

استفاده از داده‌کاوی در بهبود مدیریت شهری: موردکاوی سامانه 137 شهرداری تهران

علی محمد احمدوند¹، الهام آخوندزاده²، بهروز مینایی بیدگلی³

چکیده

مدیریت شهری یکی از مفاهیم اصلی در حوزه خدمات دولتی محسوب شده و از اهمیت بسیار زیادی برخوردار است. شهرها به عنوان سازمان های پیچیده در دنیای امروز، نیازمند شناخت دقیق و بیش از پیش هستند. دولت‌ها موظفند نیازهای شهروندان را به درستی شناسایی نموده و با پیش‌بینی آن از بروز مسائل و مشکلات احتمالی جلوگیری کنند. در این راستا، استفاده از تکنیک‌ها و ابزارهای داده‌کاوی می‌تواند بسیار مثر و ثمر واقع شود. داده‌کاوی به سازمان‌های مرتبط از جمله شهرداری کمک می‌کند که نیازها و مشکلات شهروندان را شناسایی نموده و از این طریق، مدیریت موثرتری بر روی منابع خود داشته و هماهنگی بهتری در ارتباط با شهروندان ایجاد نمایند.

سامانه 137 شهرداری تهران یکی از مراکزی است که با هدف ارائه خدمات شهری به شهروندان و رسیدگی سریع به مشکلات شهری تاسیس شده و پل ارتباطی بین شهروندان و شهرداری محسوب می‌شود. اطلاعات حاصل از این سامانه، حاوی مطالب مفیدی در مورد مشکلات و نیازهای شهروندان بوده و می‌تواند به عنوان منبعی مهم و مناسب برای پیش‌بینی نیازهای ایشان و شناخت مشکلات شهری مورد استفاده قرار بگیرد. در این مقاله با استفاده از الگوریتم *Apriori*، قوانین انجمنی موجود در بازه‌های زمانی/مکانی رخداد مشکلات به دست می‌آید. همچنین با استفاده از یک تکنیک دومارحله‌ای از این الگوریتم، مشکلاتی که با هم و به صورت همزمان در یک بازه زمانی مشخص اتفاق می‌افتند شناسایی می‌شوند. از قوانین به دست آمده می‌توان در شناسایی نیازهای شهروندان و بهبود مدیریت خدمات شهری و استفاده نمود.

کلمات کلیدی: مدیریت شهری، داده‌کاوی، قوانین انجمنی، الگوریتم *Apriori*.

Using data mining techniques to improve urban management: the case of 137 system of Tehran municipality

Ali Mohammad Ahmadvand, Elham Akhondzadeh, Behrooz Minaei Bidgoli

Abstract

Urban management is one of the principles of public services and is gaining more importance in this domain. Now days, cities as like as complex organizations need to obtain a deeper understanding. Governments should understand the needs of different citizen groups and provide respective services for each group accordingly. In this way, the application of data mining tools would be very useful to understand citizen's needs. It helps organizations such as municipalities to obtain a deeper understanding of each customer's behaviors, characteristics and to better discriminate and more effectively allocate resources for effective customer relationship management.

1- دکتری مدیریت، دانشیار گروه سیستم‌های اقتصادی - اجتماعی، دانشگاه امام حسین. alimohamad ahmadvand@gmail.com

2- کارشناس ارشد مهندسی صنایع - گروه صنایع - دانشکده فنی و مهندسی - دانشگاه شاهد، e_akhondzadeh@yahoo.com

3- دکتری علوم و مهندسی کامپیوتر - استادیار دانشکده مهندسی کامپیوتر - دانشگاه علم و صنعت ایران minaeibi@cse.msu.edu

The Tehran municipality constructed an urban service management system which was called the 137 call center to connect between the municipality and citizens. This system is focusing on managing the complaints against urban services and requirements. This call center has been established to provide appropriate and timely responses to citizens' demands and complaints. The database of this system includes proper information about citizens' needs and would be used to get a deep understanding of citizens' needs and to predict citizens' requirements. In this study, we apply the Apriori algorithm on the database of Tehran municipality to extract the association rules between complaints, time of complaints and the geographical region that a complaint occurred. Furthermore, we try to find those kinds of subjects which occur spontaneously during a special period of time by using a twostep method of the Apriori algorithm. The results help us to get a deeper understanding of citizens' needs and development of urban service management.

Keywords: Urban management, Data mining, Association rule, Apriori algorithm.

1 - مقدمه

مدیریت شهری یکی از مفاهیم اصلی در حوزه خدمات دولتی محسوب شده و از اهمیت بسیار زیادی برخوردار است. مدیریت کارآمد شهری از جمله دغدغه‌های اصلی بخش دولتی محسوب شده و پیچیده شدن نیازها و مشکلات شهری بر اهمیت این موضوع افزوده است. امروزه به موازات بزرگتر شدن شهرها و پیچیده شدن زندگی شهری، شهروندان بیش از پیش اوقات خود را در سطح شهر می‌گذرانند و از خدمات شهری استفاده می‌کنند. بنابراین شرایط کنونی می‌طلبد که برای اداره شهرها یک سیستم مدیریت شهری پایه‌ریزی شود که تمامی جنبه‌های مختلف را برای حفظ حقوق شهروندان به کار گیرد؛ و باید برای ایجاد و اجرای این سیستم تمام عناصر و سازمان‌های زیربط در این خصوص همکاری کنند. داده‌های مستقیم مدیریت شهری، خواسته‌های مردم و شهروندان هستند و ستانده‌های آن در طی فرایندی از برنامه‌ریزی، طرح و اجرا به دست می‌آیند که هدف آن توسعه کیفیت و کمیت زندگی شهری است. هدف کلی سیستم مدیریت شهری تامین، ایجاد زمینه و محیط مناسب برای زندگی امن و کارآمد شهروندان به تناسب ویژگی‌ها و خواسته‌های آنان است. بنابراین در گام اول، باید نیازها و خواسته‌های گروه‌های مختلف شهروندان را دریافت تا بتوان در گام بعدی نسبت به برنامه‌ریزی، طرح و اجرا اقدام نمود. [1 و 2].

شهرداری به عنوان یکی از کلیدی‌ترین سازمان‌های مرتبط با شهروندان و شاید اصلی‌ترین سازمان در حوزه مدیریت شهری، نقش عمده‌ای در ارائه خدمات شهری به شهروندان ایرانی ایفا می‌کند. این سازمان در رویکردی نوین با به کار بستن دانش فناوری اطلاعات و سیستم‌های مخابراتی و با بهره‌گیری از توان متخصصان داخلی و مجرب در مدیریت شهری، سیستمی تحت عنوان سامانه مدیریت شهری 137 ایجاد نموده که مرکز هدایت و کنترل، ساماندهی و رسیدگی به معضلات و مشکلات شهری است و پل ارتباطی بین شهروندان و شهرداری محسوب می‌شود. برنامه ریزی در حوزه مدیریت شهری با توجه به نیازهای شهری و شهروندی، نیازمند مدیریت هوشمند است که در این راستا راه اندازی سامانه 137، برای سرعت بخشیدن به برطرف کردن مسائل در حوزه مدیریت شهری، نقش بسزایی داشته است. مدیریت شهری در ارتباط میان بخشی و درون بخشی خود با شهروندان باید اثرگذار باشد، چرا که کلان شهر تهران با مشکلات خاص خود از یکسو و بروز و حوادث غیر مترقبه از سوی دیگر، یک سیاستگذاری صحیح و به روز را می‌طلبد که آگاهی و آشنایی به معضلات و مشکلات مناطق و محلات شهر و مهم تر از آن تشخیص نیاز و اولویت بندی رفع نیازمندی ها، از اقدامات موثر آن به شمار می‌آید.

پایگاه داده مربوط به این سامانه، حاوی مطالب مفیدی در مورد خدمات ارائه شده به شهروندان است و می‌تواند به عنوان منبعی مهم و مناسب برای شناسایی مشکلات و نیازهای مناطق مختلف شهری مورد بررسی و کاوش قرار بگیرد. در این راستا در این مقاله سعی می‌شود با استفاده از الگوریتم Apriori، قوانین انجمنی موجود در بازه‌های زمانی/مکانی رخداد مشکلات مورد بررسی و کاوش قرار گیرد. همچنین با استفاده از یک تکنیک دومرحله‌ای از این الگوریتم، مشکلاتی که با هم و به صورت همزمان در یک بازه زمانی مشخص اتفاق می‌افتند شناسایی می‌شوند. از قوانین به دست آمده می‌توان در شناسایی نیازهای شهروندان و بهبود مدیریت خدمات شهری استفاده نمود. این مقاله می‌تواند به عنوان یکی از اولین کاربردهای داده‌کاوی در مدیریت خدمات شهری مورد توجه قرار بگیرد. با توجه به تحقیقات انجام شده، مواردی از کاربرد ICT در حوزه خدمات شهری مشاهده می‌شود؛ ولی موردی از کاربرد تکنیک‌های داده‌کاوی در این حوزه مشاهده نشده است. [3، 4، 5 و 6]

در بخش 2 به مرور ادبیات مرتبط با قوانین انجمنی و الگوریتم اپریوری می‌پردازیم. در بخش 3 روش مورد نظر را ارائه نموده و پایگاه داده مربوط به سامانه 137 را به منظور کشف قوانین انجمنی مذکور مورد کاوش و بررسی قرار می‌دهیم. در بخش 4 نیز، برخی نتایج و نکات قابل ذکر آمده است.

2- مرور ادبیات

در این بخش به شرح مختصری از قوانین انجمنی و الگوریتم Apriori می‌پردازیم.

2-1- قوانین انجمنی

استخراج قوانین انجمنی⁴ یک حالت غیر نظارتی⁵ داده‌کاوی است که به جستجو برای یافتن ارتباط در میان ویژگی‌ها در مجموعه داده‌ها می‌پردازد. به عبارتی دیگر تحلیل وابستگی‌ها، مطالعه‌ی ویژگی‌ها یا خصوصیات است که با یکدیگر همراه هستند. نام دیگر روش تحلیل وابستگی، تحلیل سبد بازار است که به دنبال استخراج قوانین از میان این خصوصیات هستند. به عبارت دیگر، این روش‌ها به دنبال استخراج قوانین به منظور کمی کردن ارتباط میان دو یا چند خصوصیت هستند. قوانین وابستگی به شکل اگر و آن گاه به همراه معیارهای پشتیبان⁶ و اطمینان⁷ مربوط به قوانین هستند.

$$x \Rightarrow y(\text{support, confidence})$$

پشتیبان نشان‌دهنده‌ی درصد یا تعداد مجموعه تراکنش‌هایی در کل مجموعه است که شامل هر دو مجموعه ی X و y باشد. اطمینان نیز میزان وابستگی یک قلم کالای خاص را به دیگری بیان می‌کند و مطابق فرمول (1) محاسبه می‌شود. [8]

$$\text{Confidence}(y | x) = \text{support}(x \cup y) / \text{support}(x) \quad (1)$$

2-2-1- الگوریتم Apriori

الگوریتم اپریوری⁸ را می‌توان یکی از مهمترین یافته‌ها در تاریخ استخراج قوانین انجمنی دانست که توسط اگراوال⁹ در سال 1996 ابداع گردید. این الگوریتم از این حقیقت که همه‌ی زیر مجموعه‌های آیتم‌های مکرر، خود باید مکرر باشند و آیتم‌ها باید طبق قانون ترتیب الفبا مرتب باشند، پیروی می‌کند. در الگوریتم‌های دیگر آیتم‌های مکرر با گسترش به هر یک از اقلام مجزا که ممکن است خودشان مکرر نباشند، در هر یک از تراکنش‌ها ایجاد می‌شوند و به این دلیل پایگاه داده چند بار پیموده می‌شود، در حالی که این الگوریتم پایگاه داده را فقط یک بار می‌پیماید و بنابراین حجم محاسباتی کمتری نسبت به سایر الگوریتم‌ها دارد. اما یکی از معایب این الگوریتم این است که این الگوریتم برای محاسبه پشتیبان آیتم‌های کاندید، همه تراکنش‌ها را بررسی می‌کند و بنابراین نیازمند زمان زیادی است [7].

3- مطالعه موردی: سامانه 137 شهرداری تهران

همان‌طور که بحث شد، در مدیریت شهری، اولین گام شناسایی نیازهای شهروندان است. در این مقاله سعی ما بر این است که نیازهای شهروندان را در مناطق مختلف شهری از لحاظ زمان وقوع و نوع نیاز شناسایی کنیم. در این راستا از قوانین انجمنی و الگوریتم Apriori استفاده می‌کنیم.

پس از کشف الگوهای موجود در بازه‌های زمانی- مکانی بروز مشکلات، یکی از مسائل قابل توجه در بروز مشکلات، بررسی همزمانی وقوع آنها است. بررسی اینکه چه مشکلاتی با هم اتفاق می‌افتند و چه مشکلاتی با هم اتفاق نمی‌افتند. پیشنهاد ما برای کشف این نوع از مشکلات استفاده از یک روش دو مرحله‌ای از الگوریتم Apriori است. در این بخش روش دومرحله‌ای مورد نظر را ارائه نموده و از آن در استخراج قوانین انجمنی موجود در پایگاه داده سامانه 137 شهرداری تهران استفاده می‌کنیم.

3-1- معرفی مرکز داده 137 شهرداری

پایگاه داده‌ای 137 از نوع Ms SQL Server بوده و برای استخراج اطلاعات، از نرم‌افزار SQL Server 2005 استفاده شده است. این مجموعه داده شامل جداولی چون دسته‌بندی مشکلات، دسته‌بندی مناطق و نواحی مختلف، سازمان‌ها و واحدهای اجرایی زیر مجموعه شهرداری، ثبت پیام، پاسخ پیام، ابلاغ پیام، مشخصات شهروند، مدیران، کاربران، کارتابل پیام‌ها، مواد مصرفی، اقدام‌کنندگان، ناظران، دریافت نظرات، ثبت نظرات، بازخورد و... است.

4- Association rules
5- Undirected
6- Support
7- Confidence
8- Apriori
9- Agrawal

از این میان، جدول پیام¹⁰ شامل اطلاعات مفیدی در مورد خدمات ارائه شده به شهروندان که مبنای کار ما در این مقاله قرار گرفته است. جدول پیام، مربوط به ثبت اطلاعات تماس شهروند است و اطلاعات دقیقی را در مورد نوع مشکل، منطقه، ناحیه و زمان بروز مشکل نشان می‌دهد. تعداد رکوردهای این جدول، در حدود 2 میلیون و 900 هزار رکورد است. نمونه‌ای از آن در جدول (1) آمده است. جدول (2) نیز فیلدهای مربوط به این جدول را نشان می‌دهد.

جدول (1): نمونه‌ای از جدول پیام

| Message ID | Subject ID | Message Text | Message Address | Region | Section | Unit ID | Last state ID | Message Time | Message Solar date | Citizen Phone | Citizen Mobile |
|------------|------------|---------------------------------|------------------|--------|---------|---------|---------------|------------------|--------------------|---------------|----------------|
| 12 | 413 | حفاری توسط اداره گاز | مجیدیه جنوبی.... | 7 | 1 | 83 | 27 | 07/02/2008 10:16 | 1387/04/12 | 22502931 | |
| 13 | 420 | نسبت به جدول بندی اقدام نمایید. | میدان خراسان.... | 15 | 1 | 123 | 32 | 07/02/2008 10:16 | 1387/04/12 | | 09125168359 |

جدول (2): معرفی فیلدهای جدول پیام

| ردیف | نام فیلد | تعریف |
|------|--------------------|---|
| 1 | Message ID | شناسه پیام: به هر تماس، یک شناسه جدید اختصاص داده می‌شود. |
| 2 | Subject ID | کد مشکل مطرح شده توسط شهروند |
| 3 | Message Text | شرح مشکل ذکر شده با جزئیات آن |
| 4 | Message Address | نشانی محلی که در آن مشکل اتفاق افتاده است. |
| 5 | Message Region | منطقه مربوط به بروز مشکل |
| 6 | Message Section | ناحیه بروز مشکل هر منطقه دارای چند ناحیه است. |
| 7 | Unit ID | واحد اجرایی مسئول برای رفع مشکل |
| 8 | Last state ID | آخرین وضعیت رفع مشکل (رضایت یا عدم رضایت) |
| 9 | Message Time | شامل تاریخ و ساعت ثبت خبر |
| 10 | Message Solar Date | تاریخ وقوع مشکل به هجری شمسی |
| 11 | Citizen Phone | شماره تلفن شهروند |
| 12 | Citizen Mobile | شماره موبایل شهروند |

3-2- پیش‌پردازش

پیش‌پردازش و آماده‌سازی داده‌ها مهم‌ترین و زمان برترین مرحله در پروژه‌های داده‌کاوی است. تقریباً 60 تا 90 درصد زمان انجام یک پروژه داده‌کاوی صرف این مرحله می‌شود و 75 تا 90 درصد موفقیت پروژه‌های داده‌کاوی به آن بستگی دارد. فرایندهایی که در پیش‌پردازش انجام می‌شوند عبارت است از: ¹¹تجمیع، ¹²نمونه‌گیری، ¹³کاهش بعد، ¹⁴انتخاب زیرمجموعه ویژگی‌ها، ایجاد ویژگی و تبدیلات داده. بر اساس نوع کاربردی که عمل داده‌کاوی باید روی آن انجام شود، تکنیک‌های مختلفی برای هر یک از این اعمال مورد استفاده قرار می‌گیرد [7].

- 10- Message
- 14-Aggregation
- 15-Sampling
- 16- Dimensionality Reduction
- 17- Feature subset selection

نمونه‌ای که در این مقاله مورد استفاده قرار گرفته است، داده‌های مربوط به تابستان سال 1386 و 1387 شامل 2,326,345 رکورد است. 4 فیلد از 12 مشخصه فوق انتخاب شده و تمامی رکوردها لحاظ شده است. فیلدهای مورد نظر عبارتند از: Section, Region, Subject ID, Message Solar Date. همچنین فیلد Unit نیز در پاکسازی داده برای پرکردن داده‌های مفقوده و درست کردن داده‌های ناسازگار که در ذیل به آن اشاره شده، مورد استفاده قرار گرفته است.

در این مرحله، مشخصه‌های فوق مورد بررسی قرار گرفته و داده‌ها در قالب مورد نظر آماده شده است. همچنین، اطلاعات نویزدار و مقادیر تهی، مورد بررسی قرار گرفته است. با توجه به ارتباط بین فیلدهای Subject, Region, Section, Unit، مقادیر تهی پر شده و مقادیر ناصحیح، اصلاح شده است. در مواردی که با توجه به ارتباط فوق، امکان اصلاح وجود ندارد، رکورد مربوطه حذف گردیده است. همچنین دو فیلد Region و Section ترکیب شده و به صورت فیلد جدید Region-Section (منطقه- ناحیه) ذخیره گردیده است. همچنین دو فیلد جدید تحت عنوان "Solar Month" و "Day" که به ترتیب نشان‌دهنده ماه و روز وقوع مشکل است از فیلد Message Solar Date استخراج شده است. در ادامه، دو فیلد Solar Month و Day ترکیب شده و به صورت فیلد جدید Month-Day ذخیره گردیده است.

فیلدهای مورد نظر که در فرایند کاوش قوانین انجمنی مورد استفاده قرار می‌گیرند، در جدول (3) آمده است.

جدول (3): فیلدهای مورد استفاده

| | |
|----------------|---|
| Subject ID | 1 |
| Region | 2 |
| Section | 3 |
| Region-Section | ۴ |
| Solar month | 5 |
| Month-Day | ۶ |

3-3 - روش دومرحله‌ای مورد استفاده

برای کشف اینکه چه مشکلاتی با هم اتفاق می‌افتند، اولین راهی که به ذهن می‌رسد، استفاده از قوانین انجمنی است. اما نکته مهم تبدیل داده‌ها به فرمتی است که کشف قوانین مورد نظر امکان‌پذیر باشد. در این راستا باید داده‌ها به صورتی باشد که فیلدها نشان‌دهنده انواع مشکلات و رکوردها نماینده بازه زمانی باشد که می‌خواهیم همزمانی وقوع مشکلات را در آن به دست بیاوریم. به عنوان مثال روز یا هفته و... که در این صورت داده‌ها، به صورت دودویی¹⁵ با مقادیر صفر و یک خواهند بود؛ به این صورت که یک به این معناست که آن مشکل در بازه زمانی مورد نظر اتفاق افتاده است. در این مقاله هدف ما این است که مشکلاتی که با هم در یک روز در فصل تابستان اتفاق می‌افتند را به دست بیاوریم. در این مرحله، مسئله‌ای که با آن مواجه هستیم، مشکل بودن تبدیل داده به این فرمت و تعداد زیاد فیلدهای موجود آمده است. به عنوان مثال در فصل تابستان که در این مقاله مورد بررسی قرار گرفته است، تعداد فیلدها برابر با 268 و تعداد رکوردها برابر با 90 خواهد بود. راهی که ما در اینجا پیشنهاد می‌کنیم، استفاده از الگوریتم اپریوری به صورت دومرحله‌ای است. به این صورت که ابتدا به استخراج قوانینی می‌پردازیم که بر اساس آن بتوان به این نکته دست یافت که چه مشکلاتی در یک روز به خصوص از سال اتفاق افتاده‌اند. سپس با تبدیل دانش به دست آمده به فرمت جدید داده و استفاده مجدد از قوانین انجمنی، می‌توان به این نکته دست یافت که چه مشکلاتی با هم اتفاق می‌افتند و یا برعکس. در واقع هدف این است که با بررسی مشکلاتی که در یک روز به خصوص از تابستان با هم اتفاق افتاده‌اند، به مشکلاتی دست یابیم که با احتمال زیادی در فصل تابستان با هم اتفاق می‌افتند.

با قوانین حاصل از مرحله اول می‌توان به یک توصیف کلی از مشکلاتی که با هم اتفاق می‌افتند دست یافت. اما راه حل دقیق‌تری که پیشنهاد می‌شود، استفاده مجدد از قوانین انجمنی برای بررسی الگوهای متواتر "همزمانی وقوع مشکلات" است که از مرحله اول به دست آمده است. در واقع خروجی الگوریتم Apriori را به فرمت جدیدی از داده تبدیل کرده و به عنوان ورودی به این الگوریتم داده می‌شود. به این صورت که رکورد-های جدیدی تعریف می‌کنیم که هر یک نشان‌دهنده هر یک از الگوهای متواتر به دست آمده است. فیلدها، مشکلاتی است که در الگوهای متواتر مشاهده شده است و رکوردها، هر یک از الگوهای به دست آمده از "همزمانی وقوع مشکلات" است که مرتبط با یک روز به خصوص از فصل تابستان است که چند مشکل با هم اتفاق افتاده‌اند. در این صورت، فیلدها متغیرهای دودویی خواهند بود که در صورت وجود مشاهده در یک الگوی به دست آمده، آن متغیر برابر با یک و در غیر این صورت برابر با صفر خواهد بود. بدین ترتیب هم تعداد فیلدها بسیار کاهش یافته و همچنین با استفاده از قوانین انجمنی موجود بین این فیلدها، مشکلاتی که با هم در یک روز اتفاق می‌افتند مشخص خواهند شد. به همین صورت به دست

آوردن مشکلاتی که با هم در یک هفته، در یک ماه و یا هر بازه زمانی مورد نظر اتفاق می‌افتند، ممکن خواهد شد. با استفاده از روش فوق، تعداد رکوردها و فیلدها به نسبت خوبی کاهش می‌یابد. همچنین این روش این امکان را فراهم می‌آورد که مشکلاتی که از دید شهروندان از اولویت و اهمیت بیشتری برخوردارند، در مرحله دوم لحاظ می‌شوند.

3-3-1- کشف قوانین انجمنی موجود در بازه‌های زمانی - مکانی بروز مشکلات

همان‌طور که گفته شد به منظور شناسایی نیازها و مشکلات شهری، از الگوریتم Apriori در کشف قوانین انجمنی موجود در بازه‌های زمانی- مکانی رخداد مشکلات استفاده می‌کنیم. از آنجایی که نیازها و سلايق شهروندان در مناطق مختلف شهری متفاوت است و مردمی که در یک از منطقه زندگی می‌کنند از لحاظ سلايق و علايق مشابه یکدیگر هستند، فرایند کشف قوانین انجمنی را به تفکیک مناطق 22 گانه شهری مورد بررسی قرار می‌دهیم.

به این منظور قوانین انجمنی موجود بین فیلدهای Subject، Section و Month-Day، به منظور بررسی ارتباط بین نوع مشکلات، مکان و زمان بروز آن استخراج می‌شود. حداقل confidence، 80 درصد در نظر گرفته شده است. برخی از قوانین به دست آمده مربوط به منطقه 12 شهری در جدول 4 آمده است.

جدول (4): قوانین انجمنی موجود در بازه‌های زمانی - مکانی وقوع مشکلات در منطقه 12

| ردیف | مقدم | تالی | Support % | Confidence % |
|------|-----------------------------------|-----------------|-----------|--------------|
| ۱ | section = 11 and month-day = 5-14 | subject = 513 | 0.006 | 100 |
| ۲ | section = 8 and month-day = 5-16 | subject = 92 | 0.006 | 100 |
| ۳ | section = 10 and month-day = 5-26 | subject = 467 | 0.006 | 100 |
| ۴ | section = 9 and month-day = 5-5 | subject = 92 | 0.006 | 100 |
| ۵ | month-day = 6-2 and section = 2 | subject = 513 | 0.017 | 86.667 |
| ۶ | month-day = 6-9 and section = 2 | subject = 483 | 0.006 | 100 |
| ۷ | Section=2 and subject = 125 | month-day = 4-7 | 0.071 | 100 |
| ۸ | Section=2 and subject = 156 | month-day = 4-7 | 0.141 | 90 |
| ۹ | Section=2 and subject = 102 | month-day = 6-7 | 0.141 | 95 |
| ۱۰ | subject = 362 and section=2 | month-day = 6-7 | 0.141 | 100 |
| ۱۱ | Section=2 and month-day = 4-9 | subject = 116 | 5.87 | 80.723 |

به عنوان نمونه قانون سوم نشان می‌دهد که در منطقه 12 وناحیه 10 شهر تهران، در روز 26م از ماه دوم تابستان، مشکل "جوی نیاز به لایروبی دارد" با احتمال 100 درصد روی خواهد داد. با تحلیل قوانینی از این نوع می‌توان به یک توصیف کلی از مشکلاتی که در نواحی مختلف این منطقه روی می‌دهد، دست یافت. یا به عنوان نمونه، قانون 9 و 10 در کنار هم نشان می‌دهد که مشکلات "این معبر نیاز به شستشو دارد" و " محل در حال تغییر کاربری از مسکونی به تجاری یا اداری میباشد " با هم در روز هفتم از ماه آخر تابستان اتفاق افتاده‌اند. سوالی که در اینجا مطرح می‌شود این است که آیا این دو مشکل همیشه با هم در فصل تابستان اتفاق می‌افتند. روش دومرحله‌ای ارائه شده، راه حلی برای دستیابی به پاسخ این سوال است.

همچنین به صورت کلی در کل شهر تهران با در نظر گرفتن فیلد Region-Section، می‌توان به یک توصیف کلی از مشکلاتی که در مناطق و نواحی مختلف اتفاق افتاده‌اند دست یافت. مواردی از قوانین به دست آمده به عنوان نمونه در جدول 5 نشان داده شده است.

جدول (5): قوانین انجمنی موجود در بازه‌های زمانی - مکانی وقوع مشکلات در مناطق 22 گانه شهر تهران

| ردیف | مقدم | تالی | Support % | Confidence % |
|------|---------------------------------|---------------|-----------|--------------|
| 1 | R-S = 12_4 and month-day = 5-3 | subject = 395 | 0.063 | 96.296 |
| 4 | R-S = 6_3 and month = 5 | subject = 92 | 0.002 | 83.333 |
| 5 | R-S = 12_2 and month-day = 4-9 | subject = 116 | 0.024 | 80.723 |
| 7 | R-S = 19_4 and month-day = 6-25 | subject = 513 | 0.001 | 80 |
| 9 | R-S = 22_4 and month-day = 4-17 | subject = 89 | 0.001 | 80 |
| 10 | R-S = 19_4 and month = 6 | subject = 513 | 0.001 | 80 |
| 11 | R-S = 22_4 and month = 5 | subject = 513 | 0.001 | 80 |

3-3-2- کشف همزمانی وقوع مشکلات

در این قسمت به طور خلاصه به روش ارائه شده برای یافتن مشکلاتی که با هم در یک روز اتفاق می‌افتند اشاره می‌کنیم.

1- استخراج قوانین انجمنی موجود در فیلدهای "Month-Day"، "Section" و "Subject".

2- تبدیل نتایج به فرمت جدید به صورتی که فیلدها نشان‌دهنده انواع مشکلات و رکوردها نماینده الگوهای به دست آمده از "همزمانی وقوع مشکلات در یک روز به‌خصوص از تابستان" می‌باشند.

3- استخراج قوانین انجمنی موجود بین انواع مشکلات در داده‌های جدید.

در این قسمت به نتایجی که در منطقه 12 و ناحیه 2 به دست آمده است، اشاره می‌کنیم. جدول 6 قوانینی را نشان می‌دهد که به منظور کشف مشکلاتی که با هم در یک روز در ناحیه 2 مربوط به این منطقه اتفاق می‌افتند، استخراج شده است. حداقل confidence 95 درصد در نظر گرفته شده است. لازم به ذکر است که با استفاده از روش مذکور، تعداد فیلدها از 268 به 17 و تعداد رکوردها از 90 به 14 کاهش یافته است.

جدول 6: قوانین انجمنی موجود بین انواع مشکلات

| ردیف | مقدم | تالی | Support % | Confidence % |
|------|--------------------------------|------|-----------|--------------|
| 1 | s362 | s102 | 6.667 | 100 |
| 2 | s485 | s401 | 6.667 | 100 |
| 3 | s231 | s90 | 6.667 | 100 |
| 4 | s400 and s526 | s92 | 6.667 | 100 |
| 5 | s485 | s395 | 6.667 | 100 |
| 6 | s401 and s395 | s485 | 6.667 | 100 |
| 7 | s231 and s138 | s513 | 6.667 | 100 |
| 8 | s138 and s513 and s395 | s90 | 6.667 | 100 |
| 9 | s485 and s401 and s138 | s395 | 6.667 | 100 |
| 10 | s231 and s90 and s513 | s395 | 6.667 | 100 |
| 11 | s90 and s138 and s513 and s395 | s231 | 6.667 | 100 |

به عنوان مثال قانون 1 نشان می‌دهد که مشکلات "این معبر نیاز به شستشو دارد" و "محل در حال تغییر کاربری از مسکونی به تجاری یا اداری میباشد" به احتمال 100 درصد با هم در یک روز در منطقه 12 در فصل تابستان اتفاق می‌افتند. یا قانون 2 به این معناست که اگر در یک روز مشکلات "زباله و ضایعات به موقع جمع آوری نمی‌شود" و "اصناف در این محل سد معبر کرده اند" با هم اتفاق بیفتند، به احتمال 100 درصد مشکل "این معابر (پیاده رو و سواره رو) نیاز به نظافت دارد" نیز رخ خواهد داد.

در جدول 7 برخی از مشکلاتی که با هم اتفاق در یک روز در منطقه 12 اتفاق می‌افتند، آورده شده است.

جدول 7: همزمانی وقوع مشکلات در یک روز

| ردیف | مشکلاتی که با هم در یک روز اتفاق می‌افتند |
|------|--|
| 1 | "این معبر نیاز به شستشو دارد" و "محل در حال تغییر کاربری از مسکونی به تجاری یا اداری میباشد" |
| 2 | "این زمین بایر نیاز به نظافت دارد" و "در این محل زباله و ضایعات را دپو می‌کنند." |
| 3 | "ساخت و ساز غیر مجاز در محل صورت می‌گیرد" و "لکه گیری چاله های معابر (سواره رو)" |
| 4 | "اصناف در این محل سد معبر کرده اند" و "زباله و ضایعات به موقع جمع آوری نمی‌شود" |
| 5 | "رفع سد معبر دستفروشان و فروشندگان دوره گرد" و "مخزن زباله و ضایعات به موقع تخلیه نمی‌شود" |
| 6 | "مخزن زباله و ضایعات به موقع تخلیه نمی‌شود" و "جمع آوری زباله و سرشاخه و ضایعات" |
| 7 | "در این محل بطور غیر مجاز تفکیک و بازیافت زباله انجام می‌گیرد." و "جمع آوری خاک و نخاله" |
| 8 | "زباله و ضایعات به موقع جمع آوری نمی‌شود" و "اصناف در این محل سد معبر کرده اند" و "این معابر (پیاده رو و سواره رو) نیاز به نظافت دارد." |
| 9 | "رفع سد معبر دستفروشان و فروشندگان دوره گرد" ، "مخزن زباله و ضایعات به موقع تخلیه نمی‌شود" و "جمع آوری زباله و سرشاخه و ضایعات" و "در این محل بطور غیر مجاز تفکیک و بازیافت زباله انجام می‌گیرد" ، "مخزن زباله و ضایعات به موقع تخلیه نمی‌شود" ، "رفع سد معبر دستفروشان و فروشندگان دوره گرد" و "جمع آوری زباله و سرشاخه و ضایعات" |

بسیاری از نتایج به دست آمده در نگاه اول منطقی به نظر می‌رسد و برخی دیگر نیازمند ریشه‌یابی و تمرکز بیشتری است. با توجه به نتایج به دست آمده، می‌توان به ریشه و علت برخی مشکلات پی برد و دانش صحیح‌تری نسبت به علت و ریشه آن دست یافت. نتایجی از این نوع این امکان را فراهم می‌آورد که برخلاف تصور اولیه از علت بروز برخی از مشکلات به علت واقعی و برخی دلایل احتمالی آن پی برد و نسبت به رفع پایدار آن

اقدام نمود. همچنین می‌توان با آگاهی نسبت به همزمانی وقوع برخی از این مشکلات در یک روز نسبت به رفع آنها زودتر از تماس شهروند اقدام نمود که میزان رضایت‌مندی را به مراتب افزایش خواهد داد.

4 - نتیجه‌گیری

برنامه‌ریزی در حوزه مدیریت شهری با توجه به نیازهای شهری و شهروندی، نیازمند مدیریت هوشمند است که در این راستا راه‌اندازی سامانه 137، برای سرعت بخشیدن به رفع مسائل در حوزه شهری و شناسایی و اولویت‌بندی نیازهای شهروندان، نقش بسزایی داشته است. در این مقاله سعی شده تا با استفاده از تکنیک قوانین انجمنی، نیازهایی که از دید شهروندان اهمیت و اولویت بیشتری دارند، به تفکیک منطقه و زمان شناسایی شوند. همچنین با استفاده از یک تکنیک دومرحله‌ای از الگوریتم اپریوری، الگوهای انجمنی موجود در وقوع مشکلات در منطقه 12 و ناحیه 2 در فصل تابستان استخراج شده و سپس مشکلاتی که در یک روز با هم اتفاق می‌افتند، شناسایی شده است. روش ارائه شده، مشکلاتی را که از دید شهروندان در اولویت اول قرار دارند، لحاظ نموده و تعداد فیلدها و رکوردها را تا حد زیادی کاهش می‌دهد.

نتایج به‌دست آمده این امکان را فراهم می‌آورد که مشکلات با اهمیت و بحرانی در مناطق و نواحی مختلف شناسایی شود. همچنین با تامل و تحلیل دقیق‌تر، می‌توان به ریشه و علت برخی مشکلات پی برد و دانش صحیح‌تری نسبت به علت و ریشه آن دست یافت. نتایجی از این نوع این امکان را فراهم می‌آورد که برخلاف تصور اولیه از علت بروز برخی از مشکلات به علت واقعی و برخی دلایل احتمالی آن پی برد و نسبت به رفع پایدار آن اقدام نمود. همچنین می‌توان با آگاهی نسبت به همزمانی وقوع برخی از این مشکلات در یک روز نسبت به رفع آنها زودتر از تماس شهروند اقدام نمود که میزان رضایت‌مندی را به مراتب افزایش خواهد داد. این مقاله می‌تواند به عنوان یکی از اولین کاربردهای داده‌کاوی در حوزه مدیریت خدمات شهری مورد توجه قرار بگیرد.

فهرست مراجع

- 1- سلیمی، صدیقه، ارائه مدلی جهت اصلاح فرهنگ عمومی در راستای افزایش مارکت پویای مردم در مدیریت شهری، نخستین کنفرانس علمی- تخصصی نظارت همگانی شهرداری تهران، 1385.
- 2- نوروزی، نادر، سازوکارهای حقوقی و تاثیر آن در مدیریت شهری، نخستین کنفرانس علمی- تخصصی نظارت همگانی شهرداری تهران، 1385.
- 3-Aubrecht, C., Hollaus, M., Wagner, W., *Integrating earth observation and GIScience for high resolution spatial and functional modeling of urban land use*, Computers, Environment and Urban Systems, Volume 33, Issue 1, Pages 15-25, 2009.
- 4-Bell, M., Dean, C., Blake, M., *Forecasting the pattern of urban growth with PUP: a web-based model interfaced with GIS and 3D animation*, Computers, Environment and Urban Systems, Volume 24, Issue 6, Pages 559-581, 2000.
- 5-Barton, Jack, Plume, Jim, Parolin, Bruno, *Public participation in a spatial decision support system for housing* *public* Computers, Environment and Urban Systems, Volume 29, Issue 6, Pages 630-652, 2005.
- 6-Gouveia, Cristina, Fonseca, Alexandra, Câmara, Ant?nio, Ferreira, Francisco, *Promoting the use of environmental data collected by concerned citizens through information and communication technologies*, Journal of Environmental Management, Volume 71, Issue 2, Pages 135-154, 2004.
- 7-Han, J. and Kamber, M., *" Data Mining : Concepts and Techniques "*, Second Edition, Morgan Kaufman Publisher, 2006.
- 8-Tan, P., Steinbach, M. And Kumar, V. (2006): *Introduction to Data Mining*, Addison Wesley, ISBN: 0-321-32136-7.