρώτο παράδειγμα (Εικόνα 3) ο χρήστης θα αναζητήσει την αποθήκευση της προσωρινής μνήμης για μια απάντηση που έχει αντιστοιχιστεί με το αίτημα που έχει αυτή τη στιγμή. Εάν δεν υπάρχει η αντίστοιχη εγγραφή στην προσωρινή μνήμη cache ο χρήστης θα απευθυνθεί στον διακομιστή για να του στείλει το αίτημα. Ο διακομιστής θα επεξεργαστεί το αίτημα και θα στείλει πίσω την απάντηση με ένα μοναδικό αναγνωριστικό στην απάντηση (στις περισσότερες περιπτώσεις, αυτό θα είναι μια τιμή κατακερματισμού). Το μοναδικό αναγνωριστικό θα οριστεί στο πεδίο κεφαλίδας HTTP 'ETag'. Έτσι, η τιμή είναι γνωστή ως η τιμή ETag. Αφού ο χρήστης λάβει την απάντηση από το διακομιστή, θα την αποθηκεύσει και η τιμή ETag στην cache του πελάτη θα αντιστοιχιστεί με το σχετικό αίτημα.



Εικόνα 4: Παράδειγμα χρήσης των ETags Σενάριο 2ο (Πηγή http://dweebslair.blogspot.gr/2014/03/client-side-cache-controlling-content.html)

Στο δεύτερο παράδειγμα (Εικόνα 4) ο πελάτης θα αναζητήσει μια απάντηση στην αποθήκευση της προσωρινής μνήμης. Αυτή τη φορά βρίσκει μια αντίστοιχη απάντηση στο αίτημα. Ο πελάτης θα στείλει το αίτημα στον διακομιστή με την τιμή ETag που έχει ρυθμιστεί στο πεδίο κεφαλίδας HTTP "If-None-Match". Ο διακομιστής θα επεξεργαστεί το αίτημα και θα δημιουργήσει την απάντηση. Στη συνέχεια, θα υπολογίσει την τιμή κατακερματισμού της απόκρισης. Αν η τιμή αυτή ταιριάζει με την τιμή στο πεδίο If-None-Match της αίτησης, αυτό σημαίνει ότι η απάντηση δεν έχει αλλάξει. Έτσι, ο διακομιστής θα στείλει την απάντηση με τον

κωδικό κατάστασης HTTP 304. Εάν η τιμή αυτή δεν ταιριάζει με την τιμή στο πεδίο If-None-Match της αίτησης, αυτό σημαίνει ότι η απάντηση έχει τροποποιηθεί. Ο διακομιστής θα στείλει την ενημερωμένη απάντηση με τη νέα τιμή κατακερματισμού του στο πεδίο ETag. Εάν ο χρήστης λάβει ως απάντηση τον κωδικό κατάστασης 304 μπορεί να χρησιμοποιήσει την απάντηση που είναι αποθηκευμένη στην προσωρινή μνήμη. Αν λάβει κωδικό απάντησης 200, πρέπει να ενημερώσει την αποθήκευση της προσωρινής μνήμης με τις νέες τιμές.

Τα πεδία ETags μπορούν να φτάσουν μέχρι και 81864 bits δεδομένων και μπορούν να αποδειχθούν πολύ χρήσιμα εργαλεία για την αναγνώριση των χρηστών, μιας και αποθηκεύουν τις ίδιες πληροφορίες και έχουν την δυνατότητα να κάνουν ανασύσταση των διαγραμμένων cookies. Βέβαια πρέπει να αναφερθεί πως τα ETags δεν θεωρούνται απολύτως αξιόπιστα. Αυτό έχει να κάνει κυρίως με το γεγονός πως κάθε πρόγραμμα περιήγησης εφαρμόζει την διαδικασία προσωρινής αποθήκευσης διαφορετικά, και ο τρόπος με τον οποίο κάθε χρήστης φορτώνει μια σελίδα έχει συνέπειες για την μνήμη cache. Για παράδειγμα, εάν ένας χρήστης πληκτρολογήσει το 'ENTER' στη γραμμή διευθύνσεων URL για να φορτώσει μια ιστοσελίδα χρησιμοποιώντας το πρόγραμμα περιήγησης Chrome, το συγκεκριμένο πρόγραμμα θα φορτώσει στατιστικά πόρων από τη μνήμη cache χωρίς να τα επικυρώσει. Εάν ο χρήστης κάνει κλικ στο κουμπί Ανανέωση ή φορτώσει μια σελίδα χρησιμοποιώντας κάποιο hotkey, τότε ο Chrome θα επικυρώσει τη μνήμη cache. Από αυτό το παράδειγμα γίνεται κατανοητό ότι ακόμα και όταν το πρόγραμμα περιήγησης στείλει ένα HTTP αίτημα οι ετικέτες ETag δεν μπορούν να θεωρηθούν έγκυρες και αμετάβλητες [23].
Η πρώτη χρήση των ετικετών ETags για την παρακολούθηση των χρηστών παρατηρήθηκε το 2011, όταν ο ιστότοπος hulu.com χρησιμοποίησε την υπηρεσία KISSmetrics για την ανοικοδόμηση των διαγραμμένων cookies HTTP και HTML 5. Για να αποφύγει πλήρως την παρακολούθηση μέσω ετικετών ETags ο χρήστης πρέπει να διαγράφει την προσωρινή μνήμη από το πρόγραμμα περιήγησης πριν από κάθε επίσκεψη στον ιστότοπο ή χρησιμοποιώντας έναν διακομιστή που δεν υποστηρίζει την διαδικασία αυτή. Η παρακολούθηση είναι επίσης δυνατή κατά τη διάρκεια μίας περιόδου ιδιωτικής περιήγησης, καθώς η μνήμη cache διατηρείται μέχρι να κλείσει το τελευταίο παράθυρο του προγράμματος περιήγησης [03].

Λειτουργικές μνήμες cache

Οι λειτουργικές μνήμες cache (operational cache) είναι μνήμες που χρησιμοποιούνται για την αποθήκευση πληροφοριών που σχετίζονται με λειτουργίες των περιηγητών. Τέτοιου είδους

πληροφορίες μπορεί να περιλαμβάνουν μόνιμες ανακατευθύνσεις ή μια λίστα των domain που θα πρέπει να χρησιμοποιούνται από κοινού μέσα από το πρωτόκολλο αυστηρής ασφάλειας μεταφοράς HTTP (HTTP Strict Transport Security – HSTS) [03]:

HTTP 301 redirect cache

Οι ανακατευθύνσεις HTTP 301 χρησιμοποιούνται από τον εξυπηρετητή ιστού (web server) για να ενημερώσουν τα προγράμματα περιήγησης ότι η ζητούμενη διεύθυνση URL ανακατευθύνεται "μόνιμα" σε άλλη. Το πρόγραμμα περιήγησης έπειτα αποθηκεύει προσωρινά την ανακατεύθυνση και την χρησιμοποιεί αντί της αρχικής διεύθυνσης URL. Η ανακατεύθυνση 301 είναι το κλειδί για τη διατήρηση της αρχής του domain ενός ιστότοπου όταν η διεύθυνση URL του αλλάξει για οποιονδήποτε λόγο. Στέλνει εύκολα τους επισκέπτες και τις μηχανές αναζήτησης σε διαφορετική διεύθυνση URL από αυτήν που είχαν αρχικά ζητήσει χωρίς να χρειάζεται να πληκτρολογήσει ο χρήστης διαφορετική διεύθυνση URL, μέσω τις κεφαλίδα ανακατεύθυνσης HTTP 301. Ωστόσο, αυτός ο μηχανισμός μπορεί να χρησιμοποιηθεί για την παρακολούθηση του χρήστη από τρίτους (third parties). Όταν ο χρήστης πληκτρολογήσει το URL μιας ιστοσελίδας που θέλει να επισκεφτεί, ο κώδικας του θα εξετάσει τη διαδρομή / τις παραμέτρους διεύθυνσης URL για να διαπιστώσει εάν υπάρχει ένα μοναδικό αναγνωριστικό. Εάν η διεύθυνση URL δεν περιέχει ένα μοναδικό αναγνωριστικό τότε η σελίδα παρακολούθησης χρησιμοποιεί τις κεφαλίδες HTTP 301 για να εκχωρήσει ένα μοναδικό αναγνωριστικό στον χρήστη. Όταν γίνει η ανακατεύθυνση, το πρόγραμμα περιήγησης θα αποθηκεύσει προσωρινά τις πληροφορίες ανακατεύθυνσης, έτσι ώστε την επόμενη φορά που ο χρήστης θα συνδεθεί στη σελίδα παρακολούθησης, ο χρήστης θα μεταφερθεί στη σελίδα παρακολούθησης με το μοναδικό αναγνωριστικό του [24].

Αυτό που καθιστά την τεχνική αυτή μοναδική και δύσκολο να μπλοκαριστεί είναι ότι δεν βασίζεται σε μόνιμο μηχανισμό αποθήκευσης (stored-based) ή σε τέχνασμα βασισμένο σε JavaScript. Η ανακατεύθυνση μπορεί να γίνει σε ένα iframe, επομένως είναι διαφανής για τους χρήστες [03].

HTTP Strict Transport Security (HSTS)

Το HSTS είναι ένας μηχανισμός ασφάλειας, στόχος του οποίου είναι να εξασφαλίζει ότι η επικοινωνία μεταξύ διαδικτυακών διακομιστών και προγραμμάτων περιήγησης πραγματοποιείται μόνο με ασφαλείς συνδέσεις HTTPS και ποτέ μέσω του πρωτοκόλλου

HTTPπου δεν παρέχει καμία ασφάλεια. Για το σκοπό αυτό, το πρόγραμμα περιήγησης δημιουργεί μια βάση δεδομένων στην οποία αποθηκεύει μια λίστα των αρχικά καταχωρημένων ιστοσελίδων και στη συνέχεια, προσθέτει σταδιακά νέες ιστοσελίδες τις οποίες ο χρήστης επισκέπτεται μόνο μέσω αιτημάτων HTTPS. Τα αιτήματα HTTPS δημιουργούνται όταν στα αρχικά αιτήματα HTTP των προγραμμάτων περιήγησης επιστρέφει μια ανακατεύθυνση HTTP 301, υποδεικνύοντας ότι το συγκεκριμένο domain βρίσκεται μόνιμα διαθέσιμο σε άλλη URL, η οποία χρησιμοποιεί ασφαλή σύνδεση HTTPS [25]. Ωστόσο ένα ελάττωμα πάνω στον μηχανισμό ασφαλείας HSTS επιτρέπει στους ιστότοπους να εγκαθιστούν "super cookies" που μπορούν να χρησιμοποιηθούν για να παρακολουθούν την περιήγηση των χρηστών ακόμη και όταν είναι ενεργοποιημένη η ιδιωτική περιήγηση. Οι ερευνητές ασφάλειας αναφέρουν για το HSTS ότι επιτρέπει σε έναν ιστότοπο να υποδεικνύει ότι πρέπει η πρόσβαση να γίνεται πάντα χρησιμοποιώντας ασφαλή σύνδεση που κρυπτογραφεί την επικοινωνία με τον ιστότοπο. Αυτή η "σημαία" (flag) αποθηκεύεται στη συνέχεια από το πρόγραμμα περιήγησης, διασφαλίζοντας ότι οι μελλοντικές επισκέψεις στον ιστότοπο είναι ασφαλείς. Αλλά μέσω αυτής της ασφάλειας μπορεί επίσης να γίνει μια κατάχρηση, μέσω της οποίας μπορεί να αποθηκευτεί ένας μοναδικός αριθμός που μπορεί να χρησιμοποιηθεί για την παρακολούθηση του προγράμματος περιήγησης. Και επειδή το HSTS μεταφέρεται και σε ιδιωτική περιήγηση, σημαίνει ότι το "super cookie" μπορεί να χρησιμοποιηθεί για να εντοπίσει τους χρήστες που προσπαθούν μέσω τις ιδιωτικής περιήγησης να καλύψουν τα ίχνη τους [26]. Το πρόγραμμα περιήγησης Firefox, μετά από τις αναλύσεις των ερευνητών, έχει αναπτύξει λύση στο ζήτημα, αφού δεν μεταφέρει πλέον το HSTS σε ιδιωτικά παράθυρα. Ωστόσο ένας σύμβουλος τεχνολογίας και λογισμικού Sam Greenhalgh, μετά από αυτήν την ανακοίνωση της εταιρείας Firefox ανέφερε ότι «*Είναι μία κίνηση που ευνοεί την προστασία της ιδιωτικής ζωής θυσιάζοντας την ασφάλεια. Αν προσπαθήσει κάποιος να αγοράσει κάτι από έναν ιστότοπο χρησιμοποιώντας μια ιδιωτική περιήγηση του Firefox, και φορτώσει μια μη κρυπτογραφημένη έκδοση της ιστοσελίδας – τότε το πρόγραμμα περιήγησης μιας και δεν υποστηρίζει το HSTS, όλες οι πληροφορίες της πιστωτικής κάρτας δεν θα κρυπτογραφηθούν ποτέ*».

2.4 Μηχανισμοί Αποτυπώματος

Η δημιουργία ψηφιακών αποτυπωμάτων είναι μια ομάδα μεθόδων που χρησιμοποιούν μια ευρεία γκάμα τεχνολογιών με στόχο την παρακολούθηση του χρήστη. Στο παρελθόν τέτοια ψηφιακά αποτυπώματα αποδείχθηκαν χρήσιμα για την ανίχνευση και την πρόληψη της κλοπής ταυτότητας σε απευθείας σύνδεση και της απάτης με πιστωτικές κάρτες. Σήμερα όμως χρησιμοποιούνται όλο και περισσότερο από τις εταιρείες για να μάθουν για τους χρήστες, να μελετήσουν τη συμπεριφορά τους και για να καταφέρουν να διαφημίσουν τα προϊόντα τους. Οι οργανισμοί γίνονται ολοένα και πιο ικανοί στο να χρησιμοποιούν τα αποτυπώματα της συσκευής των χρηστών για να παρακολουθούν τις ενέργειές και να προβλέπουν τι θα κάνουν στη συνέχεια. Για τους διαφημιζόμενους, για παράδειγμα, αυτό σημαίνει ότι μπορούν να αρχίσουν να παρακολουθούν τις ενέργειες από την πρώτη φορά που ένας χρήστης δείχνει ενδιαφέρον για ένα προϊόν [27]. Υπάρχουν πολλές ιστοσελίδες που μας δείχνουν εκτεταμένες πληροφορίες που μπορούν να συλλεχθούν αυτόματα από τον υπολογιστή του χρήστη. Σύμφωνα με έρευνες που έχουν πραγματοποιηθεί, 40 από τους κορυφαίους 10 000 ιστότοπους της Alexa.com, συμπεριλαμβανομένου του skype.com, χρησιμοποιούν δέσμες αποτυπωμάτων (fingerprinting scripts) από τις βιβλιοθήκες δακτυλικών αποτυπωμάτων τριών μεγάλων εμπορικών εταιρειών: **BlueCava**, **Iovation**, **Threat Metrix** [28].

Αυτά τα διακριτά χαρακτηριστικά μπορεί να μην είναι μοναδικά, αλλά σε συνδυασμό έχουν την τάση να αναγνωρίζουν μοναδικά τους χρήστες [29]. Σε γενικές γραμμές οι μηχανισμοί αποτυπώματος μπορούν να χωριστούν σε 3 κατηγορίες:

1)Μηχανισμοί αποτυπωμάτων των προγραμμάτων συσκευής- περιήγησης

2)Μηχανισμοί αποτυπώματος λειτουργικού συστήματος

3)Μηχανισμοί αποτυπώματος δικτύου και γεωγραφικής θέσης

Μηχανισμοί αποτυπώματος συσκευής

Όπως η συλλογή των αποδείξεων που ακολουθούν οι ντετέκτιβ για να συσσωρέψουν στοιχεία σε έναν ύποπτο, έτσι και η λήψη ψηφιακών αποτυπωμάτων για συσκευές λειτουργεί συλλέγοντας τμήματα πληροφοριών για να σχηματιστεί ένα γενικό αναγνωριστικό [30].

Καθώς οι άνθρωποι είναι όλο και περισσότερο συνδεδεμένοι και εκτελούν όλο και περισσότερες ενέργειες στο διαδίκτυο, υπάρχει η τάση να χρησιμοποιούν πολλαπλές συσκευές [27]. Το 2016 η Google ανέφερε ότι το 70% των χρηστών του διαδικτύου συνδέεται

καθημερινά με τουλάχιστον δύο διαφορετικές συσκευές και αυτό καθιστά όλο και πιο δύσκολο για τις διαφημιστικές εταιρείες να συνδεθούν προσωπικά με το κοινό-στόχο τους [31]. Ο μηχανισμός αποτυπώματος συσκευής ενισχύεται επειδή τα κανονικά μέσα παρακολούθησης στο διαδίκτυο αντιμετωπίζουν όλο και περισσότερα προβλήματα. Για παράδειγμα, τα cookies, τα οποία αποτέλεσαν το επίκεντρο της ψηφιακής διαφήμισης εδώ και χρόνια, έχουν γίνει ολοένα και πιο αβάσιμα στο σημερινό περιβάλλον που σέβεται την ιδιωτικότητα. Πρώτα απ 'όλα, τα cookies δεν προσφέρουν αξιόπιστο τρόπο παρακολούθησης της χρήσης των κινητών συσκευών. Δεύτερον, τα cookies μπορούν εύκολα να διαγραφούν από τον καταναλωτή - και οι χρήστες όλο και περισσότερο ακολουθούν αυτήν την τακτική. Το αποτύπωμα της συσκευής προσφέρει μια μέθοδο δημιουργίας αντιγράφων ασφαλείας για την παρακολούθηση όταν τα cookie δεν μπορούν να αναγνωρίσουν με αξιοπιστία τους χρήστες. Οι πληροφορίες που χρησιμοποιούνται για τη δημιουργία ενός ψηφιακού αποτυπώματος συσκευής μπορούν να περιλαμβάνουν:

• Έκδοση προγράμματος περιήγησης
• Λειτουργικό σύστημα
• Στοιχεία που έχουν εγκατασταθεί (plugins / γραμματοσειρές κ.λπ.)
• Ρυθμίσεις τοποθεσίας και ζώνης ώρας.

Οι μηχανισμοί αποτυπώματος του προγράμματος περιήγησης (browser fingerprinting) αποτελούν διαδικασίες συλλογής ιδιοτήτων ή χαρακτηριστικών του προγράμματος περιήγησης του πελάτη για διάφορους λόγους, κυρίως, για την ταυτοποίηση του χρήστη. Παραδείγματα αυτών των ιδιοτήτων ή των χαρακτηριστικών αποτελούν οι διαστάσεις της οθόνης, η γραμματοσειρά του συστήματος, τα plug-in που υποστηρίζει το πρόγραμμα περιήγησης, η ζώνη ώρας, η έκδοση του προγράμματος περιήγησης, κ.λπ. Ο συνδυασμός αυτών των ιδιοτήτων ονομάζεται αποτύπωμα του προγράμματος περιήγησης (browser fingerprint) και μπορεί να χρησιμεύσει ως μοναδικό αναγνωριστικό τόσο του προγράμματος περιήγησης όσο και της συσκευής στην οποία λειτουργεί[32]. Με άλλα λόγια, είναι πολύ πιθανό να υπάρχει μόνο μία συσκευή με αυτά τα χαρακτηριστικά και, συνεπώς, η συλλογή αυτών των πληροφοριών δημιουργεί ένα μοναδικό «ψηφιακό αποτύπωμα» αυτής.

Αρκετά χαρακτηριστικά μπορούν να συλλεχθούν μέσω των κεφαλίδων HTTP. Για παράδειγμα η συμβολοσειρά User-Agent (UA) περιέχεται στις κεφαλίδες HTTP και προορίζεται για τον εντοπισμό συσκευών που ζητούν περιεχόμενο στο διαδίκτυο. Η συμβολοσειρά user agent (User-agent string), είναι ένα πεδίο της κεφαλίδας των HTTP αιτημάτων που συνήθως περιέχει

αναγνωριστικά του προγράμματος περιήγησης, της μηχανής απόδοσης (rendering engine), της έκδοσης του προγράμματος περιήγησης και του λειτουργικού συστήματος [33]. Στην Εικόνα 5 παρουσιάζονται παραδείγματα User-Agent String από τα πέντε μεγαλύτερα προγράμματα περιήγησης στο διαδίκτυο.

```
Mozilla Firefox
Mozilla/5.0 (Windows NT 6.1; Win64; x64; rv:47.0) Gecko/20100101 Firefox/47.0
Mozilla/5.0 (Macintosh; Intel Mac OS X x.y; rv:42.0) Gecko/20100101 Firefox/42.0
Chrome
Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/51.0.2704.103 Safari/537.36
Safari
Mozilla/5.0 (iPhone; CPU iPhone OS 10_3_1 like Mac OS X) AppleWebKit/603.1.30
(KHTML, like Gecko) Version/10.0 Mobile/14E304 Safari/602.1
Opera
Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/51.0.2704.106 Safari/537.36 OPR/38.0.2220.41
Internet Explorer
Mozilla/5.0 (compatible; MSIE 9.0; Windows Phone OS 7.5; Trident/5.0;
IEMobile/9.0)
```

Εικόνα 5 : Συμβολοσειρές προγραμμάτων περιήγησης (Πηγή https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/User-Agent)

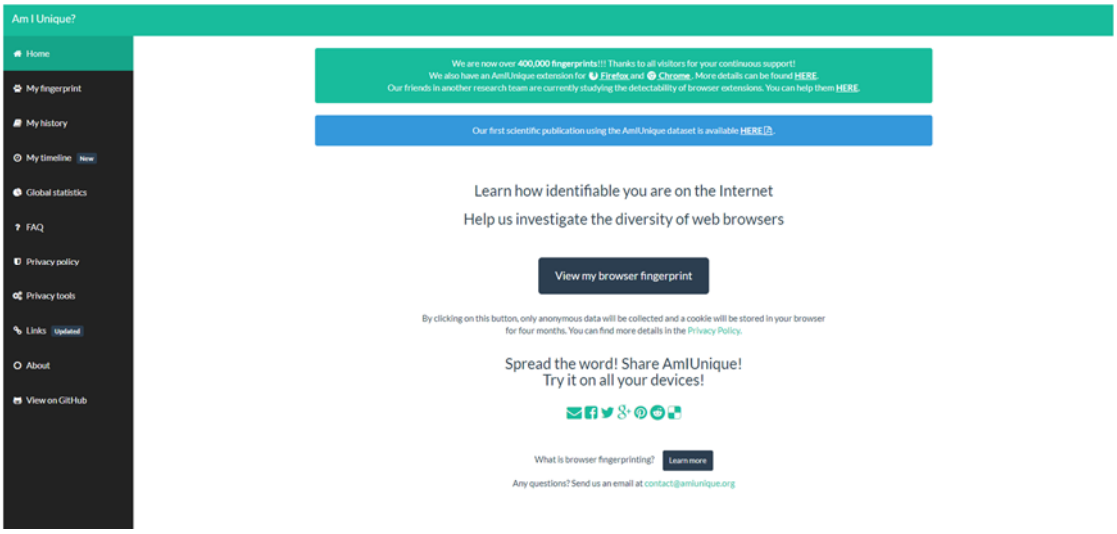
Αναλύοντας μια συμβολοσειρά User-Agent μπορούν να συλλεχθούν αρκετές πληροφορίες για τον χρήστη όπως στο παρακάτω παράδειγμα :

Mozilla/5.0 (Macintosh; Intel Mac 0S X 10_10_2) AppleWebKit/537.36 (KHTM, like Gecko) Chrome/51.0.2704.84 Safari/537.36

- Η εφαρμογή User-Agent String είναι το Mozilla έκδοση 5.0
- Το λειτουργικό σύστημα είναι OSX έκδοση 10.2.2 (και λειτουργεί σε Mac).
- Το πρόγραμμα περιήγησης είναι Chrome με έκδοση 51.0.2704.84
- Ο χρήστης βασίζεται σε έκδοση Safari 537.36
- Η μηχανή αποδοχής που είναι υπεύθυνη για την προβολή περιεχομένου σε αυτήν τη συσκευή είναι η AppleWebKit έκδοση 537.36.

Αρκετοί μηχανισμοί αποτυπώματος χρησιμοποιούν κώδικα JavaScript ή άλλο κώδικα στον τοπικό υπολογιστή του χρήστη με σκοπό την παρατήρηση και άντληση δεδομένων που αφορούν πρόσθετα χαρακτηριστικά του προγράμματος περιήγησης ή του τερματικού. Οι μηχανισμοί αυτής της κατηγορίας περιλαμβάνουν πρόσβαση στο μέγεθος της οθόνης, καταμέτρηση (enumeration) των γραμματοσειρών (fonts) ή των plug-in. Έρευνες έχουν δείξει ότι πολλές εταιρείες χρησιμοποιούν το Flash για να αποτυπώσουν το περιβάλλον ενός χρήστη. Το Adobe Flash είναι ένα ιδιόκτητο plug-in προγράμματος περιήγησης που έχει ευρεία χρήση μεταξύ των χρηστών, δεδομένου ότι παρέχει τρόπους για την παροχή πλούσιου περιεχομένου πολυμέσων που παραδοσιακά δεν θα μπορούσε να εμφανίζεται με τη χρήση HTML. Παρά το γεγονός ότι το Flash έχει επικριθεί για κακή απόδοση, έλλειψη σταθερότητας και ότι οι νεότερες τεχνολογίες, όπως το HTML5, μπορούν ενδεχομένως να παρέχουν όλα τα πλεονεκτήματα που παρέχει το Flash, εξακολουθεί εν τούτοις να είναι διαθέσιμο στη μεγάλη πλειοψηφία των επιτραπέζιων υπολογιστών [34].

Το AmIUnique.org (εικόνα 6) είναι ένας δικτυακός τόπος αφιερωμένος στα ψηφιακά αποτυπώματα συσκευών, με στόχο τόσο τη συλλογή δεδομένων σχετικά με τη διαφορετικότητα των συσκευών όσο και την ενημέρωση των χρηστών σχετικά με τις συνέπειες της ύπαρξης των ψηφιακών αυτών αποτυπωμάτων. Δημιουργήθηκε από μία ομάδα ερευνητών όπου ο κύριος σκοπός τους ήταν να ενημερώσουν τους χρήστες σχετικά με την θέση τους στο παγκόσμιο σύστημα διαδικτύου [33]. Πλέον η ιστοσελίδα περιέχει πάνω από 400.000 αποτυπώματα τα οποία έχουν χρησιμοποιηθεί κατά καιρούς από ερευνητές για εξαγωγή συμπερασμάτων. Επίσης παρέχουν και μια επέκταση προστασίας για τα προγράμματα περιήγησης **Firefox** και **Chrome.**



Εικόνα 6 : Αρχική σελίδα του ιστότοπουAmIUnique(Πηγή https://amiunique.org/)

Στον πίνακα 2 παρουσιάζεται ένα παράδειγμα δαχτυλικού αποτυπώματος από το πρόγραμμα περιήγησης ενός χρήστη στο διαδίκτυο. Τα αποτελέσματα λήφθηκαν μέσω τις εφαρμογής **AmIUnique.org**. Μέσα από το User Agent διακρίνεται εύκολα το λειτουργικό σύστημα του χρήστη (Windows 10) και το πρόγραμμα περιήγησης που χρησιμοποιεί(Chrome).Επίσης από τον κώδικα JavaScript υπάρχει δυνατότητα να διακριθούν τα plug-in που χρησιμοποιεί ο χρήστης (πχ Chrome PDF Viewer, Native Client), τα cookies αν είναι ενεργά στον περιηγητή, η ανάλυση οθόνης του χρήστη, η ζώνη ώρας και αν ο χρήστης χρησιμοποιεί επέκταση που μπλοκάρει ανεπιθύμητες διαφημίσεις (πχ AdBlock). Ακόμη μέσω του Web GLAPI εφαρμογή της JavaScriptυπάρχει η δυνατότητα ανίχνευσης της GPU του χρήστη(στο παράδειγμα AMD Radeon HD 8570D). Το WebGL χρησιμοποιείται για απόδοση γραφικών 3D σε οποιοδήποτε πρόγραμμα περιήγησης και χωρίς την χρήση plug-ins. Οι εφαρμογές WebGL αποτελούνται από κώδικα ελέγχου γραμμένο σε JavaScript και κώδικα ειδικών εφέ που εκτελείται στη GPU ενός υπολογιστή, και έχουν την δυνατότητα να παράγουν αποτυπώματα από την GPU του χρήστη [33].

Attribute	*Source*	*Value*
User Agent	HTTP header	"Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/64.0.3282.186 Safari/537.36"
Accept	HTTP header	"text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8"
Content Encoding	HTTP header	"gzip, deflate, br"
Content Language	HTTP header	"el-GR,el;q=0.9"
List of plugins	JavaScript	"Plugin 0: Chrome PDF Plugin; Portable Document Format; internal-pdf-viewer. Plugin 1: Chrome PDF Viewer;mhjfbmdgcfjbbpaeojofohoefgiehjai. Plugin 2: Native Client; ; internal-nacl-plugin. Plugin 3: Widevine Content Decryption Module; Enables Widevine licenses for playback of HTML audiovideo content. version: 1.4.9.1070; widevinecdmadapter.dll. "
Cookies Enabled	JavaScript	Yes
UseLocal/SessionStorage	JavaScript	Yes
Screen Resolution	JavaScript	"1920x1080x24"
TimeZone	JavaScript	"-120"
Web GL Vendor	JavaScript	"Google Inc."
Web GL Renderer	JavaScript	"ANGLE (AMD Radeon HD 8570D Direct3D11 vs_5_0 ps_5_0)"
Use o f Ad Block	JavaScript	Yes

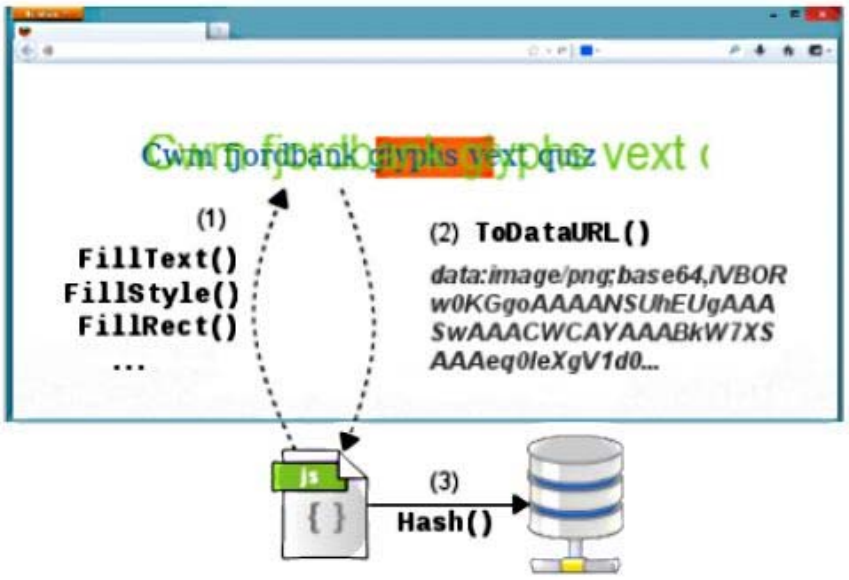
Πίνακας 2 : Παράδειγμα αποτυπώματος χρήστη μέσω της εφαρμογής AmIUnique

Canvas Fingerprinting

Τα ψηφιακά αποτυπώματα σε καμβά είναι ένας τύπος τεχνικών "ψηφιακών αποτυπωμάτων" του προγράμματος περιήγησης που επιτρέπουν στους ιστότοπους να αναγνωρίζουν και να παρακολουθούν με μοναδικό τρόπο τους επισκέπτες χρησιμοποιώντας το στοιχείο HTML 5 canvas αντί για cookies του προγράμματος περιήγησης ή άλλα παρόμοια μέσα. Το HTML Canvas είναι ένα API το οποίο χρησιμοποιείται για την σχεδίαση γραφικών και κινούμενων εικόνων σε μια ιστοσελίδα μέσω script σε JavaScript. Τα στοιχεία του HTML Canvas επιτρέπουν σε scripts να σχεδιάζουν ή να φορτώνουν εικόνες με προγραμματισμό στο χώρο που ορίζει το στοιχείο [14].

Η λήψη ψηφιακών αποτυπωμάτων σε καμβά αρχίζει όταν ένας ιστότοπος δίνει στο πρόγραμμα περιήγησης το καθήκον να σχεδιάζει ένα αντικείμενο καμβά. Οι ιστότοποι χρησιμοποιούν JavaScript για να δώσουν στα προγράμματα περιήγησης την εργασία να σχεδιάσουν μια εικόνα στο αντικείμενο του καμβά χρησιμοποιώντας ένα προκαθορισμένο

σενάριο (script). Στη συνέχεια το script καλεί τη μέθοδο ToDataURL για να πάρει τα δεδομένα pixel εικόνας του καμβά σε μορφή dataURL, τα οποία είναι κωδικοποιημένα σε base64 και τα αποθηκεύει σε μία συνάρτηση κατακερματισμού Hash Function (εικόνα 7). Τα Hashes χρησιμεύουν ως ψηφιακά αποτυπώματα και μπορούν να συνδυαστούν με άλλες ιδιότητες του προγράμματος περιήγησης, όπως η λίστα με τις προσθήκες, η λίστα με τις γραμματοσειρές [35].



Εικόνα 7 : Βασική ροή λειτουργιών δακτυλικών αποτυπωμάτων σε καμβά (Πηγή https://www.geek.com/apps/canvas-fingerprinting-is-like-a-cookie-you-cant-block-and-thousands-of-sites-are-using-it-1599967/)

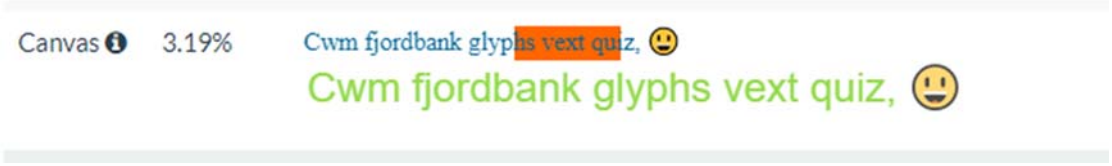
Το κύριο χαρακτηριστικό της αποτύπωσης σε καμβά και αυτό που το κάνει να ξεχωρίζει είναι ότι διαφορετικοί υπολογιστές θα «τραβήξουν» την εικόνα με έναν ελαφρώς διαφορετικό τρόπο. Ακόμη και αν οι παραγόμενες εικόνες μοιάζουν στο ανθρώπινο μάτι, υπάρχουν μικρές παραλλαγές που τους επιτρέπουν να διαφοροποιούνται, και αυτό συμβαίνει λόγω την διαφορετικής αρχιτεκτονικής του κάθε υπολογιστή [36]. Αυτή η καινοτομία μπορεί να αποδώσει μοναδικότητα αναγνωριστικού εικόνας κατά 99% (Εικόνα 8).

Your Fingerprint :

Signature	✓ 81E42236			
Uniqueness	99.28% (1855 of 258561 user agents have the same signature)			
Image File Details :	BrowserLeaks.com <canvas> 1.0			
File Size	4689 bytes			
Number of Colors	209			
PNG Hash	B2522922E7860375635A13B552A960FD			
PNG Headers	Chunk :	Length :	CRC :	Content :
	IHDR	13	477A703E	PNG image header: 220x30, 8 bits/sample, truecolor+alpha, noninterlaced
	IDAT	4632	81E42236	PNG image data
	IEND	0	AE426082	end-of-image marker

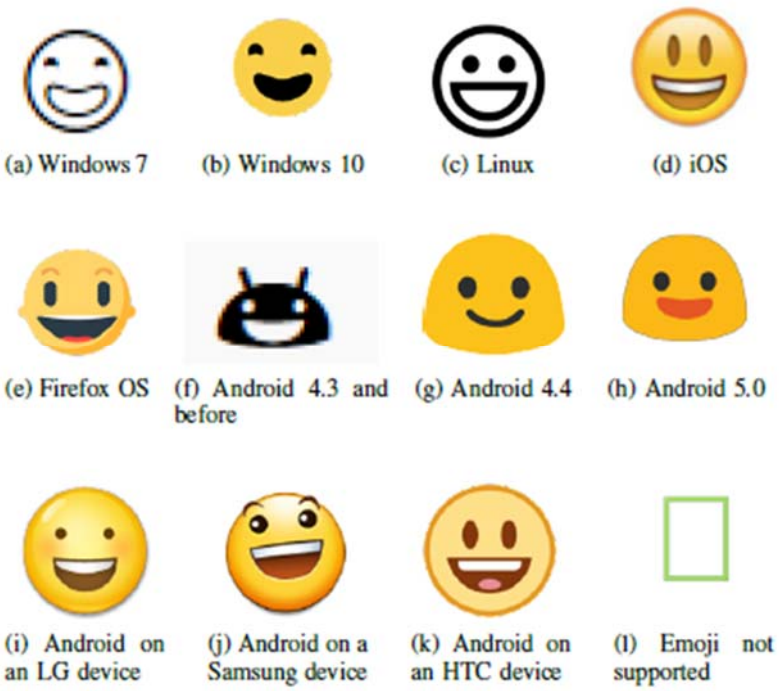
Εικόνα 8: Αποτύπωμα εικόνας καμβά υπολογιστή χρήστη (Πηγή https://browserleaks.com/canvas#further-reading)

Σύμφωνα με έρευνες η λήψη ψηφιακών αποτυπωμάτων σε καμβά είναι η πιο συχνά χρησιμοποιούμενη μέθοδος δημιουργίας ψηφιακών αποτυπωμάτων, η οποία είναι παρούσα σε πάνω από το 5,5% των κορυφαίων 100000 ιστότοπων όπως προσδιορίζονται στο Alexa.com [14]. Τα σενάρια δημιουργίας ψηφιακών αποτυπωμάτων από το **addthis.com** ήταν υπεύθυνα για την πλειοψηφία (95%) των προσπαθειών δημιουργίας ψηφιακών αποτυπωμάτων στους ιστότοπους. Στην εικόνα 9 παρουσιάζεται ένα αποτύπωμα σε καμβά από τον ιστότοπο **AmIUnique.com**.



Εικόνα 9 : Αποτύπωμα εικόνας καμβά (Πηγή https://amiunique.org/fp)

Η εικόνα 10 δείχνει παραστάσεις του emoji "Smiling face with open mouth" σε διάφορα λειτουργικά συστήματα και κινητές συσκευές. Η χρήση του emojis μπορεί να είναι μια ισχυρή τεχνική για την αποκάλυψη πληροφοριών, ειδικά σε κινητές συσκευές όπου οι κατασκευαστές τηλεφώνων παρέχουν τις δικές τους σειρές emojis [33].



Εικόνα 10 : Σύγκριση του "Smiling face with open mouth" emoji σε διαφορετικές συσκευές και λειτουργικά συστήματα [33]

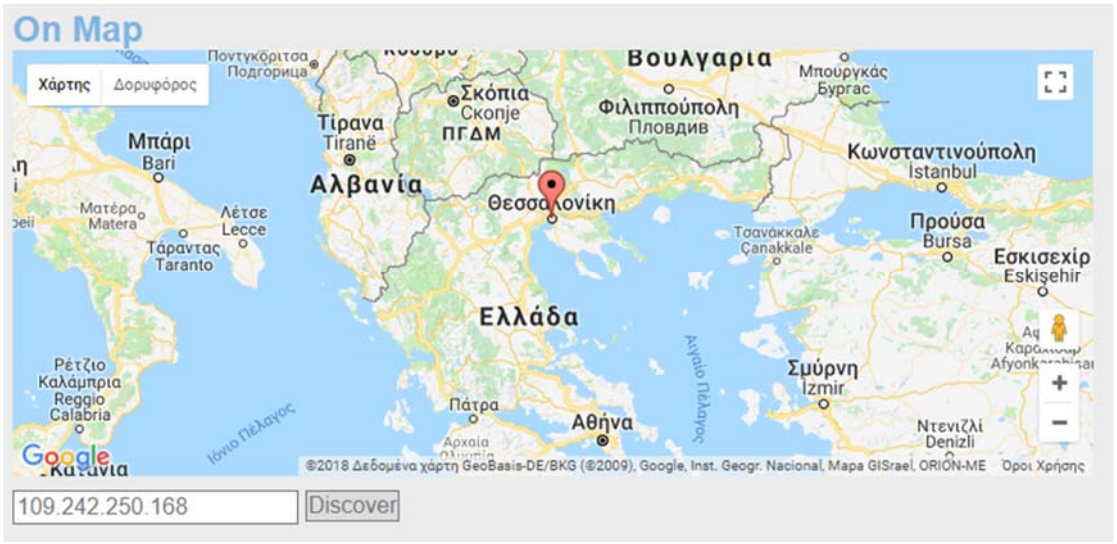
Αποτύπωμα λειτουργικού συστήματος

Η έκδοση και η αρχιτεκτονική (32/64 bit) του λειτουργικού συστήματος μπορούν να αναγνωριστούν τόσο από το JavaScript όσο και από το Flash. Το JavaScript επίσης διευκολύνει την αναγνώριση της γλώσσας του συστήματος και της γλώσσας του χρήστη για το σύστημα, όπως επίσης της τοπικής ζώνης ώρας και της τοπικής ημερομηνίας. Τα βάθη χρώματος και οι διαστάσεις της οθόνης μπορούν να ανιχνευθούν τόσο από το JavaScript όσο και από το Flash. Το Flash μπορεί επίσης να ανιχνεύσει αν το σύστημα διαθέτει δυνατότητες ήχου, αν η πρόσβαση στην κάμερα και το μικρόφωνο του χρήστη έχει απαγορευτεί ή επιτραπεί, εάν το σύστημα υποστηρίζει ή δεν υποστηρίζει εκτύπωση και εάν έχει απαγορευτεί ή επιτρέπεται η πρόσβαση ανάγνωσης στον σκληρό δίσκο του χρήστη [03]. Επίσης αναγνώριση του λειτουργικού συστήματος μπορεί να γίνει και μέσω της στοίβας του TCP πρωτόκολλου. Η αποτύπωση αυτή περιλαμβάνει τη συγκέντρωση πληροφοριών σχετικά με τις ρυθμιστικές ιδιότητες του επιπέδου TCP, παράμετροι που όταν συνδυαστούν μπορεί να χρησιμεύσουν και για την δημιουργία ψηφιακού αποτυπώματος του λειτουργικού συστήματος του υπολογιστή του χρήστη. Ενδεικτικά κάποιες τιμές που βοηθούν σε αυτόν τον σκοπό είναι το μέγεθος του πακέτου, το TTL, τα flags του πακέτου και το μέγεθος του παραθύρου cwnd [37].

Αποτύπωμα δικτύου και τοποθεσίας

Η διεύθυνση IP της συσκευής ή του δικτύου στο οποίο βρίσκεται αποτελεί ένα από τα πιο κοινά χρησιμοποιούμενα δεδομένα στην προσπάθεια εντοπισμού των χρηστών του διαδικτύου. Ένας "παραδοσιακός" τρόπος για τον εντοπισμό συσκευών είναι μέσω της διεύθυνσης IP. Αυτό μπορεί να μην είναι πάντα εφικτό, καθώς οι συσκευές ενδέχεται να αλλάζουν IP με την πάροδο του χρόνου (πχ δυναμική IP). Επιπλέον, ένας χρήστης μπορεί να χρησιμοποιήσει έναν ανώνυμο/διακομιστή μεσολάβησης για να αποκρύψει την πραγματική IP διεύθυνση του (Tor browser). Ωστόσο, η γνώση της διεύθυνσης IP ενός χρήστη μπορεί να αποδειχθεί πολύτιμη κατά την διαδικασία της λήψης ψηφιακών αποτυπωμάτων. Οι διευθύνσεις IP γενικά δεν αλλάζουν ταχύτατα, γεγονός που θα απλοποιήσει την παρακολούθηση ενός χρήστη τουλάχιστον κατά τη διάρκεια μιας μόνο συνεδρίας και μπορεί παράλληλα να παρέχει ενδείξεις ως προς τη γεωγραφική θέση ενός χρήστη (Εικόνες 11,12) [28].

Εικόνα 11 : Απόκτηση ιδιωτικής διεύθυνσης IP μέσω του WEB RTC (Πηγή http://net.ipcalf.com/)



Εικόνα 12 : Απόκτηση δημόσιας IP και αποτύπωση στον χάρτη (Πηγή http://www.ipfingerprints.com/)

Οι διευθύνσεις IP στα δίκτυα υπολογιστών δεν αντιπροσωπεύουν ακριβώς συγκεκριμένες γεωγραφικές τοποθεσίες. Ωστόσο, είναι θεωρητικά δυνατό να προσδιοριστεί η φυσική θέση των διευθύνσεων IP σε πολλές περιπτώσεις. Τα λεγόμενα συστήματα γεωγραφικής τοποθέτησης προσπαθούν να αντιστοιχίσουν διευθύνσεις IP σε γεωγραφικές τοποθεσίες χρησιμοποιώντας μεγάλες βάσεις δεδομένων υπολογιστή. Ορισμένες βάσεις δεδομένων γεωγραφικού εντοπισμού διατίθενται προς πώληση, ενώ μερικές μπορούν επίσης να αναζητηθούν δωρεάν στο διαδίκτυο. Οι διαφημιζόμενοι μπορούν να χρησιμοποιήσουν μια υπηρεσία γεωγραφικού εντοπισμού για να παρακολουθήσουν τη γεωγραφική κατανομή των επισκεπτών στον ιστότοπο τους. Εκτός από την ικανοποίηση της γενικής περιέργειας, οι προηγμένες τοποθεσίες Web μπορούν επίσης να αλλάξουν δυναμικά το περιεχόμενο που εμφανίζεται σε κάθε επισκέπτη με βάση την τοποθεσία τους. Αυτοί οι ιστότοποι ενδέχεται επίσης να αποκλείουν την πρόσβαση σε επισκέπτες από ορισμένες χώρες ή τοπικές τοποθεσίες [38].

Τα αποτυπώματα των διευθύνσεων IP μπορούν να συγκεντρωθούν από την ανάλυση της κυκλοφορίας δεδομένων ανάμεσα στον ιστότοπο και το χρήστη και μέσω των κεφαλίδων HTTP. Επίσης πληροφορίες σχετικές με τις διευθύνσεις IP των χρηστών μπορεί ακόμα να προέρχονται από τους διακομιστές μεσολάβησης (proxy servers), λειτουργίες JavaScript και μικροεφαρμογές Flash [03].

Κεφάλαιο 3
Real Time Bidding (RTB)

Η άτυπη δημοπρασία διαφημιστών σε πραγματικό χρόνο (Real Time Bidding, RTB) έχει αναδειχθεί από το 2009 και αποτελεί μια πολλά υποσχόμενη τακτική για τη διαφήμιση στο διαδίκτυο. Σε αντίθεση με τη συμβατική διαφημιστική αναζήτηση ή τη συμφραζόμενη διαφήμιση, όπου ένας διαφημιζόμενος προκαθορίζει μια τιμή προσφοράς για κάθε επιλεγμένη λέξη-κλειδί για την καμπάνια του, η RTB επιτρέπει σε έναν διαφημιζόμενο να χρησιμοποιεί αλγόριθμους υπολογιστών για να υποβάλει προσφορά για κάθε εμφάνιση μέσα σε πολύ σύντομο χρονικό διάστημα, της τάξης των 100ms [39]. Οι πληροφορίες αυτές βασίζονται στα χαρακτηριστικά της εμφάνισης, όπως τα cookies του χρήστη και οι πληροφορίες του περιεχομένου.

Μια δημοπρασία σε πραγματικό χρόνο φιλοξενείται από ένα μεσάζοντα ο οποίος αποτελεί τον ανταλλαγέα διαφημίσεων και ο οποίος επιλέγει τη διαφήμιση με την υψηλότερη προσφορά για να την εμφανίσει στον χρήστη [40].

Η RTB έχει αλλάξει θεμελιωδώς το τοπίο της διαφήμισης στο Διαδίκτυο και οδήγησε στην επίλυση των προβλημάτων που εμφανίζονται από τη συμβατική διαφήμιση για δύο λόγους: 1) η δυνατότητα συναλλαγών ανά εμφάνιση κλιμακώνει τη διαδικασία αγοράς σε μεγάλο αριθμό διαθέσιμων αποθεμάτων διαφημίσεων, 2) τα δεδομένα ακροαματικότητας σε πραγματικό χρόνο ενθαρρύνουν τη στόχευση συμπεριφοράς και κάνουν μια σημαντική στροφή προς την αγορά που επικεντρώνεται στα δεδομένα του χρήστη και όχι σε συμφραζόμενα δεδομένα. Με τη λελογισμένη στόχευση των χρηστών και τον μηχανισμό δημοπράτησης, η RTB βελτίωσε σημαντικά την επένδυση στην καμπάνια και έγινε ένα βασικό παράδειγμα διαφήμισης στο Διαδίκτυο. Αυτό βέβαια, με τη σειρά του, εγείρει ζητήματα προστασίας της ιδιωτικότητας των χρηστών, αφού δημιουργούνται ερήμην τους προφίλ αυτών, συχνά αρκετά λεπτομερές, με απώτερο σκοπό το κέρδος.

3.1 Δίκτυα Διαφημίσεων – online διαφήμιση

Η online διαφήμιση μεταμορφώθηκε σε μια παγκόσμια βιομηχανία πολλών δισεκατομμυρίων δολαρίων, που εκτείνεται από τις παραδοσιακές μεθόδους όπως το μάρκετινγκ ηλεκτρονικού ταχυδρομείου, και το μάρκετινγκ μηχανών αναζήτησης (SEM).

Κάποτε οι διαδικτυακές διαφημίσεις επικεντρώνονταν αποκλειστικά στο ηλεκτρονικό ταχυδρομείο. Η αύξηση της διαθεσιμότητας του διαδικτύου τη δεκαετία του 1990 δημιούργησε μια νέα ευκαιρία για τους διαφημιστές στο διαδίκτυο και άνοιξε το δρόμο για έναν ακόμα πιο αποτελεσματικό τρόπο προσέγγισης του κοινού-στόχου τους – την λεγόμενη online διαφήμιση προβολής (online advertising).

Στις πρώτες ημέρες της ηλεκτρονικής διαφήμισης προβολής (online advertising), η ανταλλαγή μεταξύ των διαφημιζόμενων (advertisers) και του εκδότη (ο κάτοχος του ιστότοπου-publisher) ήταν μια διαδικασία απευθείας πωλήσεων. Οι διαφημιζόμενοι έρχονταν σε επαφή με τον εκδότη και αγόραζαν χώρο διαφημίσεων στον ιστότοπο τους με κόστος βασισμένο στο σύστημα CPM (cost per mille). Αυτό το σύστημα σήμαινε ότι οι διαφημιζόμενοι θα πληρώνουν μια συγκεκριμένη τιμή για κάθε 1.000 εμφανίσεις της διαφήμισης (εικόνα 13).

Εικόνα 13: Η αγορά και πώληση των αποθεμάτων πραγματοποιήθηκε απευθείας και πωλήθηκε με βάση το σύστημα CPM (Πηγή https://clearcode.cc/blog/real-time-bidding-online-display-advertising/)

Οι δυσκολίες εμφανίστηκαν όταν ο αριθμός των ιστότοπων, και επομένως οι εκδότες, άρχισαν να αυξάνονται. Ενώ οι online διαφημίσεις έβγαζαν αρκετά κέρδη για τους διαφημιζόμενους, οι εκδότες σύντομα διαπίστωσαν ότι δεν είχαν πληρωθεί πολλά από τα αναμενόμενα αποτελέσματα και ένιωθαν πως έπεσαν θύματα υπερεκμετάλλευσης. Για να αντιμετωπιστεί αυτό το πρόβλημα, άρχισαν να εμφανίζονται τα διαφημιστικά δίκτυα (advertising networks) [41]. Ένα δίκτυο διαφήμισης ή ένα δίκτυο διαφημίσεων συνδέει επιχειρήσεις που θέλουν να προβάλλουν διαφημίσεις με ιστότοπους που επιθυμούν να τους φιλοξενήσουν. Το κύριο χαρακτηριστικό ενός διαφημιστικού δικτύου είναι η συγκέντρωση διαφημιστικού χώρου και η αντιστοίχιση του με τις ανάγκες του διαφημιζόμενου. Τα δίκτυα διαφημίσεων συνεργάζονται με εκδότες σε όλο τον ιστό, βοηθώντας οποιονδήποτε έχει απούλητα αποθέματα ή διαφημιστικό χώρο και επιθυμεί να αποκομίσει κέρδος από προσφορές διαφημίσεων. Στη συνέχεια, τα δίκτυα διαφημίσεων συγκεντρώνουν αυτόν τον ελεύθερο χώρο, το συσκευάζουν και το πωλούν στους διαφημιζόμενους (εικόνα 14).

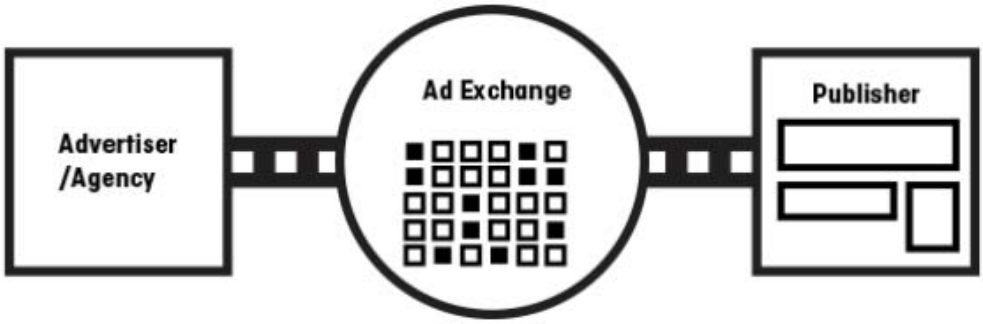
Εικόνα 14: Τα δίκτυα διαφημίσεων και ο τρόπος λειτουργίας τους (Πηγή https://clearcode.cc/blog/real-time-bidding-online-display-advertising/)

Έτσι τα δίκτυα διαφημίσεων έφεραν νέα μοντέλα – συστήματα τιμολόγησης πέραν του CPM. Άρχισε να εμφανίζεται το σύστημα CPC (cost-per-click) όπου οι εμφανίσεις διαφημίσεων που προβάλλονται σε διαφημιστικό χώρο δεν επηρεάζουν το κόστος, αντ 'αυτού, οι διαφημιζόμενοι πληρώνουν για κάθε κλικ που λαμβάνουν από μια διαφημιστική καμπάνια. Επίσης εμφανίστηκε το σύστημα CPA (cost-per-action) όπου ο διαφημιζόμενος πλήρωνε για μια συγκεκριμένη απόκτηση. Για παράδειγμα μια καμπάνια που κοστίζει 250 δολάρια και είχε ως αποτέλεσμα 25 πωλήσεις σημαίνει ότι το CPA ήταν $ 10. Εάν το κέρδος ανά πώληση υπερβαίνει τα $ 10 τότε η εκστρατεία ήταν επιτυχημένη [42].

Παρόλο που η εισαγωγή δικτύων διαφημίσεων πρόσθεσε ευελιξία στις διαδικασίες αγοράς και πώλησης διαφημίσεων, οι εκδότες σύντομα ανακάλυψαν ότι δεν ήταν σε θέση να πουλήσουν όλο το υπόλοιπο απόθεμά τους μέσω ενός διαφημιστικού δικτύου, οπότε άρχισαν να πωλούν το απόθεμά τους μέσω πολλαπλών δικτύων διαφημίσεων. Η αύξηση του αριθμού των εταιρειών ad-network υποδήλωνε επίσης ότι οι εκδότες ξόδεψαν πολύ περισσότερο χρόνο για να βρουν το καλύτερο δίκτυο. Αυτό σήμαινε ότι δυσκολεύονταν να εντοπίσουν τους διαφημιζόμενους με τις καλύτερες επιδόσεις και συχνά έπρεπε να πληρώνουν προμήθειες σε διάφορα δίκτυα.

Για τους διαφημιστές, αυτό δημιούργησε επίσης προκλήσεις. Σύντομα διαπίστωσαν ότι δεν μπορούν να φτάσουν στο κοινό-στόχο τους χρησιμοποιώντας ένα μόνο διαφημιστικό δίκτυο, έτσι άρχισαν να αγοράζουν αποθέματα από πολλά δίκτυα διαφημίσεων. Ωστόσο, η αγορά αποθέματος από πολλά δίκτυα διαφημίσεων οδήγησε τελικά στο ότι «αγόραζαν» συχνά το

ίδιο κοινό περισσότερες από μία φορές και δεν είχαν σαφή εικόνα για την αποτελεσματικότητα των διαφημίσεών τους. Έτσι η επόμενη επαναστατική μέθοδος αγοράς ήταν η εισαγωγή των Ανταλλαγών Διαφημίσεων (ad Exchanges) και των Προσφορών σε Πραγματικό Χρόνο (real time bidding, εικόνα 15).



Εικόνα 15: Η ανταλλαγή διαφημίσεων διεξάγει δημοπρασίες με βάση την εμφάνιση (Πηγή https://clearcode.cc/blog/real-time-bidding-online-display-advertising/)

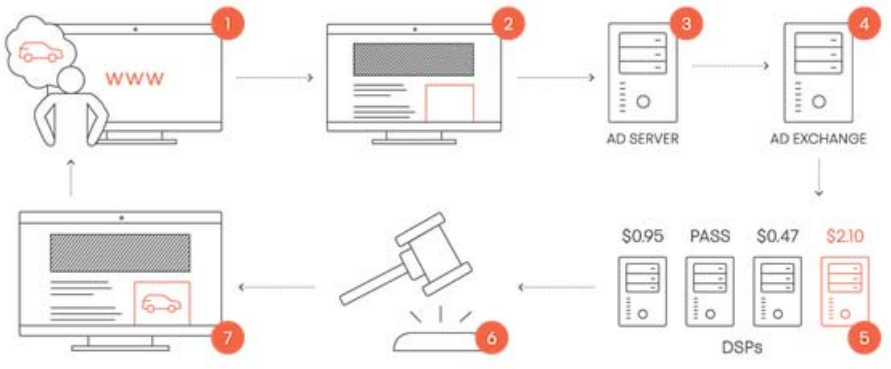
Η εισαγωγή των ανταλλαγών διαφημίσεων (ad Exchanges) έχει αλλάξει εντελώς τον τρόπο με τον οποίο αγοράζονται και πωλούνται οι διαφημίσεις προβολής. Αντί να αγοράζουν εμφανίσεις με βάση το κόστος CPM, οι διαφημιζόμενοι μπορούν πλέον να αγοράζουν διαφημίσεις με βάση την εμφάνιση (impression-by-impression). Υποβάλλοντας προσφορές μόνο για τις εμφανίσεις (ιστότοπους) που σχετίζονται με αυτές, με βάση το κοινό που θέλουν να προσεγγίσουν, οι διαφημιζόμενοι μπορούν να στοχεύσουν πιο άμεσα στο επιθυμητό για αυτούς κοινό. Αυτό παρέχει επίσης στους εκδότες υψηλότερα έσοδα από διαφημίσεις, καθώς τα αποθέματά τους εμφανίζονται στους σωστούς χρήστες που θέλουν να στοχεύσουν οι διαφημιζόμενοι [39]. Με την πάροδο του χρόνου δημιουργήθηκαν αρκετές εταιρείες ad Exchanges (πχ Open X, Admeld, Content web), αλλά δεν αποτελεί έκπληξη το γεγονός ότι οι τρεις μεγάλες εταιρείες (Google, Yahoo και Microsoft) αναλαμβάνουν τον ηγετικό ρόλο στην ανάπτυξη της υποδομής του Ad Exchange. Η μεγάλη οικονομική τους ευρωστία επιτρέπουν στην επένδυση των απαραίτητων πόρων, για την λειτουργία κέρδους, όπως είναι το Real Time Bidding.

3.2 Λειτουργία του Real Time Bidding

Η τυπική λειτουργία της άτυπης δημοπρασίας σε πραγματικό χρόνο μπορεί να περιγραφεί ως εξής:

Στο επίκεντρο της λειτουργίας του RTB βρίσκονται τέσσερεις κατηγορίες παικτών:
1)**The publisher**: Στο πλαίσιο της διαφήμισης προβολής, ο εκδότης είναι ο ιστότοπος που επισκέπτονται οι χρήστες. Ένα παράδειγμα ενός εκδότη μπορεί να είναι κάτι τόσο απλό όσο ένα ταξιδιωτικό blog ή ένας δικτυακός τόπος ειδήσεων, αλλά μπορεί επίσης να καλύπτει εφαρμογές ιστού όπως το Facebook.
2)**Ad Server**: Ένας διακομιστής διαφημίσεων είναι λογισμικό που αποθηκεύει δεδομένα σχετικά με διαφημιστικό περιεχόμενο και παρέχει διαφημίσεις σε ιστότοπους και εφαρμογές. Ανάλογα με το συγκεκριμένο προϊόν ή υπηρεσία, ένας διακομιστής διαφημίσεων μπορεί επίσης να παρακολουθεί προβολές και κλικ διαφημίσεων και διαφημίσεις στόχευσης βάσει προκαθορισμένων κριτηρίων.
3)**Ad exchanges (ADX)**: Οι ADX είναι δυναμικές τεχνολογικές πλατφόρμες που διευκολύνουν τη διαδικασία αγοράς και πώλησης των διαθέσιμων εμφανίσεων μεταξύ των διαφημιζόμενων (αγοραστών) και των εκδοτών (πωλητών) - μοιάζει πολύ με τον τρόπο με τον οποίο τα χρηματιστήρια διαχειρίζονται την αγορά και πώληση μετοχών μεταξύ επενδυτών και εταιρειών. Πάροχοι που προσφέρουν τέτοιες πλατφόρμες είναι η ADX της Google και ADX Yahoo.
4)**The Demand-Side Platform (DSP)**: Οι πλατφόρμες ζήτησης αφορούν παρόχους που αναλαμβάνουν για τους διαφημιζόμενους την αγορά των δημοπρατούμενων διαφημίσεων. Τέτοιοι πάροχοι είναι π.χ. η doubleclick (της Google), η adform, η d3media, η spree7, η appnexus και η Dataxu.

Στην παρακάτω εικόνα (εικόνα 16) βλέπουμε ένα παράδειγμα του RTBκαι πώς αυτό υλοποιείται σε πραγματικό χρόνο.



Εικόνα 16: Η διαδικασία του RTB (Πηγή https://districtm.net/en/blog/detail/real-time-bidding-rtb-101/)

1: Η όλη διαδικασία RTB ξεκινά όταν ένας χρήστης του Διαδικτύου αποκτήσει πρόσβαση σε έναν ιστότοπο ή μια εφαρμογή στο διαδίκτυο. Πληροφορίες, όπως το φύλο, η ηλικία, τα ενδιαφέροντα και η συμπεριφορά στο διαδίκτυο, είναι διαθέσιμες μέσω των τεχνικών παρακολούθησης που είδαμε στο κεφάλαιο 2.

2: Το πρόγραμμα περιήγησης στέλνει αίτημα στον διακομιστή διαφημίσεων (ad server) κατά τη διάρκεια του χρόνου φόρτωσης.

3: Ο διακομιστής διαφημίσεων ελέγχει αν υπάρχουν εκ των προτέρων καταχωρισμένες εμφανίσεις διαφημίσεων, μέσω προηγούμενων συμφωνιών. Εάν όχι, γίνεται μια κλήση στο AD exchange για να προσφέρει το χώρο διαφήμισης σε πιθανούς διαφημιζόμενους.

4: Το AD exchange στη συνέχεια στέλνει το αίτημα διαφήμισης στις πλατφόρμες ζήτησης (DSPs).

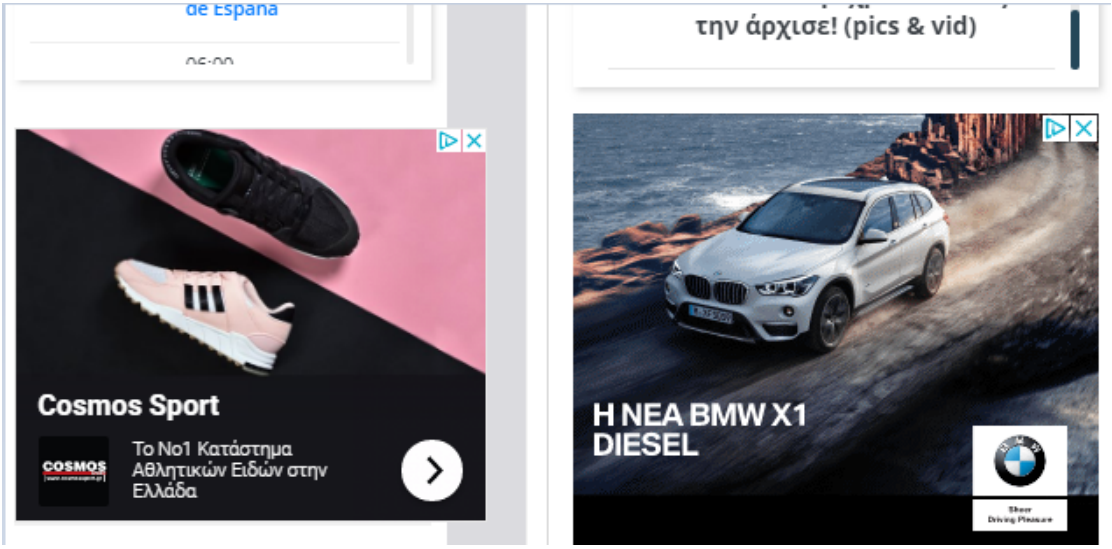
5: Ξεκινάει η δημοπρασία σε πραγματικό χρόνο. Οι DSPs θα εισέλθουν στη δημοπρασία ή όχι ανάλογα με το εάν οι προκαθορισμένες παράμετροι της καμπάνιας που εκπροσωπούν, ταιριάζουν με τις διαθέσιμες πληροφορίες για το χρήστη που έχει πρόσβαση στη σελίδα. Με αυτόν τον τρόπο εξασφαλίζουν ότι η διαφήμισή τους είναι σχετική με τον θεατή.

6: Σε δημοπρασία πρώτης τιμής, ο πλειοδότης θα κερδίσει την εμφάνιση της διαφήμισης και η δημοπρασία θα εκκαθαριστεί στην υψηλότερη τιμή προσφοράς.

7: Η νικήτρια διαφήμιση προβάλλεται στον ιστότοπο.

Η διαδικασία της συγκεκριμένης λειτουργίας περιλαμβάνει την πραγματοποίηση της δημοπρασίας και την εμφάνιση της διαφήμισης και ολοκληρώνεται σε πολύ μικρά διαστήματα της τάξης των 10-100ms και για αυτό ονομάζεται και άτυπη δημοπρασία σε

πραγματικό χρόνο. Στην εικόνα 17 φαίνεται ένα παράδειγμα που πραγματοποιήσαμε για να δούμε πως υποτυπώνεται στην πραγματικότητα το Real Time Bidding. Οι εικόνες έχουν συλλεχθεί από δύο διαφορετικούς υπολογιστές. Στον έναν υπολογιστή κάναμε αναζήτηση για αθλητικά παπούτσια (αριστερά) και στον άλλον υπολογιστή αναζητήσαμε αυτοκίνητα (δεξιά). Στην συνέχεια επισκεφτήκαμε τον ιστότοπο www.gazzetta.gr. Χαρακτηριστικό, πέρα από τις παραπλήσιες διαφημίσεις που λάβαμε, είναι ότι οι διαφημίσεις εμφανίστηκαν σε λιγότερο από 0.2 sec.

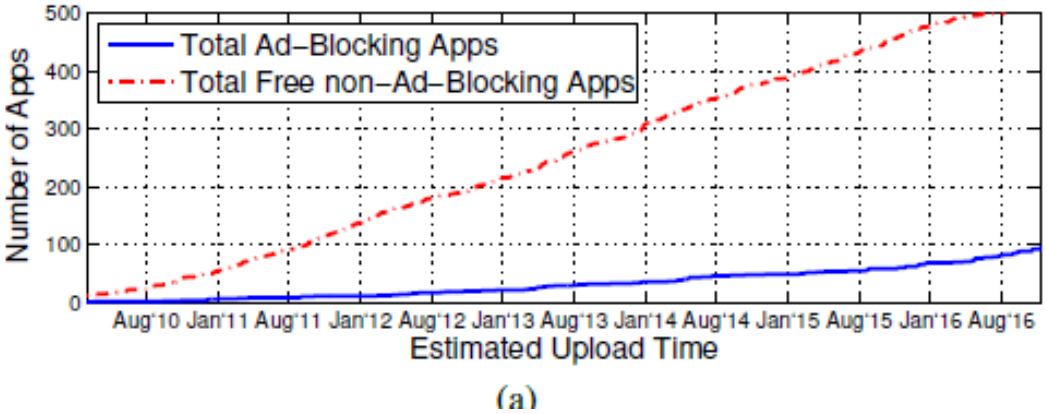


Εικόνα 17: Αποτύπωση διαφημίσεων με την λειτουργία του RTB στην ιστοσελίδα www.gazzetta.gr

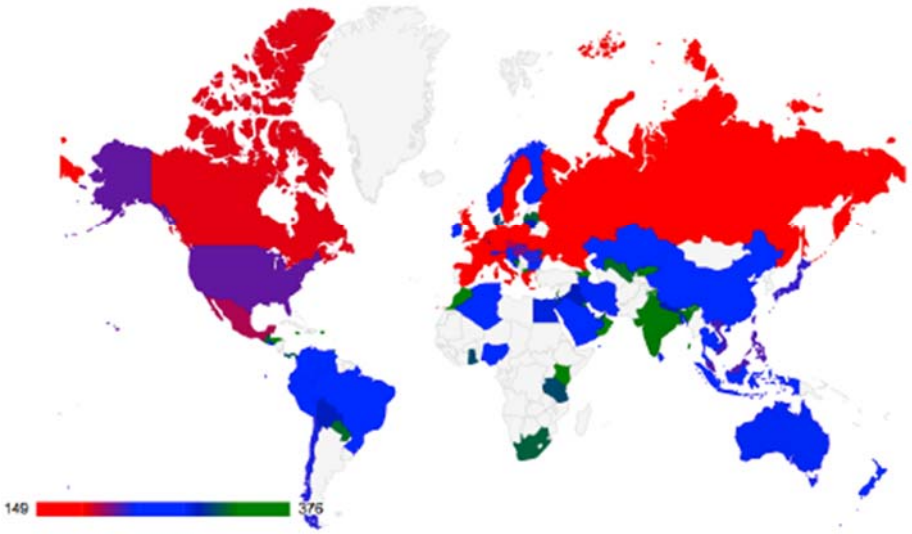
Κεφάλαιο 4
Ανάλυση εφαρμογών αποκλεισμού διαφημίσεων

Η διαδικτυακή διαφήμιση είναι πανταχού παρούσα στη σημερινή ψηφιακή οικονομία. Οι μηχανισμοί παρακολούθησης χρησιμοποιούνται τόσο σε ιστότοπους όσο και σε κινητές εφαρμογές. Οι εφαρμογές κινητών από την άλλη πλευρά, επιτρέπουν αυτήν την στοχευμένη διαφήμιση από τρίτους με σκοπό την δημιουργία εσόδων από υπηρεσίες που θέλουν να προβάλλουν τις διαφημίσεις τους. Επιπροσθέτως, η δημοτικότητα των εφαρμογών για κινητά όλο και αυξάνεται λόγω της πανταχού παρούσας χρήσης των κινητών τηλεφώνων και αυτό έχει αυξήσει κατά πολύ την διαφήμιση μέσω εφαρμογών. Αυτή η σημαντική αύξηση των εφαρμογών έχει δημιουργήσει προβλήματα όμως στην ιδιωτικότητα και την προστασία των προσωπικών δεδομένων των χρηστών. Είναι γνωστό ότι τα προσωπικά δεδομένα χρηστών, συμπεριλαμβανομένων των επαφών, των τοποθεσιών, των SMS και του ιστορικού ιστού, είναι εύκολα προσβάσιμα σε κινητές συσκευές για εφαρμογές που σχετίζονται με τα

δικαιώματα και αποτελούν σημαντικά στοιχεία για διαφημιζόμενους και «ιχνηλάτες τρίτων μερών». Για αυτόν τον λόγο, το λεγόμενο «οικοσύστημα εφαρμογών για κινητές συσκευές» έχει πρόσφατα γίνει μάρτυρας της εμφάνισης μιας νέας κατηγορίας εφαρμογών, γνωστών ως **εφαρμογές αποκλεισμού διαφημίσεων (ad blocking apps).** Η εικόνα 18 δείχνει την ραγδαία αύξηση αυτών των εφαρμογών στο διάστημα 2011 μέχρι 2016. Επιπλέον στην εικόνα 19 μπορούμε να δούμε και την ζήτηση των ad blocking apps σε παγκόσμιο επίπεδο ανά γεωγραφική περιοχή ή χώρα. Το σκούρο κόκκινο χρώμα, αντιπροσωπεύει την μεγάλη ζήτηση για τους Ad-Blockers στην αντίστοιχη γεωγραφική περιοχή ή χώρα.



Εικόνα 18: Αύξηση των εφαρμογών Ad Blocking από το 2010 μέχρι 2016 [43]



Εικόνα 19: Μια επισκόπηση της δημοτικότητας των Ad-Blockers ανά χώρα [43]

Οι εφαρμογές αυτές αποτελούνται από έναν περιηγητή με εγκατεστημένο έναν Ad Blocker στο σύστημα του. Σκοπός των συγκεκριμένων εφαρμογών είναι η απαγόρευση της προβολής των διαφημίσεων στην οθόνη του χρήστη. Παράλληλα όμως οι εφαρμογές αυτές ισχυρίζονται ότι παρέχουν προστασία προσωπικών δεδομένων χρηστών και ασφαλή περιήγηση κρατώντας μακριά τους third party trackers. Πολύ πρόσφατες έρευνες έδειξαν πως οι συγκεκριμένες εφαρμογές εγείρουν αρκετά θέματα κατά της ιδιωτικότητας [43]. Επιπρόσθετα έρευνες που εστίασαν στην μελέτη συγκεκριμένων Ad blockers (Adblock plus, Ghostery δείχνοντας ότι η χρήση ενός ad blocker μπορεί πράγματι να αυξήσει το επίπεδο απορρήτου αλλά όχι στο σημείο έτσι ώστε να αποφευχθεί πλήρως η διαρροή πληροφοριών σχετικά με την περιήγηση του χρήστη προς τους ιχνηλάτες τρίτων. Ακόμη και όταν οι ad blockers είναι ρυθμισμένοι στο επίπεδο μέγιστης προστασίας, διάφοροι trackers είναι δυνατόν να καταγράφουν τα δεδομένα των χρηστών [44]. Στον πίνακα 3 μπορούμε να δούμε πως ορισμένοι third party trackers που ανήκουν στην Google και Facebook, μπορούν να παρακολουθούν τις κινήσεις των χρηστών, ακόμη και όταν οι ad blockers είναι ρυθμισμένοι στο μέγιστο επίπεδο προστασίας.

Third-Party Domain	Legal Entity	TPD Degree		
		None	Ghostery	AdblockPlus
doubleclick.net	Google Inc.	486	0	1
google-analytics.com	Google Inc.	476	4	0
google.com	Google Inc.	383	93	144
facebook.com	Facebook Inc.	318	5	164
gstatic.com	Google Inc.	308	226	235
googlesyndication.com	Google Inc.	204	0	0
google.ch	Google Inc.	189	0	0
fonts.googleapis.com	Google Inc.	185	145	141
adnxs.com	AppNexus Inc.	159	0	0
facebook.net	Facebook Inc.	157	0	140

Πίνακας 3: Third Party Trackers που εξακολουθούν να παρακολουθούν τους χρήστες, κατά την λειτουργία των ad blockers [44]

Στο κεφάλαιο αυτό επιχειρούμε μια διεξοδική ανάλυση πέντε εφαρμογών Ad blocking και συγκεκριμένα των: **Free Ad Blocker browser, CM browser, Ad Blocker browser, Brave browser** και **Dolphin browser**. Η συγκεκριμένη έρευνα παρουσιάζει διαφορές σε σχέση με την [43]. Σε αντίθεση με την προαναφερθείσα εργασία που επικεντρώθηκε περισσότερο στην ανάλυση του πηγαίου κώδικα κάθε εφαρμογής, με στόχο την εύρεση ιών και βιβλιοθηκών

από ιχνηλάτες τρίτων μερών, κύριος στόχος μας στην έρευνά μας είναι να εκτελέσουμε τις εφαρμογές σε εικονικό περιβάλλον και να αναλύσουμε την συμπεριφορά τους κατά την εκτέλεση τους, πραγματοποιώντας έτσι δυναμική ανάλυση αυτών. Με την ανάλυση των εφαρμογών επιχειρούμε να αναδείξουμε τι δεδομένα συλλέγουν οι εφαρμογές και κατά πόσο αυτές αποτελούν απειλές για την ιδιωτικότητα χρήστη. Στην συνέχεια προσπαθήσαμε να διαλευκάνουμε αν αυτά τα δεδομένα μερικά ή όλα, διαρρέονται προς τρίτες εταιρείες εν αγνοία των χρηστών

Αξίζει να αναφερθεί πως η επιλογή των πέντε αυτών εφαρμογών έγινε λαμβάνοντας υπ' όψιν την προτίμησή τους από τους χρήστες στο Google Play Store (downloads, κριτικές, βαθμολογίες).

4.1 Λειτουργικό σύστημα Android

Το Android είναι ένα λειτουργικό σύστημα για κινητές συσκευές με οθόνες αφής όπως smartphones και tablets και είναι βασισμένο στην αρχιτεκτονική του πυρήνα Linux. Ιδρύθηκε το 2007, και μέσα σε αυτήν την δεκαετία έχει καταφέρει να αποτελεί το κυρίαρχο λειτουργικό σύστημα για φορητές συσκευές. Συμφώνα με έρευνα της Strategy Analytics (https://www.strategyanalytics.com/), την δεδομένη χρονική στιγμή, το Android "τρέχει" στις περισσότερες συσκευές με ποσοστό πάνω από 85 % παγκοσμίως. Η ραγδαία αυτή εξέλιξη του Android οφείλεται στο γεγονός ότι πρόκειται για ένα ελεύθερο λειτουργικό ανοιχτού κώδικα (μπορεί οποιοσδήποτε να χρησιμοποιήσει / αλλάξει τον κώδικα για δικό του σκοπό), καθώς επίσης στο ότι υπάρχουν διάφορες εκδόσεις του Android διαθέσιμες ανάλογα με τις προτιμήσεις των χρηστών. Ένα ακόμη στοιχείο που κάνει το λειτουργικό σύστημα Android ξεχωριστό, είναι η τεράστια ποικιλία από εφαρμογές, που βρίσκονται διαθέσιμες στους χρήστες μέσω του Google Play Store, καθώς η πλειοψηφία αυτών είναι δωρεάν.

4.2 Αναγνωριστικά συσκευών

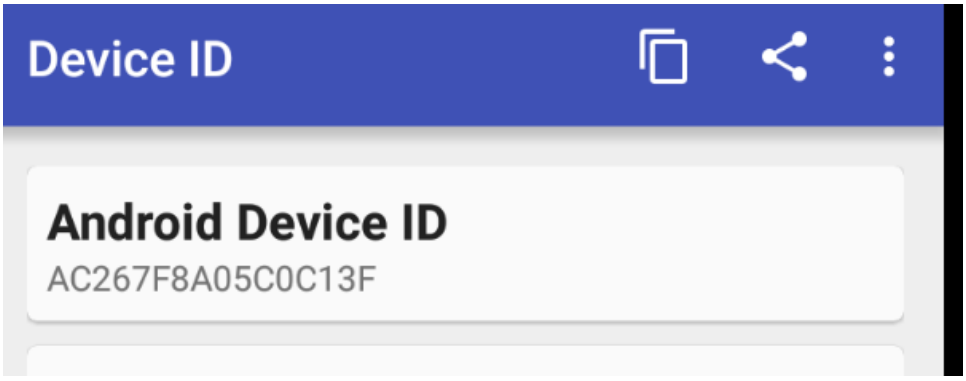
Τα αναγνωριστικά συσκευών (Ids) είναι μοναδικοί αριθμοί οι οποίοι μπορούν να χρησιμοποιηθούν για να αναγνωριστεί μία κινητή συσκευή. Ο παρακάτω πίνακας (πίνακας 4)

αποτυπώνει κάποια αναγνωριστικά μιας συσκευής, που χρησιμοποιεί το λειτουργικό σύστημα Android.

Τα αναγνωριστικά συσκευών διαδραματίζουν βασικό ρόλο τόσο για την καταμέτρηση των κλικ των χρηστών όσο και για την προβολή διαφημίσεων στον χρήστη. Τα αναγνωριστικά βοηθούν τις εταιρείες και τους διαφημιστές να εξακριβώσουν εάν έχουν ήδη στείλει μια διαφήμιση σε συγκεκριμένο χρήστη. Στην εικόνα 20 μπορούμε να δούμε το Android ID μιας συσκευής. Όπως θα δούμε και στην συνέχεια του κεφαλαίου, το Android ID έχει σημαντικό ρόλο στις εφαρμογές Ad Blocking.

Identifier	Description	Attribute
GAID	User-resettable 32-digit alphanumeric identifier	Pseudonymous
Android ID	64-bit number randomly generated when device is set up for the first time [5]	Semi-permanent
IMEI	15-digit decimal identifier representing GSM or LTE device	Permanent
IMSI	15-digit decimal identifier representing mobile subscriber identity	Permanent
MAC address	48-bit number assigned to the device's Wi-Fi network interface	Permanent

Πίνακας 4: Τα αναγνωριστικά κινητού με λειτουργικό σύστημα Android [45]



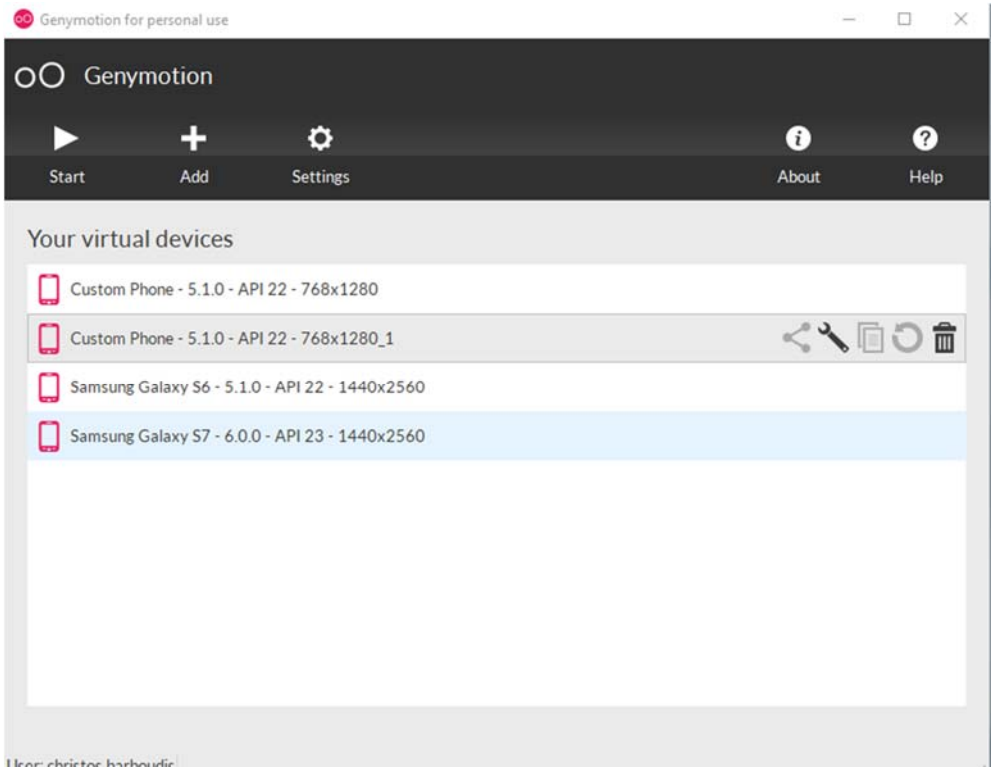
Εικόνα 20: Το Android ID συσκευής κινητού

4.3 Δημιουργία περιβάλλοντος δοκιμών

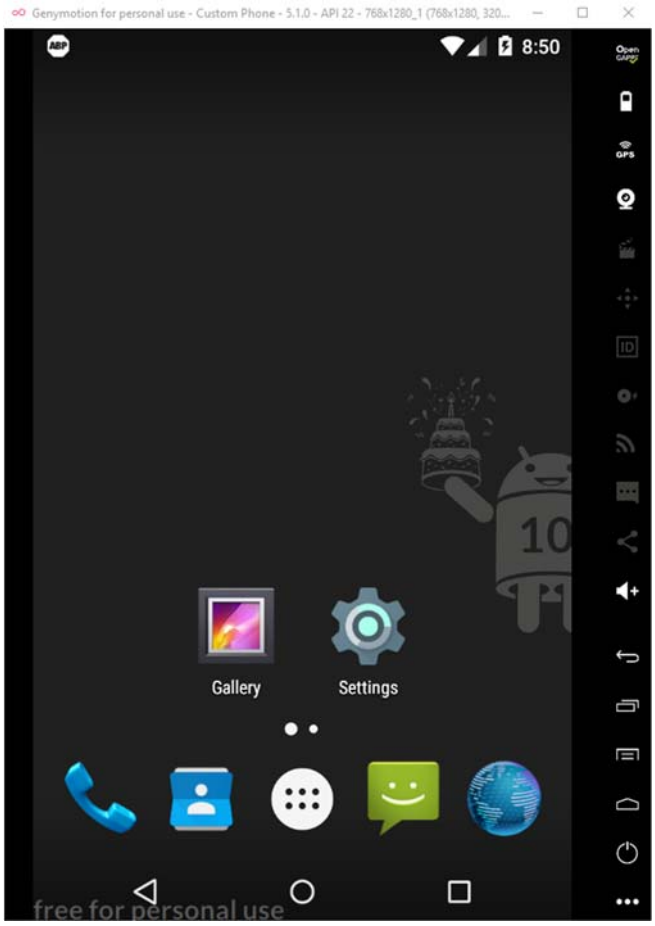
Για την ανάλυση των εφαρμογών αποκλεισμού διαφημίσεων, αναπτύχθηκε ένα εικονικό περιβάλλον εργασίας με την χρήση του Genymotion, πάνω στο οποίο εγκαταστάθηκε λειτουργικό σύστημα Android 5.1. Στην συνέχεια έγινε εγκατάσταση του Xposed Framework στο εικονικό μηχάνημα, όπου μέσω αυτού έγινε προσθήκη του module Inspeckage, ένα εργαλείο που μας βοηθά στην δυναμική ανάλυση των εφαρμογών. Στη συνέχεια, μέσω της εφαρμογής Google Play Store, έγινε εγκατάσταση των εφαρμογών Free Ad Blocker browser, CM browser, Ad Blocker browser, Brave browser, Dolphin - Best Web browser και Lumen Privacy Monitor.

Genymotion

Το Genymotion είναι ένας εξομοιωτής ο οποίος έρχεται με προκαθορισμένες εκδόσεις Android, και αποτελεί ένα ιδανικό εργαλείο για έλεγχο εφαρμογών κινητού στον υπολογιστή του χρήστη. Με την βοήθεια του Genymotion μπορούμε εύκολα να εγκαταστήσουμε ένα Image μιας έκδοσης Android σε περιβάλλον Windows, Linux, IOS. Στο παράδειγμα μας εγκαταστήσαμε ένα Custom Phone με έκδοση Android 5.1 και API 22 και δώσαμε δικαιώματα διαχειριστή Root.



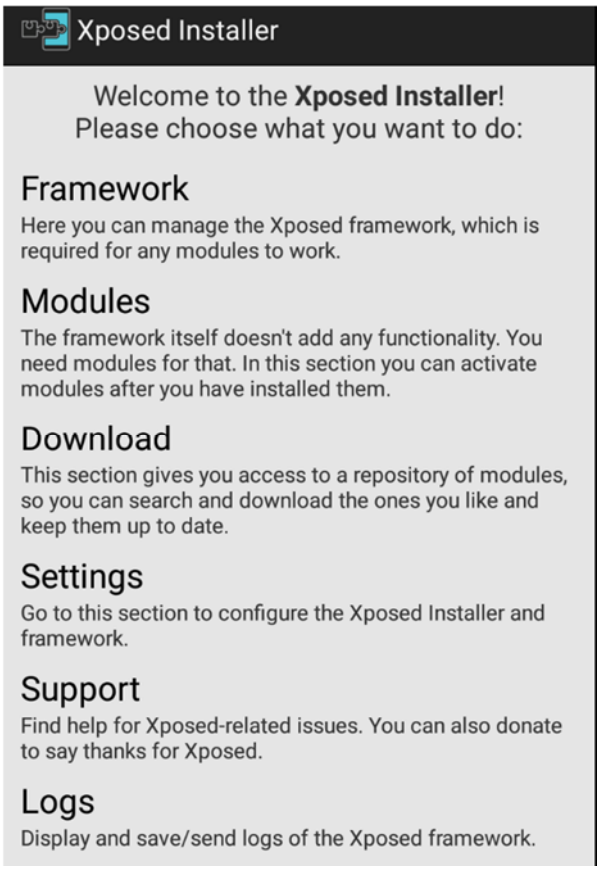
Εικόνα 21: Αρχικό περιβάλλον του Genymotion



Εικόνα 22: Η αρχική οθόνη android 5.1 μέσα από το Image του Genymotion

Xposed Framework

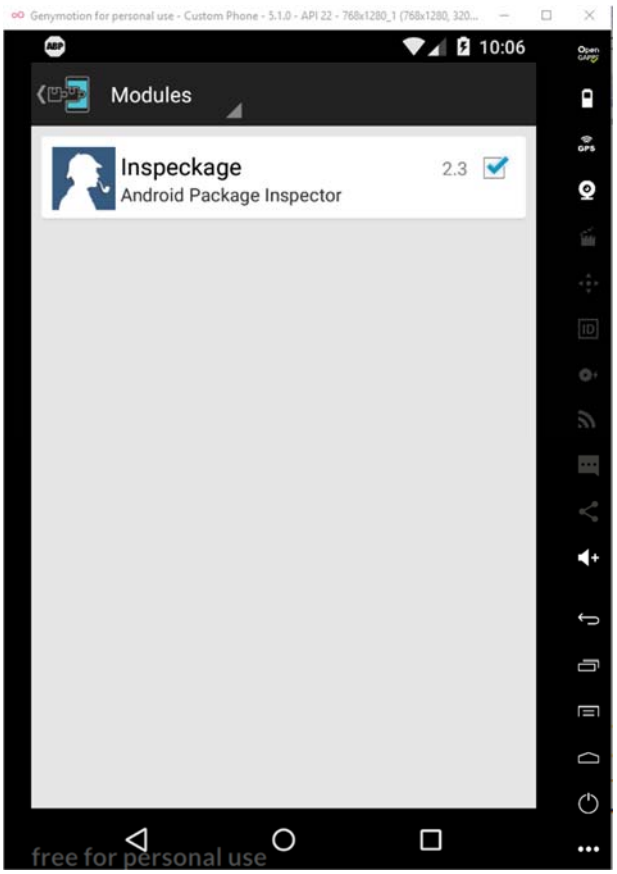
Το Xposed είναι μια πλατφόρμα που επιτρέπει την εγκατάσταση προγραμμάτων στη συσκευή Android που μπορούν να προσαρμόσουν την εμφάνιση και τη λειτουργικότητά του κινητού τηλεφώνου. Το Xposed Framework διαθέτει αρκετά modules που δίνουν την δυνατότητα στους χρήστες να αλλάξουν αρκετά την λειτουργία του τηλεφώνου τους με ασφάλεια. Μέσω του Xposed Framework εγκαταστήσαμε το module Inspeckage που θα μας βοηθήσει στην ανάλυση των εφαρμογών αποκλεισμού διαφημίσεων.



Εικόνα 23: Αρχικό περιβάλλον του Xposed Framework

Inspeckage

Το Inspeckage είναι ένα εργαλείο που παρέχει δυναμική ανάλυση σε εφαρμογές Android. Εφαρμόζοντας σαν module μέσα από το Genymotion, το Inspeckage θα μας βοηθήσει να καταλάβουμε τι ακριβώς κάνει μια εφαρμογή Android κατά το χρόνο εκτέλεσης. Το Inspeckage συνοδεύεται με ένα πολύ φιλικό προς τον χρήστη GUI Interface.



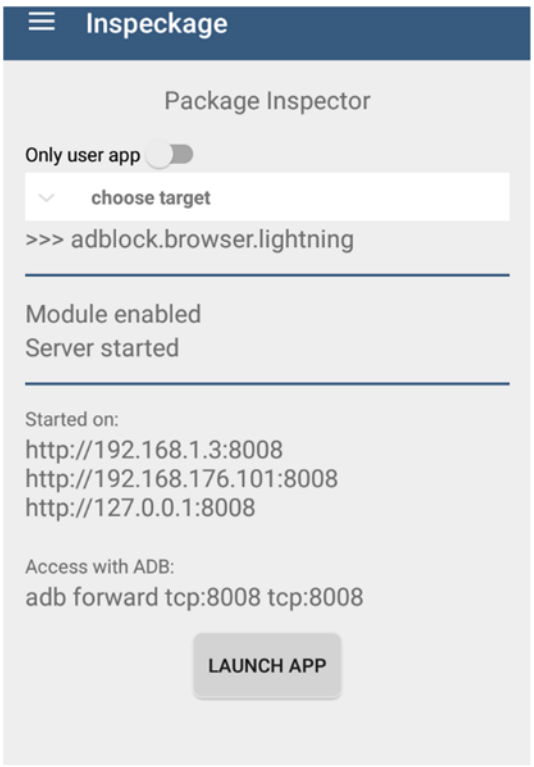
Εικόνα 24: Το module Inspeckage μέσα από το περιβάλλον του Xposed Framework

Για να ενεργοποιηθεί το interface του Inspeckage, θα πρέπει ο χρήστης να πληκτρολογήσει την εντολή Adb forward tcp: 8008 tcp: 8008 στο τερματικό του υπολογιστή (εικόνα 25). Το μόνο που έχει να κάνει μετά ο χρήστης είναι να πληκτρολογήσει σε μία κενή σελίδα σε κάποιον περιηγητή την IP διεύθυνση που υποδηλώνει η εφαρμογή Inspeckagesε συνδυασμό με την πόρτα 8008. Στην εικόνα 26 μπορούμε να δούμε το αρχικό περιβάλλον του Inspeckage, και σε ποιες IP διευθύνσεις βρίσκουμε το interface. Με την επιλογή choose target επιλέγουμε την εφαρμογή που θέλουμε για ανάλυση.

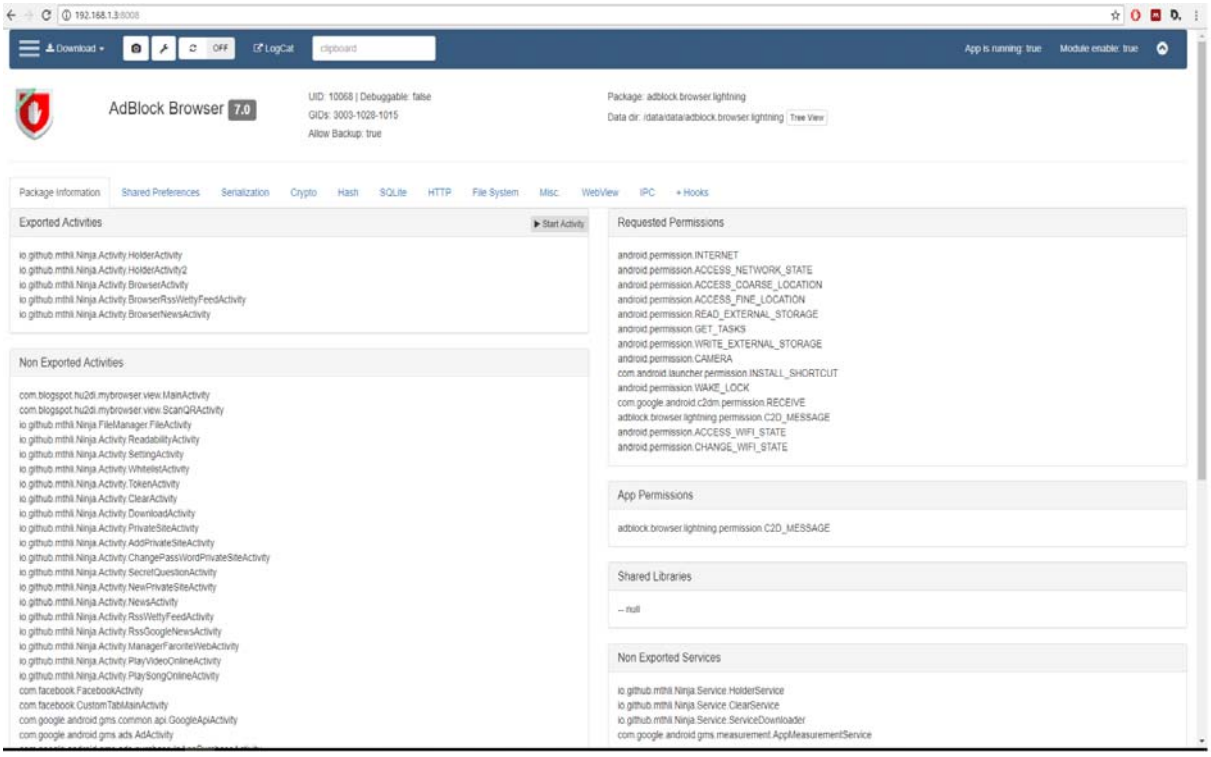
```
C:\Program Files\Genymobile\Genymotion\tools>Adb forward tcp:8008 tcp:8008

C:\Program Files\Genymobile\Genymotion\tools>
```

Εικόνα 25: Η εντολή που χρησιμοποιείται για να ανοίξει το γραφικό περιβάλλον της εφαρμογής Inspeckage



Εικόνα 26: Αρχικό περιβάλλον Inspeckage



Εικόνα27: Γραφικό περιβάλλον Inspeckage

Lumen Privacy Monitor

Η εφαρμογή Lumen Privacy Monitor είναι μια εφαρμογή Android που αναλύει την κυκλοφορία δικτύου των κινητών τηλεφώνων και βοηθά τους χρήστες να αναγνωρίσουν τις διαρροές προσωπικών δεδομένων που προκαλούν οι εφαρμογές, καθώς επίσης και ποιοι οργανισμοί συλλέγουν αυτές τις πληροφορίες. Το Lumen εκτελείται τοπικά στη συσκευή, και παρακολουθεί όλη την κυκλοφορία του δικτύου, συμπεριλαμβανομένων των κρυπτογραφημένων ροών, καθώς εισάγεται ως ενδιάμεσο λογισμικό μεταξύ των εφαρμογών και κινητού. Η εφαρμογή έχει αναπτυχθεί υπό την καθοδήγηση ανεξάρτητων ακαδημαϊκών ερευνητικών κέντρων, όπως το ICSI--UC και MDEA Networks, σε συνεργασία με το Stony Brook University. Το Lumen Privacy Monitor έχει αποδειχθεί πως είναι μια αρκετά αξιόπιστη εφαρμογή καθώς έχει ήδη χρησιμοποιηθεί σε αρκετές περιπτώσεις ερευνών [46]. Αρκετές έρευνες εξ αυτών έχουν δείξει πως σχεδόν το 70% των εφαρμογών Android διαρρέουν προσωπικά δεδομένα σε υπηρεσίες τρίτων, όπως οι υπηρεσίες ανάλυσης και δίκτυα διαφημίσεων. Η εφαρμογή είναι διαθέσιμη στην ιστοσελίδα https://haystack.mobi/ καθώς επίσης και μέσω της εφαρμογής Play Store της Google.

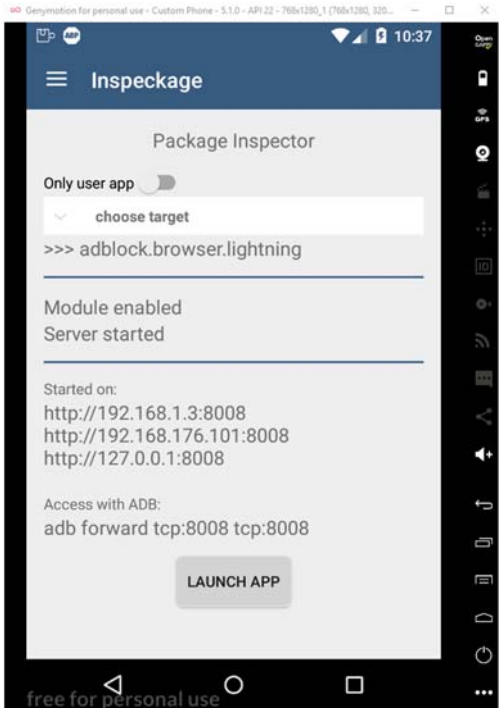
Εικόνα28:Η εφαρμογή Lumen Privacy Monitor διαθέσιμη μέσω του Google Play Store

4.4 Free Ad Blocker browser

Το Free Ad Blocker browser είναι ένα πρόγραμμα περιήγησης στο διαδίκτυο για λειτουργικό Android. Το πρόγραμμα περιήγησης διαθέτει ενσωματωμένη λειτουργία adblock που αποκλείει αυτόματα όλες τις διαφημίσεις και τους ιχνηλάτες ακόμη και προτού ληφθούν στη συσκευή του χρήστη. Σύμφωνα με τον επίσημο ιστότοπο https://FreeAdBlocker browser.com/, το Free Ad Blocker σαν πρόγραμμα περιήγησης προστατεύει την ιδιωτική ζωή των χρηστών, και έτσι κανείς δεν μπορεί να παρακολουθεί τις δραστηριότητές στο διαδίκτυο. Τα κύρια χαρακτηριστικά που προσφέρει το Free Ad Blocker browser είναι: Αποκλεισμός διαφημίσεων, γρηγορότερη περιήγηση, προστασία προσωπικών δεδομένων και ασφαλή περιήγηση. Την εφαρμογή την βρίσκουμε διαθέσιμη μέσω του Google Play Store.

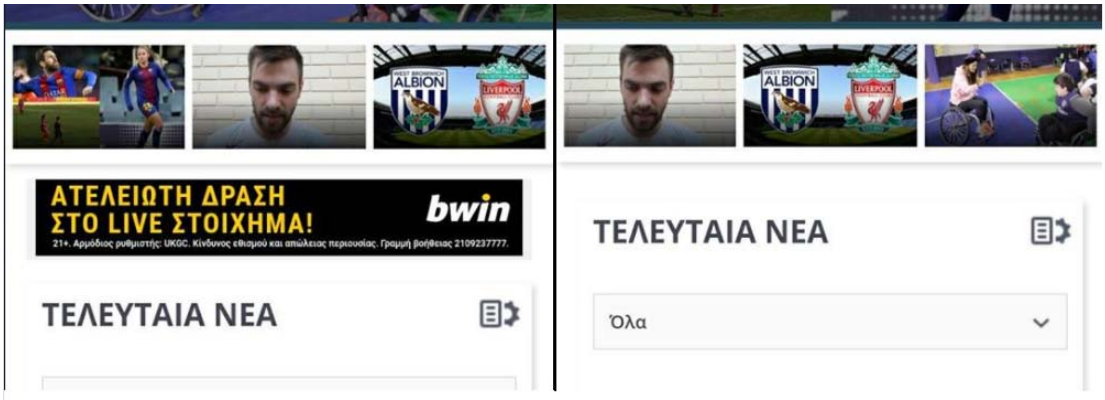
Εικόνα29: Η εφαρμογή Free Ad Blocker browser διαθέσιμη μέσω του Google Play Store

Με την χρήση του Inspeckage θα εκτελέσουμε ανάλυση στην εφαρμογή με σκοπό να δούμε τι δεδομένα επεξεργάζεται. Στην επόμενη εικόνα φαίνεται η παραμετροποίηση που πραγματοποιείται για να αναλυθεί η εφαρμογή μέσω του Inspeckage.



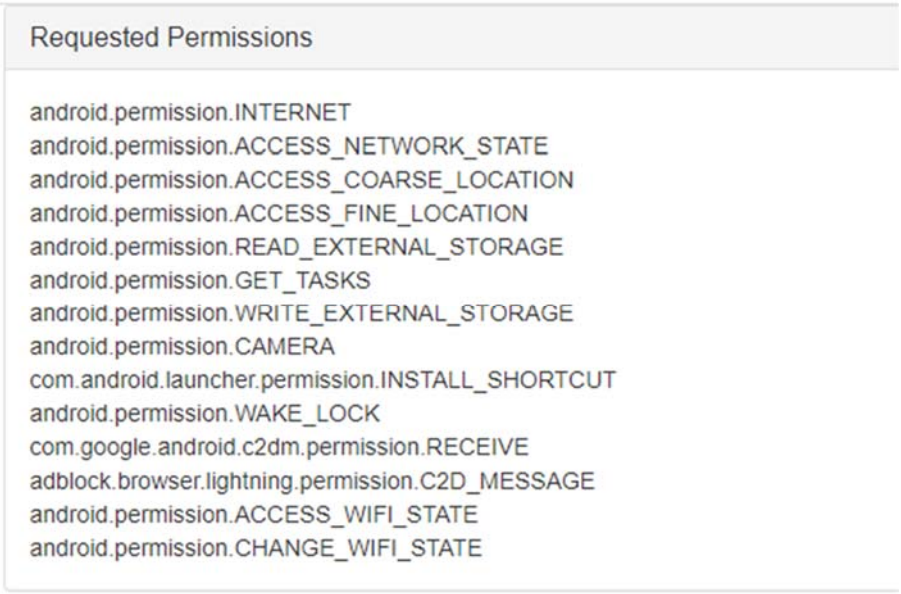
Εικόνα 30: Το κεντρικό μενού του Inspeckage και η επιλογή ανάλυσης της εφαρμογής Free Ad Blocker browser

Με την επιλογή Launch App η εφαρμογή αρχίζει να καταγράφει την κίνηση και την συμπεριφορά του Free Ad Blocker browser στο κινητό. Αμέσως μετά ξεκινήσαμε την περιήγηση σε διάφορους ιστότοπους παγκοσμίως. Κατά την περιήγηση μέσω του Free Ad Blocker browser παρατηρήσαμε ότι δεν εμφανίζονται banner με διαφημίσεις στους ιστότοπους που επιλέξαμε. Στην εικόνα 31 μπορούμε να δούμε την διαφορά ανάμεσα στο Ad Blocker browser και σε έναν άλλον περιηγητή που δεν αποκλείει τις διαφημίσεις.



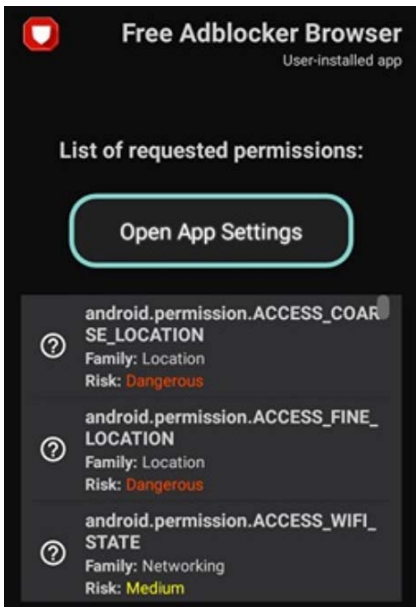
Εικόνα 31: Ο ιστότοπος http://www.gazzetta.gr/ από περιηγητή χωρίς αποκλεισμό (αριστερά) και από το Free Ad Blocker browser (δεξιά).

Από αυτό μπορούμε να συμπεράνουμε ότι το Free Ad Blocker browser πετυχαίνει τον αποκλεισμό διαφημίσεων. Αυτό όμως που έχει ιδιαίτερο ενδιαφέρον έρχεται μέσα από τις εφαρμογές Inspeckage και Lumen,οι οποίες φανερώνουν ότι ο Free Ad Blocker browser ζητά πρόσβαση σε προσωπικά δεδομένα του χρήστη. Μέσω του Inspeckage βλέπουμε σε ποιες κατηγορίες του κινητού έχει δικαιώματα η εφαρμογή Free Ad Blocker browser στο κινητό τηλέφωνο του χρήστη (εικόνα 32).

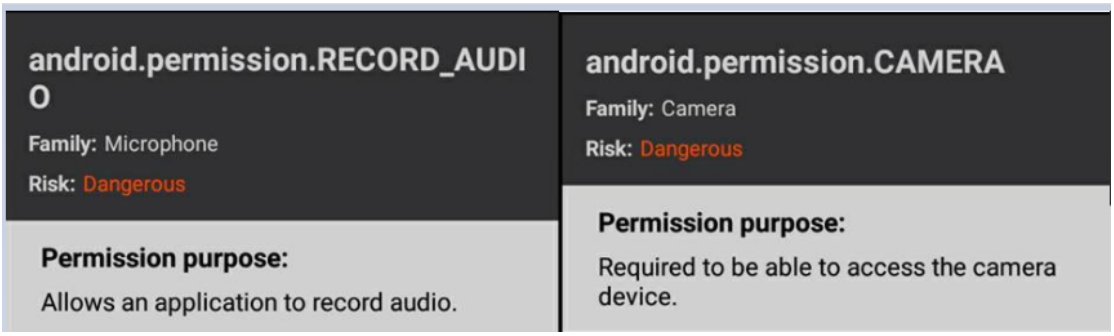


Εικόνα 32: Άδειες σε πληροφορίες που αποκτά η εφαρμογή Free Ad Blocker browser

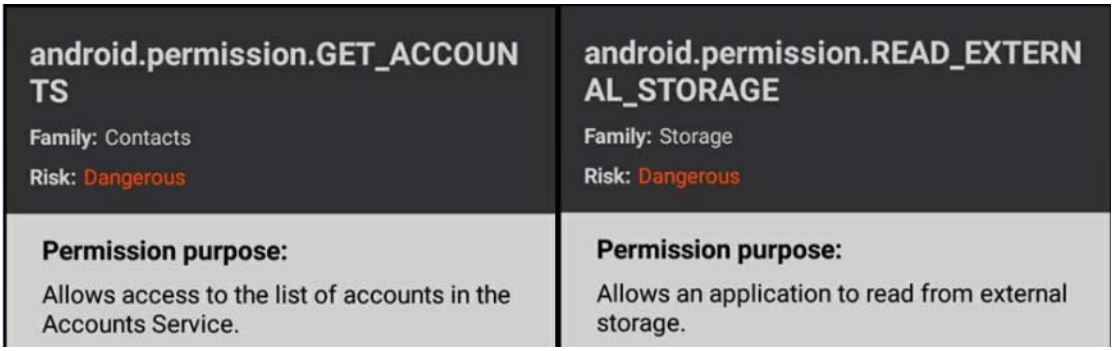
Παρατηρούμε ότι το Free Ad Blocker browserαποκτά πρόσβαση στην ακριβή τοποθεσία του χρήστη, σε πληροφορίες για την κατάσταση του δικτύου, και σε διάφορες πληροφορίες που αφορούν το κινητό τηλέφωνο. Τρέχοντας την εφαρμογή Lumen Privacy Monitor μερικές από αυτές τις προσβάσεις, χαρακτηρίζονται από υψηλού βαθμού επικινδυνότητας (εικόνες 33, 34, 35). Παρατηρούμε επίσης ότι το Free Ad Blocker browser λαμβάνει πρόσβαση και στην κάμερα, στους λογαριασμούς, στην αποθήκευση και στην εγγραφή ήχων του χρήστη.



Εικόνα 33: Οι προσβάσεις στην τοποθεσία του χρήστη από την εφαρμογήFree Ad Blocker browser



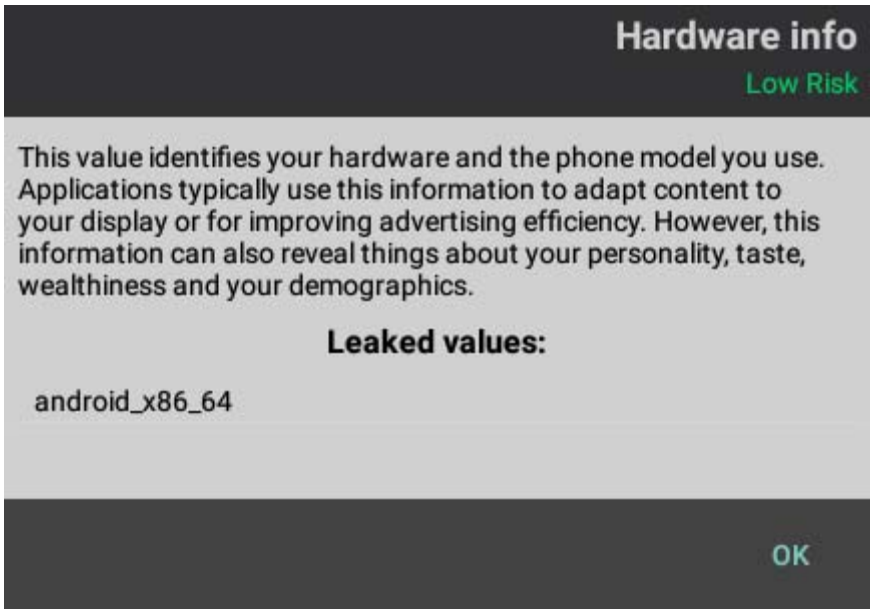
Εικόνα 34: Η πρόσβαση του Free Ad Blocker browser σε πληροφορίες χωρίς την συγκατάθεση του χρήστη σύμφωνα με την εφαρμογή Lumen



Εικόνα 35: Η πρόσβαση τουFree Ad Blocker browser σε πληροφορίες χωρίς την συγκατάθεση του χρήστη σύμφωνα με την εφαρμογή Lumen

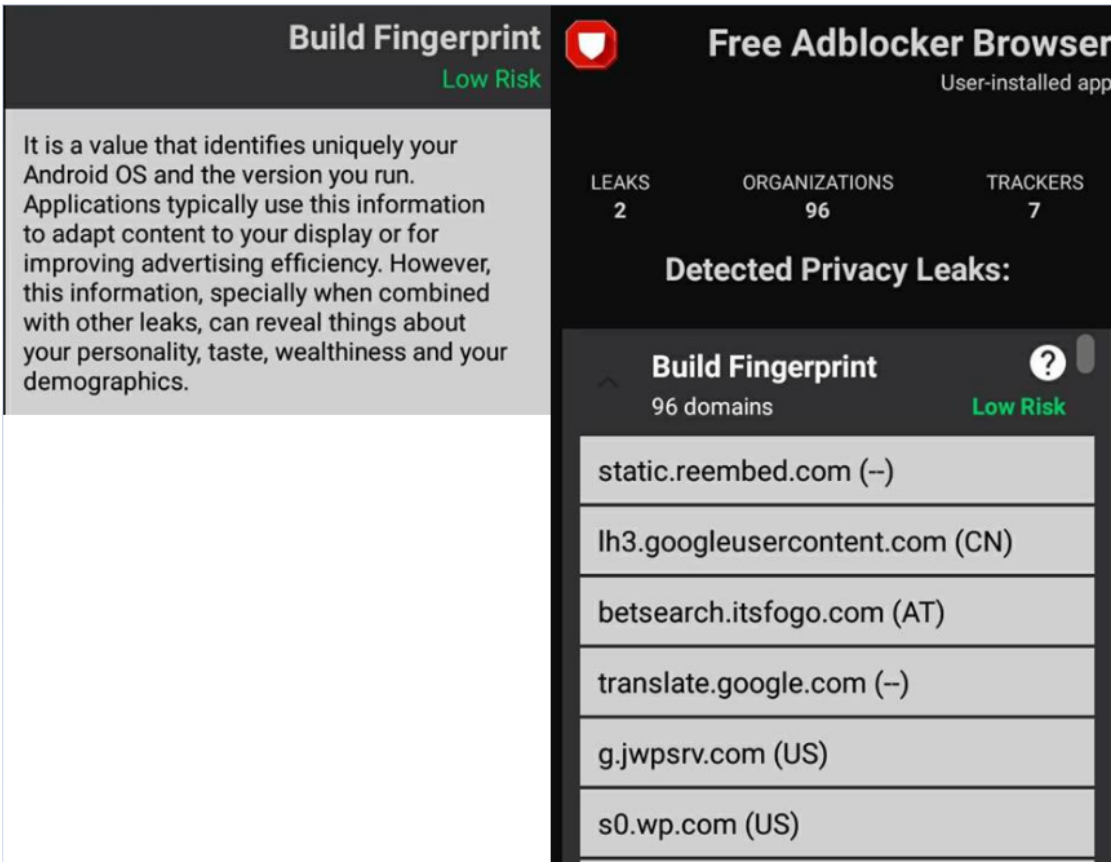
Επιπλέον ένα ακόμη στοιχείο σύμφωνα με την εφαρμογή Lumen, μέσω του Free Ad Blocker browser είναι πιθανόν να γίνουν διαρροές, που αφορούν πληροφορίες της κινητής συσκευής προς άλλες υπηρεσίες. Στις εικόνες 36,37 μπορούμε να δούμε τις πληροφορίες αυτές, και σε ποίες υπηρεσίες διαρρέονται. Παρατηρούμε ότι τα δεδομένα συλλέγονται για λογαριασμό της Google και της Adjust, μιας εταιρείας που συλλέγει δεδομένα για λογαριασμό διαφημιστών (ανίχνευση συσκευών, διευθύνσεις URL, παρακολούθηση τοποθεσιών).

Εικόνα 36: Πληροφορίες του λογισμικού που διαρρέονται σε τρίτες υπηρεσίες



Εικόνα 37: Ο σκοπός απόκτησης πληροφοριών Hardware

Ένα ακόμη στοιχείο που διαρρέεται μέσω του Free Ad Blocker browser αφορά το αποτύπωμα την συσκευής(Build Fingerprint). Το αποτύπωμα αυτό είναι μια τιμή που προσδιορίζει μοναδικά το λειτουργικό σύστημα Android και την έκδοση που εκτελείται στο κινητό τηλέφωνο. Σύμφωνα με την εφαρμογή Lumen, αλλά και σύμφωνα με όσα είδαμε στο κεφάλαιο 2, οι εταιρείες χρησιμοποιούν αυτές τις πληροφορίες για την προσαρμογή του περιεχομένου στην οθόνη ή για τη βελτίωση της αποτελεσματικότητας της διαφήμισης (εικόνα 38). Στην καρτέλα Fingerprint του Inspeckage μπορούμε να δούμε το build fingerprint της συσκευής, μια συμβολοσειρά που προσδιορίζει με μοναδικό τρόπο την κατασκευή του κινητού. Το αποτύπωμα φαίνεται στην εικόνα 39.



Εικόνα 38: Διαρροή προσωπικών δεδομένων και πληροφοριών κινητού όταν χρησιμοποιείται η εφαρμογή Free Ad Blocker browser

☐	BUILD	HOST	39ef4e4d060f
☐	BUILD	ID	LMY47D
☐	BUILD	MANUFACTURER	unknown
☐	BUILD	MODEL	Custom Phone - 5.1.0 - API 22 - 768x1280_1
☐	BUILD	PRODUCT	vbox86p

Εικόνα 39: Απεικόνιση αποτυπώματος συσκευής Android μέσω του Inspeckage

Τα συμπεράσματα λοιπόν είναι τα εξής :

1)Πράγματι ο Free Ad Blocker browser πραγματοποιεί αποκλεισμό διαφημίσεων, παρόλα αυτά λαμβάνει πρόσβαση σε πολύ προσωπικά δεδομένα χρήστη. Φαίνεται ότι η εφαρμογή κρατάει δεδομένα χρήστη που σύμφωνα με το Lumen Privacy Monitor χαρακτηρίζονται υψηλού κινδύνου, και δεν μπορεί να αποκλειστεί το ενδεχόμενο εκμετάλλευσης των πληροφοριών αυτών από τρίτους.

2)Η εφαρμογή δεν παράγει πλήρη ιδιωτικότητα, μιας και δεδομένα συσκευής διαμοιράζονται προς τρίτες υπηρεσίες είτε για σκοπούς ανάλυσης είτε για σκοπούς παρακολούθησης χρηστών. Τα δεδομένα μπορεί να μην είναι προσωπικά για τον χρήστη και να χαρακτηρίζονται από την εφαρμογή Lumen Privacy Monitor ως κίνδυνος «χαμηλού» επιπέδου, είναι όμως ικανά να σχηματίσουν ένα διαδικτυακό προφίλ χρήστη.

4.5 CM BROWSER

Το CM browser είναι μία εφαρμογή από την εταιρεία Cheetah Mobile και έρχεται ενσωματωμένη με έναν adblocker με στόχο την αποφυγή διαφημίσεων. Κύριο χαρακτηριστικό της εφαρμογής είναι ότι απαιτεί πολύ μικρό χώρο (2 MB), και ότι προσφέρει ασφαλή και απόρρητη περιήγηση χωρίς να αφήνει διαδικτυακά ίχνη κατά την περιήγηση, κρατώντας μακριά της εταιρείες tracking. Η εφαρμογή έχει μέχρι στιγμής πάνω από 50 εκατομμύρια λήψεις και είναι διαθέσιμη μέσω του Google Play Store.

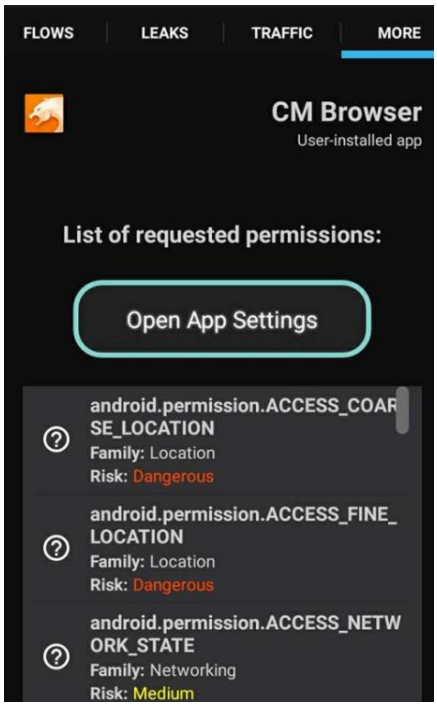
Εικόνα40: Η εφαρμογή CM browser διαθέσιμη μέσω του Google Play Store

Στην εικόνα 41 παρουσιάζονται οι προσβάσεις του CM browser μέσω του εργαλείου Inspeckage. Εκτελώντας την εφαρμογή Lumen παρατηρούμε ότι η εφαρμογή ζητά δικαιώματα σε προσωπικές πληροφορίες, ορισμένες με μεγάλου βαθμού επικινδυνότητας (εικόνα 42).

Requested Permissions

android.permission.ACCESS_COARSE_LOCATION
android.permission.ACCESS_FINE_LOCATION
android.permission.ACCESS_NETWORK_STATE
android.permission.ACCESS_WIFI_STATE
com.android.launcher.permission.INSTALL_SHORTCUT
com.android.launcher.permission.UNINSTALL_SHORTCUT
com.android.launcher.permission.READ_SETTINGS
android.permission.INTERNET
android.permission.WAKE_LOCK
android.permission.WRITE_EXTERNAL_STORAGE
android.permission.VIBRATE
android.permission.MOUNT_UNMOUNT_FILESYSTEMS
android.permission.WRITE_SETTINGS
com.android.browser.permission.READ_HISTORY_BOOKMARKS
android.permission.SYSTEM_ALERT_WINDOW
android.permission.GET_ACCOUNTS
com.google.android.c2dm.permission.RECEIVE
com.ksmobile.cb.permission.C2D_MESSAGE
android.permission.READ_EXTERNAL_STORAGE
com.google.android.finsky.permission.BIND_GET_INSTALL_REFERRER_SERVICE

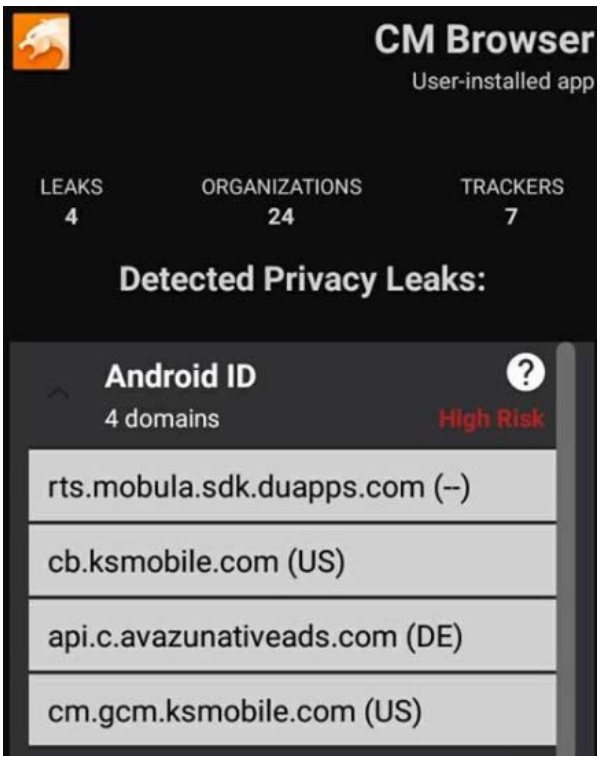
Εικόνα 41: Δικαιώματα που αποκτά η εφαρμογή CM browser



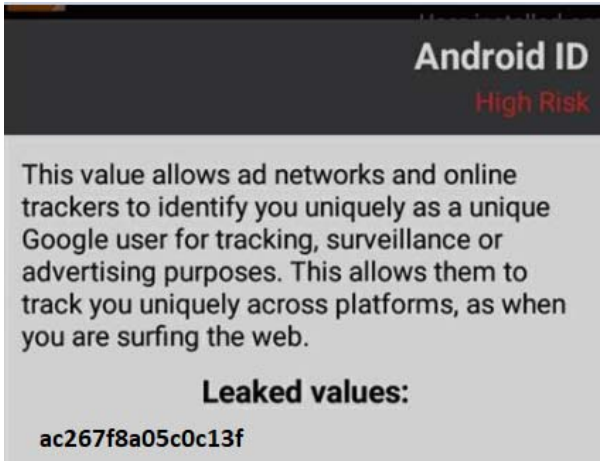
Εικόνα 42:Οι προσβάσεις στην τοποθεσία του χρήστη από την εφαρμογή CM browser, σύμφωνα με την εφαρμογή Lumen χαρακτηρίζοντας τις προσβάσεις υψηλού κινδύνου

Στην συνέχεια, μέσω της εφαρμογής Lumen, προσπαθήσαμε να παρακολουθήσουμε την εφαρμογή CM browser για ενδεχόμενες διαρροές δεδομένων σε τρίτους. Η ανάλυση μας έδειξε ότι το αναγνωριστικό Android ID διαβιβάζεται σε domain τρίτων, και συγκεκριμένα στα **rts.mobula.sdk.duapps.com**, **cd.ksmobile.com**, **api.c.avazunativeads.com,** και χαρακτηρίζεται από την εφαρμογή υψηλού βαθμού επικινδυνότητας (εικόνα 43)**.** Αναζητώντας τα domain στο διαδίκτυο παρατηρήσαμε ότι αφορούν servers της εταιρείας Baidu, εταιρεία που παρέχει διαδικτυακές λύσεις στου διαφημιζόμενους, καθώς επίσης και για servers της ίδιας της εταιρείας Cheetah Mobile.

Το Android ID, ένα μοναδικό αναγνωριστικό της συσκευής επιτρέπει στα δίκτυα διαφημίσεων και στους ιχνηλάτες να εντοπίζουν αποκλειστικά ως μοναδικό χρήστη για σκοπούς παρακολούθησης ή διαφήμισης. Αυτό τους επιτρέπει να παρακολουθήσουν μοναδικά τους χρήστες σε όλες τις πλατφόρμες, όπως όταν πραγματοποιούν πλοήγηση στο διαδίκτυο (εικόνα 44).

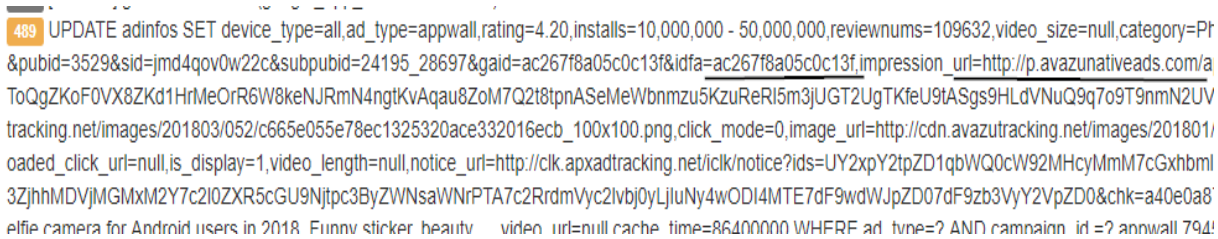


Εικόνα 43: Διαρροή Android ID προς τρίτους, σύμφωνα με την εφαρμογή Lumen Privacy Monitor



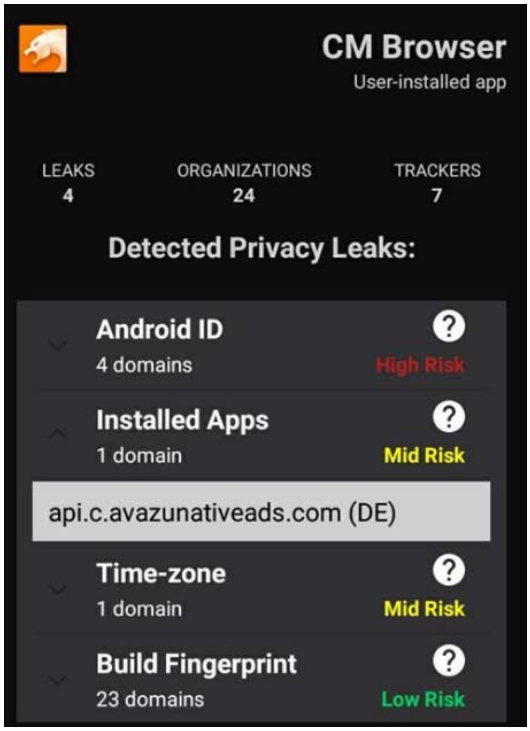
Εικόνα 44: Το Android ID της συσκευής

Η διαρροή του Android ID επιβεβαιώνεται και μέσω του εργαλείου Inspeckage.Στην παρακάτω εικόνα (εικόνα 45) μέσω της καρτέλας SQL Lite βλέπουμε την διαρροή του Android ID προς το domain **avazunativeads.com,** domain που κατέγραψε και το Lumen Privacy Monitor.



Εικόνα 45: Διαρροή Android ID προς το Domain avazunativeads.com

Υπάρχει επίσης μία παράβαση δεδομένων με μέτριο βαθμό επικινδυνότητας, και αφορά τις υπόλοιπες εγκατεστημένες εφαρμογές του χρήστη, και φαίνεται ότι ο CM browser ενημερώνεται για ποιες εφαρμογές είναι εγκατεστημένες στην συσκευή του χρήστη (εικόνα 46) με σκοπό την αναγνώριση της προσωπικότητας («προφίλ») του χρήστη και τις προτιμήσεις του. Αυτές οι πληροφορίες μπορούν να βοηθήσουν τους ιχνηλάτες και την εταιρείες διαφημίσεων για μελέτες μάρκετινγκ και δημογραφικούς σκοπούς. Τέλος αναφέρονται ακόμη δυο παραβιάσεις (χαμηλού βαθμού) που αφορούν το timezone που βρίσκεται το κινητό του χρήστη και το build Fingerprint, που αναλύθηκε και για το Free Ad Blocker browser.



Εικόνα 46: Διαρροή των Installed Apps, Time –zone, Build Fingerprint προς τρίτους, συμφώνα με την εφαρμογή Lumen Privacy Monitor.

Τα συμπεράσματα λοιπόν είναι τα εξής:

1) Το CM browser, κατά την εγκατάσταση του αποκτά πρόσβαση σε πολλές πληροφορίες του κινητού του χρήστη. Η εφαρμογή επομένως κρατάει ένα μεγάλο όγκο από πληροφορίες που ενδέχεται να διατεθούν προς χρήση από τρίτες εταιρείες.

2) Η διαρροή του Android ID, ενός μοναδικού αναγνωριστικού της συσκευής, προς τρίτους αποτελεί ένα σημαντικό παράπτωμα κατά της ιδιωτικότητας του χρήστη. Με το Android ID οι τρίτες εταιρείες μπορούν να παρακολουθούν τον χρήστη μοναδικά για διαφημιστικούς σκοπούς.

3)Η ενημέρωση του CM browser για τις υπόλοιπες εγκατεστημένες εφαρμογές, μπορεί να σχηματίσει ένα διαδικτυακό προφίλ για τον χρήστη και να βγουν συμπεράσματα για την προσωπικότητα του και τις προτιμήσεις του και αυτές η πληροφορίες είναι δυνατόν να μοιραστούν προς τρίτους.

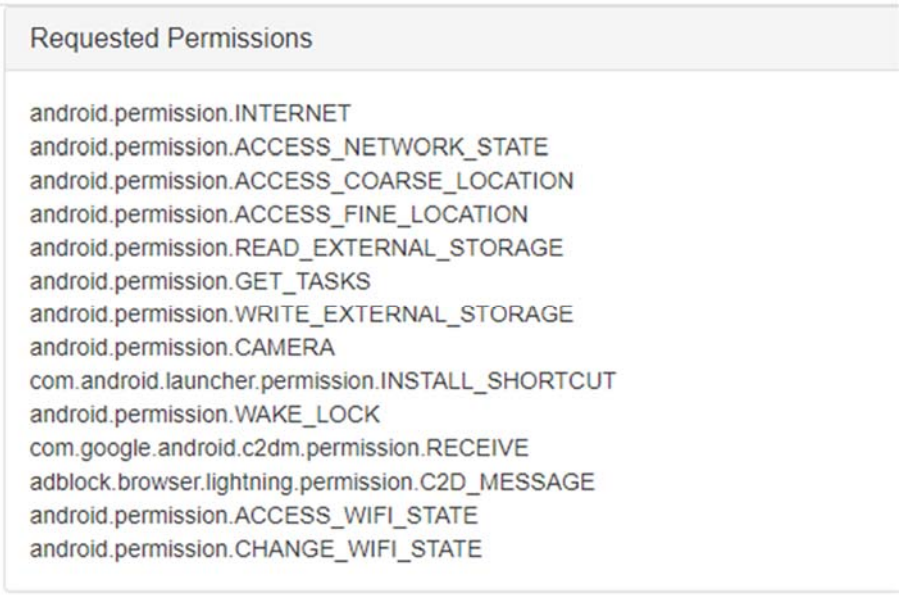
4.6 Ad Blocker browser

Το Ad Blocker browser είναι μία πολύ διαδεδομένη εφαρμογή με πάνω από 1εκατομμύριο λήψεις. Συμφώνα με την εταιρεία παρέχει αποφυγή διαφημίσεων με την χρήση ενός Ad Blocker καθώς επίσης και ανώνυμη περιήγηση, καθιστώντας την εφαρμογή ασφαλής για ιδιωτική περιήγηση. Η εφαρμογή είναι διαθέσιμη μέσω του Google Play Store.

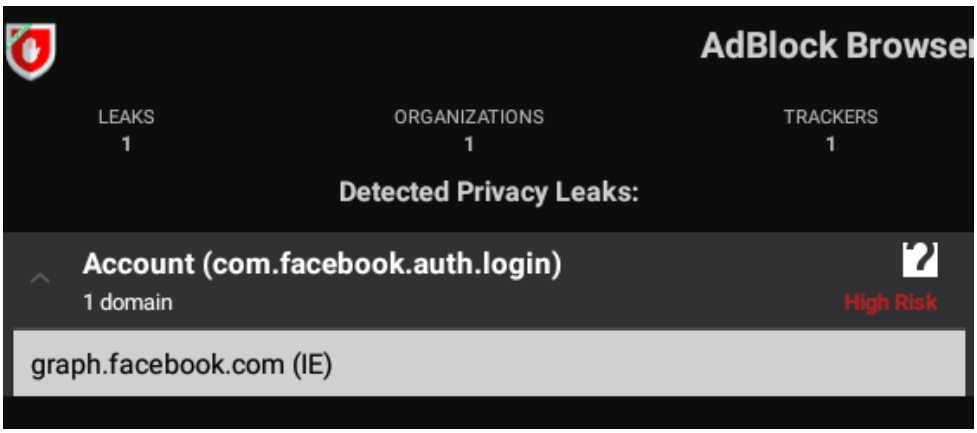
Εικόνα47: Η εφαρμογή Ad Blocker browser διαθέσιμη μέσω του Google Play Store

Η διαδικασία που έλαβε χώρα είναι ίδια με τις δύο πρώτες εφαρμογές. Μέσω του εργαλείου Inspeckage έγινε δυναμική ανάλυση της εφαρμογής και στην συνέχεια έγινε χρήση του Lumen Privacy Monitor. Στην παρακάτω εικόνα (εικόνα 48) βλέπουμε την προσβάσεις που αποκτά η εφαρμογή Ad Blocker στο κινητό του χρήστη.

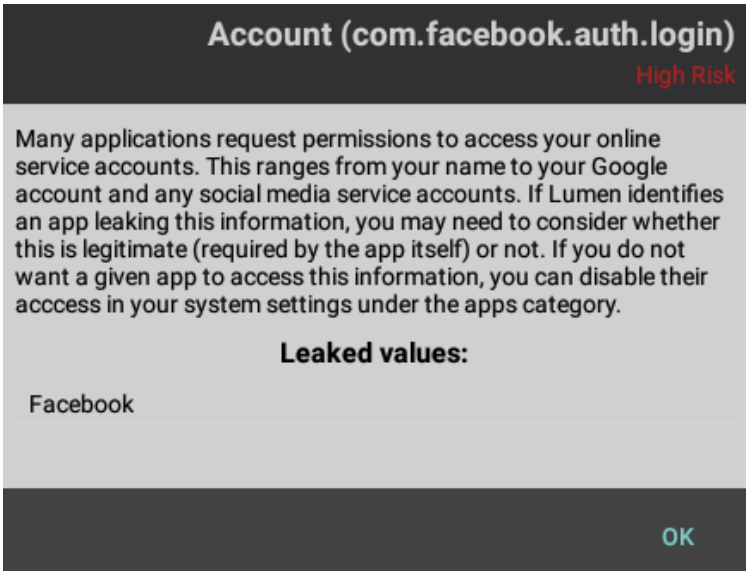
Στην συνέχεια, και χωρίς να περιηγηθούμε καθόλου στο Ad Blocker browser, χρησιμοποιήσαμε την εφαρμογή Lumen Privacy Monitor για να δούμε τυχόν διαρροές δεδομένων από την εφαρμογή προς τρίτους. Το πιο σημαντικό αποτέλεσμα της ανάλυσης του Lumen, που καθορίστηκε από διαρροή υψηλού κινδύνου εφαρμογής, είναι ότι η εφαρμογή Ad Blocker browser αποκτά πρόσβαση στην εφαρμογή Facebook στη συσκευή μας, (εικόνα 49). Όπως εξηγεί και το Lumen Privacy Monitor, πολλές εφαρμογές προσπαθούν να έχουν πρόσβαση σε λογαριασμούς χρηστών. Ο χρήστης εάν δει αυτήν την διαρροή στο σύστημα του, τότε θα πρέπει να δει προσεκτικά εάν αυτό έγινε με την έγκριση του η όχι (εικόνα 50).



Εικόνα 48: Άδειες σε πληροφορίες που αποκτά η εφαρμογή Ad Blocker browser



Εικόνα 49 : Διαρροές προσωπικών δεδομένων κατά τη χρήση του Ad Blocker browser, με βάση την παρακολούθηση Lumen



Εικόνα 50: Επεξηγηματικό μήνυμα από το Lumen Privacy Monitor

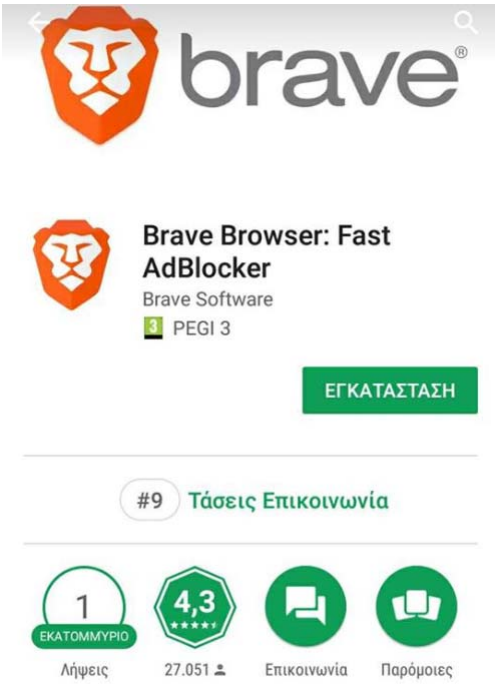
Τα συμπεράσματα οπότε είναι τα εξής :

1)Η εφαρμογή λαμβάνει πρόσβαση σε προσωπικά δεδομένα χρήστη, χωρίς την ρητή συγκατάθεση του, με κίνδυνο την διαρροή αυτών σε τρίτους.

2)Είναι πιθανό να υπάρχουν διαρροές προσωπικών δεδομένων σε τρίτους -- η πρόσβαση της εφαρμογής Ad Blocker browser στην εφαρμογή Facebook, που χαρακτηρίζεται ως διαρροή δεδομένων υψηλού κινδύνου από το εργαλείο παρακολούθησης Lumen, πρέπει να διερευνηθεί περαιτέρω, καθώς μπορεί να επιτρέψει στους διακομιστές του Adblocker να συνδεθούν με τους λογαριασμούς χρηστών στο Facebook.

4.7 Brave browser

Το Brave browser είναι ένας περιηγητής για κινητές συσκευές Android, με εγκατεστημένο Adblocker. Παράλληλα, σύμφωνα με το επίσημο site (https://brave.com/), η εφαρμογή εκτός από αποκλεισμό διαφημίσεων, παρέχει «απαράμιλλη ιδιωτικότητα και ασφάλεια» προς τον χρήστη. Η εφαρμογή Brave βρίσκεται διαθέσιμη στο Google Play Store, και μέχρι στιγμής έχει πάνω από 1 εκατομμύριο ¨κατεβάσματα¨.



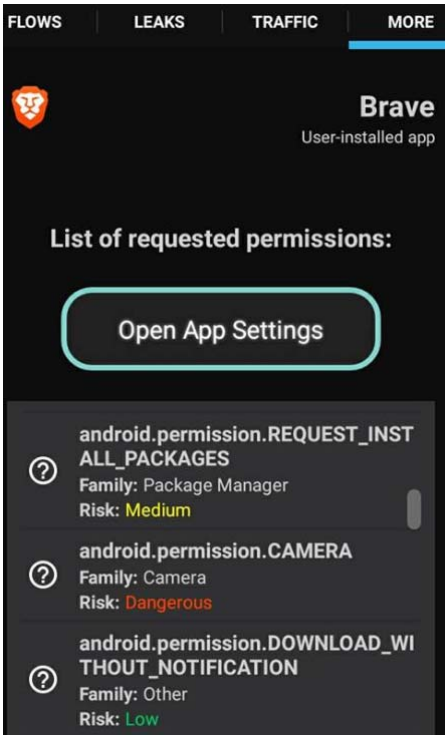
Εικόνα 51: Η εφαρμογή Brave browser διαθέσιμη στο Google Play Store

Στην εικόνα 52 παρουσιάζονται οι προσβάσεις του CM browser μέσω του εργαλείου Inspeckage. Τρέχοντας την εφαρμογή Lumen παρατηρούμε ότι η εφαρμογή ζητά δικαιώματα σε πληροφορίες, ορισμένες με μεγάλου βαθμού επικινδυνότητας (εικόνα 53, 54).

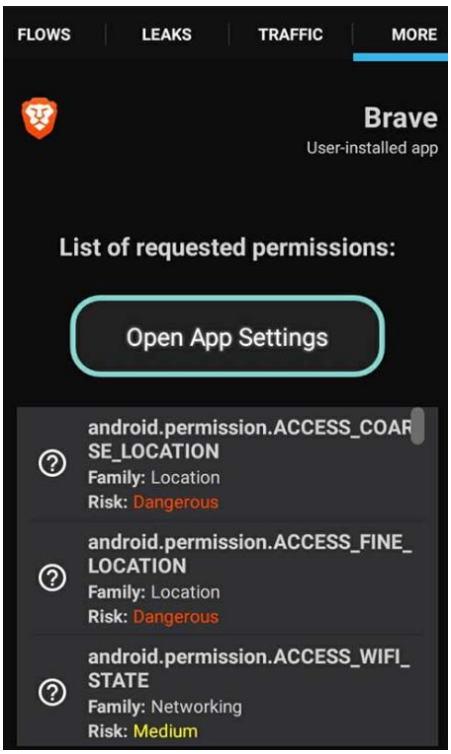
Requested Permissions

android.permission.ACCESS_COARSE_LOCATION
android.permission.ACCESS_FINE_LOCATION
android.permission.ACCESS_NETWORK_STATE
android.permission.CAMERA
android.permission.DOWNLOAD_WITHOUT_NOTIFICATION
android.permission.INTERNET
android.permission.MODIFY_AUDIO_SETTINGS
android.permission.NFC
android.permission.READ_EXTERNAL_STORAGE
android.permission.RECEIVE_BOOT_COMPLETED
android.permission.RECORD_AUDIO
android.permission.VIBRATE
android.permission.WAKE_LOCK
android.permission.WRITE_EXTERNAL_STORAGE
com.brave.browser.permission.C2D_MESSAGE
com.brave.browser.permission.READ_WRITE_BOOKMARK_FOLDERS
com.brave.browser.TOS_ACKED
com.chrome.permission.DEVICE_EXTRAS
com.android.browser.permission.READ_HISTORY_BOOKMARKS
com.android.browser.permission.WRITE_HISTORY_BOOKMARKS
com.google.android.c2dm.permission.RECEIVE
com.android.launcher.permission.INSTALL_SHORTCUT

Εικόνα 52: Δικαιώματα που αποκτά η εφαρμογή CM browser, σύμφωνα με το Inspeckage

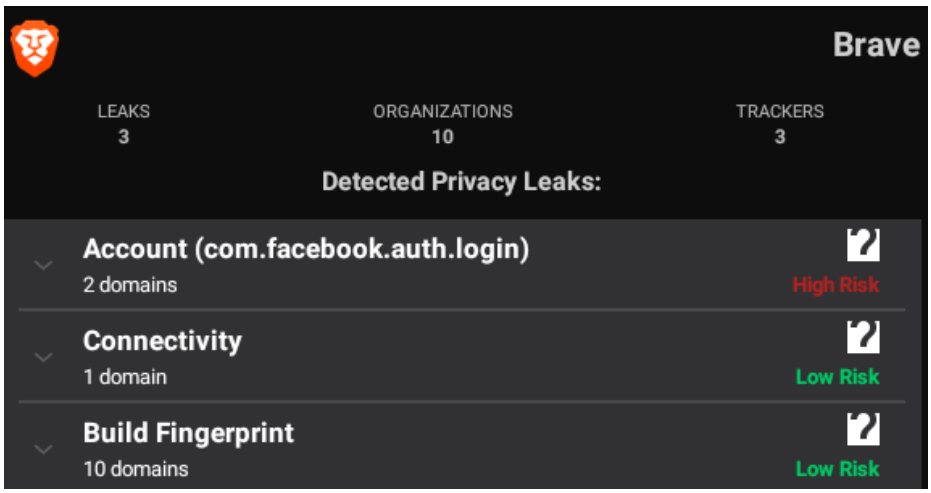


Εικόνα 53: Οι πληροφορίες που παίρνει πρόσβαση η εφαρμογή, και οι βαθμοί επικινδυνότητας

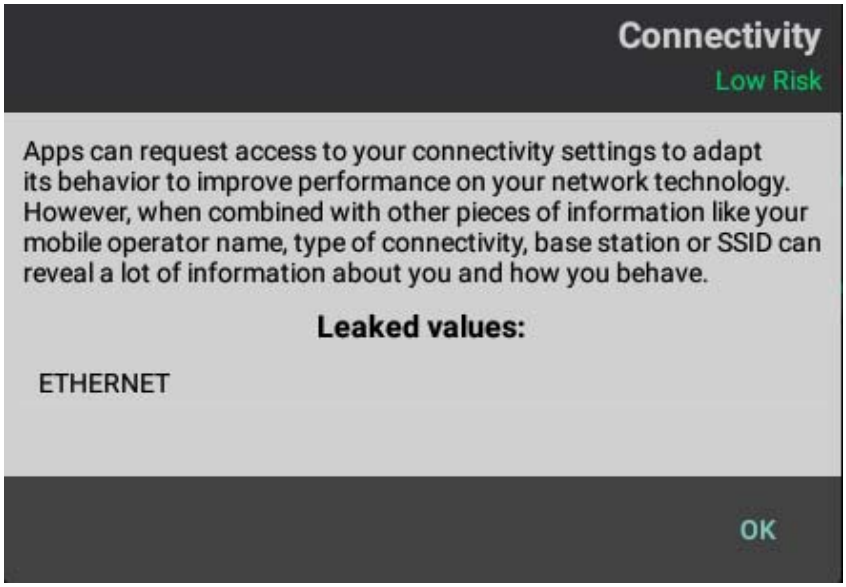


Εικόνα 54: Οι πληροφορίες που παίρνει πρόσβαση η εφαρμογή, και οι βαθμοί επικινδυνότητας

Στην συνέχεια, και χωρίς να περιηγηθούμε καθόλου στο Brave browser, χρησιμοποιήσαμε την εφαρμογή Lumen Privacy Monitor για να δούμε τυχόν διαρροές δεδομένων από την εφαρμογή προς τρίτους. Το πιο σημαντικό αποτέλεσμα της ανάλυσης του Lumen, που καθορίστηκε από διαρροή υψηλού κινδύνου εφαρμογής, είναι ότι η εφαρμογή Ad Blocker browser αποκτά πρόσβαση στην εφαρμογή Facebook στη συσκευή μας, (εικόνα 55). Παρατηρούμε επίσης μία διαρροή που αφορά την συνδεσιμότητα της συσκευής (connectivity), που δεν είχαμε συναντήσει μέχρι τώρα σε άλλη εφαρμογή. Σύμφωνα με το πρόγραμμα Lumen η πληροφορίες της συνδεσιμότητας της συσκευής, χρησιμοποιούνται από τρίτες υπηρεσίες με σκοπό την δημιουργία διαδικτυακού προφίλ χρήστη (εικόνα 56). Η πληροφορία αυτή φαίνεται να διαρρέεται για λογαριασμό της Amazon (εικόνα 57).



Εικόνα 55: Διαρροές δεδομένων μέσω του Brave browser, σύμφωνα με το Lumen Privacy Monitor



Εικόνα 56: Επεξηγηματικό μήνυμα από το Lumen Privacy Monitor

Εικόνα 57: Πληροφορίες για την συνδεσιμότητα της συσκευής, που διοχετεύονται στην υπηρεσία της Amazon

Τα συμπεράσματα λοιπόν είναι τα εξής:

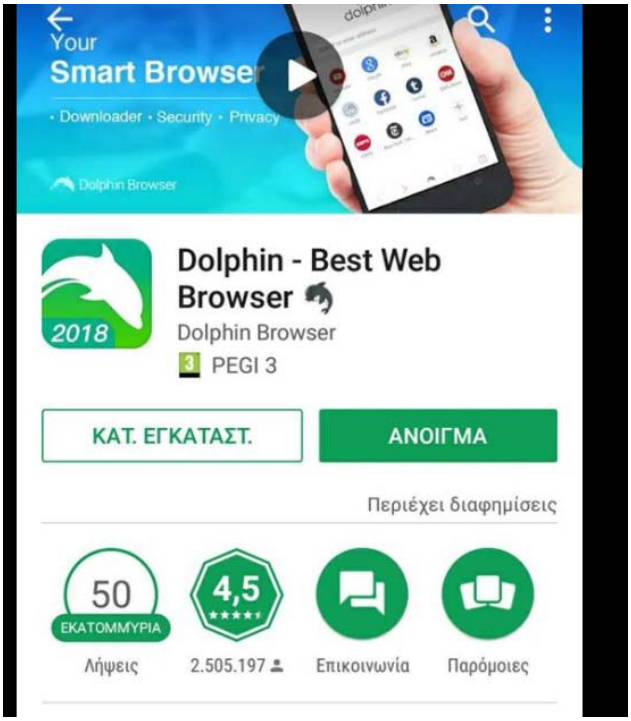
1)Η εφαρμογή λαμβάνει πρόσβαση σε προσωπικά δεδομένα χρήστη, όπως το ψηφιακό αποτύπωμα συσκευής, με κίνδυνο διαρροής αυτού σε τρίτους.

2)Η ενδεχόμενη πρόσβαση της εφαρμογής Ad Blocker browser στην εφαρμογή Facebook, που χαρακτηρίζεται ως διαρροή δεδομένων υψηλού κινδύνου από το εργαλείο παρακολούθησης Lumen, πρέπει να διερευνηθεί περαιτέρω, καθώς μπορεί να επιτρέψει στους διακομιστές του Adblocker να συνδεθούν με τους λογαριασμούς χρηστών στο Facebook.

3) Η διαρροή της συνδεσιμότητας του κινητού προς την υπηρεσία της Amazon, μπορεί να χαρακτηρίζεται ως χαμηλού κινδύνου, αποδεικνύει όμως πως ορισμένοι ιχνηλάτες δεν αποκλείονται από την εφαρμογή, όπως η ίδια ισχυρίζεται.

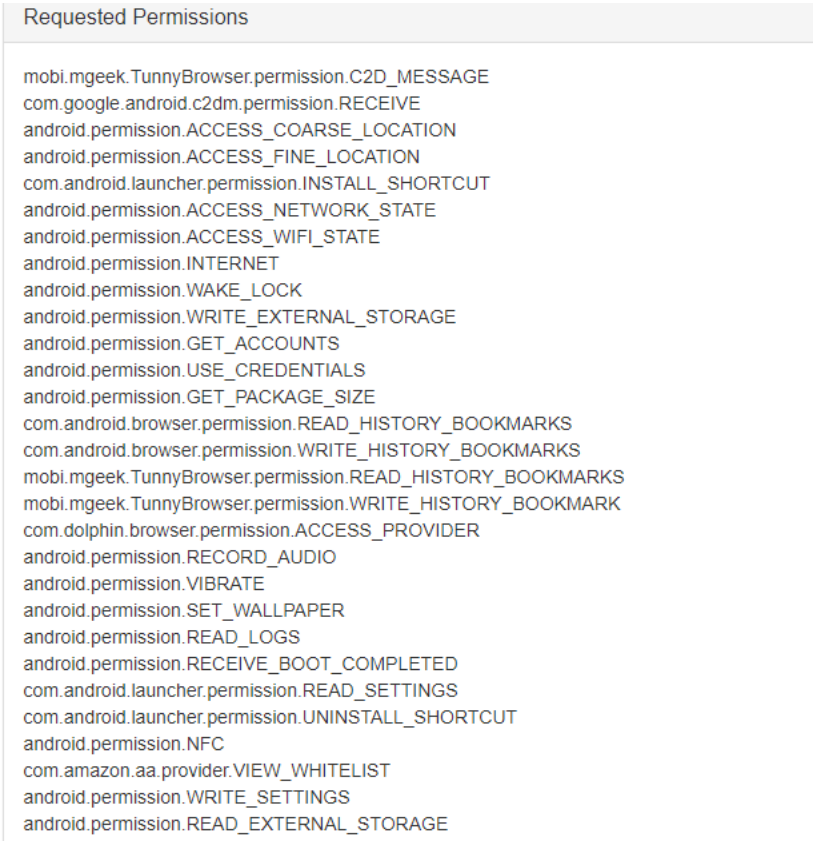
4.8 Dolphin –Best Web Browser

Το Dolphin – Best Web browser είναι μία πολύ διαδεδομένη εφαρμογή με πάνω από 50 εκατομμύρια λήψεις. Συμφώνα με την εταιρεία παρέχει αποφυγή διαφημίσεων με την χρήση ενός Ad Blocker καθώς επίσης και ανώνυμη περιήγηση, καθιστώντας την εφαρμογή ασφαλή για ιδιωτική περιήγηση. Η εφαρμογή είναι διαθέσιμη μέσω του Google Play Store.



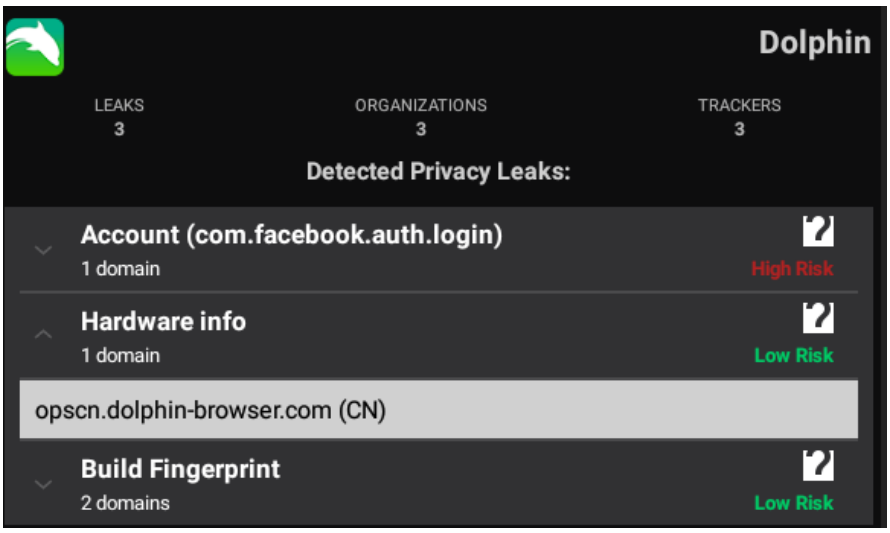
Εικόνα58: Η εφαρμογή Dolphin - Best Web browser διαθέσιμη μέσω του Google Play Store

Η διαδικασία που έλαβε χώρα είναι ίδια με τις παραπάνω εφαρμογές. Μέσω του εργαλείου Inspeckageέγινε δυναμική ανάλυση της εφαρμογής και στην συνέχεια έγινε χρήση του Lumen Privacy Monitor. Στην παρακάτω εικόνα (εικόνα 59) βλέπουμε την προσβάσεις που αποκτά η εφαρμογή Dolphin στο κινητό του χρήστη. Μπορούμε εύκολα να διαπιστώσουμε ότι είναι η εφαρμογή με τις περισσότερες προσβάσεις στην μέχρι τώρα ανάλυση μας.



Εικόνα 59: Άδειες σε πληροφορίες που αποκτά η εφαρμογή Dolphin – Best Web browser

Ανοίγοντας την εφαρμογή Lumen Privacy Monitor, παρατηρούμε την ενδεχόμενη πρόσβαση του Dolphin – Best Web browser στην εφαρμογή Facebook του χρήστη, όπως συμβαίνει και με τις προηγούμενες εφαρμογές, που χαρακτηρίζεται ως υψηλού κινδύνου (εικόνα 60). Επίσης παρατηρούμε διαρροή πληροφοριών που αφορούν το Hardware του κινητού, καθώς επίσης του ψηφιακού αποτυπώματος για λογαριασμό υπηρεσιών του περιηγητή Dolphin.



Εικόνα 60: Διαρροή πληροφοριών προς υπηρεσίες του Dolphin- browser

Τα συμπεράσματα λοιπόν είναι τα εξής:

1)Η εφαρμογή λαμβάνει πρόσβαση σε προσωπικά δεδομένα χρήστη, με κίνδυνο διαρροής αυτών σε τρίτους.

2)Η ενδεχόμενη πρόσβαση της εφαρμογής Dolphin – Best Web browser στην εφαρμογή Facebook, που χαρακτηρίζεται ως διαρροή δεδομένων υψηλού κινδύνου από το εργαλείο παρακολούθησης Lumen, πρέπει να διερευνηθεί περαιτέρω, καθώς μπορεί να επιτρέψει στους διακομιστές του Ad blocker να συνδεθούν με τους λογαριασμούς χρηστών στο Facebook.

3) Η διαρροή πληροφοριών hardware και fingerprint του κινητού προς την υπηρεσία της Dolphin – Best Web browser, μπορεί να χαρακτηρίζεται ως χαμηλού κινδύνου, αποδεικνύει όμως πως η ίδια εφαρμογή συλλέγει δεδομένα χρηστών για δικούς τους σκοπούς.

Κεφάλαιο 5
Επίλογος

Η παρούσα μεταπτυχιακή διατριβή ασχολήθηκε με το θέμα της συμπεριφορικής διαφήμισης και δημιουργίας προφίλ στο διαδίκτυο. Παρουσιάστηκαν λεπτομερώς οι τεχνικές παρακολούθησης χρηστών, που χρησιμοποιούνται ευρέως στις μέρες μας από υπηρεσίες/εταιρείες, για σκοπούς συλλογής δεδομένων χρηστών και στοχευμένης διαφήμισης. Επιπλέον παρουσιάστηκε η εξέλιξη των δικτύων διαφημίσεων με την πάροδο του χρόνου και αναλύθηκε λεπτομερώς η τελευταία μέθοδος που χρησιμοποιείται κατά κόρον (Real Time Bidding), και που προσφέρει τεραστία έσοδα στο διαδικτυακό μάρκετινγκ. Επιπροσθέτως αναφερθήκαμε στο γεγονός της ραγδαίας αύξησης εφαρμογών Ad Blocking που χρησιμοποιούν οι χρήστες στην προσπάθεια τους για αποφυγή παρακολούθησης και στοχευμένης διαφήμισης από υπηρεσίες /εταιρείες. Οι πολύ πρόσφατες έρευνες που έχουν γίνει για τις συγκεκριμένες εφαρμογές, έδειξαν ότι εγείρουν βασικά θέματα κατά της ιδιωτικότητας των χρηστών. Αυτό το στοιχείο αποτέλεσε την βάση για την υλοποίηση κατάλληλου περιβάλλοντος εργασίας για την έρευνα μας. Με την ανάλυση των εφαρμογών

επιχειρούμε να αναδείξουμε τι δεδομένα συλλέγουν οι εφαρμογές και κατά πόσο αυτές αποτελούν απειλές για την ιδιωτικότητα του χρήστη. Στην συνέχεια προσπαθήσαμε να διαλευκάνουμε αν αυτά τα δεδομένα μερικά ή όλα, διαρρέονται προς τρίτες εταιρείες εν αγνοία των χρηστών.

Η ανάλυση των εφαρμογών μας έδειξε ότι εφαρμογές πετυχαίνουν τον αποκλεισμό διαφημίσεων, παρόλα αυτά λαμβάνουν πρόσβαση σε πολύ προσωπικά δεδομένα χρήστη, και δεν μπορεί να αποκλειστεί το ενδεχόμενο εκμετάλλευσης των πληροφοριών αυτών από τρίτες οντότητες. Επιπλέον η ανάλυση φανέρωσε ότι καμία εφαρμογή δεν παράγει πλήρη ιδιωτικότητα, μιας και τα δεδομένα που συλλέγονται διαμοιράζονται προς τρίτες υπηρεσίες είτε για σκοπούς ανάλυσης είτε για σκοπούς παρακολούθησης χρηστών και είναι ικανά να σχηματίσουν ένα διαδικτυακό προφίλ χρήστη. Το εύρημα αυτό επιβεβαιώνεται τόσο από την άντληση του Android ID της συσκευής, το οποίο βρέθηκε να διαμοιράζεται από μία εφαρμογή Ad Blocking σε τρίτους, όσο και από το αποτύπωμα της συσκευής (build fingerprinting) και πληροφοριών hardwaretης συσκευής, που βρέθηκε να διαμοιράζεται προς τρίτους σε τέσσερις από τις πέντε εφαρμογές Ad blocking. Επιπλέον τρεις από τις πέντε εφαρμογές αποκτούν πρόσβαση στην εφαρμογή Facebook του χρήστη, χωρίς την συγκατάθεση του, κάτι που πρέπει να διερευνηθεί περαιτέρω, καθώς μπορεί να επιτρέψει στους διακομιστές των εφαρμογών Ad Blocking να συνδεθούν με τους λογαριασμούς χρηστών στο Facebook.

Τα ευρήματα αυτά δείχνουν πως η ιδιωτικότητα των χρηστών εξακολουθεί να αποτελεί ένα μεγάλο ζήτημα και είναι πολύ δύσκολο να επιτευχθεί πλήρως. Το γεγονός ότι πολλοί χρήστες, χρησιμοποιούν Ad blockers με σκοπό την αποφυγή παρακολούθησης και διαφημίσεων, θεωρώντας τα απόλυτα ασφαλή, καθιστά τις εν λόγο εφαρμογές ένα μη ασφαλές περιβάλλον που ενδεχομένως να επιφέρει δυσάρεστες επιπτώσεις προς τον χρήστη.

Το κύριο συμπέρασμα είναι το ότι οι χρήστες των «έξυπνων» εφαρμογών, στο βαθμό που οι προγραμματιστές αυτών τελικά προβαίνουν είτε άμεσα είτε έμμεσα – μέσω των «βιβλιοθηκών» τρίτων μελών (third party libraries) – σε επεξεργασίες προσωπικών δεδομένων ερήμην τους, δεν έχουν τελικά πολλές δυνατότητες στο να αντιταχθούν σε αυτήν την επεξεργασία. Για την αντιμετώπιση αυτού του ζητήματος θα πρέπει όλοι όσοι εμπλέκονται στην ανάπτυξη εφαρμογών – από τους παρόχους λειτουργικών συστημάτων μέχρι τις εταιρείες που αναπτύσσουν βοηθητικές «βιβλιοθήκες» (libraries) λογισμικού – να

καταβάλλουν κάθε προσπάθεια έτσι ώστε να υιοθετούνται τεχνικές φιλικές προς την ιδιωτικότητα. Για παράδειγμα, προς σε αυτήν την κατεύθυνση, στην έκδοση Android-οτου Android, οι εφαρμογές δεν μπορούν να έχουν απευθείας πρόσβαση στο μοναδικό Androidαναγνωριστικό (Android ID) και, μάλιστα, η κάθε διαφορετική εφαρμογή «βλέπει» διαφορετικό Android ID[1].

Σε κάθε περίπτωση, θα πρέπει να ληφθεί υπόψη και το νέο ευρωπαϊκό νομοθετικό πλαίσιο για την προστασία των προσωπικών δεδομένων που τίθεται σε εφαρμογή στις 25 Μαΐου 2018, το Γενικό Κανονισμό Προστασίας Δεδομένων (General Data Protection Regulation - GDPR), ο οποίος θα έχει εφαρμογή σε κάθε επεξεργασία δεδομένων ευρωπαίων πολιτών, ακόμα και αν αυτός που κάνει την επεξεργασία δεν βρίσκεται στην Ευρωπαϊκή Ένωση. Ο Γενικός Κανονισμός θέτει πρόσθετες υποχρεώσεις για τη νόμιμη επεξεργασία προσωπικών δεδομένων, μεταξύ των οποίων μεγαλύτερη διαφάνεια και λήψη αδιαμφισβήτητης ρητής και ειδικής συγκατάθεσης για κάθε επεξεργασία της οποίας η συγκατάθεση είναι απαραίτητη προϋπόθεση. Κατά συνέπεια, και οι επεξεργασίες των ad blockers που περιγράφηκαν στην παρούσα διατριβή, εφόσον αξιοποιούνται από Ευρωπαίους πολίτες, φαίνεται ότι εμπίπτουν στο πεδίο εφαρμογής του GDPR, καθιστώντας πιο επιτακτική την ανάγκη συμμόρφωσής τους με αυτόν.

[1] https://android-developers.googleblog.com/2017/04/changes-to-device-identifiers-in.html

Βιβλιογραφία

[01] S. C. Boerman, S. Kruikemeier, and F. J. Zuiderveen Borgesius, “Online Behavioral Advertising: A Literature Review and Research Agenda,” *J. Advert.*, vol. 46, no. 3, pp. 363–376, 2017.

[02] Y. Yuan, F. Wang, J. Li, and R. Qin, “A survey on real time bidding advertising,” *Proc. 2014 IEEE Int. Conf. Serv. Oper. Logist. Informatics, SOLI 2014*, pp. 418–423, 2014.

[03] T. Bujlow, V. Carela-Espanol, B. R. Lee, and P. Barlet-Ros, “A Survey on Web Tracking: Mechanisms, Implications, and Defenses,” *Proc. IEEE*, vol. 105, no. 8, pp. 1476–1510, 2017.

[04] K. McKinley, “Cleaning Up After Cookies,” *Communications*, pp. 1–12, 2010.

[05] N. Schmucker, “Web Tracking,” *SNET2 Semin. Pap.*, vol. 57, pp. 1–34, 2011.

[06] S. E. Peacock, “Prized assets for web tracking,” *Proc. 2015 Int. Conf. Soc. Media Soc. - SMSociety ’15*, no. July, pp. 1–5, 2015.

[07] B. Information, “What are cookies ? What are the differences between them (session vs . persistent)?,” pp. 2–3, 2014.

[08] N. Muhammad, “Tracking and Identifying Individual Users in a Web Surfing Session.”

[09] A. Afolayan, “Web security : ways to identify & track user of a web surfing session , His activities and the sites visited .”

[10] K. Michkova, “Should we be worried about HTTP cookies and the tracking of the people ' s online movement ? The comparison of the US and the EU law directives and what other precautions should be done for the future .,” 2014.

[11] G. Fleischer, “Implementing Web Tracking,” pp. 1–37.

[12] A. Soltani, S. Canty, Q. Mayo, L. Thomas, and C. J. Hoofnagle, “Flash Cookies and Privacy,” *Names*, pp. 1–8, 2009.

[13] Aleecia M. McDonald and Lorrie Faith Cranor, “A Survey of the Use of Adobe Flash Local Shared Objects to Respawn HTTP Cookies,” *I/S A J. Law Policy Inf. Soc.*, vol. 7, pp. 639–721, 2012.

[14] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, “The Web Never Forgets: Persistent Tracking Mechanisms in the Wild,” *Proc. 2014 ACM SIGSAC Conf. Comput. Commun. Secur. - CCS ’14*, pp. 674–689, 2014.

[15] B. T. Vega, H. M. Language, and J. Cox, “Privacy Risks,” pp. 10–13, 2018.

[16] P. Verleg, “Cache Cookies : searching for hidden browser storage,” 2014.

[17] C. Scientist and T. Italia, “Flash Cookies and Privacy II :,” *World Wide Web Internet Web Inf. Syst.*, 2009.

[18] S. Z. Naseem and F. Majeed, “Extending HTML5 local storage to save more data; Efficiently and in more structured way,” *8th Int. Conf. Digit. Inf. Manag. ICDIM 2013*, pp. 337–340, 2013.

[19] W. Workers and I. Api, “IndexedDB API,” pp. 1–7, 2018.

[20] S. Kimak and J. Ellman, “HTML5 IndexedDB Encryption: Prevention against Potential Attacks,” *Int. J. Intell. Comput. Res.*, vol. 6, no. 4, pp. 621–629, 2015.

[21] J. Ellingwood, “Web Caching Basics: Terminology, HTTP Headers, and Caching Strategies,” *DigitalOcean*, pp. 1–11, 2015.

[22] A. Juels, M. Jakobsson, and T. N. Jagatic, “Cache cookies for browser authentication (extended abstract),” *Proc. - IEEE Symp. Secur. Priv.*, vol. 2006, pp. 301–305, 2006.

[23] R. Fielding, U. C. Irvine, and J. Gettys, “16/3/2018 www.ietf.org/rfc/rfc2616.txt,” pp. 1–143, 2018.

[24] E. Bursztein, “Tracking users that block cookies with a HTTP redirect,” no. July 2011, pp. 1–6, 2011.

[25] D. Goodin, “Browsing in privacy mode? Super Cookies can track you anyway,” *Ars Tech.*, p. 1, 2014.

[26] R. O. B. Price, “New ‘ Super Cookies ’ Can Track Your Private Web Browsing — And Apple Users Can 't Get Rid Of Them,” pp. 1–3, 2018.

[27] M. Zawadziński, “How Does It Work ? Successful,” pp. 1–7, 2018.

[28] N. Nikiforakis, A. Kapravelos, W. Joosen, C. Kruegel, F. Piessens, and G. Vigna, “Cookieless monster: Exploring the ecosystem of web-based device fingerprinting,” *Proc. - IEEE Symp. Secur. Priv.*, no. 20837680, pp. 541–555, 2013.

[29] C. J. C. Hoofnagle, A. Soltani, N. Good, D. J. D. D. J. Wambach, and M. Ayenson, “Behavioral Advertising: The Offer You Cannot Refuse,” *Harv. L. Pol’y Rev.*, vol. 6, p. 273, 2012.

[30] G. Retscher, “FUSION of LOCATION FINGERPRINTING and TRILATERATION BASED on the EXAMPLE of DIFFERENTIAL WI-FI POSITIONING,” *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 4, no. 2W4, pp. 377–384, 2017.

[31] B. Marr, “How Businesses Can Use Device Fingerprinting To Identify And Track Customers,” pp. 2017–2018, 2018.

[32] P. Eckersley, “How unique is your web browser?,” *Lect. Notes Comput. Sci. (including*

Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 6205 LNCS, pp. 1–18, 2010.

[33] P. Laperdrix, B. Baudry, and W. Rudametkin, "Beauty and the Beast : Diverting modern web browsers to build unique browser fingerprints," *IEEE Symp. Secur. Priv.*, 2016.

[34] N. Nikiforakis, A. Kapravelos, W. Joosen, C. Kruegel, F. Piessens, and G. Vigna, "Cookieless monster: Exploring the ecosystem of web-based device fingerprinting," *Proc. - IEEE Symp. Secur. Priv.*, pp. 541–555, 2013.

[35] K. Mowery and H. Shacham, "Pixel Perfect : Fingerprinting Canvas in HTML5," *Web 2.0 Secur. Priv. 20*, pp. 1–12, 2012.

[36] U. H. Functions, "Everything You Need to Know About Canvas Fingerprinting How Do Websites Read Canvas," pp. 1–10, 2018.

[37] R. B. Posts, "Passive OS Fingerprinting," pp. 3–5, 2018.

[38] B. Mitchell, "Does IP Address Location (Geolocation) Really Work?," *About Tech*, pp. 1–8, 2015.

[39] S. Yuan, J. Wang, and X. Zhao, "Real-time Bidding for Online Advertising: Measurement and Analysis," 2013.

[40] S. Muthukrishnan, "Ad Exchanges: Research Issues."

[41] K. W. De Bock and D. Van den Poel, "WORKING PAPER Predicting web site audience demographics for web advertising targeting using multi-web site clickstream data Predicting web site audience demographics for web advertising," pp. 1–32, 2009.

[42] B. Edelman, M. Ostrovsky, and M. Schwarz, "Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords," *Am. Econ. Rev.*, vol. 97, no. 1, pp. 242–259, 2007.

[43] M. Ikram and M. A. Kaafar, "A First Look at Ad Blocking Apps on Google Play," arXiv:1709.02901, September 2017.

[44] A. Gervais, A. Filios, V. Lenders, and S. Capkun, "Quantifying web adblocker privacy," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10493 LNCS, pp. 21–42, 2017.

[45] S. Son, D. Kim, and V. Shmatikov, "What Mobile Ads Know About Mobile Users," *Proc. 2016 Netw. Distrib. Syst. Secur. Symp.*, 2016.

[46] A. Razaghpanah *et al.*, "Apps , Trackers , Privacy , and Regulators," no. February, 2018.