

Big Data & Bots

Αναγνωστοπούλου Νίκη – Σακκίδου Μαριτίνα Β'2

Τα «Μεγάλα Δεδομένα» είναι σύνολα δεδομένων, που είναι πολύ μεγάλα ή πολύπλοκα για να αντιμετωπιστούν από το παραδοσιακό λογισμικό εφαρμογών επεξεργασίας δεδομένων. Τα δεδομένα με πολλές εγγραφές (γραμμές) προσφέρουν μεγαλύτερη στατιστική ισχύ, ενώ δεδομένα με υψηλότερη πολυπλοκότητα (περισσότερα χαρακτηριστικά ή στήλες) μπορεί να οδηγήσουν σε υψηλότερο ποσοστό ψευδούς ανακάλυψης. Οι προκλήσεις της ανάλυσης μεγάλων δεδομένων περιλαμβάνονται από τη συλλογή δεδομένων, την αποθήκευση δεδομένων, την ανάλυση δεδομένων, την αναζήτηση, την κοινή χρήση, τη μεταφορά, την αναζήτηση, την ενημέρωση, το απόρρητο πληροφοριών και την πηγή δεδομένων. Τα μεγάλα δεδομένα συνδέθηκαν αρχικά με τρεις βασικές έννοιες: όγκος, ποικιλία και ταχύτητα. Η τρέχουσα χρήση του όρου *μεγάλα δεδομένα* τείνει να αναφέρεται στη χρήση προγνωστικών αναλυτικών στοιχείων, ανάλυσης συμπεριφοράς χρήστη ή ορισμένων άλλων προηγμένων μεθόδων ανάλυσης δεδομένων που εξαγάγουν αξία από μεγάλα δεδομένα και σπάνια σε συγκεκριμένο μέγεθος συνόλου δεδομένων. «Δεν υπάρχει αμφιβολία ότι οι ποσότητες των δεδομένων που είναι τώρα διαθέσιμες είναι πράγματι μεγάλες, αλλά αυτό δεν είναι το πιο σχετικό χαρακτηριστικό αυτού του νέου οικοσυστήματος δεδομένων».

-Αναγνωστοπούλου Νίκη

Πηγές: https://en.wikipedia.org/wiki/Big_data





Το «**Διαδικτυακό Ρομπότ**» είναι ένα πρόγραμμα που εκτελεί αυτοματοποιημένες εργασίες μέσω του διαδικτύου. Στις περισσότερες περιπτώσεις τα Bots εκτελούν σχετικά απλές λειτουργίες που θα πρέπει να επαναληφθούν εκατοντάδες ή χιλιάδες φορές. Μία κλασσική εφαρμογή των Bots είναι οι αράχνες του διαδικτύου, οι οποίες περιφέρονται από ιστοσελίδα σε ιστοσελίδα και χρησιμοποιούνται για την ανάλυσή της σε ρυθμό πολλαπλάσιο απ' ό,τι θα μπορούσε

ένας άνθρωπος. Οι μηχανές αναζήτησης χρησιμοποιούν τέτοιες αράχνες για την ανάλυση και ταξινόμηση των ιστοσελίδων σύμφωνα με διάφορες λέξεις-κλειδιά, ούτως ώστε στην συνέχεια να μπορούν να παρουσιάσουν στον χρήστη τα αποτελέσματα της αναζήτησης σε πολύ μικρό χρονικό διάστημα. Τέλος, οι διάφοροι web servers μπορούν να δημιουργήσουν ένα απλό αρχείο κειμένου με το όνομα "robots.txt", το οποίο θα περιέχει κανόνες που θα πρέπει να τηρήσουν τα Bots που επισκέπτονται την σελίδα. Παρόλα αυτά όμως, μπορεί κάποιο bot να αγνοήσει αυτούς τους κανόνες. Παρότι τα Διαδικτυακά ρομπότ είναι πολύ χρήσιμα για την καθημερινότητά μας, χρησιμοποιούνται πολλές φορές από χάκερς. Γενικά μπορεί κανείς να διαχωρίσει τις κακόβουλες χρήσεις των bots σε διάφορες κατηγορίες:

- *Spambots*: Ψάχνουν κάθε σελίδα του διαδικτύου για διευθύνσεις email, ούτως ώστε αυτές στην συνέχεια να χρησιμοποιηθούν για spamming.
- Bots που κατεβάζουν ολόκληρα sites και χρησιμοποιούν όλο το διαθέσιμο bandwidth του web server με αποτέλεσμα να μην μπορούν να εξυπηρετηθούν άλλοι χρήστες (επιθέσεις άρνησης υπηρεσιών, αγγλ. DoS - Denial of Service).
- Bots που χρησιμοποιούνται από διάφορους ιούς υπολογιστών, δούρειους ίππους και σκουλήκια.
- Οι υπολογιστές zombie χρησιμοποιούνται πολλές φορές για την δημιουργία Bots και την διεξαγωγή διαδικτυακής επίθεσης σε έναν server.

-Σακκίδου Μαριτίνα

Πηγές: https://en.wikipedia.org/wiki/Internet_bot