



## روش کشف تقلب در پایگاه داده گرافی

الهام عبدالهی<sup>۱</sup>، فرشته آزادی پرند\*<sup>۱</sup>، علی اصغر صفائی<sup>۲</sup>

گروه رایانه، دانشکده علوم ریاضی رایانه، دانشگاه علامه طباطبائی<sup>۱</sup>

گروه انفورماتیک پزشکی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، تهران، ایران<sup>۲</sup>

### چکیده

امروزه با توجه به گسترش ارتباطات بایستی حجم عظیمی از اطلاعات تولید، مبادله و در نهایت ذخیره‌سازی شود. سیستم‌های پایگاه داده‌ی رابطه‌ای، قابلیت ذخیره‌سازی این حجم از اطلاعات را ندارند. مفهوم NoSQL برای مواجهه با این مشکل در سال ۱۹۹۸ توسط کارلو استروزی ظهور پیدا کرد. از انواع پایگاه داده‌های NoSQL می‌توان به پایگاه داده‌های سندگرا، کلید مقدار، ستون محور و در نهایت پایگاه‌های گرافی اشاره کرد. در این کار تمرکز ما بر روی پایگاه داده‌های گرافی خواهد بود. پرکاربردترین پایگاه داده‌های گرافی، پایگاه داده neo4j است. پایگاه داده مذکور توسط شرکت‌های معروفی از جمله hp , cisco , Microsoft , adobe و نظایر آن استفاده می‌شود. از مهم‌ترین کاربردهای پایگاه داده مذکور می‌توان به مساله کشف تقلب و کلاهبرداری، کار با اینترنت اشیا و سیستم‌های توصیه‌گر، اشاره کرد. در این کار به بررسی تقلب افزایش قیمت سهم در یکی از شرکت‌های وابسته به سازمان بورس پرداخته خواهد شد. روند کار به این شکل است که عوامل دخیل در افزایش قیمت سهم شناسایی شده و داده‌های مربوطه در پایگاه داده گرافی neo4j ذخیره شده است. در نهایت زبان پرس و جوی cypher جهت بازیابی زنجیره‌های تقلب به کار گرفته شده است.

کلمات کلیدی: پایگاه داده غیررابطه‌ای گرافی Neo4j، زبان پرس و جوی Cypher، کشف تقلب، NoSQL