

عنوان فارسی مقاله :

برچسب گذاری موازی داده های انبوه XML با مپ ردیوس

عنوان انگلیسی مقاله :

Parallel labeling of massive XML data with MapReduce



توجه !

این فایل تنها قسمتی از ترجمه میباشد.

برای تهیه مقاله ترجمه شده کامل با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی

مقاله، [اینجا](#) کلیک نمایید.

The key contributions of our approach are summarized as follows:

1. We provide an efficient method to parallelize conventional tree labeling algorithms with MapReduce. In other words, we provide an efficient way to tailor conventional tree labeling algorithms to fit the MapReduce programming model.
2. In our approach, elements are naturally grouped by tag name. In addition, elements with the same tag name are sorted in an ascending order of labels after the labeling process. This is just the same as the input type used in most of holistic twig join algorithms such as TwigStack [10]. Therefore, the results of our labeling algorithms are directly used for twig pattern joins with no more efforts.
3. In the MapReduce programming model, an input is partitioned into many equal-sized blocks with no semantic knowledge about its input data. This makes it difficult to label XML elements in parallel. We provide a method for keeping structural information of an XML document during the labeling process.



نکات کلیدی شیوه مطرح شده عبارتند از:

1. در اینجا یک روش کارآمد برای موازی سازی الگوریتم های برچسب گذاری درختی معمولی با مپ ردیوس مطرح می کنیم. به عبارت دیگر، راهی کارآمد برای تطبیق الگوریتم های برچسب گذاری درختی معمولی با مدل برنامه نویسی مپ ردیوس ارائه می کنیم .
2. در شیوه معرفی شده، عناصر به طور طبیعی برحسب نام برچسب (تگ) گروه بندی شده اند. به علاوه، عناصری با نام برچسب یکسان به ترتیب صعودی برچسب ها بعد از فرایند برچسب گذاری طبقه بندی شده اند. این امر درست شبیه به تیپ ورودی بکاررفته در اکثر الگوریتم های پیوند و اتصال توپیگ کل نگر نظیر توپیگ استک می باشد. بنابراین، از نتایج الگوریتم های برچسب گذاری مستقیماً برای اتصالات و پیوند الگوی توپیگ بدون تلاشهای بیشتر استفاده می گردد.
3. در مدل برنامه نویسی مپ ردیوس، یک ورودی به چندین بلوک هم اندازه بدون دانش معنایی راجع به داده های ورودی اش تقسیم می شود. این امر برچسب گذاری عناصر XML به صورت موازی را سخت و دشوار می باشد. در اینجا روشی برای حفظ اطلاعات ساختاری یک سند XML در طول فرایند برچسب گذاری ارائه می کنیم .

توجه !

این فایل تنها قسمتی از ترجمه می باشد.

برای تهیه مقاله ترجمه شده کامل با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

همچنین برای مشاهده سایر مقالات این رشته [اینجا](#) کلیک نمایید.