A NEW EFFICIENT METHOD FOR MINING FREQUENT ITEMSETS IN MARKET BASKET DATA ANALYSIS

S.M. Fakhrahmad  
mfakhrahmad@cse.shirazu.ac.ir

M.H. Sadredini  
sadredin@shirazu.ac.ir

M. Zolghadri Jahromi  
zjahromi@shirazu.ac.ir

Abstract: Discovery of hidden and valuable knowledge from large data warehouses is an important research area and has attracted the attention of many researchers in recent years. Most of Association Rule Mining (ARM) algorithms start by searching for frequent itemsets by scanning the whole database repeatedly and enumerating the occurrences of each candidate itemset. In data mining problems, the size of data is often too large to fit in main memory. However, in some cases such as records of sales of a large supermarket, the probability of a particular item to be present in a transaction is often very low. This is due to the fact that a large number of items are usually available for purchase and also the fact that a small set of items is purchased by a customer in a shopping. In this paper, we make use of these facts to propose an efficient method for mining frequent itemsets. In our approach, the database is scanned just once, and data is encoded into a compressed form and held in a proper data structure in main memory. In each iteration, the time required to measure the frequency of itemsets, is reduced further (i.e., enumerating n-dimensional candidate itemsets is much faster than (n-1)-dimensional itemsets). We evaluate the efficiency of our technique using both synthetic and real-life datasets and compare it with other ARM methods proposed in past research.

روشی کارا برای کاوش مجموعه اقلام پرکرار در تحلیل داده‌های
سید خرید
سیدمحمد فخراحمد، محمدهادی صدرالدینی و منصور ذوالقدیر جهرمی

چکیده: کشف الگوهای پنهان و ارزشمند از درون حجم ویژعی از داده‌های خام اخیراً توجه بسیاری از حرفه‌جویان را به خود جلب کرده‌است. اغلب روش‌های کاوش قوانین در مرحله اول کار خود کلی قائم پرکرار (ساده و ترکیبی) را از بین تأمین اقلام موجود در داده‌ها جستجو می‌کنند که این امر نیازمند به خواندن مکرر کل داده‌ها از دیسک است. در مسائل داده‌های واحد، حجم پایگاه داده‌های تراکنش معمولاً بالا می‌رود. زبان کامل بدین شکننده در حافظه اصلی نمی‌باشد. اما در برخی موارد مانند پایگاه داده‌های تحلیلی مربوط به سیدهای خرید یک فروشگاه، با توجه به تعداد سنین زیاد نظام ممکن (کل اجناس فروشگاه) و نیز محدودیت نسبی اندازه تراکنش‌ها (اقلام خریداری شده در هر سید), احتمال رخ دادن یک قلم باد (خریداری شدید که کلیه خاص، پایین است. در این مقاله از مکاله‌های از این ویژگی روشن کارا برای کاوش قائم پرکرار در مجموعه داده‌های از این قبیل این دیسک ارزش می‌دهیم. در روش پیشنهادی، داده‌ها تکی بار از دیسک خوانده می‌شوند و بعد از آن به یک ساختار رمز شده و خلاصه تبدیل می‌گردد. پیش‌تریکه اول‌اً قابل نگهداری در حافظه اصلی می‌باشد و تابی به توجه به ساختار خاصی که

تاریخ وصول: 1098/12/20
تاریخ تصویب: 1097/3/26

mfakhrahmad@cse.shirazu.ac.ir
sadredin@shirazu.ac.ir
zjahromi@shirazu.ac.ir
1. مقدمه

کاوش فواید تداومی یک زمینه تحقیقاتی در محیط داده‌کاوی و در راستای کشف ارتباطات جالب و با همبستگی اقلام بزرگ و یا بایگان داده‌های تراکنشی می‌باشد. ای در اولین راه‌حل آن را با استفاده از دو گروه داده‌های ساختمانی و واقعی از بایگان کردن و بی‌بندی کردن داخل آن را تاکنون ارائه داده‌های مقایسه‌ی کمیم.

2. تحقیقات مربوط

الگوریتم‌های زیادی برای کشف فواید تداومی تاکنون ارائه‌شده‌اند. بعضی از این الگوریتم‌ها مربوط به جمله: Apriori، روش پایه‌ای و معروف Fast FPM است.

3. نتایج

در پایان نتیجه‌گیری می‌کنیم باید فواید قابل قبولی از ارائه داده‌های تراکنشی می‌باشد.

4. کلید واژه‌ها

داده‌کاوی، فواید تداومی، اقلام بزرگ، تراکنش، تحلیل سبد خرید.
Milk, Butter
Milk, Bread, Butter
Bread, Beer
Milk Bread Butter Beer
1 1 1 0
1 0 1 0
1 1 0 1
0 1 0 1

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)

\(\text{FastFPM} \quad \text{FPM} \quad \text{FP-tree} \quad \text{FP-growth} \quad \text{FP-mine} \quad \text{FP-growth} \quad \text{FP-growth}

\(\text{VPER} \quad \text{ARMOR} \quad \text{Oracle} \quad \text{VIPER} \quad \text{APriori} \quad \text{FP-Growth} \quad \text{Closet} \quad \text{ARMOR}\)
رابطه شكل 1 (الف) کمتری می‌شود. شرط اصلی برای کارایی الگوریتم پیشنهادی و جدید آن خصوصاً در مجموعه داده‌است.

گام نخست از الگوریتم، داده‌ها را با روش ایجاد ایفی چندین نظر کرده‌اند. با استفاده از تعدادی مکالمه، که هر یک از این مکالمات که هر یک از این مکالمات در مورد مقدار مناسب برای این یکی از مکالمات آن باشد، به بررسی آن مکالمه یکی از مکالمات مورد استفاده از الگوریتم پیشنهادی کل مکالمه و داده‌های خود شامل برای هرکدام این پیشنهادی شد.

رابطه شکل 2 مجموعه داده یک مشتمل بر 34 تراکنش.

رابطه قطعه‌نامی شده فقط یک بار با پیشنهادات پیشنهادی کلیه اقلام تکی (1 بعدی) محسوب می‌گردد تا اقلام تکی پیرترک در ضریب این عمل مقدار دهنده غیر از صفر (اعداد سنجش محلی به‌طور خلاصه) از تماشای قطعه‌ها استخراج می‌شوند. پایانه‌می‌شود از رابطه در صورتی که قلم داده مداوم که پیرترک باشد، یک جدول ستونی مقداری نشان می‌دهند، داده در جدول، داده‌های لازم ایجاد می‌کنند. مقداری ذکرشده در این جدول، همان‌طور که اعداد صحیح غیر صفر هستند فلزی که در قطعه‌های مختلف استخراج خواهد. یکی

دسترسی مستقیم به هرکدام از مقدار همان اعداد صحیح به عنوان (الف) شکل 2 مجموعه داده یک مشتمل بر 34 تراکنش.

رابطه قطعه‌نامی شده فقط یک بار با پیشنهادات پیشنهادی کلیه اقلام تکی (1 بعدی) محسوب می‌گردد تا اقلام تکی پیرترک در ضریب این عمل مقدار دهنده غیر از صفر (اعداد سنجش محلی به‌طور خلاصه) از تماشای قطعه‌ها استخراج می‌شوند. پایانه‌می‌شود از رابطه در صورتی که قلم داده مداوم که پیرترک باشد، یک جدول ستونی مقداری نشان می‌دهند، داده در جدول، داده‌های لازم ایجاد می‌کنند. مقداری ذکرشده در این جدول، همان‌طور که اعداد صحیح غیر صفر هستند فلزی که در قطعه‌های مختلف استخراج خواهد. یکی

## تراکنش

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### شکل 2

4. جدول‌های درهم زمان مربوط به (الف) (ب) A, B

جهت محاسبه پیشنهادات یک ترکیب از اقلام مانند AB دیگر نیازی به پیشنهاد کل داده‌ها نیست و این کار با استفاده از جدول‌های در هم زمان با اقلام و باین (ب) و (ب) انجام می‌شود. برای یک انتظار مربوط به اقلام در کمترین (شامل عناصر کمتر) شروع کنید. بازه کلی دسترسی موقتی از این جدول در صورتی که کلیه در جدول دوم می‌تواند باشد، مقدار صحیح مربوط به آن AND از دو جدول در هم زمان با این می‌کنند. بنابراین عمل AND می‌تواند به عنوان دسترسی دیگری که می‌کنند در موارد 2 نشان دهنده عمل و تعداد رخداده‌های همان‌نام عناصر در B و A می‌توانند در قطعه مربوطه است.

### محبوب‌ترین نشان‌دهنده این است که در AND صفر بودن حاصل از

## Hash Table

1.majid hash
روش‌کا برای کاوش مجموعه اقلام پرترکر در تحلیل داده‌های سبد خرید

3.1- تحلیل داده‌های سبد خرید

اگرچه کاوش‌های اقلام پرترکر در داده‌های سبد خرید می‌تواند به روش‌های مختلفی انجام شود، اما یکی از روش‌های محبوبی که برای این‌که کاربرد دارد، تحلیل‌های اقلام پرترکر است. این روش بر اساس تعداد بارهای که یک سبد خرید را در یک مدت زمانی به طور متوسط خرید کرده‌است، به طور تصادفی گزارش می‌شود. در این روش، سبد خریدی که به عنوان یک سبد خرید می‌باشد، می‌تواند از مجموعه‌ای از اقلام پرترکر در داده‌های سبد خرید مشخص شود. این روش بر اساس تعداد بارهای که یک سبد خرید را در یک مدت زمانی به طور متوسط خرید کرده‌است، به طور تصادفی گزارش می‌شود. در این روش، سبد خریدی که به عنوان یک سبد خرید می‌باشد، می‌تواند از مجموعه‌ای از اقلام پرترکر در داده‌های سبد خرید مشخص شود.
مقدار مناسب \( k \) می‌باشد. نتایج آزمایش‌هایی که بر روی مجموعه داده‌های واقعی عمل شده‌اند، نشان می‌دهند، برخی از این معیارهای ارزیابی از روی داده‌ها مناسب‌ترین مقدار مناسب می‌باشد.

\[ k \text{ مقدار مناسب} \]

در دقت‌آمیزی جداگانه، منظور می‌شود که QP در این محور باید جایگاه داشته باشد. در این محور، مشخص کردن داده‌های مختلفی از دقت‌آمیزی جداگانه می‌باشد. این نتایج نشان می‌دهد، که QP در این محور باید جایگاه داشته باشد.

\[ k \text{ مقدار مناسب} \]

نتایج آزمایش‌هایی که بر روی مجموعه داده‌های واقعی عمل شده‌اند، نشان می‌دهند، برخی از این معیارهای ارزیابی از روی داده‌ها مناسب‌ترین مقدار مناسب می‌باشد.

\[ k \text{ مقدار مناسب} \]

در دقت‌آمیزی جداگانه، منظور می‌شود که QP در این محور باید جایگاه داشته باشد. در این محور، مشخص کردن داده‌های مختلفی از دقت‌آمیزی جداگانه می‌باشد. این نتایج نشان می‌دهد، که QP در این محور باید جایگاه داشته باشد.

\[ k \text{ مقدار مناسب} \]

نتایج آزمایش‌هایی که بر روی مجموعه داده‌های واقعی عمل شده‌اند، نشان می‌دهند، برخی از این معیارهای ارزیابی از روی داده‌ها مناسب‌ترین مقدار مناسب می‌باشد.

\[ k \text{ مقدار مناسب} \]
4. نتایج عملی مقایسه الگوریتم‌ها

برای ارزیابی الگوریتم پیشنهادی و مقایسه آن با الگوریتم‌های موجود در مدل آزمایش ترتیب دادیم. در حالی که، آزمایش بر روی داده‌های ساختگی و در حالت دوم روش‌های داده‌های واقعی انجام شد. هم‌اکنون که ما دانیم الگوریتم‌های کاوش قوانین تداعی عموماً روال مشخصی را نشان می‌دهند و در شرایط مشابه خروجی‌های رئوسکی دارند.

بنابراین، معیار مهم در مقایسه روش‌های مختلف کارایی آنها از لحاظ زمان پاسخ و حافظه مورد نظر ما پایین‌ترین (k = 32) باشد. در این آزمایش، کارایی الگوریتم پیشنهادی (با در نظر گرفتن C++ ) را با الگوریتم‌های ARMOR، Apriori و FP_Growth مقایسه کردیم. الگوریتم معروف دیگری به نام FastFPM نیز وجود دارد که بدلیل اینکه برای مجموعه داده‌های خیلی حجم بزرگ از آن استفاده نمی‌شود، ما آن را نمونه‌برداری نمودیم. از مقایسه این مقدمات با انتخاب مجموعه داده اولیه (تریپلاس) 100 میلیون تا 10 میلیارد مورد بررسی گردید. این مقایسه از وضوح مدل‌های ISC و X-MAKER در مقایسه با FastFPM الگوریتم آن سنجش صحت بود. نتایج در مقایسه با ARMOR نسبی، در رابطه با الگوریتمی مطلبی که که برای مقایسه VIPER و يک برتری FastFPM و یک برتری Apriori از الگوریتم در مقایسه با FastFPM موجود دارد.

شکل 5. کارایی روشهای مختلف بر روی داده‌های تصادفی با اختلاف‌های مختلف برای بیت 1 در داده‌ها

الف) C++/0/0/1 (ب) 0/0/1 (ت) 0/0/5/0/1 (ش)
مقدمه

ازمیشی دوم ما بر روی چند مجموعه داده واقعی به نام‌های BMS-WebView-1 و BMS-WebView-2 کار می‌کنیم. مجموعه‌هایی که به دریاچه واقعی برای اولین بار از داده‌های اینترانس را شکل می‌دهند. هدف از جمع آوری این داده‌ها یافتن ارتباطات بین اقلام خریدی در تأثیر احتمال حفظ مدل آزمایشی در اجرا. 

جدول 1. مقادیر حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

زمان‌سنجی گرفته و میزان اثر را در ارزیابی کارایی الگوریتم‌ها اعمال کردیم. زمان اجرای الگوریتم‌های مختلف برای کشف مجموعه اقلام پرتکردار و پرتکردار بروز می‌باشد. این مقدار مناسب‌ترین توجه زمانی که به مجموعه‌های فکری می‌باشد. در نتیجه این، روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.

همچنین، در مقدمه حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.

ازمیشی دوم ما بر روی چند مجموعه داده واقعی به نام‌های BMS-WebView-1 و BMS-WebView-2 کار می‌کنیم. مجموعه‌هایی که به دریاچه واقعی برای اولین بار از داده‌های اینترانس را شکل می‌دهند. هدف از جمع آوری این داده‌ها یافتن ارتباطات بین اقلام خریدی در تأثیر احتمال حفظ مدل آزمایشی در اجرا. 

جدول 1. مقادیر حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.

ازمیشی دوم ما بر روی چند مجموعه داده واقعی به نام‌های BMS-WebView-1 و BMS-WebView-2 کار می‌کنیم. مجموعه‌هایی که به دریاچه واقعی برای اولین بار از داده‌های اینترانس را شکل می‌دهند. هدف از جمع آوری این داده‌ها یافتن ارتباطات بین اقلام خریدی در تأثیر احتمال حفظ مدل آزمایشی در اجرا. 

جدول 1. مقادیر حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.

ازمیشی دوم ما بر روی چند مجموعه داده واقعی به نام‌های BMS-WebView-1 و BMS-WebView-2 کار می‌کنیم. مجموعه‌هایی که به دریاچه واقعی برای اولین بار از داده‌های اینترانس را شکل می‌دهند. هدف از جمع آوری این داده‌ها یافتن ارتباطات بین اقلام خریدی در تأثیر احتمال حفظ مدل آزمایشی در اجرا. 

جدول 1. مقادیر حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.

ازمیشی دوم ما بر روی چند مجموعه داده واقعی به نام‌های BMS-WebView-1 و BMS-WebView-2 کار می‌کنیم. مجموعه‌هایی که به دریاچه واقعی برای اولین بار از داده‌های اینترانس را شکل می‌دهند. هدف از جمع آوری این داده‌ها یافتن ارتباطات بین اقلام خریدی در تأثیر احتمال حفظ مدل آزمایشی در اجرا. 

جدول 1. مقادیر حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.

ازمیشی دوم ما بر روی چند مجموعه داده واقعی به نام‌های BMS-WebView-1 و BMS-WebView-2 کار می‌کنیم. مجموعه‌هایی که به دریاچه واقعی برای اولین بار از داده‌های اینترانس را شکل می‌دهند. هدف از جمع آوری این داده‌ها یافتن ارتباطات بین اقلام خریدی در تأثیر احتمال حفظ مدل آزمایشی در اجرا. 

جدول 1. مقادیر حاویت مصرفی در مراحل مختلف آزمایش 1

<table>
<thead>
<tr>
<th>حاویت مصرفی (مگابایت)</th>
<th>0.5</th>
<th>0.1</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>احتمال بیت 1 در داده‌ها</td>
<td>2/3</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

روش‌های جدیدی شامل در سه مرحله فرزشی که در انتهای آن خودرسیده می‌باشد. به‌طور کلی، ترتیب عملیات محاسبه پیشینستی اقلام نیز از نظر مقداری می‌باشد. تجربه نشان داده شده در جدول 1 را می‌توان کواهی این ادعا دانست.
مراجع


