Mining Ordinance Data from the Web for Smart City Development

Xu Du, Aparna S. Varde and Robert W. Taylor

Abstract—**In this research, we aim to discover knowledge from ordinances, i.e., local laws on urban policy. This is useful in policy assessment which we address especially with respect to smart cities. To analyze the publicly available ordinance data from websites guided by human judgment, we use common sense knowledge from a repository called WebChild and its domain-specific knowledge bases in relevant areas, e.g., town planning. Much of this ordinance data maps to smart city characteristics, e.g., smart environment. Hence, based on mining using association rules and other methods, we give feedback to urban agencies for decision support, particularly in a smart city context. To the best of our knowledge, this is among the first works to conduct ordinance mining.**

Keywords --- Association Rules; Classification; Common Sense Knowledge; Decision Support; Urban Policy

I. INTRODUCTION

Public policy has produced many laws that support the goals of environmental management. *Ordinances*, i.e., local laws at municipal levels, are direct policy tools developed by urban management agencies and passed by local-level jurisdictions. The legislation and amendment of those laws are interactive and related with local public opinions [1, 2].

Analyzing the relationship between ordinances and conducting related studies would thus support efficient urban management. We address this issue, particularly with the intention of heading towards the development of smart cities. A smart city is typically expected to have the characteristics [3] as shown in Fig. 1.

These characteristics are smart governance, smart environment, smart mobility, smart living, smart people, and smart economy [3, 4]. *Smart Governance* pertains to government effectiveness, including transparency and public participation in decisions. *Smart Environment* is concerned with energy efficiency, pollution control, sustainable resources etc. *Smart Mobility* focuses on transport issues such as local accessibility with sustainable and safe systems. *Smart Economy* is concerned mainly with competitiveness, innovative spirit, productivity and maintaining cost savings while meeting imperative demands. *Smart Living* deals with public health, safety, housing quality etc. The S*mart People* characteristic entails social and human capital, qualification, creativity and related aspects.

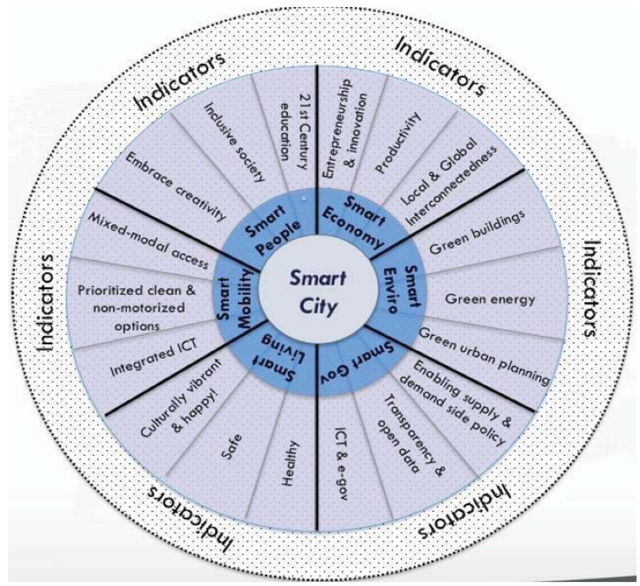

Fig. 1. Typical smart city characteristics

There are many smart cities all over the world. Fig. 2 shows an example of a smart city Amsterdam in the Netherlands where street lamps allow municipal councils to dim and brighten lights based on pedestrian usage [4]. This would certainly enhance transportation by providing more sustainable systems, thus catering to the smart mobility characteristic of smart cities.

Fig. 2. Smart city example – Amsterdam

The work in this paper is funded by a doctoral research assistantship for Xu Du from Montclair State University. During the earlier part of this work, Dr. Aparna Varde was a visiting senior researcher at Max Planck Institute for Informatics, Saarbruecken, Germany. The relevant work of Dr. Robert Taylor on technologies for climate change adaptation is supported by the Vietnam Educational Foundation and has a global perspective.

Xu Du is a Doctoral Candidate in the Department of Earth and Environmental Studies in their Environmental Management PhD Program at Montclair State University, NJ, USA (email: dux3@montclair.edu)

Aparna Varde is an Associate Professor in the Department of Computer Science at Montclair State University and Doctoral Faculty Member in their Environmental Management PhD Program (email:vardea@montclair.edu)

Robert Taylor is a Full Professor in the Department of Earth and Environmental Studies at Montclair State University and Doctoral Faculty Member in their Environmental Management PhD Program (email: taylorr@montclair.edu)

Given this overall framework, the issues in urban policy can also be divided into these different categories in order to address a smart city context. Considering this, our problem goals are as follows.

- Investigate ordinances passed by urban agencies in a given location over multiple time spans based on enactment, initialization and other relevant aspects.
- Gauge the effectiveness of ordinances with respect to urban policy considering the respective smart city characteristics they address.

Our source of data for ordinances is public websites. We consider these ordinances over multiple time spans. We aim to conduct mining over the data that can be used to answer questions of interest to urban agencies, e.g., *"Which ordinances in a given year cater to smart environment?"; "What is the average time span of an ordinance legislation in a given session?"; "Which smart city characteristic has received the greatest attention over all the years?"; "What is the relationship between the initialization and enactment of an ordinance over a certain time period?"; "How have ordinances on smart mobility changed in the last five years?"* etc. This would help urban agencies investigate their overall performance and also assess where they stand in developing a smart city, i.e., which characteristics are considered and how they are addressed.

The rest of this paper is organized as follows. Section II describes our approach on mining of ordinances. Section III summarizes the experimental results we obtain. Section IV outlines related work in the area. Section V states the conclusions and ongoing research.

II. APPROACH FOR ORDINANCE MINING

A. *Overview of Approach*

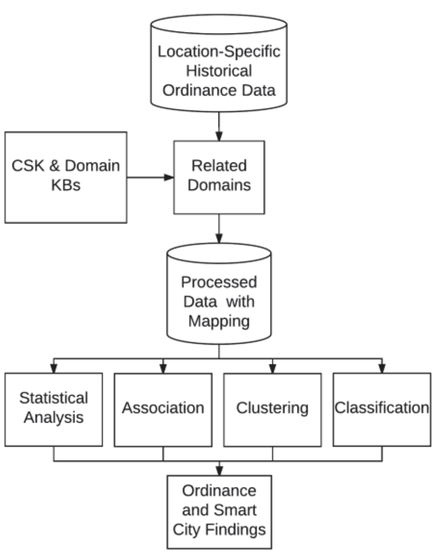

Fig. 3. Illustration of ordinance mining approach

The approach we deploy for mining location-specific temporal ordinance data is illustrated in Fig. 3. We propose the use of common sense knowledge (CSK) since it helps in mapping the ordinances to smart city characteristics. For instance, an ordinance on an aspect of energy consumption may not have the words "smart environment" or its terms in Fig. 1. By using CSK, it is possible to find this mapping and hence enhance the mining on ordinances pertaining to smart city characteristics.

Data on ordinances obtained from the Web is subjected to processing guided by common sense knowledge. Hence, this is mapped to relevant smart city characteristics. It is then subject to data mining using statistical approaches, association rules, clustering and classification. The knowledge discovered is reported as ordinance and smart city findings and can be used to answer questions useful to urban management agencies for decision support. We now explain its detailed steps.

B. *Harnessing Common Sense Knowledge*

In order to utilize CSK, we use WebChild, a huge common sense knowledge base built from Web contents [5]. It has a browser with which users can search information on real-world concepts, their common properties and related terms with pictures. Fig. 4 is a snapshot of the WebChild browser [5]. This has been used to create domain-specific knowledge bases (domain KBs) in relevant areas [6] with ground truth constituting common sense concepts on urban policy. Given a file with terms from a probabilistic domain classifier, relevant domains are selected and concepts in those domains entered as elaborated in [6] and briefly illustrated in Figs. 5 and 6 respectively.

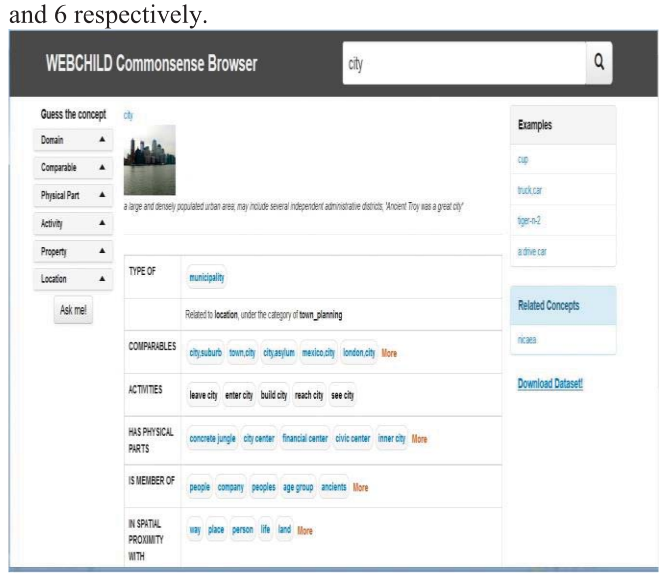

Fig. 4. Snapshot of WebChild browser

Fig. 5. Relevant domains selected in KB

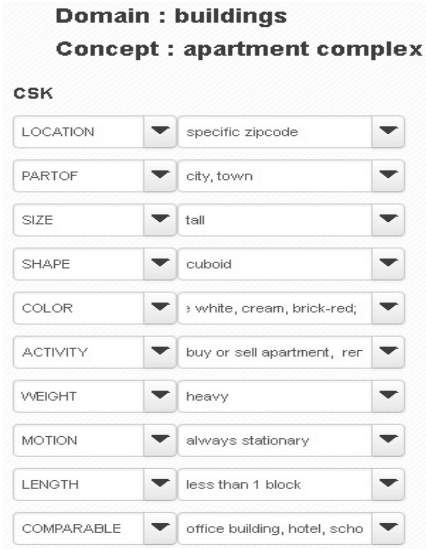

Fig. 6. Concepts entered in domains

C. Data Processing and Smart City Mapping

Considering CSK and domain KBs, we proceed as follows. As shown in Fig. 3, we start with the publicly available ordinance data comprising the location-specific historical information over different time spans. This raw data on ordinances is processed by our program using the terms in WebChild [5] and its related domain KBs [6].

Note that common sense knowledge and related domain-specific knowledge bases play a twofold role here. First, they drastically reduce the data set size for mining by selecting only pertinent data on "Related Domains" (see Fig. 3) and filtering out the rest, thus making the process more effectual. This is done by incorporating relevant CSK and domain KB terms into the programs that access the respective websites to produce the ordinance data in a format suitable for mining. Second, they also map the ordinances to their smart city characteristics analogous to a human, but more efficiently. For example, consider the following ordinance found on a website [2]: "A local law to amend the New York City building code in relation to requiring carbon monoxide detectors in certain apartments is hereby passed." This would be found relevant to the domain KB on "buildings" with its specific concept being "apartment complex'. Hence, the concerned program would map this to the smart city characteristic of smart living which comprises public health, safety, housing quality etc., taking into account CSK which relates "building" to "housing". Details of all mappings are not shown herewith but occur on similar lines, thus helping to generate processed data in the format shown in TABLE I.

TABLE I
PROCESSED DATA ON ORDINANCES

Time	Ordinance	Related Policies	Smart City Characteristics
MM/DD/YY	Actual content in the website	Energy/ Transportation/ ……..	Smart Environment/ Smart Mobility
…	…	…	…

This table depicts a data set for a specific location over multiple time spans. It exemplifies the processed data stored as intermediate output from Web based ordinances and is maintained in databases for further analysis.

D. Deployment of Mining Methods

The processed data sets are subjected to exploratory data mining with statistical analysis [7] including temporal factors, median calculations, minimum and maximum value observations and other aspects. Association rules, clustering and classification are also conducted over the data [7]. We select these data mining techniques for the following reasons.

Since association rule mining finds relationships of the type A => B, it is expected to be useful in identifying how one feature of the urban policy relates to another. Clustering places items in groups based on their similarity and hence is likely to help in finding similarities among the ordinances by grouping the relevant ones together. Classification predicts a target based on analysis of existing data and thus is found potentially suitable with respect to categorization. In other words, it would help to specifically categorize an ordinance based on its smart city characteristic addressed.

Outputs provided by data mining are therefore expected to help in understanding relationships between various aspects of the ordinances passed by urban management agencies and in assessing them with reference to smart city characteristics. This would guide decision support for these urban agencies.

III. Experimental Results

We conducted experiments using our approach for ordinance mining as described herewith in the subsections on statistical analysis with clustering; association rules and classification respectively. In this paper, we focused on New York City as the location and considered its council data [2]. We chose NYC since it is the most populous city in the USA, has systematic ordinance data publicly available on the Web and also has many urban policy issues addressed. An example of the NYC council data used in our work appears in Fig. 7.

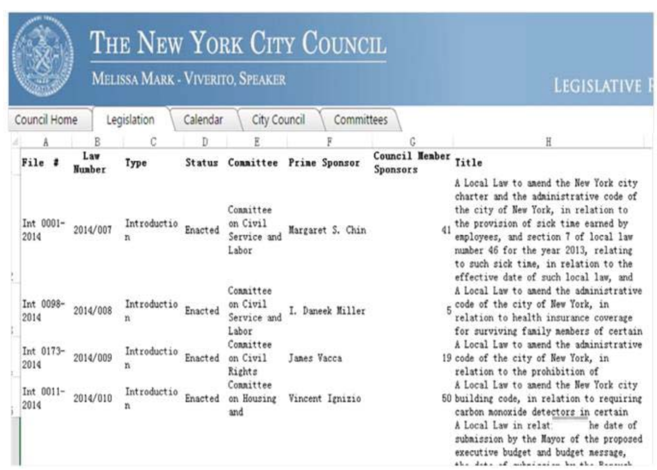

Fig. 7. Example of NYC ordinance data

A. Statistical Analysis with Clustering

We conducted exploratory data mining on the legislative activity related to the ordinances by the NYC council from 2006 to 2013 using statistical analysis [7] taking into account the time factor. The time period corresponded to the two latest full NYC city council sessions. We thereby analyzed the distribution of the ordinances by different urban committees over time. We also calculated the days that the city council spent to enact each ordinance. The results of this analysis are summarized in Fig. 8. Here the dotted line indicates the ordinances initialized in the respective years and the solid line indicates the ones enacted that year.

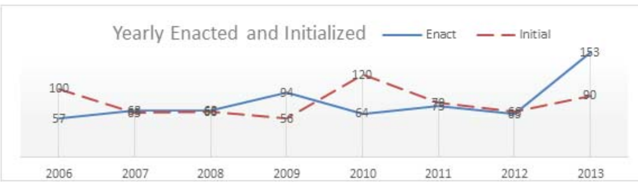

Fig. 8. Statistical plot of enacted and initialized ordinances

Furthermore, these ordinances were subjected to a simple clustering [7] by grouping them with respect to sessions between 2006-2009 and 2010-2013. The results of this process are visualized in Figs. 9 and 10 respectively. The dotted and solid lines represent the same aspects as in Fig. 8.

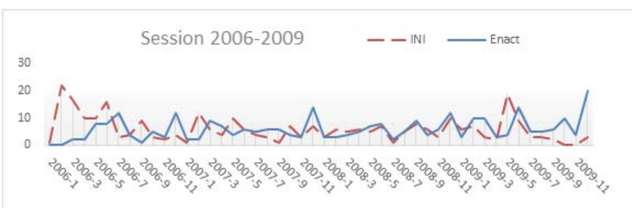

Fig. 9. Visualization of ordinances clustered from 2006-2009

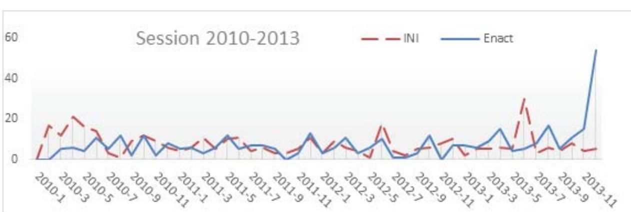

Fig. 10. Visualization of ordinances clustered from 2010-2013

From this temporal statistical analysis and basic session-related clustering we found that the first year of each session had the highest number of initialized ordinances while the last year had the highest number of enacted ordinances. Also, some additional observations from the statistical analysis revealed that the total number of ordinances increased from 287 in earlier the time period of 2006-2009 to 358 in the later time period of 2010-2013. *This indicated that urban agencies passed more ordinances as time progressed*.

Temporal statistical analysis of the legislation time span presented other interesting results. It was found that the average time span of ordinance legislation increased from 204 to 222 days in the two respective sessions. The median number had a drastic change from 89 to 131 days. *These results indicated that, on the whole, ordinance legislation took a longer time in the later sessions.*

Considering the distribution of ordinances by committees, the top three committees of session 2006-2009 were found to be the Committee on Finance, Committee on Housing and Buildings and Committee on Environmental Protection, the percentages being 17.07%, 16.72% and 9.41% respectively. The top three committees of session 2010-2013 were the Committee on Housing and Buildings, Committee on Finance and Committee on Transportation, the percentages being 16.76%, 12.57% and 12.29% respectively. *This showed an interesting change, i.e., an increase in the number of transportation ordinances being passed in the 2010-2013 session*.

The average legislation time span of the top three committees from 2006-2009 were 89, 192 and 406 days respectively while the time span of the 2010-2013 respective committees were 176, 133 and 283 days. We noticed that the average time span of all the ordinances in those two sessions were 204 and 222 days respectively. This indicated that the ordinances of the Committee on Finance had a much shorter time span of legislation (89d&204d, 133d&222d) respectively while those of the Committee of Environment Protection had a longer time span (406d&204d, 355d&222d) respectively. *Thus, financial legislations were found to be faster while the environmental ones were much slower, probably indicating that there was a significantly greater demand to speed up financial policies.*

B. Association Rule Mining

We conducted association rule mining using the classical Apriori algorithm [7] which follows the principle of frequent item sets and their supersets. The mining was done in two steps. We first mined the plain committee-based data. This was done by discretizing the time span into ten bins based on equal width and using the Apriori algorithm with suitable parameters. Examples of rules obtained are shown below.

1. Committee=Finance 94 => TimeSpan='(-Inf-143.6]' 77 <conf:(0.82) > lift: (1.4) Lev: (0.03) [22] Conv: (2.17)
2. EnactM=7 45 => TimeSpan='(-inf-143.6]' 35 <conf:(0.78)> lift:(1.33) lev:(0.01) [8] conv:(1.7)
3. EnactM=5 49 => TimeSpan='(-inf-143.6]' 34 <conf:(0.69)> lift:(1.19) lev:(0.01) [5] conv:(1.27)
4. EnactM=10 58 => TimeSpan='(-inf-143.6]' 40 <conf:(0.69)> lift:(1.18) lev:(0.01) [6] conv:(1.27)
5. EnactM=6 67 => TimeSpan='(-inf-143.6]' 46 <conf:(0.69)> lift:(1.17) lev:(0.01) [6] conv:(1.27)
6. IniM=12 56 => TimeSpan='(-inf-143.6]' 35 <conf:(0.63)> lift:(1.07) lev:(0) [2] conv:(1.06)
7. IniM=5 76 => TimeSpan='(-inf-143.6]' 46 <conf:(0.61)> lift:(1.04) lev:(0) [1] conv:(1.02)
8. IniM=3 65 => TimeSpan='(-inf-143.6]' 38 <conf:(0.58)> lift:(1) lev:(0) [0] conv:(0.96)
9. IniM=6 111 => TimeSpan='(-inf-143.6]' 63 <conf:(0.57)> lift:(0.97) lev:(-0) [-1] conv:(0.94)
10. IniM=4 58 => TimeSpan='(-inf-143.6]' 32 <conf:(0.55)> lift:(0.94) lev:(-0) [-1] conv:(0.89)

The useful knowledge extracted from here was that the ordinances initialized and enacted during the middle of the year had a relatively short legislation time span. The rules 2, 3, 5, 7 and 8 seemed to support this fact. Also rule 1 supported our previous inference from the statistical analysis that the ordinances passed by the Committee of Finance had a shorter legislation time span.

As a next step, we added more dimensions to the data for association rule mining, i.e., the relevance to smart city characteristics (pertaining to Table 1). Accordingly, we conducted association rule mining on the processed data incorporating the respective smart city terms. It generated a different set of rules, examples of which are shown next.

11. Committee=Transportation 68 => Concept=Mobility 60 <conf:(0.88)> lift:(5.17) lev:(0.08) [48] conv:(6.27)
12. Committee=Environmental Protection 42 => Concept=Environment 36 <conf:(0.86)> lift:(5.58) lev:(0.05) [29] conv:(5.08)
13. TimeSpan='(-inf-143.6]' Committee=Housing and Buildings 59 => Concept=Living 47 <conf:(0.8)> lift:(3.06) lev:(0.05) [31] conv:(3.36)
14. Committee=Finance Concept=Economy 51 => TimeSpan='(-inf-143.6]' 40 <conf:(0.78)> lift:(1.34) lev:(0.02) [10] conv:(1.77)
15. Committee=Housing and Buildings 108 => Concept=Living 83 <conf:(0.77)> lift:(2.95) lev:(0.09) [54] conv:(3.07)
16. TimeSpan='(-inf-143.6]' Concept=Economy 61 => Committee=Finance 40 <conf:(0.66)> lift:(4.5) lev:(0.05) [31] conv:(2.37)
17. Concept=Economy 84 => Committee=Finance 51 <conf:(0.61)> lift:(4.17) lev:(0.06) [38] conv:(2.11)
18. Concept=Living 168 => TimeSpan='(-inf-143.6]' 101 <conf:(0.6)> lift:(1.03) lev:(0) [2] conv:(1.03)
19. Concept=Governance 162 => TimeSpan='(-inf-143.6]' 92 <conf:(0.57)> lift:(0.97) lev:(-0) [-2] conv:(0.95)
20. Committee=Housing and Buildings Concept=Living 83 => TimeSpan='(-inf-143.6]' 47 <conf:(0.57)> lift:(0.97) lev:(-0) [-1] conv:(0.93)
21. Concept=Mobility 110 => TimeSpan='(-inf-143.6]' 61 <conf:(0.55)> lift:(0.95) lev:(-0.01) [-3] conv:(0.91)

These rules depicted the relationships between the respective committees and smart city characteristics. It was thus found that the Committee on Transportation, Committee on Environmental Protection, Committee on Housing and Building and Committee on Finance had a strong correlation with smart mobility, smart environment, smart living and smart economy respectively, which was not surprising. These rules also showed the connection between time span as observed in the statistical analysis with respect to the relevant smart city concepts. It was found that the ordinances related to smart economy and smart living took a shorter time to enact. The ordinances related to smart governance and smart mobility, on the other hand, had a relatively longer time span. *It could thus be inferred that ordinances related to smart economy and smart living were probably found to be more demanding and thus needed faster legislation.*

C. *Decision Tree Classification*

We conducted classification analysis of the data using J4.8 decision tree classifiers [7]. As is well known in the data mining community, decision trees provide a stem and leaf structure with the stems representing the paths based on attributes of the data and the leaves representing the decisions or the classification targets. The J4.8 algorithm for classification is a Java based extension of C4.5 which follows the principle of entropy in inducing a decision tree given a data set [7]. In our data sets, the classification targets were designed to be the smart city characteristics. A summary of the findings is listed below as observed.

Committee = Committee on Finance: Economy (94.0/43.0)
Committee = Committee on Housing and Buildings: Living (108.0/25.0)
Committee = Committee on Sanitation and Solid Waste Management: Environment (34.0/8.0)
Committee = Committee on Contracts: Governance (10.0/3.0)
Committee = Committee on Parks and Recreation: Mobility (33.0/8.0)
Committee = Committee on Standards and Ethics: Governance (2.0)
Committee = Committee on Governmental Operations: Governance (34.0/3.0)
Committee = Committee on Transportation: Mobility (68.0/8.0)
Committee = Committee on Mental Health, Developmental Disability, Alcoholism, Substance Abuse and Disability Services: Living (2.0)

The numbers here can be interpreted as follows. Consider the first finding. Here, among the 94 ordinances passed by the Committee on Finance, 43 were classified as addressing the smart economy characteristic of smart cities. Likewise, it can be inferred from the results of the overall classification analysis seen here that various smart city characteristics were addressed to some extent in the urban policy ordinance data mined herewith.

D. *Summary of Observations*

Based on the ordinance data mining conducted so far, we tabulated the results with reference to the smart city characteristics as shown next. TABLE II herewith depicts an overall distribution of the ordinances addressed in each session with respect to the characteristics of smart cities.

TABLE II
ORDINANCE DISTRIBUTION W.R.T. SMART CITY

Smart City Concept	Session 2006-2009	Percent	Session 2010-2013	Percent
Governance	61	21.25%	101	28.21%
Living	91	31.71%	78	21.79%
People	3	1.05%	18	5.03%
Economy	47	16.38%	37	10.34%
Mobility	42	14.63%	68	18.99%
Environment	43	14.98%	56	15.64%

Thus, we found that the smart city characteristic achieving the greatest attention in the 2006-2009 session was "smart living" while that with the least attention was "smart people". Likewise, in 2010-2013, the maximum ordinances passed were on the "smart governance" characteristic, while the minimum ordinances were on "smart people". *Thus, the urban management agencies can potentially be provided with the suggestion that they need to focus more on urban policy issues related to the "smart people" characteristic such as social and human capital, 21st century education etc.* Note that this characteristic did receive somewhat more attention in the later time period than the earlier one, though still significantly less compared to the others.

Based on our experiments we can conclude the following. The data mining on the ordinances does reveal useful information. First, it helps to explore the statistical aspects of the ordinances with respect to time, e.g., trends in the enacted versus initialized ordinances over the years, maximum number of ordinances in a given session etc. Second, it helps to determine how much the urban policy issues head towards developing a smart city. This mining would thus help to answer some questions useful to urban management agencies as stated in the introduction. It would help the agencies assess their effectiveness and enable them to gauge how close they are in catering to smart city characteristics. It also would have the future impact of making them pass ordinances that head towards making their city smarter. Thus, data mining on the ordinances would potentially guide decision support for the urban management agencies in the overall development of smart cities.

IV. Related Work

The paradigm of smart cities is receiving tremendous attention today. The city of Melbourne in Australia is considered to be one of the finest “knowledge cities” as gathered from the literature, e.g., [4, 8]. The term knowledge city is often used synonymously with smart city, however, it does have subtle differences [8], the focus being more on ubiquitous knowledge dissemination in the first case versus several smart city characteristics in the second one. Many smart cities are found in Europe catering to several characteristics. For example, buses in Barcelona operate on routes designed to optimize energy efficiency [4]. Recycling is encouraged in some cities by giving customers money back for returning recyclable items such as empty plastic bottles. Solar panels are installed on rooftops in many places as an energy efficiency mechanism.

Consequently smart city research is certainly motivated. A framework for automating implicit requirements in software engineering has been built [9] based on common sense knowledge along with text mining and ontology and has an application in the development of smart city tools. Since these requirements are implicit as opposed to explicit ones, they take into account subtle aspects that users often desire but not state upfront, therefore they are found to be crucial in the adequate functioning of software systems and would be particularly useful in smart city applications, catering to various characteristics. Global sustainability has been addressed from an agricultural perspective [10]. Issues such as food security; urban versus rural agriculture; and carbon footprints are discussed with the important conclusion that under-used roof space in large global cities can be used to grow food. This heads towards making a city smarter by more effective use of resources for meeting population needs. In our work in this paper we are in line with these general paradigms, addressing the specific issue of urban policy ordinances.

Computational analysis and data mining have helped significantly in geographic studies. Nagy et al. [11] analyze urban and rural gradients in the USA. They consider various social, economic and environmental aspects along with some relevant responses from an ecological perspective. Many of these have been found useful in geographic data analysis. Pampoore-Thampi et al. address the issue of predicting urban sprawl based on data in geographic information systems (GIS) [12]. They estimate factors causing urban sprawl considering the state of NY with sprawl affected areas over different time spans. They consider factors such as population, employment and transportation with respect to the bidirectional impact on sprawl. In [13], assessing air quality by mining data on fine particle pollutants and related attributes is conducted, especially with respect to public health and safety standards recommended by EPA, the Environmental Protection Agency of the USA. Our work in this paper falls under the same broad realm. We conduct mining on Web based temporal and location-specific urban policy ordinance data with respect to the characteristics of smart cities, use common sense knowledge in the overall process and aim to provide inputs for urban management agencies based on the results.

Researchers have conducted several studies to test the capability of social media mining and sentiment analysis, such as preferences for candidates, interests on certain topics or goods and political opinions [14]. Often, positive correlations have been found between the mining results and reality outcomes. On the other hand, some researchers criticized the method [15] indicating that social media mining and sentiment analysis is not perfect and still has room for improvement.

However, the critics have still asserted that this will be a very good complement to the traditional methods. The supporters as well as the critics agree that full-fledged user surveys in the real world are extremely time-consuming. Therefore, extracting useful knowledge from public opinion expressed in cyberspace seems a better alternative. We intend to address this in our future work by mining social media data on public reaction to ordinances. This would help to assess public satisfaction on urban policy issues through opinion mining, thus providing additional suggestions for decision support.

V. Conclusions

In this paper, we mine data on location-specific ordinances over different time spans in order to assess the effectiveness of urban policy in a smart city context. We deploy common sense knowledge along with related domain-specific knowledge bases for selecting pertinent ordinances, and also for mapping them to the concerned smart city characteristics.

The analysis conducted in our work would potentially help in answering questions to guide urban management agencies in decision support for urban policy in general and especially for the development of smart cities. To the best of our understanding, this paper is among the first works to perform data mining on urban policy ordinance data in particular, thereby presenting interesting applied research.

A few interesting findings from this research are listed herewith with respect to NYC ordinance mining.

- Urban agencies passed more ordinances during the 2010-2013 time span than during 2006-2009, hence indicating an increase in the need for urban policies as time progressed.
- Finance-related ordinances were passed in the shortest span of time, thus implying a greater focus on speeding up policies in "smart economy" so far.
- Ordinances initialized and enacted around the middle of a year seemed to progress faster in legislation, thereby providing a potentially useful suggestion to urban agencies to pass more ordinances around that time to ensure faster progress in the future.
- The smart city characteristic receiving the least attention in both sessions was on "smart people", which could serve as an input to urban agencies to give greater attention to its aspects such as 21st century education.
- The characteristic of "smart living" got the maximum attention in the 2006-2009 session, but dropped to 2nd place in the 2010-2013 session having fewer ordinances on it passed than in the earlier session, thus offering a potential suggestion to give it more importance unless the public seems really satisfied.
- The "smart governance" characteristic topped the list overall, receiving greater attention in the 2010-2013 session with 101 ordinances passed, which is a good observation and should be well-maintained by urban agencies henceforth.

Note that these observations and the related suggestions are intended to support the future decisions of the urban management agencies while helping them assess their current performance. The disclaimer is that the analysis in this paper does not actually translate to making decisions for these agencies, it would be up to their discretion. However, these suggestions would help in heading more towards smart cities, further corroborated with public opinion as needed.

It is also to be noted that the methods in our analysis here relate each ordinance with just one smart city characteristic. While this discovers interesting knowledge, it presents some limitations, as many ordinances could potentially relate to multiple smart city characteristics. This would be addressed in future work. An integration of clustering and association rule mining could probably be helpful here.

Future work would also include mining social media data to discover knowledge from public opinion on ordinances. This would constitute sentiment analysis to assess the satisfaction of the public on urban policy in a general context and with particular emphasis on smart city characteristics. All this work is geared towards decision support for urban management agencies, especially in providing inputs to build and enhance smart cities.

References

[1] *Law vs Ordinance | Difference Between*, Differencebetween.info, 2016

[2] *New York City Council*, Legistar.council.nyc.gov, 2016

[3] *Smart Cities Council | Definitions and Overviews*, Smartcitiescouncil.com, 2015.

[4] Vienna University of Technology (TU Wien), *European Smart Cities*, Technical Report, Vienna, Austria, 2015.

[5] N. Tandon, G. de Melo, F. Suchanek, G. Weikum, "WebChild: Harvesting and Organizing Commonsense Knowledge from the Web", *ACM WSDM*, Feb 2014, pp. 523-532.

[6] A. Varde, N. Tandon, S. Nag Chowdhury, G. Weikum, *Common Sense Knowledge in Domain-Specific Knowledge Bases*, Technical Report, Max Planck Institute for Informatics, Saarbruecken, Germany, Aug 2015.

[7] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques,* Morgan Kaufmann, 3rd Edition, © Elsevier 2012.

[8] M. de Jong, S. Joss, D. Schraven, C. Zhan and M. Weijnen, "Sustainable-Smart-Resilient-Low Carbon Eco-Knowledge Cities; Making sense of a multitude of concepts promoting sustainable urbanization", *Journal of Cleaner Production,* Elsevier, Dec 2015, 109:25-38.

[9] O. Emebo, A. Varde, D. Olawande, "Common Sense Knowledge, Ontology and Text Mining for Implicit Requirements", *DMIN*, Jul 2016, ISBN: 1-60132-431-6, CSREA Press, pp. 146 – 152.

[10] R. Taylor, J. Carandang, C. Alexander, J. Callega, "Making Global Cities Sustainable: Urban Rooftop Hydroponics for Diversified Agriculture in Emerging Economies", *OIDA International Journal of Sustainable Development*, Dec 2012, 5(7):11-28.

[11] R. Nagy, B. Lockaby, "Urbanization in the Southeastern United States: Socioeconomic Forces and Ecological Responses along an Urban-Rural Gradient". *Journal of Urban Ecosystems*, 2010, 14(1): 71-86.

[12] A. Pampoore-Thampi, A. Varde, D. Yu, "Mining GIS Data to Predict Urban Sprawl", *ACM KDD (Bloomberg Track)* Aug 2014, pp. 118-125.

[13] X. Du, O. Emebo, A. Varde, N. Tandon, S. Nag Chowdhury, G. Weikum, "Air Quality Assessment from Social Media and Structured Data", *IEEE ICDE Workshops*, May 2016, pp. 54-59.

[14] S. Baccianella, A. Esuli, F. Sebastiani, "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining", *LREC*, May 2010, 10:2200-2204.

[15] F. Diaz, M. Gamon, J. Hofman, E. Kacaman, D. Rothschild, "Online and Social Media Data as an Imperfect Continuous Panel Survey", *PLOS ONE*, 2016, 11(1):e0145406.