

عنوان مقاله:

معرفی الگوریتم زمانبندی MK-ESAMR برای تحلیل و بررسی داده های حجیم

محل انتشار:

نخستین کنفرانس بین المللی فناوری اطلاعات (سال: 1394)

تعداد صفحات اصل مقاله: 11

نویسندگان:

نوا اژدری

محمدرضا بابایی

پریوش شعبانی

خلاصه مقاله:

؛ MapReduce یک مدل برنامه نویسی است که برای حل مسایل محاسباتی و تحلیل داده ها در مقیاس وسیع و نیز به صورت توزیعی، مورد استفاده قرار می گیرد. ایده اصلی در این سیستم، استفاده از MapReduce به عنوان یک لایه ارتباطی و هماهنگ کننده فعالیت های بین چندین گره می باشد که بدین منظور از Hadoop که یک نمونه پیاده سازی شده معروف بر اساس روش MapReduce می باشد استفاده کرده است. از Hadoop برای تحلیل داده ها در مقیاس وسیع استفاده می شود. در شبیه سازی ها دیده شده که، کارایی Hadoop به واسطه در نظر گرفتن برخی فرضیات بسیار محدود شده. به عنوان مثال دیده شده که گره های خوشه ها به صورت همگن در نظر گرفته می شود و فرض می شود که کار ها به صورت خطی پردازش می شوند. چنین فرضیاتی در عمل کارایی مناسبی ندارند. در تحلیل داده های بزرگ الگوریتم های مختلفی نظیر SAMR، LATE و غیره وجود دارد، هر یک از این الگوریتم ها در عمل نقاط ضعفی دارند. از این رو در این مقاله الگوریتم جدیدی (MK-ESAMR) پیشنهاد می شود که تا حد زیادی این نقاط ضعف را برطرف کرده است. در ابداع الگوریتم پیشنهادی از عملکرد الگوریتم های SAMR به همراه K-Medoids الهام گرفته شده است و تا حدود زیادی کارایی الگوریتم K-Medoids به منظور استفاده در الگوریتم پیشنهادی بهبود داده شده است. نتایج شبیه سازی نشان می دهد که الگوریتم MK-ESAMR در میان الگوریتم های موجود کمترین میزان خطا و سرعت بالاتری در اجرای کارها داشته است و همچنین توانسته کارهای با مدت زمانی طولانی را به خوبی تشخیص و مسیر جایگزین مناسبی برای اجرای بهتر آن ها پیدا کند.

کلمات کلیدی:

الگوریتم خوشه بندی، تحلیل داده های حجیم MapReduce، Hadoop، Big Date

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/478128>

