

Inwazyjna analiza sieci społecznych

Sieć społeczna często oznacza strukturę społeczną między podmiotami, które są na ogół osoby lub poszczególne organizacje. Pokazuje zależności różnych rodzajów, począwszy od przypadkowej znajomości w ścisłym związku lub do obiektu (na przykład przepływ informacji, towarów, pieniędzy, sygnały, pośrednich w cyklu produkcyjnym, itp) między członkami społeczności.

Analiza sieci społecznych (SNA) koncentruje się na przewidywaniu własności sieci społecznej w przyszłości, co ma liczne zastosowania w badaniach naukowych (np. w socjologii) oraz przemyśle (np. marketing). Szczególne znaczenie mają tu sieci o strukturze dwu-modalnej, jak np. sieci klient-kupiony produkt. Z uwagi na ich potencjalne zastosowania, wielu badaczy proponowało modele wzrostu takich sieci, starając się włączyć możliwości modelowania wielu zjawisk, obserwowanych w rzeczywistych sieciach. W chwili obecnej najbardziej zaawansowanym modelem wydaje się być model CSUIM (Cold Start User-Item Model) zaproponowany przez dra Chojnackiego, uwzględniający efekt wpływu rekomendacji na rozwój sieci w warunkach tzw. zimnego startu (niewielkiej ilości produktów nabywanych przez poszczególnych klientów), przy czym rekomendacje są efektem analizy SNA. Stąd w niniejszej pracy określono go mianem „inwazyjnej analizy sieci”.

Z praktycznego punktu widzenia cenną własnością modelu byłoby możliwość odtworzenia jego parametrów z rzeczywistej obserwowanej sieci społecznej tak, aby móc przewidywać jej przyszłe własności. Niestety stopień złożoności modelu Chojnackiego powodował, iż dotychczas nie udało się uzyskać metody jego identyfikacji z sieci.

W niniejszej rozprawie podjęto wyzwanie identyfikacji modelu CSUI z sieci. Celem było w pierwszej kolejności określenie z wystarczającą precyzją parametrów wzrostu dwudzielnej sieci społecznej w tym modelu stosując mieszaninę symulacji, regresji i metod analitycznych, a następnie jakościowe uchwycenie faktu, że w pewnym momencie parametry modelu zostały zmienione. Osiągając te cele wykazano, że CSUIM jest przydatna do analizy ilościowej rzeczywistych jak i syntetycznych dwudzielnych grafów sieci społecznych.

Metoda badawcza polegała na symulacyjnym badaniu populacji sieci generowanych z tego samego modelu CSUIM i opracowaniu metody odzyskiwania modelu z sieci przez odwrócenie relacji między oryginalnymi parametrami modelu i obserwowalnymi metrykami sieci. Koncentrowano się na wykrywaniu parametrów z pojedynczej migawki sieci.

Rezultatem prac jest pogłębienie teoretycznego rozumienia modelu CSUI przez wykrycie matematycznych relacji między podstawowymi parametrami modelu CSUI. Ponadto opracowano analityczną metodę identyfikacji parametrów modelu CSUI oraz wskazano na ograniczenia takiego podejścia. Zbadano możliwości identyfikacji modelu metodami uczenia maszynowego uzyskując lepsze, ale nadal niesatysfakcjonujące wyniki identyfikacji. Wreszcie opracowano autorskie wieloetapowe podejście do odkrywania parametrów modelu z sieci stosując mieszaninę metod analitycznych, regresyjnych oraz symulacyjnych. To ostatnie podejście zbadano pod kątem możliwości wykrycia zmian parametrów w trakcie rozwoju sieci uzyskując satysfakcjonujące wyniki.

W pracy przedstawiono także wyniki badań eksperymentalnych potwierdzające skuteczność opracowanych metod identyfikacji. Wskazano na ich ograniczenia oraz kierunki dalszych badań.