

Abstract

Tematem pracy jest badanie problemu znajdowania najbardziej wpływowych wierzchołków w sieciach społecznościowych bazując na modelu zaproponowanym w 2011 [1]. Artykuł ten wykorzystuje techniki data mining do znalezienia zbioru wierzchołków, który maksymalizuje tzw. spread informacji w sieci. W naszej pracy implementujemy algorytm z [1], noszący nazwę RankedReplace, i testujemy go w stworzonej przez nas aplikacji "Most Desired". Nasza aplikacja wizualizuje zadaną na wejściu sieć, ilustruje zbiór najbardziej wpływowych wierzchołków i pozwala na ich analizę.

Drugim głównym celem pracy jest ocena efektywności działania algorytmu RankedReplace w przypadku próby ochrony sieci przed spreadem. Dokładniej, w naszej aplikacji zaprojektowaliśmy doświadczenie, w którym "szczepimy" wybrany zbiór wierzchołków przeciw "chorobie" i obliczamy liczbę wierzchołków "zarażonych". W pierwszym kroku wybranym zbiorem wierzchołków jest zbiór wierzchołków najbardziej wpływowych. W drugim kroku, zbiór wierzchołków o największych stopniach wyjścia. Porównując oba kroki eksperymentu stwierdzamy, że wybór wierzchołków metodą RankedReplace nie powinien być używany w przypadku, gdy w grę wchodzi ochrona sieci.

Keywords: Data Mining, Network Mining, Rozprzestrzenianie Się Wirusa, Najbardziej Wpływowe Wierzchołki, Ochrona Sieci, Model Flow Authorities