

عنوان مقاله:

بهبود کارایی اندیس گذاری باروز-ویلر برای همترازسازی توالی های خوانش کوتاه با انتخاب زیرمجموعه ای از ماتریس پسوند

محل انتشار:

هشتمین کنفرانس بین المللی فناوری اطلاعات و دانش (سال: 1395)

تعداد صفحات اصل مقاله: 7

نویسندگان:

ربابه شریفی - گروه مهندسی کامپیوتر، دانشکده ی مهندسی، دانشگاه گیلان

اسدالله شاه بهرامی - گروه مهندسی کامپیوتر، دانشکده ی مهندسی، دانشگاه گیلان

خلاصه مقاله:

همترازسازی توالی ها یکی از وظایف مهم در بیوانفورماتیک است. الگوریتم های همترازسازی توالی ها در دو دسته ی کلی مبتنی بر برنامه نویسی پویا و الگوریتم های ابتکاری قرار می گیرند. در الگوریتم های نوع دوم، اندیس گذاری ژنوم ها یک مرحله ی پیش نیاز مهم است. تبدیل باروز-ویلر یک روش اندیس گذاری پرکاربرد است که علاوه بر مصرف حافظه ی کم، ساختار مناسبی برای جستجوی سریع و دقیق در توالی ها فراهم می کند. این اندیس در سه مرحله ساخته می شود؛ ساختن ماتریس پسوند، مرتب سازی پسوندها و ساختن داده های کمکی مربوط به اندیس. بررسی ها نشان می دهد که مرحله ی مرتب سازی پسوندها دارای بیشترین زمان اجرا است بطوریکه برای یک توالی به طول 25600 نماد، بیش از 3 ساعت طول می کشد. در این مقاله یک روش برای بهبود زمان اندیس گذاری باروز-ویلر با استفاده از تغییری کوچک در مرتب سازی ماتریس پسوندها معرفی شده که بر اساس ویژگی های الگوریتم جستجوی دقیق عقبگرد پیشنهاد شده است. این الگوریتم جستجو یکی از الگوریتم های ابزار همترازسازی باروز-ویلر است که برای جستجوی توالی های خوانش کوتاه تولید شده توسط فناوری های تعیین توالی جدید (حداکثر 100 نماد)، در ژنوم ها به کار می رود. ایده ی اصلی، کاهش اندازه ی مسئله ی مرتب سازی با انتخاب پیشوندی از تمام سطرهای ماتریس پسوند بر اساس نیازهای الگوریتم جستجوی دقیق عقبگرد است؛ به طوریکه، در درستی الگوریتم جستجو تأثیر منفی نداشته باشد. نتایج حاصل از اجرای الگوریتم نشان می دهد که با انتخاب طول 100 برای پیشوندها، زمان اندیس گذاری یک توالی 25600 نمادی از حدود 3 / 4 ساعت به 3 / 5 دقیقه کاهش می یابد. با توجه به اینکه فناوری های تعیین توالی جدید، خوانش هایی با طول کوتاه تولید می کنند، می توان با انتخاب طول پیشوند متناسب با این فناوری ها، روش پیشنهادی را بدون از دست دادن درستی الگوریتم جستجو به کار برد.

کلمات کلیدی:

اندیس گذاری ژنوم، تبدیل باروز-ویلر، بهبود کارایی اندیس گذاری، جستجوی دقیق توالی، همترازسازی توالی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/548764>

