

Szeregowanie zadań w sieciach zbierających dane

Popularnonaukowe streszczenie projektu

Zbieranie danych z odległych urzędów jest ważnym elementem wielu współczesnych aplikacji. Złożone obliczenia muszą być wykonywane w systemach rozproszonych, jak gridy czy chmury obliczeniowe, w celu zapewnienia wystarczającej mocy obliczeniowej. W niektórych przypadkach otrzymane wyniki mogą być przechowywane w rozproszonym systemie plików, jednak często należy je zebrać w określonej lokalizacji. Zbierające dane bezprzewodowe sieci czujnikowe są szeroko stosowane w celach militarnych, w monitorowaniu środowiska, automatyzacji w domach, monitorowaniu pacjentów i w wielu innych obszarach. Serwisy internetowe gromadzą informacje o użytkownikach, ich działaniach, zwyczajach i upodobaniach. W związku z tym zbieranie danych jest ważnym i ciekawym przedmiotem badań.

Zbierająca dane sieć o topologii *gwiazdy* składa się ze zbioru *węzłów z danymi*, które otrzymują pewne zbiory danych w drodze obliczeń lub monitorowania środowiska, oraz *stacji bazowej*, do której dane powinny zostać przesłane w celu dalszego przetworzenia. Sieć o topologii *drzewa* zawiera ponadto *węzły pośrednie* odpowiedzialne za przekazywanie danych między węzłami z danymi i stacją bazową. W niektórych przypadkach węzły pośrednie mogą wstępnie przetwarzać lub kompresować dane. Szeregowanie zadań w sieciach zbierających dane polega na określeniu sposobu przesyłania danych i prowadzenia obliczeń, biorąc pod uwagę rozmaite parametry węzłów sieci, w taki sposób, żeby cały ten proces był jak najwydajniejszy. Wykorzystanie dobrych algorytmów szeregowania sprawia, że można gromadzić dane w krótszym czasie, przy mniejszych kosztach i niższym zużyciu energii.

W tym projekcie przeanalizowane zostaną pewne zagadnienia szeregowania zadań w zbierających dane sieciach o topologii gwiazdy i drzewa. Głównym przedmiotem zainteresowania są następujące problemy:

- minimalizacja czasu zbierania danych w sieci z ograniczoną pamięcią stacji bazowej;
- minimalizacja czasu zbierania danych w sieci z terminami gotowości zbiorów danych;
- minimalizacja maksymalnego spóźnienia w sieci z pożądanymi terminami przetworzenia zbiorów danych.

Pierwszym celem badań jest określenie złożoności obliczeniowej rozważanych problemów. Stawiamy hipotezę, że są one NP-zupełne, co oznacza, że wydajne algorytmy konstruujące optymalne uszeregowania dla tych problemów prawdopodobnie nie istnieją. Planujemy jednak wskazać specjalne przypadki, w których problemy te można rozwiązać w czasie wielomianowym. Ponieważ dokładne algorytmy rozwiązujące problemy NP-zupełne są zbyt wolne, by wykorzystywać je w praktyce, kolejnym celem projektu jest skonstruowanie wydajnych algorytmów heurystycznych, tworzących dobre, choć nieoptymalne uszeregowania. Pierwszym etapem będzie implementacja kilku prostych algorytmów dla każdego problemu i przeanalizowanie ich za pomocą eksperymentów obliczeniowych dla różnych kombinacji wartości parametrów sieci. W ten sposób dowiemy się, które algorytmy i w jakich sytuacjach generują dobre rozwiązania oraz jaką strukturę mają dobre uszeregowania. Zgromadzona wiedza zostanie wykorzystana do zaprojektowania zaawansowanych heurystyk, które zostaną przeanalizowane eksperymentalnie i teoretycznie. Jeśli będzie to możliwe, zostaną podane gwarancje dotyczące jakości otrzymywanych rozwiązań. Następnie rozważane będą sieci o topologii drzewa. Ponieważ zawierają one dodatkowe węzły w porównaniu z sieciami o topologii gwiazdy, problemy szeregowania stają się w tym przypadku jeszcze trudniejsze do rozwiązania. Algorytmy zaprojektowane dla sieci gwiazdzystych zostaną uogólnione na przypadek topologii drzewiastej i wzbogacone o elementy zaprojektowane specjalnie dla tej topologii. Jakość otrzymywanych rozwiązań zostanie także w tym przypadku przeanalizowana teoretycznie i za pomocą eksperymentów obliczeniowych.