Υλοποίηση Αλγορίθμων με χρήση του Pregel Programming Framework

Ημερομηνία Παράδοσης: ~30/06/2024

Η συγκεκριμένη εργασία είναι το **πολύ δύο ατόμων** και θα δοθεί **μόνο σε μία ομάδα**. Είναι **απαλλακτική** για το σύνολο των εργασιών που θα δοθούν στο μάθημα.

Το Pregel είναι μία κλιμακώσιμη και ανθεκτική στα λάθη πλατφόρμα που αναπτύχθηκε από την Google με ένα κατάλληλο API που να επιτρέπει τον εύκολο προγραμματισμό αλγορίθμων σε γραφήματα. Στοχεύει κυρίως σε πολύ μεγάλα γραφήματα που προϋποθέτουν λόγω του όγκου τους κατανεμημένη επεξεργασία. Υπάρχουν αρκετές υλοποιήσεις ανοικτού κώδικα αυτής της πλατφόρμας (από τις πρώτες είναι η Giraph).

Τα ζητούμενα είναι τα εξής:

1. Περιγραφή του προγραμματιστικής προσέγγισης Pregel, χρησιμοποιώντας κατά προτίμηση το API GraphX. Οι φοιτητές μπορούν επίσης να χρησιμοποιήσουν άλλες βιβλιοθήκες ή βάσεις δεδομένων για μεγάλα γραφήματα όπως Giraph, Neo4J ή άλλα.
2. Υλοποίηση επιλεγμένου(-ων) αλγορίθμου(-ων) σε αυτό το προγραμματιστικό περιβάλλον και πειραματισμός.
3. Συγγραφή μίας αναφοράς που εν δυνάμει θα αποτελέσει ένα εγχειρίδιο για μελλοντικούς φοιτητές στο πως να χρησιμοποιούν αυτή την προγραμματιστική προσέγγιση – πάντα σε συνδυασμό με την προσέγγιση που έχετε υιοθετήσει σχετικά με τον χειρισμό των γραφημάτων.

Ο διδάσκων και η ομάδα του δεν έχει εμπειρία σε Pregel – πέρα από τη θεωρητική θεμελίωση της “thinking like a vertex” προσέγγισης. Ως εκ τούτου, οι φοιτητές θα πρέπει να ψάξουν αρκετά αν και δεν είναι τόσο δύσκολο όσο ακούγεται (δείτε άρθρο).

Η συγκεκριμένη εργασία θα ανανεώνεται όσον αφορά τα ζητούμενα (ποιοι αλγόριθμοι θα υλοποιούνται) καθώς προχωρά η ύλη και μετά από συνεννόηση μεταξύ της ομάδας και του διδάσκοντα. Αρχικά θα πρέπει οι φοιτητές να κατανοήσουν τις βασικές αρχές του Pregel και έπειτα να επιλέξουν την πλατφόρμα διαχείρισης των δικτύων. Αφού επιτευχθούν αυτοί οι στόχοι, η ομάδα θα πρέπει να υλοποιήσει τον τυχαιοποιημένο αλγόριθμο MIS. Η ομάδα που θα αναλάβει αυτή την εργασία θα ασχοληθεί με τεχνολογίες αιχμής όσον αφορά τη διαχείριση, ανάλυση και επεξεργασία μεγάλων δικτύων.

Οι ομάδες που ενδιαφέρονται παρακαλούνται να επικοινωνήσουν με τον διδάσκοντα μέσω email.