

طراحی و پیاده سازی الگوریتم ترازبندی چند گانه دنباله ها با استفاده از بلوم فیلتر ماتریسی بر روی GPU

سمیرا رنجبر تپه کلی^۱، سمیه رنجبر تپه کلی^{*}، محمود احمدی^۲

^۱ کارشناسی ارشد کامپیوتر، دانشگاه آزاد اسلامی کرمانشاه

^۲ کارشناسی ارشد کامپیوتر، دانشگاه آزاد اسلامی کرمانشاه

^۳ استادیار مهندسی کامپیوتر، دانشگاه رازی کرمانشاه

چکیده

تحقیق حاضر با عنوان «ترازبندی چند گانه توالی با استفاده از بلوم فیلتر ماتریسی» در سال ۱۳۹۷ انجام شده است. رویکردهای جدید برای ترازبندی چندگانه توالی ها، به طور گسترده ای رو به پیشرفت اند. ترازبندی های چندگانه برای چند صد یا حتی چندین هزار توالی، با ابزارهای معروف فعلی مانند ClustalW نیازمند زمان زیادی برای انجام محاسبات می باشد. در این پژوهش الگوریتمی ایجاد می شود که به صورت موازی ترازبندی چندگانه ی توالی ها را انجام می دهد و برای اینکار از سیستم خط لوله برای موازی سازی استفاده می کند. بررسی ها نشان می دهند که برای دیتاست بزرگ توالی پروتئین و DNA (که برای دانلود در دسترس عموم قرار دارد)، افزایش سرعت چشمگیری در مقایسه با ClustalW بدست می آید. این پژوهش به مطالعه ی بلوم فیلتر ها، مساله ی ترازبندی چند گانه توالی و طراحی الگوریتمی برای حل این مساله با استفاده از تکنولوژی GPU می پردازد. همچنین در این پژوهش به معرفی رابط های برنامه نویسی مرسوم مانند ClustalW و ClustalW-MPI پرداخته شده است. نتایج بدست آمده نشان می دهد که برای مرحله ۱، ۲ و ۳ افزایش سرعت به ترتیب ۳۶،۹۱، ۱۸،۷۴ و ۱۱،۲۷ نسبت به ClustalW-MPI می باشد.

کلمات کلیدی

ClustalW-MPI، ترازبندی چند گانه دنباله ها، GPU

۱- مقدمه

توالی دنباله ها یک فعالیت مهم در زمینه ی بیوانفورماتیک است که با استفاده از آن می توان رشته های DNA، RNA و پروتئین را از هم تفکیک کرد و به این ترتیب بیماری هایی که به صورت ژنتیکی منتقل می شوند، بهتر درک می شوند. ترازبندی چندگانه توالی، مبحثی است که برای تحلیل و توالی یابی پروتئین ها و ژن ها از کامپیوتر استفاده می شود، چرا که یافتن الگوهای مشابه و تطابق الگو نیاز به محاسبات بسیاری دارد. در تراز بندی چند گانه ی توالی، یک دنباله ی

* s.ranjbar1397@gmail.com