

0619 ინფორმაციისა და კომუნიკაციის ტექნოლოგიები
INFORMATION AND COMMUNICATION TECHNOLOGIES (ICTS)

მოთხოვნის დამუშავება გენეტიკური პროგრამირების მეშვეობით

ბადრი მეფარიშვილი

საქართველოს ტექნიკური უნივერსიტეტი

Georgian Technical University

E-mail: meparishvilibadri08@gtu.ge

გიორგი მურადოვი

საქართველოს ტექნიკური უნივერსიტეტი

Georgian Technical University

E-mail: muradovi.giorgi22@gtu.ge

რეზიუმე

მონაცემთა ბაზის მართვის სისტემებში მოთხოვნის დამუშავება არის კვლევისა და განვითარების კრიტიკული მიმართულება, რომელიც ფოკუსირებულია მოთხოვნის შესრულების ოპტიმიზაციაზე მასშტაბურობისა და სიზუსტის გათვალისწინებით. გენეტიკური პროგრამირება (GP) პერსპექტიული მიდგომაა რთული ოპტიმიზაციის პრობლემების გადასაჭრელად, მათ შორის მოთხოვნის დამუშავებისას. სტატიაში განხილულია გენეტიკური პროგრამირების გამოყენების შესაძლებლობები მოთხოვნის დამუშავებაში. წარმოდგენილია აღნიშნული მიდგომის პოტენციური სარგებელი, გამოწვევები და კვლევის მიმდინარე ტენდენციები. განხილულია GP-ის ძირითადი კონცეფციები, მისი ინტეგრაცია მოთხოვნის ოპტიმიზაციის ტრადიციულ მეთოდებთან და წარმოდგენილია აღნიშნული მიმართულებით ჩატარებული კვლევები, რომლებიც აჩვენებენ მის ეფექტურობას მოთხოვნის შესრულების გაუმჯობესებაში.

საკვანძო სიტყვები: მოთხოვნის ოპტიმიზაცია, გენეტიკური პროგრამირება (GP), ევოლუციური ალგორითმები.

მოთხოვნის დამუშავება მოიცავს საფეხურების სერიას მაღალი დონის მოთხოვნის ეფექტურად განხორციელების გეგმად გარდაქმნისთვის, რომელიც საჭირო მონაცემებს ამოიღებს მონაცემთა ბაზიდან. მოთხოვნის ოპტიმიზაციის ევრისტიკულ მეთოდში შეკითხვის გეგმის არჩევის გადაწყვეტილება მიიღება წინასწარ შემოთავაზებული წესების საფუძველზე. მნიშვნელოვანია, რომ ეს წესები ეფუძნება ზოგად იდეებს და ყოველთვის არ იძლევა ეფექტურ შედეგებს. მაგრამ ეს მიდგომა საკმაოდ მარტივია რეალიზების თვალსაზრისით და ხშირად გამოიყენება წამყვან მონაცემთა ბაზების მართვის სისტემებშიც კი. თუმცა, ეს მეთოდები შეიძლება შეზღუდული იყოს რთული მოთხოვნების დამუშავებისა და მონაცემთა სტრუქტურების სირთულის შემთხვევებში. გენეტიკური პროგრამირება გვთავაზობს ახალ მიდგომას ბუნებრივი გადარჩევისა და ევოლუციის პრინციპების გამოყენებით ოპტიმიზებული გადაწყვეტილებების გენერირებისთვის [1][4].

გენეტიკური პროგრამირება, როგორც ევოლუციური გამოთვლების მეთოდი, რომელიც შთაგონებულია ბუნებრივი გადარჩევის თეორიით, მოიცავს გადაწყვეტილებებისთვის პოპულაციის წარმოქმნას, რასაც მოჰყვება შერჩევის განმეორებითი პროცესები, კროსოვერი, მუტაცია და ფიტნეს შეფასება. მოთხოვნის დამუშავების კონტექსტში, GP მიზნად ისახავს მოთხოვნის შესრულების გეგმების განვითარებას, რომლებიც ოპტიმიზებულია კონკრეტული მეტრიკებისთვის, როგორიცაა შესრულების დრო, რესურსების გამოყენება ან მეხსიერების მოხმარება [1, 4].

კლასიკური გენეტიკური პროგრამირების (GP) ალგორითმი მსგავსია ბუნებრივი ევოლუციის პროცესის პრობლემის გადაწყვეტის ოპტიმიზაციისთვის. ის ავითარებს კომპიუტერული პროგრამების პოპულაციას ან ფუნქციებს გენეტიკური ოპერაციების მეშვეობით, როგორიცაა შერჩევა, თანაკვეთა, მუტაცია. ქვემოთ მოცემულია გენეტიკური პროგრამირების ტიპური ალგორითმის აღწერა [4, 5]:

ინიციალიზაცია

პოპულაციის გენერაცია: იქმნება შემთხვევითი ხეების საწყისი პოპულაცია (ხეზე დაფუძნებული GP-ის კონტექსტში), რომელიც წარმოადგენს პრობლემის შესაძლო გადაწყვეტას.

შეფასება

ფიტნეს შეფასება: პოპულაციის შეფასებისთვის გამოიყენება ფიტნეს ფუნქცია, რათა დადგინდეს მისი ხარისხი ან ვარგისიანობა საუკეთესო შედეგის მისაღებად. აღნიშნული ფუნქციის მეშვეობით შესაძლებელია ცალკეულ პოპულაციას ჰქონდეს შესაბამისი რეიტინგი.

შერჩევა

მშობლის შერჩევა: შეირჩევა ინდივიდები პოპულაციიდან, რომლებიც იქნებიან მშობლები მომავალი თაობისთვის. შესაძლებელია ზოგიერთი ინდივიდის პირდაპირი ასლის გადატანა მომავალ თაობაში მრავალფეროვნების უზრუნველსაყოფად.

კროსოვერი (რეკომბინაცია)

Crossover Operation: შემთხვევითი სახით დაწყვილებულ მშობლებში ხდება შეჯვარების ოპერაციის შესრულება, რაც წარმოქმნის ახალ შთამომავლებს. მსგავს რეკომბინაციას შემოაქვს მრავალფეროვნება და იძლევა ახალი გადაწყვეტილებების შესწავლის საშუალებას. ამავდროულად საჭიროა თანაკვეთის სიჩქარის კონტროლი, რათა შენარჩუნებულ იქნას ბალანსი მრავალფეროვნებასა და კონვერგენციას შორის.

მუტაცია

მუტაციის ოპერაცია: შთამომავლებში შემთხვევითი ცვლილებების განხორციელება, დამატებითი ცვალებადობის დასაწერად. ეს შეიძლება მოიცავდეს კვანძების შეცვლას, კავშირების შეცვლას ან პროგრამის სემენტების შეცვლას. საჭიროა დარეგულირდეს მუტაციის სიხშირე, რათა თავიდან აცილებულ იქნას ნაადრევი კონვერგენცია და უზრუნველყოფილ იქნას საკმარისი მრავალფეროვნება.

ახალი პოპულაციის მიღება

შთამომავლებისგან ახალი პოპულაცია ფორმირდება, რომლებიც წარმოიქმნება თანაკვეთისა და მუტაციის შედეგად, რითაც ახალი პოპულაცია ჩაანაცვლებს ძველს. აქაც საჭიროა პოპულაციის ზომის შენარჩუნება სტაბილური ევოლუციის უზრუნველსაყოფად.

შეწყვეტის კრიტერიუმი

შეწყვეტის კრიტერიუმად ჩაითვლება მდგომარეობა, როდესაც ახალ პოპულაციაში ცვლილებები აღარ ხდება ანუ შედეგი გაჯერდება რაიმე ნიშნულზე. ალგორითმი იძლევა საუკეთესო შედეგს საბოლოო პოპულაციიდან [4][5].

როგორც ვხედავთ, კლასიკური გენეტიკური პროგრამირების ალგორითმში, პოპულაცია ვითარდება განმეორებითი პროცესის მეშვეობით, რომელიც მოიცავს შერჩევას, თანაკვეთას, მუტაციას და შეფასებას. ეს პროცესი გრძელდება მანამ, სანამ შეწყვეტის პირობა არ დაკმაყოფილდება, რა დროსაც საუკეთესო გამოსავალი მიიღება საბოლოო პოპულაციისგან. თითოეული ნაბიჯი გადამწყვეტ როლს ასრულებს ალგორითმის ოპტიმალური ან დამაკმაყოფილებელი გადაწყვეტისკენ წარმართვაში. გენეტიკური პროგრამირების კლასიკური ალგორითმის ბლოკ-სქემა მოცემულია ნახ.1-ზე.

განვიხილოთ გენეტიკური პროგრამირების ძირითადი პროცესები მოთხოვნის დამუშავებაში

1. მოთხოვნის გეგმების კოდირება

პირველი ნაბიჯი GP-ის გამოყენებისას მოთხოვნის დამუშავებაში არის მოთხოვნის შესრულების გეგმების კოდირება გენეტიკური ოპერაციებისთვის შესაფერ ფორმატში. ეს ხშირად გულისხმობს შესრულების გეგმების ხეების ან სხვა იერარქიული სტრუქტურების სახით წარმოდგენას, სადაც კვანძები წარმოადგენს მოთხოვნის ოპერაციებს (მაგ., შეერთებას, ფილტრებს, პროგნოზებს) და წიბოები წარმოადგენს მონაცემთა ნაკადს [2][4].

2. ფიტნეს ფუნქციის განსაზღვრა

ფიტნეს ფუნქცია არის GP-ის მნიშვნელოვანი კომპონენტი, რომელიც განსაზღვრავს თითოეული წევრის გადაწყვეტის ხარისხს. მოთხოვნის დამუშავებისას, ფიტნეს ფუნქციამ შეიძლება გაითვალისწინოს ისეთი ფაქტორები, როგორიცაა შესრულების დრო, მეხსიერების გამოყენება ან I/O ოპერაციები. ეფექტური ფიტნეს ფუნქცია წარმართავს ევოლუციურ პროცესს ოპტიმალური მოთხოვნის გეგმებისკენ [1][2][4].

3. შერჩევა, კროსოვერი და მუტაცია

შერჩევა გულისხმობს საუკეთესო შედეგის მქონე წევრების არჩევას ახალი თაობის შესაქმნელად. კროსოვერი და მუტაცია მრავალფეროვნებას ნერგავს პოპულაციაში, რაც შესაძლებელს ხდის სხვადასხვა შესრულების გეგმების შესწავლას. კროსოვერი აერთიანებს ორი მშობლის გადაწყვეტის ნაწილებს

შთამომავლობის შესაქმნელად, ხოლო მუტაცია იწვევს შემთხვევით ცვლილებებს ცალკეულ წევრებში [1][2][4].

4. შეწყვეტის კრიტერიუმები

GP პროცესი გრძელდება მანამ, სანამ არ დაკმაყოფილდება შეწყვეტის კრიტერიუმი, როგორცაა თაობების მაქსიმალური რაოდენობა ან შესრულების სამიზნე დონე. პროცესის დროს ნაპოვნი საუკეთესო გამოსავალი, როგორც წესი, შერჩეულია, როგორც მოთხოვნის შესრულების ოპტიმიზებული გეგმა [2][4].

გენეტიკური პროგრამირების უპირატესობები და გამოწვევები მოთხოვნის დამუშავებაში

მოქნილობა: GP-ს შეუძლია შეისწავლოს მოთხოვნის შესრულების გეგმების ფართო სპექტრი, მათ შორის არატრადიციული გადაწყვეტილებები.

ადაპტაცია: GP-ს შეუძლია მოერგოს მონაცემთა ბაზის გარემოს და მოთხოვნის შაბლონებს.

ოპტიმიზაციის პოტენციალი: GP-მ აჩვენა შესაძლებლობა, რომ ვიპოვოთ უფრო ოპტიმიზებული შესრულების გეგმები ტრადიციულ მეთოდებთან შედარებით [2][4][5].

სირთულე: GP-ის გამოთვლითი ზედნადები შეიძლება იყოს მნიშვნელოვანი, რაც მოითხოვს პარამეტრების ფრთხილად რეგულირებას და ეფექტურ განხორციელებას.

კონვერგენცია: გლობალურ ოპტიმალთან დაახლოების უზრუნველყოფა შეიძლება იყოს რთული, ლოკალური რისკებით.

ინტეგრაცია ტრადიციულ სისტემებთან: GP-ის ჩართვა მონაცემთა ბაზის მართვის არსებულ სისტემებში მოითხოვს თავსებადობას და მთლიანობას [2][5].

ჩატარებული კვლევები და შედეგები

რამდენიმე კვლევამ აჩვენა GP-ის გამოყენება მოთხოვნის დამუშავებაში, რაც ხაზს უსვამს GP-ის მრავალმხრივობას და მის პოტენციალს მოთხოვნის ოპტიმიზაციის გასაუმჯობესებლად [2][4][6].

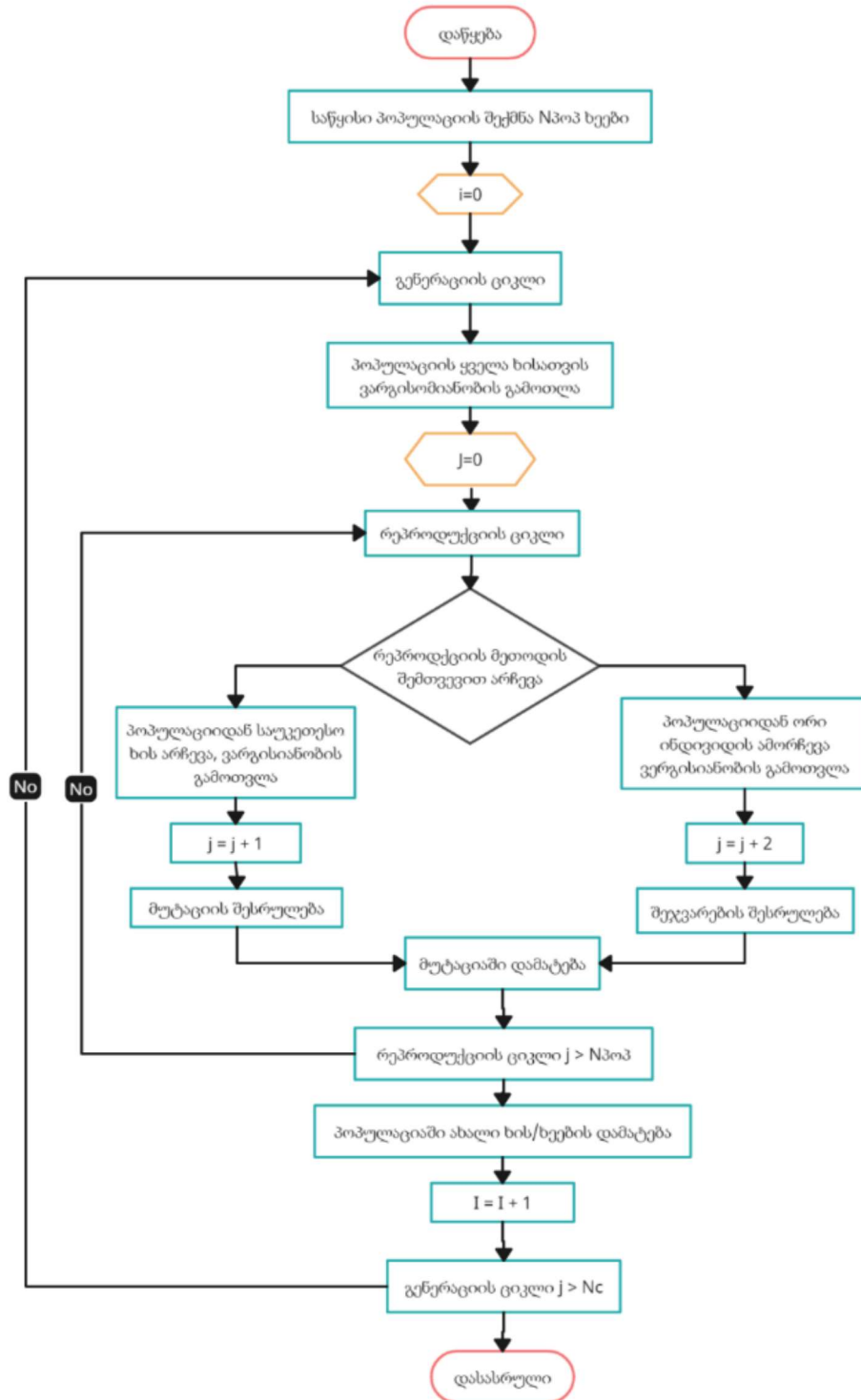
შეერთების შეკვეთების ოპტიმიზაცია

ამ კვლევაში GP გამოიყენებოდა კომპლექსურ SQL მოთხოვნებში შეერთების შეკვეთების ოპტიმიზაციისთვის. შედეგებმა აჩვენა, რომ GP-ს შეეძლო უფრო ეფექტური შეერთების ბრძანებების პოვნა, რაც გამოიწვევს შესრულების დროის შემცირებას ტრადიციულ მეთოდებთან შედარებით [4][2].

ადაპტური მოთხოვნის ოპტიმიზაცია

ამ კვლევამ შეისწავლა GP-ის გამოყენება ადაპტური მოთხოვნის ოპტიმიზაციისთვის, სადაც სისტემა დინამიურად არეგულირებდა მოთხოვნის შესრულების გეგმებს მონაცემთა განაწილების შეცვლაზე დაყრდნობით. აღნიშნულმა მიდგომამ განაპირობა გაუმჯობესებული შესრულება დროთა განმავლობაში [2][4].

ამდენად, გენეტიკური პროგრამირების გამოყენება მოთხოვნის დამუშავებაში არის პერსპექტიული მიდგომა, რომელსაც აქვს განუსაზღვრელი კვლევის შესაძლებლობები. მომავალი სამუშაო ფოკუსირებული იქნება GP-ის ეფექტურობის გაუმჯობესებაზე, ფიტნესის ახალი ფუნქციების შესწავლაზე და GP-ის სხვა ოპტიმიზაციის მეთოდებთან ინტეგრირებაზე.



სურათი 1: გენეტიკური პროგრამირების კლასიკური ალგორითმის ბლოკ-სქემა

ლიტერატურა

1. Goldberg, D. E. (1989). Genetic algorithms in search, optimization and machine learning.
2. Li, C., et al. (2019). Query Optimization Techniques in Big Data Analytics: A Survey.
3. Bellman, R. (1957). Dynamic programming.
4. Koza, J. R. (1992). Genetic Programming: On the Programming of Computers by Means of Natural Selection.
5. Beyer, H.-G., & Schwefel, H.-P. (2002). Evolution Strategies.
6. Fogel, L. J., et al. (1966). Evolutionary Programming.
7. Rumelhart, D. E., et al. (1986). Learning Representations by Back-Propagating Errors.

Request Processing By Genetic Programming Method

Badri Meparishvili, Giorgi Muradovi

Abstract

Query processing in database management systems is a critical area of research and development, focusing on optimizing query performance with scalability and accuracy in mind. Genetic programming (GP) is a promising approach for solving complex optimization problems, including query processing. This article discusses the possibilities of using genetic programming in query processing. It presents the potential benefits, challenges, and current research trends of this approach. The main concepts of GP, its integration with traditional methods of query optimization, and studies demonstrating its effectiveness in improving query performance are discussed.

Keywords: query optimization, genetic programming (GP), evolutionary algorithms.