



## الگوریتم کاوش قوانین انجمنی مکانی تعمیم یافته

در پایگاه داده ی تصاویر

سهیلا صفری، کیوان برنا

۱- دانشجوی کارشناسی ارشد دانشگاه آزاد اسلامی واحد قزوین

۲- استادیار دانشکده علوم ریاضی و کامپیوتر، دانشگاه خوارزمی، تهران

نویسنده رابط: (نویسنده دوم) [borna@khu.ac.ir](mailto:borna@khu.ac.ir)

نام ارائه دهنده: (نویسنده اول) سهیلا صفری

### خلاصه

در این مقاله ما یک الگوریتم کشف قوانین انجمنی مکانی جدید به نام 17DLT را پیشنهاد می کنیم که این الگوریتم منجر به کاهش قوانین انجمنی مکانی کشف شده غیر مفید در پایگاه های داده مربوط به تصاویر می گردد. در طول اجرای این الگوریتم ابتدا به کشف الگوهای یک تکراری می پردازیم و بعد با استفاده از الگوهای کاندید چند تایی الگوهای پرتکرار بعدی را بدست می آوریم. همچنین نسبت به الگوریتم 9DLT که در سال ۲۰۰۷ توسط Anthony J.T. Lee و همکارانش مطرح شد تعداد تراکشن های کمتر و زمان اجرای کوتاهتری خواهد داشت.

کلمات کلیدی: داده کاوی مکانی، قوانین انجمنی مکانی، رشته 17DLT، پایگاه داده ی عکس.

### ۱- مقدمه

در اوایل دهه ۱۹۹۰ پیشرفت تکنولوژی اطلاعات و به خصوص شبکه هایی مانند اینترنت منجر به سهولت در ثبت داده های بدست آمده سازمان ها شد. بعد از آن حجم زیادی از داده ها اعم از عکس ها و تصاویر مختلف جمع آوری شده و در پایگاه های داده ذخیره گردید. با توجه به اینکه اطلاعات مهم و ارزشمندی در میان این حجم انبوه داده های نهفته وجود دارد سازمان ها به ابزاری برای کشف این الگوها نیاز دارند. بنابراین ابزارهای داده کاوی که با تنوع زیاد در تکنیک های تحلیلی که اغلب از روابط کلاسیک آماری به دست آمده اند توسعه پیدا کردند. تکنیک کاوش قوانین انجمنی یکی از همین تکنیک های مهم و کاربردی است. [1] در این میان حجم بسیار بالای داده های مکانی، پیچیدگی انواع داده های مکانی و روشهای دسترسی به آنها، کارایی را به یک چالش تحقیقاتی مهم در زمینه الگوریتمهای داده کاوی مکانی مبدل می سازد. که در این مقاله ما بر روی پیدا کردن قوانین انجمنی رابطه های مکانی میان شیء ها در یک عکس تمرکز کرده ایم. بنابراین ما یک الگوریتم جدید را پیشنهاد می کنیم که کاوش قوانین انجمنی مکانی را بر اساس نمایش 17DLT و تصویر موجود در پایگاه داده تصاویر انجام می دهد. روش پیشنهادی ما شامل دو مرحله می باشد. در مرحله اول ما ابتدا تمام الگوهای تکرار پذیر به طول یک را پیدا کرده و در مرحله دوم با استفاده از الگوهای k تکرار ( $k \geq 1$ )، به طوری که k برابر طول الگو می باشد همه ی الگوهای ( $k+1$ ) تکرار پذیر را ایجاد و سپس با اسکن پایگاه داده و با کنترل مقدار support اگر آن الگو تکرار پذیر بود به تولید الگوی ( $k+1$ ) می پردازد. با تکرار مراحل موجود در فاز دو تا زمانی که الگوهای با طول بیشتر یافت نگردد