

Quicksort

Vojtěch Bartl

1. listopadu 2016

1 Úvod

Quicksort (česky „rychlé řazení“) je jeden z nejrychlejších běžných algoritmů řazení založených na porovnávání prvků. Jeho průměrná časová složitost je pro algoritmy této skupiny nejlepší možná ($O(N \log N)$), v nejhorším případě (kterému se ale v praxi jde obvykle vyhnout) je však jeho časová náročnost $O(N^2)$. Další výhodou algoritmu je jeho jednoduchost.

Objevil jej Sir Charles Antony Richard Hoare v roce 1961.

2 Algoritmus

Základní myšlenkou quicksortu je rozdělení řazené posloupnosti čísel na dvě přibližně stejné části (quicksort patří mezi algoritmy typu rozděl a panuj). V jedné části jsou čísla větší a ve druhé menší, než nějaká zvolená hodnota (nazývaná *pivot* – anglicky „střed otáčení“). Pokud je tato hodnota zvolena dobře, jsou obě části přibližně stejně velké. Pokud budou obě části samostatně seřazeny, je seřazené i celé pole. Obě části se pak rekurzivně řadí stejným postupem, což ale neznamená, že implementace musí taky použít rekurzi. Volba pivotu

2.1 Volba Pivotu

Největším problémem celého algoritmu je volba pivotu. Pokud se daří volit číslo blízké mediánu řazené části pole, je algoritmus skutečně velmi rychlý. V opačném případě se jeho doba běhu prodlužuje a v extrémním případě je časová složitost $O(N^2)$. Přirozenou metodou na získání pivotu se pak jeví volit za pivot medián. Hledání mediánu (a obecně k -tého prvku) v posloupnosti běží v lineárním čase vzhledem k počtu prvků, tím dostaneme složitost $O(N \log N)$ quicksortu v nejhorším případě. Nicméně tato implementace není příliš rychlá z důvodu vysokých konstant schovaných v O notaci. Proto existuje velké množství alternativních způsobů, které se snaží efektivně vybrat pivot co nejbližší mediánu. Zde je seznam některých metod:

- První prvek – popřípadě kterákoli jiná fixní pozice. (Fixní volba prvního prvku je velmi nevýhodná na částečně seřazených množinách.)

- Náhodný prvek – často používaná metoda. Průměr přes každá data je $O(N \log N)$, přičemž zde se průměr bere přes všechny možné volby pivotů (rozděleno rovnoměrně). Nejhorší případ zůstává $O(N^2)$, protože pro každá data může náhoda nebo Velmi Inteligentní Protivník vybírat soustavně nevhodného pivotu, např. druhé největší číslo. V praxi většinou není dostupný generátor skutečně náhodných čísel, proto se používá pseudonáhodný výběr.
- Metoda mediánu tří – případně pěti či jiného počtu prvků. Pomocí pseudonáhodného algoritmu (také se používají fixní pozice, typicky první, prostřední a poslední) se vybere několik prvků z množiny, ze kterých se vybere medián, a ten je použit jako pivot.

Pokud by bylo zaručeno, že pivotu volíme vždy z 98 % prvků uprostřed a ne z 1 % na některé straně, algoritmus by stále měl nejhorší asymptotickou složitost $O(N \log N)$, byť s poněkud větší konstantou v O -notaci.

Praktické zkušenosti a testy ukazují, že na pseudonáhodných nebo reálných datech je Quicksort nejrychlejší ze všech obecných řadicích algoritmů (tedy i rychlejší než Heapsort a Mergesort, které jsou formálně rychlejší). Rychlost Quicksortu však není zaručena pro všechny vstupy. Maximální časová náročnost $O(N^2)$ Quicksort diskvalifikují pro kritické aplikace.