

## عنوان مقاله:

خوشه بندی داده های جریانی با توجه به تاریخچه تکامل خوشه ها

## محل انتشار:

سومین کنفرانس ملی مهندسی برق و کامپیوتر (سال: 1398)

تعداد صفحات اصل مقاله: 12

## نویسندگان:

خمیله حسینی - دانشجوی مقطع کارشناسی ارشد دانشگاه امام رضا (ع)

عادل قاضی خانی - استادیار گروه کامپیوتر دانشگاه امام رضا (ع)

## خلاصه مقاله:

جریان داده به دنباله ای نامحدود از داده های گفته می شود که به طور پیوسته در حال تولید است و از سیستم های بلادرنگ جمع آوری و مورد تجزیه و تحلیل قرار می گیرند. خوشه بندی از مهم ترین مسائل یادگیری بدون نظارت از مرسوم ترین تکنیک های داده کاوی می باشد که هدف آن پیدا کردن خوشه های مشابه در بین نمونه های ورودی داده می باشد الگوریتم  $K\_Means$  نیز یکی از معروف ترین الگوریتم خوشه بندی است که با وجود مزایای زیادی که دارد یک محدودیت عمده دارد آن هم مشخص کردن تعداد خوشه ها یعنی  $k$  می باشد. در الگوریتم های  $K\_Means$  یک الگوریتم تکاملی سریع برای خوشه بندی جریان داده ها می باشد که به طور خودکار اجازه برآورد  $k$  از داده ها را به صورت آنلاین می دهد. در این پژوهش ما یک الگوریتم تکاملی سریع را با استفاده از تاریخچه تکامل خوشه ها برای خوشه بندی جریان داده ها بهبود دادیم. در این روش ما از یک بار برای بهبود صحت استفاده کردیم. نقطه قوت الگوریتم پیشنهادی در خوشه بندی در ترکیب با الگوریتم ژنتیک و حفظ مقادیر به صورت بهترین بهترین ها در متغیری مانند بافر است و برای ارزیابی الگوریتم پیشنهادی از سه دیتاست واقعی استفاده کردیم و میزان بهبود صحت که همان معیار  $SS$  هست را بر روی آن سنجیدیم که به بهبود قابل ملاحظه ای دست یافتیم. در نتیجه، این تجزیه و تحلیل می تواند برای درک رفتار هر خوشه از آغاز داده های جریانی مفید باشد.

## کلمات کلیدی:

خوشه بندی جریان داده ها، الگوریتم های تکاملی، رانش مفهوم، جریان داده، خوشه بندی  $k\_means$

## لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1005926>

