

عنوان مقاله:

خوشه بندی داده های جریان داده دارای برچسب

محل انتشار:

مجله علوم رایانشی، دوره 3، شماره 3 (سال: 1397)

تعداد صفحات اصل مقاله: 14

نویسندگان:

زهرا بیات - دانشجوی کارشناسی ارشد دانشکده ریاضی، آمار و علوم کامپیوتر، پردیس علوم، دانشگاه تهران

هدیه ساجدی - استادیار دانشکده ریاضی، آمار و علوم کامپیوتر، پردیس علوم، دانشگاه تهران

خلاصه مقاله:

امروزه مسئله پیدا کردن الگوریتم های خوشه بندی نظارت شده برای داده های پویا و جریان داده اهمیت زیادی دارد. محققان سعی می کنند برای حل کردن این مسئله الگوریتم های جدید ارائه نمایند و یا الگوریتم های موجود را بهبود دهند. در میان این الگوریتم ها، روش SAIC برای خوشه بندی داده های پویا با خوشه های با اندازه و شکل دلخواه ارائه شده است. در این روش، تعداد خوشه ها به طور خودکار توسط الگوریتم مشخص می شود اما این الگوریتم قادر به تشخیص صحیح خوشه های مسائل تک دسته ای نیست. این امر موجب می شود که بعضا در خوشه بندی جریان داده اختلال ایجاد شود. در این مقاله ضمن توضیح علت ایجاد این مشکل، الگوریتم ISAIC برای بهبود الگوریتم SAIC پیشنهاد می شود. همچنین عملکرد الگوریتم ISAIC با الگوریتم SAIC روی چند مجموعه داده مورد مقایسه قرار گرفته و نتایج ارایه شده است. میزان بهبود دقت دسته بندی بر روی مجموعه داده های مورد آزمایش حداقل صفر و حداکثر حدود ۶۵٪ است.

کلمات کلیدی:

توده جریان داده، جریان داده خوشه بندی، خوشه بندی نظارت شده

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1900945>

