

عنوان مقاله:

ارائه روشی جهت خوشه بندی داده ها مبتنی بر الگوریتم k-means با حفظ حریم خصوصی

محل انتشار:

بیست و هفتمین کنفرانس بین المللی کامپیوتر انجمن کامپیوتر ایران (سال: 1400)

تعداد صفحات اصل مقاله: 6

نویسندگان:

مینا معصومی - دانشجوی کارشناسی ارشد، گروه نرم افزار، دانشکده مهندسی کامپیوتر، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران

نگین دانشپور - دانشیار، گروه نرم افزار، دانشکده مهندسی کامپیوتر، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران

معصومه صفحانی - دانشیار، گروه نرم افزار، دانشکده مهندسی کامپیوتر، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران

خلاصه مقاله:

همزمان با رشد فضای ابری و کارایی استفاده از آن، برون سپاری داده ها برای انجام عملیات مختلف داده کاوی از جمله خوشه بندی، رشد سریعی داشته است. چالش اصلی در این گونه از سامانه ها، حفظ امنیت داده ها و حفظ دقت و کارایی الگوریتم های خوشه بندی است. در این راستا، الگوریتم های رمزنگاری و روش های مبتنی بر گمنام سازی داده ها ارائه شده اند که در هر کدام از آنها سعی شده است موازنه ای بین امنیت و دقت ایجاد شود. در این مقاله روشی برای خوشه بندی داده های توزیع شده با الگوریتم k-means و استفاده از رمزنگاری همومورفیک ارائه شده است که ضمن رمزنگاری داده های حساس، سرعت و هزینه ی محاسباتی آن را کاهش میدهد. دقت روش پیشنهادی برای خوشه بندی داده های غیر متمرکز روی مجموعه داده iris و seeds به ترتیب برابر %۸۹.۳۴ و %۸۸.۵۳ است. در حالتی که داده های این دو مجموعه در سرور متمرکز باشند و خوشه بندی انجام شود، دقت خوشه بندی به ترتیب برابر %۹۰.۳۰ و %۸۹.۰۰ است. نتایج حاصل نشان میدهد دقت روش پیشنهادی به حالتی که سرور خوشه بندی را به صورت متمرکز انجام میدهد، بسیار نزدیک است. همچنین زمان اجرای روش پیشنهادی چون حجم داده ی رمز شده کم میشود، روی مجموعه داده iris و seeds به ترتیب برابر %۷.۵ و %۱۲.۷ ثانیه است که نسبت به حالتی که مالکان داده، تمام داده های خود را رمز میکنند که برابر %۱۳.۵ و %۳۲.۷ ثانیه است کاهش قابل توجهی داشته است.

کلمات کلیدی:

حفظ حریم خصوصی در داده کاوی، داده کاوی، رمزنگاری همومورفیک، الگوریتم k-means

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1452924>

