Comparison of Pattern Matching Methods

مقدمه

هدف این مطالعه اینه که ببینیم الگوریتم‌های Pattern Matching چجوری کار می‌کنن و عملکردشون رو توی سه حالت مختلف مقایسه کنیم. ما سه تا الگوریتم معروف رو برای این مقایسه انتخاب کردیم: الگوریتم KMP، Boyer-Moore و Rabin-Karp. این گزارش توضیح می‌ده که این الگوریتم‌ها چجوری کار می‌کنن، پیچیدگی زمانی‌شون چقدره و عملکردشون بر اساس نتایج آزمایش مقایسه شده.

توضیحات الگوریتم‌ها

**KMP**: الگوریتم KMP قبل از شروع جستجو، الگو رو پردازش می‌کنه و یه prefix table درست می‌کنه که باعث می‌شه از مقایسه‌های اضافی جلوگیری بشه. این روش از بررسی دوباره کاراکترهایی که از قبل می‌دونه تطابق دارن، صرف نظر می‌کنه.

**Boyer-Moore**: الگوریتم Boyer-Moore الگو رو از راست به چپ مقایسه می‌کنه و از دو تکنیک به اسم bad character و good suffix استفاده می‌کنه تا بخش‌های اضافی متن رو رد کنه و جهش‌های بزرگی داشته باشه.

**Rabin-Karp**: الگوریتم Rabin-Karp از هش کردن برای پیدا کردن الگو توی متن استفاده می‌کنه. مقدار هش الگو رو حساب می‌کنه و با هش زیررشته‌های متن مقایسه می‌کنه. اگه مقدارها یکی بودن، کاراکترها رو بررسی می‌کنه تا مطمئن بشه که برخورد تصادفی نبوده.

مقایسه عملکرد

الگوریتم‌ها توی سه سناریوی مختلف تست شدن و عملکردشون نسبت به متد جستجوی پیش‌فرض Java اندازه‌گیری شد. نتایج توی چارت‌های زیر اومده:

**Case 1**: فقط الگوریتم Boyer-Moore سریع‌تر از متد جستجوی Java بود، با 6.25 برابر سرعت بیشتر. در حالی که KMP و Rabin-Karp به ترتیب 4 و 6.6 برابر کندتر بودن.

**Case 2**: همه الگوریتم‌ها از متد Java بهتر عمل کردن و Boyer-Moore بهترین عملکرد رو با 333 برابر سرعت بیشتر داشت. KMP، 18برابر سریع‌تر و Rabin-Karp، 9.4 برابر سریع‌تر بودن.

**Case 3**: همه الگوریتم‌ها از متد Java بهتر عمل کردن و KMP بهترین عملکرد رو با 15.6 برابر سرعت بیشتر داشت. Boyer-Moore هم با 2.3 برابر سرعت بیشتر خوب عمل کرد و Rabin-Karp، 7.5 برابر سریع‌تر بود.

نتیجه

به طور کلی، نتایج آزمایش نشون می‌ده که الگوریتم‌های تخصصی Pattern Matching می‌تونن عملکرد بهتری نسبت به جستجوی پیش‌فرض Java داشته باشن، مخصوصاً توی شرایط خاص. این نتایج اهمیت انتخاب الگوریتم مناسب رو بر اساس ویژگی‌های داده‌های ورودی و کاربردهای مورد انتظار نشون می‌ده.