

عنوان مقاله:

خوشه بندی توزیع شده فازی متون مبتنی بر آنتولوژی

محل انتشار:

نهمین سمپوزیوم بین المللی پیشرفتهای علوم و تکنولوژی (سال: 1393)

تعداد صفحات اصل مقاله: 11

نویسندگان:

سیدمحمدجعفر راشدی - دانشجوی کارشناسی ارشد نرم افزار، دانشگاه آزاد اسلامی واحد علوم و تحقیقات قزوین

حسن رشیدی - هیات علمی گروه رایانه دانشگاه علامه طباطبائی

امید سجودی - مدیر گروه مهندسی نرم افزار دانشگاه آزاد اسلامی واحد علوم و تحقیقات قزوین

خلاصه مقاله:

سراسر وب از انواع تعداد زیادی اسناد متنی تشکیل یافته است قالب عمومی نگهداری داده ها در بانک های اطلاعاتی بصورت متن است. خوشه بندی اسناد تکنولوژی قدرتمندی است که به طور گسترده میتواند در ساماندهی داده ها در هسته های کوچکتر مدیریت پذیر استفاده گردد. برای حل مساله روشهای متعدد پیشنهادی از مشکلات متعددی چون ابعاد بالا، مترادف بودن، ابهام فقدان محتوای توصیفی رنج میبرند. با استفاده از یک دانش پس زمینه مانند شبکه معنایی کلمات فارسی(فارس نت) میتواند استاندارد الگوریتمهای خوشه بندی را افزایش داد. در این مقاله تلفیقی از فارس نت الگوریتم خوشه بندی کلمات تکرار شونده با نام PFMDC ارایه گردیده که به صورت رویکردی فازی موازی پیاده سازی شده است. نتایج بدست آمده نشان میدهد که این رویکرد بهبود قابل توجهی در کیفیت خوشه بندی متون بهبود حدود پنجاه درصدی در زمان اجرا داشته است.

کلمات کلیدی:

خوشه بندی فازی متون، خوشه بندی کلمات تکرار شونده، فارس نت، آنتولوژی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/841468>

