

ΠΤΥΧΙΑΚΕΣ ΕΡΓΑΣΙΕΣ

Επίκουρου Καθηγητή Ιωάννη Ε. Βενέτη

10 Οκτωβρίου 2024

Γενική περιγραφή

- Οι προσφερόμενες Πτυχιακές Εργασίες (Π.Ε.) κινούνται σε δύο βασικούς άξονες:
 - Μελέτη, αξιολόγηση, σχεδίαση και ανάπτυξη μεθόδων για την αξιοποίηση Παράλληλου Υπολογισμού σε εφαρμογές με αυξημένες υπολογιστικές απαιτήσεις, όπως εφαρμογές Τεχνητής Νοημοσύνης, προσομοιώσεις (simulations), προβλήματα μηχανικής, κλπ. Στην περίπτωση αυτή η αξιοποίηση Παράλληλου Υπολογισμού γίνεται τυπικά με την χρήση κατάλληλων βιβλιοθηκών για την χρησιμοποιούμενη γλώσσα προγραμματισμού, οι οποίες υποστηρίζουν την παράλληλη εκτέλεση του κώδικα των συναρτήσεων τους. Επικουρικά είναι πιθανό να χρειαστεί η ανάπτυξη τμημάτων κώδικα με την χρήση κατάλληλων παράλληλων προγραμματιστικών μοντέλων (OpenMP ή/και MPI ή/και CUDA).
 - Μελέτη, αξιολόγηση, σχεδίαση και ανάπτυξη μεθόδων για την υποστήριξη αποδοτικής παράλληλης εκτέλεσης εφαρμογών στο πλαίσιο παράλληλων προγραμματιστικών μοντέλων. Ζητήματα αυτού του είδους περιλαμβάνουν αλγόριθμους αποδοτικής δημιουργίας και διαχείρισης νημάτων και έργων (tasks), αλγόριθμους εξισορρόπησης φόρτου εργασίας (workload balancing algorithms), αλγόριθμους χρονοδρομολόγησης βρόχων (loop scheduling algorithms), παράλληλους αλγόριθμους για την υλοποίηση δομών δεδομένων (π.χ. ουρών), κλπ.
- Αφετηρία σε πολλές από τις παρακάτω εκφωνήσεις είναι το γεγονός ότι η Υπολογιστική Υψηλών Επιδόσεων (HPC = High Performance Computing) έχει καταστεί απαραίτητη υποδομή σε μυριάδες εφαρμογών. Με βάση αυτό, αλλά και πρόσφατες μελέτες που αναδεικνύουν την σημασία των επενδύσεων στο αντικείμενο, η απόκτηση ειδικής εμπειρίας στο αντικείμενο μέσω μαθημάτων και της εκπόνησης της Π.Ε., εφόσον βέβαια συνοδεύεται και από ενδιαφέρον για τα συναφή ζητήματα, μπορεί να αποτελέσει πολύ χρήσιμο εφόδιο.
- Βασική προϋπόθεση για την εκπόνηση καλής Π.Ε. στα παρακάτω είναι να σας προκαλεί το ενδιαφέρον και να σας ευχαριστεί η ενασχόληση με το θέμα που θα επιλέξετε (ελλείψει καλύτερου τρόπου, κρίνετε με βάση το ενδιαφέρον σας για τα συναφή μαθήματα). Επίσης να έχετε τη διάθεση να επενδύσετε τον απαιτούμενο χρόνο για την εκπόνηση της εργασίας. Όλες οι εργασίες θα χρειαστούν μελέτη της σχετικής βιβλιογραφίας καθώς και τη συγγραφή κώδικα και υλοποιήσεις σε ειδικά περιβάλλοντα.
- Σημαντικό τμήμα κάθε Π.Ε. είναι οι υλοποιήσεις με ανάπτυξη λογισμικού. Το περιβάλλον προγραμματισμού και υλοποίησης εξαρτάται από την Π.Ε.
- Οι παρακάτω εκφωνήσεις παρέχουν ένα γενικό πλαίσιο το οποίο όμως μπορεί να τροποποιηθεί ή να αναθεωρηθεί μετά από συνεννόηση. Αν υπάρχει ενδιαφέρον, ορισμένες από τις Π.Ε. μπορούν να μετεξελιχθούν σε εργασίες μεταπτυχιακού επιπέδου.
- Για περισσότερες πληροφορίες παρακαλείστε να επικοινωνείτε με τον κ. Βενέτη (venetis@unipi.gr).

Προϋποθέσεις για την εκπόνηση καλής Π.Ε.

Η Π.Ε. είναι συνήθως η πρώτη φορά που ένας φοιτητής αναλαμβάνει να εκπονήσει μια εκτενή εργασία που περιλαμβάνει επισκόπηση και αξιολόγηση βιβλιογραφίας, συστηματική έρευνα γύρω από ένα αντικείμενο, μελέτη θεωρίας, ανάπτυξη συνοδευτικών προγραμμάτων, διεξαγωγή πειραμάτων και σύνδεσης με τη θεωρία, αξιολόγηση αποτελεσμάτων, ενίοτε συνεισφορά νέων θεωρητικών αποτελεσμάτων, και εντέλει, συγγραφή συνοδευτικού κειμένου και παρουσίασης των αποτελεσμάτων. Επιπλέον, η Π.Ε. αποτελεί βασικό στοιχείο αναφοράς για το βιογραφικό σας σημείωμα. Κατά συνέπεια, θα υπάρχει ενθάρρυνση να ετοιμάσετε μία αξιολογη Π.Ε. Επομένως, είναι απαραίτητο να σας ενδιαφέρει η περιοχή που θα επιλέξετε, ώστε να διερευνήσετε με ευχαρίστηση τις τρέχουσες εξελίξεις (state-of-the-art) στο αντίστοιχο θέμα. *Αξίζει να σημειώσετε ότι τα θέματα μπορούν να τροποποιηθούν ή να υπάρξει μεγαλύτερη εξειδίκευση εφόσον υπάρχει αμοιβαίο ενδιαφέρον.*

1 Μελέτη Μοντέλων Παράλληλης Εκτέλεσης για την Εκπαίδευση Μοντέλων Βαθιάς Μάθησης σε GPUs

Περιγραφή του Προβλήματος

Η εκπαίδευση μοντέλων βαθιάς μάθησης είναι υπολογιστικά απαιτητική και χρονοβόρα διαδικασία, ειδικά για μοντέλα, όπως τα συνελκτικά νευρωνικά δίκτυα (CNNs) και τα transformers. Καθώς η διαστατικότητα των δεδομένων και η πολυπλοκότητα των μοντέλων αυξάνει, ο χρόνος που απαιτείται για την εκπαίδευση των μοντέλων μετατρέπεται σε κώλυμα για την αξιοποίηση των αποτελεσμάτων. Η Π.Ε. στοχεύει να αντιμετωπίσει αυτό το πρόβλημα εξετάζοντας τεχνικές παράλληλου υπολογισμού για την εκπαίδευση μοντέλων βαθιάς μάθησης σε GPUs, όπως ο παραλληλισμός δεδομένων, ο παραλληλισμός μοντέλων και υβριδικές προσεγγίσεις. Ο στόχος είναι να αξιολογηθεί πως αυτές οι τεχνικές μπορούν να χρησιμοποιηθούν για τη βελτιστοποίηση του χρόνου εκπαίδευσης χωρίς να θυσιάζεται η ακρίβεια ή η σταθερότητα του μοντέλου.

Μεθοδολογία

Η εργασία θα περιλαμβάνει υλοποίηση και πειραματισμό με διάφορες τεχνικές παραλληλισμού, χρησιμοποιώντας δημοφιλή πλαίσια βαθιάς μάθησης, όπως το TensorFlow ή το PyTorch. Θα εξεταστεί ο παραλληλισμός δεδομένων (κατανομή του συνόλου δεδομένων σε πολλές GPUs), ο παραλληλισμός μοντέλων (κατανομή διαφορετικών μερών του μοντέλου σε πολλές GPUs) και υβριδικές προσεγγίσεις που συνδυάζουν και τις δύο μεθόδους. Αυτές οι τεχνικές θα δοκιμαστούν σε πρότυπα μοντέλα βαθιάς μάθησης και σύνολα δεδομένων αναφοράς, μετρώντας την απόδοση με βάση τον χρόνο εκπαίδευσης, την ακρίβεια του μοντέλου και τη χρήση των πόρων GPU.

Αναμενόμενα Αποτελέσματα

Αναμένεται η εργασία να προσδιορίσει την πιο αποτελεσματική τεχνική παραλληλισμού για συγκεκριμένα είδη μοντέλων βαθιάς μάθησης και σύνολα δεδομένων. Θα πρέπει να επιδειχθεί σημαντική μείωση του χρόνου εκπαίδευσης και αποτελεσματική χρήση των πόρων GPU, ενώ θα διατηρηθεί συγκρίσιμη ακρίβεια με την εκπαίδευση σε μια GPU. Επιπλέον, η εργασία μπορεί να αποκαλύψει πληροφορίες σχετικά με τα συμβιβαστικά στοιχεία μεταξύ των διαφορετικών μεθόδων παραλληλισμού, όπως το κόστος επικοινωνίας, τις απαιτήσεις μνήμης και τη δυνατότητα κλιμάκωσης του μοντέλου.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Τεχνητή Νοημοσύνη και Έμπειρα Συστήματα, Θέματα Επιστήμης Δεδομένων

2 Παράλληλοι Εξελικτικοί Αλγόριθμοι για Βελτιστοποίηση Υπερπαραμέτρων σε Μοντέλα Τεχνητής Νοημοσύνης

Περιγραφή του Προβλήματος

Η βελτιστοποίηση υπερπαραμέτρων είναι κρίσιμη για τη βελτίωση της απόδοσης των μοντέλων τεχνητής νοημοσύνης, αλλά είναι μια χρονοβόρα και υπολογιστικά απαιτητική διαδικασία, καθώς περιλαμβάνει την εξερεύνηση ενός μεγάλου χώρου πιθανών τιμών. Οι εξελικτικοί αλγόριθμοι προσφέρουν μια βιολογικά εμπνευσμένη προσέγγιση για την αναζήτηση βέλτιστων λύσεων, αλλά η εφαρμογή τους σε μεγάλα μοντέλα μπορεί να είναι αργή. Αυτή η εργασία προτείνει τη χρήση παράλληλων εξελικτικών αλγορίθμων για την επιτάχυνση της βελτιστοποίησης υπερπαραμέτρων σε μοντέλα τεχνητής νοημοσύνης, εξετάζοντας μεθόδους κατανομής του υπολογιστικού φόρτου σε πολλαπλές επεξεργαστικές μονάδες.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει την ανάπτυξη και υλοποίηση παραλλαγών εξελικτικών αλγορίθμων, όπως γενετικοί αλγόριθμοι και αλγόριθμοι σμήνους, σε παράλληλα περιβάλλοντα. Θα εξεταστούν διάφορες προσεγγίσεις παραλληλισμού, όπως η εκτέλεση πολλαπλών πληθυσμών ή η κατανομή της αξιολόγησης των υποψηφίων λύσεων σε πολλαπλές μονάδες επεξεργασίας (π.χ. GPUs). Οι προτεινόμενοι αλγόριθμοι θα εφαρμοστούν στη βελτιστοποίηση υπερπαραμέτρων μοντέλων βαθιάς μάθησης και θα αξιολογηθούν σε όρους χρόνου εκτέλεσης και ακρίβειας των μοντέλων, σε σύγκριση με σειριακές προσεγγίσεις.

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι οι παράλληλοι εξελικτικοί αλγόριθμοι θα επιτύχουν σημαντική μείωση του χρόνου βελτιστοποίησης υπερπαραμέτρων, χωρίς να μειώνεται η ποιότητα των λύσεων. Επίσης, αναμένεται να προκύψουν καλύτερα μοντέλα με υψηλότερη ακρίβεια λόγω της ικανότητας των εξελικτικών αλγορίθμων να εξερευνούν αποτελεσματικά μεγάλους χώρους υπερπαραμέτρων. Η σύγκριση μεταξύ των διαφόρων παραλλαγών παραλληλισμού θα προσδιορίσει τις πιο αποδοτικές τεχνικές για συγκεκριμένα προβλήματα.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Τεχνητή Νοημοσύνη και Έμπειρα Συστήματα, Θέματα Επιστήμης Δεδομένων

3 Παράλληλος Υπολογισμός στην Εκπαίδευση Μοντέλων Επεξεργασίας Φυσικής Γλώσσας

Περιγραφή του Προβλήματος

Η επεξεργασία φυσικής γλώσσας (Natural Language Processing - NLP) περιλαμβάνει την εκπαίδευση πολύπλοκων μοντέλων, όπως τα transformers, τα οποία απαιτούν μεγάλο πλήθος δεδομένων και υπολογιστικών πόρων. Η παραδοσιακή εκπαίδευση σε ένα μόνο επεξεργαστή (CPU) ή GPU δεν επαρκεί για την ταχεία εκπαίδευση τέτοιων μοντέλων, οδηγώντας σε εκτεταμένους χρόνους εκπαίδευσης. Η Π.Ε. στοχεύει να εξετάσει τεχνικές παραλληλισμού για την επιτάχυνση της εκπαίδευσης μοντέλων NLP, χρησιμοποιώντας παράλληλο υπολογισμό σε πολλαπλές GPUs, προκειμένου να μειωθεί ο χρόνος εκπαίδευσης χωρίς να μειωθεί η ακρίβεια του μοντέλου.

Προτεινόμενη Μεθοδολογία

Η εργασία θα περιλαμβάνει την εφαρμογή και αξιολόγηση τεχνικών παραλληλισμού, όπως παραλληλισμός δεδομένων και μοντέλων, για την εκπαίδευση NLP μοντέλων όπως το BERT και το GPT. Θα χρησιμοποιηθούν βιβλιοθήκες και εργαλεία όπως το PyTorch και το TensorFlow για την υλοποίηση παραλλαγών παραλληλισμού σε διαφορετικά υπολογιστικά περιβάλλοντα. Οι επιδόσεις των μοντέλων θα μετρηθούν σε σχέση με τον χρόνο εκπαίδευσης, τη χρήση πόρων, και την ακρίβεια του μοντέλου. Θα γίνει επίσης ανάλυση της απόδοσης σε διαφορετικά μεγέθη δεδομένων και διαφορετικές αρχιτεκτονικές.

Αναμενόμενα Αποτελέσματα

Αναμένεται η μελέτη να δείξει σημαντική μείωση του χρόνου εκπαίδευσης σε μεγάλα NLP μοντέλα μέσω της χρήσης παραλληλισμού, διατηρώντας ταυτόχρονα την ακρίβεια και την απόδοση του μοντέλου. Τα αποτελέσματα θα προσδιορίσουν ποιοι τύποι παραλληλισμού είναι πιο αποδοτικοί για συγκεκριμένα NLP προβλήματα και μεγέθη δεδομένων, συμβάλλοντας στην ανάπτυξη βελτιστοποιημένων διαδικασιών εκπαίδευσης για πραγματικές εφαρμογές NLP.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Τεχνητή Νοημοσύνη και Έμπειρα Συστήματα, Θέματα Επιστήμης Δεδομένων

4 Παράλληλη Εκτέλεση Πολλαπλών Προσομοιώσεων Δικτύων Βιολογικών Νευρώνων με Μετατοπισμένα Μητρώα Συνδεσιμότητας

Περιγραφή του Προβλήματος

Η προσομοίωση δικτύων βιολογικών νευρωνικών είναι μια εξαιρετικά υπολογιστικά απαιτητική διαδικασία, καθώς προσπαθεί να αναπαραστήσει τη λειτουργία του εγκεφάλου μέσω μοντέλων που προσεγγίζουν τη συμπεριφορά των νευρώνων. Η εκτέλεση πολλαπλών προσομοιώσεων είναι συχνά απαραίτητη για την ανάλυση παραμέτρων ή τη μελέτη διαφορετικών σεναρίων. Ωστόσο, ο υψηλός υπολογιστικός φόρτος απαιτεί αποδοτική παράλληλη εκτέλεση αυτών των προσομοιώσεων, ιδίως σε συστήματα HPC (High Performance Computing). Η Π.Ε. θα εξετάσει την αποδοτική εκτέλεση πολλαπλών προσομοιώσεων του ίδιου δικτύου νευρώνων με μετατοπισμένες (shifted) τιμές. Τα σεναρία αυτά αντιστοιχούν στο ίδιο μητρώο γειτνίασης (connectivity matrix), με τις μη μηδενικές τιμές μετατοπισμένες κάθε φορά κατά μια σταθερή τιμή.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει την αξιοποίηση των κοινών υπολογισμών μεταξύ των μετατοπισμένων μητρώων για την μείωση του απαιτούμενου πλήθους υπολογισμών, σε σύγκριση με την περίπτωση που τα σεναρία εκτελούνταν ξεχωριστά. Θα αξιοποιηθεί επίσης η υπολογιστική ισχύς πολλαπλών επεξεργαστών ή/και GPUs. Θα εκτελεστούν πειράματα με διαφορετικές τοπολογίες δικτύων και παραμέτρους, για να μελετηθούν οι επιδράσεις του πλήθους των κοινών υπολογισμών καθώς και του παραλληλισμού στον χρόνο εκτέλεσης. Το μοντέλο προσομοίωσης βιολογικών νευρώνων που θα χρησιμοποιηθεί είναι το Leaky Integrate-and-Fire (LIF).

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι η χρήση παράλληλων υπολογισμών θα επιταχύνει σημαντικά την εκτέλεση πολλαπλών προσομοιώσεων δικτύων βιολογικών νευρώνων, επιτρέποντας ταχύτερη ανάλυση μεγάλου αριθμού παραμέτρων και σεναρίων. Τα αποτελέσματα θα δείξουν πως οι παραλληλοποιημένοι αλγόριθμοι επηρεάζουν την απόδοση, καθώς και τον τρόπο με τον οποίο οι υπολογιστικοί πόροι μπορούν να αξιοποιηθούν αποδοτικά σε συστήματα HPC. Η εργασία θα συμβάλει στην κατανόηση της βελτιστοποίησης των προσομοιώσεων για τη μελέτη πολύπλοκων βιολογικών συστημάτων.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Εφαρμοσμένη Άλγεβρα

5 Παράλληλη Εκτέλεση Προσομοιώσεων Δικτύων Βιολογικών Νευρώνων σε Πολλαπλές GPUs

Περιγραφή του Προβλήματος

Η προσομοίωση δικτύων βιολογικών νευρώνων χρησιμοποιώντας το μοντέλο Leaky Integrate-and-Fire (LIF) είναι υπολογιστικά απαιτητική και απαιτεί σημαντικούς πόρους μνήμης και επεξεργαστικής ισχύος. Η χρήση πολλαπλών GPUs σε έναν υπολογιστικό κόμβο μπορεί να μειώσει τον χρόνο εκτέλεσης και να αυξήσει το μέγεθος των προσομοιούμενων δικτύων. Ωστόσο, απαιτείται η κατάλληλη κατανομή του υπολογιστικού φόρτου και η διαχείριση της επικοινωνίας μεταξύ των GPUs στον ίδιο κόμβο για τη μεγιστοποίηση της απόδοσης. Η Π.Ε. θα εξετάσει την παραλληλοποίηση του μοντέλου LIF σε υπολογιστικό σύστημα με πολλαπλές GPUs, με στόχο την επίτευξη βέλτιστης απόδοσης και χρήσης των διαθέσιμων πόρων.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα επικεντρωθεί στην παραλληλοποίηση του μοντέλου Leaky Integrate-and-Fire (LIF) με χρήση εργαλείων όπως η CUDA για την κατανομή του υπολογιστικού φόρτου μεταξύ των GPUs στον ίδιο κόμβο. Θα αναπτυχθούν τεχνικές κατανομής των νευρώνων και των συνάψεων σε διαφορετικές GPUs, με στόχο την αποδοτική διαχείριση της μνήμης και τη μείωση της επικοινωνίας μεταξύ των GPUs. Θα πραγματοποιηθούν πειράματα για τη μέτρηση της απόδοσης του συστήματος σε διαφορετικά σενάρια, εξετάζοντας την κλιμάκωση της απόδοσης καθώς αυξάνεται ο αριθμός των νευρώνων που προσομοιώνονται.

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι η παράλληλη εκτέλεση του μοντέλου LIF σε υπολογιστικό σύστημα με πολλαπλές GPUs θα βελτιώσει σημαντικά την απόδοση, μειώνοντας τον χρόνο εκτέλεσης και επιτρέποντας την προσομοίωση μεγαλύτερων δικτύων βιολογικών νευρώνων. Η εργασία θα δείξει πως η αποδοτική κατανομή των πόρων μπορεί να βελτιστοποιήσει τις προσομοιώσεις σε συστήματα με πολλαπλές GPUs, επιτυγχάνοντας καλύτερη κλιμάκωση της απόδοσης.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Εφαρμοσμένη Άλγεβρα

6 Βελτιστοποίηση Παραλληλισμού Εργασιών σε Ετερογενή Υπολογιστικά Συστήματα με Χρήση Τεχνητής Νοημοσύνης

Περιγραφή του Προβλήματος

Σε ετερογενή υπολογιστικά συστήματα, τα οποία περιλαμβάνουν διαφορετικούς τύπους επεξεργαστικών μονάδων, όπως CPUs, GPUs, και FPGAs, η βέλτιστη κατανομή εργασιών παράλληλου προγράμματος είναι πολύπλοκη λόγω των διαφορετικών επιδόσεων και αρχιτεκτονικών των υποσυστημάτων. Η αναποτελεσματική κατανομή μπορεί να οδηγήσει σε καθυστερήσεις και χαμηλή απόδοση. Η Π.Ε. στοχεύει στη χρήση τεχνητής νοημοσύνης για τη βελτιστοποίηση του προγραμματισμού παράλληλων εργασιών σε τέτοια ετερογενή συστήματα, με σκοπό την εξισορρόπηση του φόρτου εργασίας και την αύξηση της συνολικής απόδοσης.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει την ανάπτυξη μοντέλων τεχνητής νοημοσύνης, όπως ενισχυτική μάθηση ή γενετικοί αλγόριθμοι, για τη βελτίωση του προγραμματισμού παράλληλων εργασιών σε ετερογενή συστήματα. Οι προτεινόμενοι αλγόριθμοι θα εκπαιδευτούν σε προσομοιωμένα και πραγματικά δεδομένα προγραμματισμού, λαμβάνοντας υπόψη τις διαφορετικές δυνατότητες των συστημάτων (π.χ. ταχύτητα επεξεργασίας, ενεργειακή κατανάλωση). Θα χρησιμοποιηθούν εργαλεία ανάλυσης για τη σύγκριση των AI-driven προσεγγίσεων με παραδοσιακές μεθόδους προγραμματισμού ως προς τον χρόνο εκτέλεσης, την ισορροπία φόρτου και την αποδοτικότητα των πόρων.

Αναμενόμενα Αποτελέσματα

Η εργασία αναμένεται να αναδείξει πως η τεχνητή νοημοσύνη μπορεί να επιτύχει πιο αποδοτική κατανομή παράλληλων εργασιών σε ετερογενή υπολογιστικά συστήματα, σε σύγκριση με παραδοσιακές τεχνικές. Αναμένεται να βελτιωθούν ο χρόνος εκτέλεσης και η χρήση πόρων, ενώ τα συστήματα θα γίνουν πιο αποδοτικά σε εφαρμογές υψηλής απόδοσης. Τα αποτελέσματα θα αποδείξουν τη δυνατότητα γενίκευσης των AI-driven τεχνικών σε διάφορες αρχιτεκτονικές και περιβάλλοντα.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Τεχνητή Νοημοσύνη και Έμπειρα Συστήματα, Θέματα Επιστήμης Δεδομένων

7 Μελέτη για την Ελαχιστοποίηση των Μη Μηδενικών Στοιχείων σε Δυνάμεις Αραιού Μητρώου

Περιγραφή του Προβλήματος

Η διαχείριση αραιών μητρών είναι κρίσιμη σε πολλές εφαρμογές υπολογιστικών συστημάτων, όπως η επεξεργασία γραφημάτων και η ανάλυση μεγάλων δεδομένων. Η διαδικασία υπολογισμού των δυνάμεων ενός αραιού μητρώου συχνά οδηγεί σε αύξηση των μη μηδενικών στοιχείων, γεγονός που ενδέχεται να προκαλέσει αυξημένες απαιτήσεις αποθήκευσης και υπολογιστικής ισχύος. Η Π.Ε. στοχεύει στη μελέτη και την ανάπτυξη μεθόδων για την ελαχιστοποίηση του αριθμού των μη μηδενικών στοιχείων στις δυνάμεις αραιών Μητρώων, εστιάζοντας σε τεχνικές όπως η αναδιάταξη, η αραιώση γραφημάτων, η διάσπαση σε block, κλπ.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει τη διεξαγωγή πειραμάτων με διάφορα αραιά μητρώα και την εφαρμογή τεχνικών βελτιστοποίησης για την ελαχιστοποίηση των μη μηδενικών στοιχείων στις δυνάμεις τους. Θα χρησιμοποιηθούν αλγόριθμοι αναδιάταξης, όπως οι Cuthill-McKee και Reverse Cuthill-McKee, καθώς και τεχνικές αραιώσης. Τα αποτελέσματα θα αξιολογηθούν με τη μέτρηση του αριθμού των μη μηδενικών στοιχείων πριν και μετά την εφαρμογή των μεθόδων, καθώς και την επίδραση στην υπολογιστική απόδοση και στην αποθήκευση.

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι οι προτεινόμενες μέθοδοι θα οδηγήσουν σε σημαντική μείωση του αριθμού των μη μηδενικών στοιχείων στις δυνάμεις αραιών μητρώων, χωρίς να επηρεάζεται αρνητικά η ακρίβεια των υπολογισμών. Η εργασία θα προσφέρει επίσης χρήσιμες κατευθύνσεις για την επιλογή κατάλληλων τεχνικών βελτιστοποίησης ανάλογα με την αρχιτεκτονική των μητρώων και τις εφαρμογές τους.

Συναφή Μαθήματα

Δομές Δεδομένων, Αλγόριθμοι, Εφαρμοσμένη Άλγεβρα

8 Παραλληλοποίηση Πολλαπλασιασμού Αραιών Μητρώων

Περιγραφή του Προβλήματος

Ο πολλαπλασιασμός αραιών μητρώων είναι ένας βασικός αλγόριθμος σε εφαρμογές τεχνητής νοημοσύνης, όπως τα νευρωνικά δίκτυα γράφων (Graph Neural Networks - GNNs) και η επεξεργασία μεγάλων δεδομένων. Ωστόσο, η παραλληλοποίηση αυτών των πράξεων σε σύγχρονα υπολογιστικά συστήματα, όπως GPUs και πολυπύρηνους επεξεργαστές, παρουσιάζει προκλήσεις, λόγω της άνισης κατανομής των μη μηδενικών στοιχείων στα αραιά μητρώα και της ανάγκης για αποτελεσματική εκμετάλλευση των υπολογιστικών πόρων. Η Π.Ε. στοχεύει στη μελέτη και βελτιστοποίηση του παραλληλισμού του πολλαπλασιασμού αραιών μητρώων, εστιάζοντας στη μείωση του χρόνου υπολογισμού και στην αποδοτικότερη χρήση της μνήμης.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει την υλοποίηση αλγορίθμων παραλληλοποίησης για τον πολλαπλασιασμό αραιών μητρώων, χρησιμοποιώντας προσεγγίσεις όπως η κατανομή των δεδομένων και η χρήση τεχνικών επεξεργασίας με βάση την αραιότητα. Η αξιολόγηση θα γίνει μέσω πειραμάτων σε πραγματικά δεδομένα και σύγκρισης με παραδοσιακές μεθόδους πολλαπλασιασμού μητρώων. Επιπλέον, θα εξεταστούν διαφορετικές μέθοδοι αποθήκευσης (Compressed Sparse Row, Compressed Sparse Column, κλπ), για τη βελτίωση της απόδοσης.

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι η προτεινόμενη εργασία θα προσφέρει βελτιωμένες λύσεις για τον πολλαπλασιασμό αραιών μητρώων, μειώνοντας σημαντικά τον χρόνο υπολογισμού και βελτιώνοντας την αποδοτικότητα της χρήσης μνήμης, ειδικά σε εφαρμογές τεχνητής νοημοσύνης που χειρίζονται μεγάλα και αραιά δεδομένα.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Εφαρμοσμένη Άλγεβρα

9 Ανάλυση Νέων Μετροπρογραμμάτων για Υπολογιστικά Συστήματα Υψηλής Απόδοσης (HPC)

Περιγραφή του Προβλήματος

Τα παραδοσιακά μετροπρογράμματα (benchmarks), όπως το LINPACK και το SPEC, χρησιμοποιούνται εδώ και χρόνια για την αξιολόγηση των συστημάτων υψηλής απόδοσης (HPC). Ωστόσο, με την εισαγωγή νέων τεχνολογιών όπως οι GPUs, οι επιταχυντές και τα ετερογενή συστήματα, αναδεικνύεται η ανάγκη για πιο αντιπροσωπευτικά μετροπρογράμματα που αντανακλούν τις σύγχρονες απαιτήσεις των HPC. Πρόσφατα έχουν προταθεί νέα μετροπρογράμματα όπως το HPCG (High Performance Conjugate Gradient), το AI Benchmark για HPC και το Graph500, που στοχεύουν στην καλύτερη αναπαράσταση πραγματικών φορτίων εργασίας και την αξιολόγηση της απόδοσης σε συνθήκες πιο κοντινές στις σύγχρονες απαιτήσεις, όπως η επεξεργασία μεγάλων δεδομένων και η εκπαίδευση μοντέλων τεχνητής νοημοσύνης. Η Π.Ε. θα αναλύσει αυτά τα νέα μετροπρογράμματα, εστιάζοντας στην καταλληλότητά τους για τις σύγχρονες αρχιτεκτονικές HPC.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει την ανάλυση και συγκριτική αξιολόγηση παραδοσιακών μετροπρογραμμάτων, όπως το LINPACK, σε σχέση με νέα εργαλεία όπως το HPCG, που αξιολογεί την αποδοτικότητα συστημάτων σε πραγματικές συνθήκες με λιγότερη έμφαση στην απόδοση κορυφής των CPUs, το AI Benchmark, που εστιάζει σε εφαρμογές τεχνητής νοημοσύνης, και το Graph500, που εστιάζει σε εφαρμογές που βασίζονται σε επεξεργασία γραφημάτων, οι οποίες είναι εξαιρετικά σημαντικές για τη μελέτη μεγάλων συνόλων δεδομένων. Μέσω πειραμάτων σε σύγχρονες HPC υποδομές, θα γίνει αξιολόγηση των νέων μετροπρογραμμάτων και σύγκριση των μετρήσεων που παράγουν σε σχέση με τα παραδοσιακά.

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι τα νέα μετροπρογράμματα θα προσφέρουν πιο ρεαλιστική αξιολόγηση της απόδοσης των HPC συστημάτων για σύγχρονες εφαρμογές, όπως η ανάλυση μεγάλων δεδομένων και η τεχνητή νοημοσύνη, σε αντίθεση με τα παραδοσιακά εργαλεία που επικεντρώνονται σε αλγόριθμους με μεγαλύτερη θεωρητική σημασία. Τα αποτελέσματα θα επισημάνουν την καταλληλότητα των νέων μετροπρογραμμάτων και θα προτείνουν βελτιώσεις για την αξιολόγηση των συστημάτων με τις νέες τεχνολογίες.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Τεχνητή Νοημοσύνη και Έμπειρα Συστήματα, Αρχιτεκτονική Υπολογιστών

10 Ανάλυση της Επίδρασης Μικτής Αριθμητικής σε Παράλληλους Αλγορίθμους Τεχνητής Νοημοσύνης

Περιγραφή του Προβλήματος

Η μικτή αριθμητική, δηλαδή η χρήση αριθμητικών υπολογισμών με διαφορετικά επίπεδα ακρίβειας (όπως διπλής, μονής ή ακόμα και μισής ακρίβειας πρόσφατα), μπορεί να βελτιώσει την ταχύτητα εκτέλεσης παραλληλοποιημένων αλγορίθμων Τεχνητής Νοημοσύνης (TN) χωρίς σημαντική απώλεια στην ακρίβεια των αποτελεσμάτων. Οι σύγχρονες εφαρμογές TN, όπως η εκπαίδευση και η βελτιστοποίηση νευρωνικών δικτύων, συχνά βασίζονται σε επαναληπτικούς αλγορίθμους, όπου η ακρίβεια μπορεί να θυσιαστεί μερικώς για την αύξηση της απόδοσης. Η Π.Ε. θα μελετήσει την επίδραση της μικτής αριθμητικής σε επαναληπτικούς αλγορίθμους που χρησιμοποιούνται στην τεχνητή νοημοσύνη, όπως οι Stochastic Gradient Descent (SGD), Adam, και k-Means Clustering. Η εργασία θα αξιολογήσει την απόδοση αυτών των αλγορίθμων όταν χρησιμοποιούνται διάφορες μορφές αριθμητικής ακρίβειας, με έμφαση στην ταχύτητα και την αριθμητική σταθερότητα.

Προτεινόμενη Μεθοδολογία

Η μεθοδολογία θα περιλαμβάνει την παραλληλοποίηση των αλγορίθμων Stochastic Gradient Descent (SGD), Adam και k-Means Clustering, με υλοποίηση διαφορετικών συνδυασμών αριθμητικής ακρίβειας. Τα πειράματα θα εκτελεστούν σε σύγχρονες υποδομές πολλαπλών πυρήνων ή GPUs για να εξεταστεί η απόδοση αυτών των αλγορίθμων σε διαφορετικά περιβάλλοντα TN, όπως η εκπαίδευση νευρωνικών δικτύων και η ανάλυση μεγάλων δεδομένων. Θα μετρηθεί η ταχύτητα εκτέλεσης, η χρήση πόρων (π.χ. μνήμης) και η ακρίβεια των αποτελεσμάτων για να εκτιμηθεί η επίδραση της μικτής αριθμητικής στην απόδοση των αλγορίθμων.

Αναμενόμενα Αποτελέσματα

Αναμένεται ότι η χρήση μικτής αριθμητικής θα οδηγήσει σε σημαντικές βελτιώσεις στην απόδοση των παραλληλοποιημένων αλγορίθμων, ειδικά σε επαναληπτικούς αλγόριθμους που χρησιμοποιούνται στην εκπαίδευση νευρωνικών δικτύων. Παράλληλα, η ανάλυση θα αποκαλύψει τα όρια της αποδεκτής μείωσης ακρίβειας ώστε να διασφαλιστεί η αριθμητική σταθερότητα και η ποιότητα των αποτελεσμάτων. Τα ευρήματα θα προσφέρουν χρήσιμες πληροφορίες για τη βελτιστοποίηση αλγορίθμων TN σε υπολογιστικά περιβάλλοντα υψηλής απόδοσης.

Συναφή Μαθήματα

Παράλληλος Υπολογισμός, Τεχνητή Νοημοσύνη και Έμπειρα Συστήματα, Θέματα Επιστήμης Δεδομένων

11 Άλλα θέματα

Υπάρχει η δυνατότητα ανάληψης Π.Ε. σε άλλα θέματα που θα προκύψουν μετά από συζήτηση με τον επιβλέποντα.