



# A lessons mining system for searching references to support decision making towards sustainable urbanization

DOI:

[10.1016/j.jclepro.2018.10.244](https://doi.org/10.1016/j.jclepro.2018.10.244)

## Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Wang, J., Shen, L., Ren, Y., Ochoa, J. J., Guo, Z., Yan, H., & Wu, Z. (2019). A lessons mining system for searching references to support decision making towards sustainable urbanization. *Journal of Cleaner Production*. <https://doi.org/10.1016/j.jclepro.2018.10.244>

## Published in:

Journal of Cleaner Production

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [openresearch@manchester.ac.uk](mailto:openresearch@manchester.ac.uk) providing relevant details, so we can investigate your claim.



Manuscript Number: JCLEPRO-D-18-09141

Title: A lessons mining system for searching references to support decision making towards sustainable urbanization

Article Type: Original article

Keywords: Case-based Reasoning (CBR); Similarity matching; Lessons Mining System (LMS); Lessons learnt; Decision making; Sustainable urbanization.

Corresponding Author: Professor Liyin Shen, Ph.D.

Corresponding Author's Institution:

First Author: Jinhuan Wang

Order of Authors: Jinhuan Wang; Liyin Shen, Ph.D.; Yitian Ren; J.Jorge Ochoa; Zhenhua Guo; Hang Yan; Zezhou Wu

Abstract: The recurrence of similar problems caused by human errors in urbanization process is common throughout the world. However, the knowledge learnt from these problems should become lessons as references for decision-making to avoid the re-currence of these problems, thus the results of urbanization are sustainable. It is considered of imperative importance to incorporate the lessons experienced into the decision-making process in a way that can help foresee the potential problems and take proper measures for addressing the problems. There is little existing study on how previous lessons are mined and incorporated in foreseeing the potential problems in future. The lack of this mining mechanism presents a significant barrier for decision makers to learn from the existing lessons thus to have references of how to make better decisions for future urbanization practices. This paper presents a Lessons Mining System (LMS) to assist in mining lessons experienced from previous practices. The system includes five components, namely, Lessons-case Representation, Lessons-case Store, Lessons-case Retrieval, Lessons-case Application, and Lessons-case Update. LMS can facilitate decision makers to understand what potential problems might occur from their current actions by referring to the lessons experienced previously in similar circumstances. This understanding can help decision makers take preventive measures to mitigate the potential problems. In other words, the use of LMS can send alarming messages to decision makers about what possible problematic consequence may occur thus they can modify their actions before too late. The establishment of LMS is based on Case-based Reasoning (CBR) theory and the similarity matching principles. A demonstration of Yangwu Town is presented to show the application of the system, and the result shows that the lessons mined can provide valuable references for the government of Yangwu Town to improve their decision-making quality.

Suggested Reviewers: Chimay J Anumba Ph.D  
Florida, USA, School of Sustainable Built Environment, University of  
Florida  
anumba@ufl.edu

David G. Carmichael Ph.D  
Sydney NSW 2052 Australia, School of Civil and Environmental Engineering,  
The University of New South Wales  
d.Carmichael@unsw.edu.au

Hongqing Fan Ph.D  
Hong Kong, Department of Building and Real Estate, Hong Kong Polytechnic  
University  
hongqin.fan@polyu.edu.hk

LOW Sui Pheng Ph.D  
Singapore 117566, School of Design and Environment, National University  
of Singapore  
bdglowsp@nus.edu.sg

Vivian Tam Ph.D  
Australia., School of Computing, Engineering and Mathematics, University  
of WesternSydney  
v.tam@westernsydney.edu.au

## **A lessons mining system for searching references to support decision making towards sustainable urbanization**

**Jinhuan Wang<sup>a,b</sup>, Liyin Shen<sup>a,b,\*</sup>, Yitian Ren<sup>a,b,c</sup>, J.Jorge Ochoa<sup>d</sup>, Zhenhua Guo<sup>a,b</sup>, Hang Yan<sup>a,b</sup>, Zezhou Wu<sup>e</sup>**

<sup>a</sup> *School of Construction Management and Real Estate, Chongqing University, Chongqing, PR China*

<sup>b</sup> *International Research Center for Sustainable Built Environment, Chongqing University, Chongqing, PR China*

<sup>c</sup> *Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong*

<sup>d</sup> *School of Natural and Built Environments, University of South Australia, Adelaide, South Australia 5001, Australia*

<sup>e</sup> *Department of Construction Management and Real Estate, College of Civil Engineering, Shenzhen University, Shenzhen, PR China*

**Keywords:** Case-based Reasoning (CBR), Similarity matching, Lessons Mining System (LMS), Lessons learnt, Decision making, Sustainable urbanization.

### **Abstract**

The recurrence of similar problems caused by human errors in urbanization process is common throughout the world. However, the knowledge learnt from these problems should become lessons as references for decision-making to avoid the recurrence of these problems, thus the results of urbanization are sustainable. It is considered of imperative importance to incorporate the lessons experienced into the decision-making process in a way that can help foresee the potential problems and take proper measures for addressing the problems. There is little existing study on how previous lessons are mined and incorporated in foreseeing the potential problems in future. The lack of this mining mechanism presents a significant barrier for decision

---

makers to learn from the existing lessons thus to have references of how to make better decisions for future urbanization practices. This paper presents a Lessons Mining System (LMS) to assist in mining lessons experienced from previous practices. The system includes five components, namely, Lessons-case Representation, Lessons-case Store, Lessons-case Retrieval, Lessons-case Application, and Lessons-case Update. LMS can facilitate decision makers to understand what potential problems might occur from their current actions by referring to the lessons experienced previously in similar circumstances. This understanding can help decision makers take preventive measures to mitigate the potential problems. In other words, the use of LMS can send alarming messages to decision makers about what possible problematic consequence may occur thus they can modify their actions before too late. The establishment of LMS is based on Case-based Reasoning (CBR) theory and the similarity matching principles. A demonstration of Yangwu Town is presented to show the application of the system, and the result shows that the lessons mined can provide valuable references for the government of Yangwu Town to improve their decision-making quality.

## **1. Introduction**

Urbanization generally refers to both the process of population concentration in urban areas and the transformation of rural areas into urban areas (Street, 1997; UN, 2010), and it becomes the driving engine for development in 21<sup>st</sup> century particularly in those developing countries. The world is in the midst of the largest wave of urbanization in history (Hodson, 2016; Liu et al., 2017). The reason for this is that urbanization can bring great scale of benefits in multiple dimensions, such as more job opportunities and incomes, better education and health conditions, better social integration, and others (Dye, 2008; Dying, 2009). The presence of these benefits has been attracting the constant flow of people from rural to urban areas. According to UNDESA (2015), the proportion of urban population in the world has increased from 29% in 1950 to 55% in 2015. It has been projected that this figure will reach to 70%

---

by 2050, indicating that 6.3 billion people will be living in cities (Ochoa et al., 2018). However, it has been widely appreciated that rapid urbanization has already induced various problems, such as air and water pollution (Li et al., 2017; Carrascal Incera et al., 2017; Sun et al., 2018), traffic congestion (Wang et al., 2017), depletion of cultivated land (Marselis et al., 2017), habitat destruction (Malico et al., 2016), the inadequate provision of urban shelter, the irrational rise of housing price (Ge, 2017), and various types of urban crimes (Alves et al., 2018). These problems present huge challenges to sustainable urban development.

Nevertheless, it is not these problems themselves but their recurrences for concerns. Air pollution, for example, is a typical recurring event, which has been resulting in multi-type health problems and death of a large number of people around the world. Major air pollution events reported include the Meuse Valley fog incident caused by exhaust fume emissions from numbers of factories in Belgium in 1930 (Yin and Yin, 2013), the Los Angeles photochemical smog episode caused by exhaust gas emissions from large number of cars and factories from 1943 to 1970 (Rani et al., 2011), the American Donora smog incident caused by harmful industrial emissions from factories in 1948 (Kamrin, 2014), the Great Smog incident of London caused by fume emissions from millions of coal stoves and local factories in 1952 (Zhang et al., 2014). In fact, the problem of air pollution becomes worse rather than improved in many cities around the world.

Water pollution, for another example, is also a typical recurring problem and has been the major reason for multiple types of diseases and ecological problems. Major water pollution events reported include the river pollution of Rasht City in Northern Iran caused by untreated domestic and industrial wastes (Noorhosseini et al., 2017), the water pollution in Bangladesh caused by the lack of effective treatments (Haque, 2017), the water pollution incidents in China caused by industrial pollutants and wastewater emissions (Wen et al., 2016). There are still other types of urban problems occurring repeatedly to threaten the sustainable development of cities, such as traffic

---

congestion caused by improper urban planning, which has induced huge social, economic, and environmental costs in many cities around the world (Christidis and Rivas, 2012; Alam and Ahmed, 2013; Yu et al., 2017).

The recurrence of these urban problems is highly associated with the poor decision quality caused by human errors. They have costed hugely to our societies, and the damages from some of these problems are irreversible. Unless these errors are learnt as lessons by decision makers and professionals and incorporated to decision-making process for addressing new problems, the recurrence of problems cannot be reduced. Therefore, it is considered valuable to foresee whether a particular type of problem may reoccur in a specific urbanization process by referring to the existing lessons received in previous similar circumstances. The decision-making quality will be improved when these lessons are referred properly, consequently, the huge scale of costs for addressing the recurrences of these problems can be saved.

Previous studies have presented various techniques for understanding how past lessons can be used as valuable references to make better decisions. Reason (2000) proposed a model called Swiss Cheese Model to assist decision makers in foreseeing potential aviation problems by learning previous accident lessons. Servos et al. (2013) investigated the lessons from environmental problems by reviewing 58 Global Environmental Facility (GEF) projects and suggested measures for decision makers to foresee the recurrence of these problems. Mannan and Waldram (2014) pointed out that it is important to understand previous the problems and the lessons received in the process of decision making when similar types of decisions are to be made, thus the recurrence of the problems can be avoided. Zhao et al. (2014) appreciated the importance of lessons learnt from past chemical accidents to help the chemical industry reduce the risk of catastrophic accidents in future. Ferjencik and Dechy (2016) concluded that the lessons learnt from the previous accidents in dynamite manufacturing plants could help prevent or mitigate those accidents that followed.

There are several typical research works contributed in examining how lessons

---

can be learnt for foreseeing the potential problems in the context of construction project management. Paranagamage et al. (2012) pointed out that there is a lack of mechanism in contractors' practices for recording and reviewing lessons to foresee the mistakes of project management. Carrillo et al. (2013) opined that learning and capturing previous lessons can help understand potential problems and contribute to improvements in project management performance. Ferrada et al. (2014) argued that the potential problems in construction project can be better foreseen if people communicate and share effectively the lessons gained from unsuccessful projects, thus performance of future projects can be improved. Duffield and Whitty (2015) investigated the adaptation of Swiss Cheese Model by project management organizations and demonstrated how past project lessons can be learnt to help foresee the problems in implementing the current projects. Ferrada et al. (2016b) introduced the concept of lessons-learning in the discipline of construction management and suggested that construction companies can better understand the possible problems in committing future projects by learning previous lessons. There are still other studies investigating the effectiveness of foreseeing potential problems by employing previous lessons (Drupsteen and Hasle, 2014; Labib and Harris, 2015; Eken et al., 2015; Labib and Read, 2015; Ferrada et al., 2016a; Love et al., 2016; Dash et al., 2016; Eric Stemna, 2017; Kim and Rhee, 2017; Suraji, 2003).

The above discussions demonstrate that the importance of incorporating the previous lessons for foreseeing the potential problems has been well appreciated, as better understanding on potential problems can help improve the quality of decision making. However, little study has been conducted in examining the method of how to mine previous lessons and how to incorporate these lessons into the decision-making process for improving the decision quality. Therefore, this paper aims to develop a Lessons Mining System (LMS) which can assist decision makers to mine previous lesson cases and help foresee the potential problems in the context of urbanization. The system does not seek to inform decision-makers 'what to do' but rather provide



---

decision makers with an avenue for understanding ‘what not to do’ in the practice of pursuing sustainable urbanization.

## 2. Research method

In order to develop a Lessons Mining System (LMS) as described in previous introduction section, the Case-based Reasoning (CBR) technique will be used as a reference tool. CBR is a method that helps extract the effective solutions adopted previously in addressing certain type of problems and apply these solutions as decision-making references to solve a new problem (Yang and Wang, 2009). The method is widely used in many fields for decision making (Chen et al., 2016a; Zhang and Dai, 2018), problems diagnosis (Gu et al., 2017; Tung et al., 2010), products design (Yang and Chen, 2011; Shen et al., 2017a), and business management (Carmona et al., 2013; Sartori et al., 2016).

CRB is not in a form of a quantitative model, instead, it is composed of five functional actions, namely, Represent, Retrieve, Reuse, Revise, and Retain (5R), which form a cyclical model (Finnie and Sun, 2003), as shown in Figure 1.

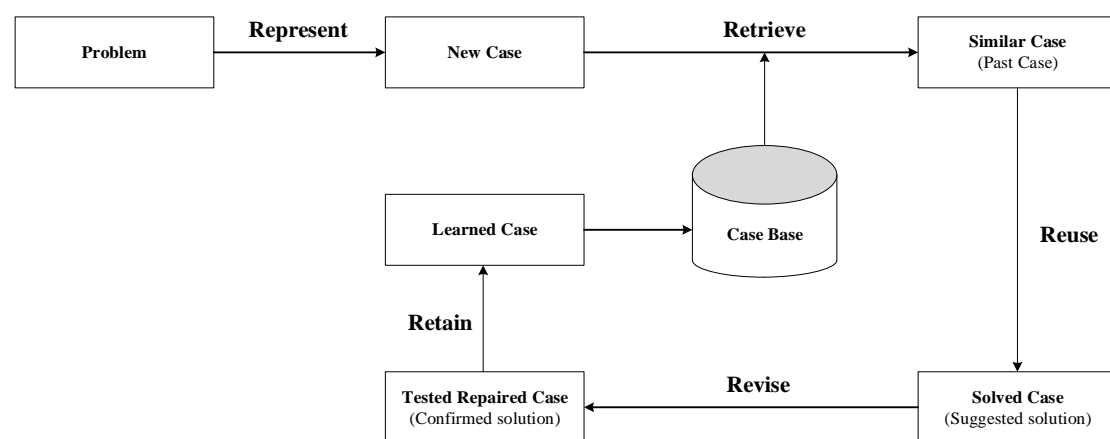


Figure 1. Case-based Reasoning (CBR) cycle

“Represent” is to structure the information about an existing case into three elements: problem, solutions and outcome. The information about these three elements needs to be organized and stored in a structured case base.

---

“Retrieve” is to mine the cases from the case base that are most relevant to a concerned case (target case) through a matching process.

“Reuse” is to identify the solutions embodied in the retrieved cases, which will be adopted as most valuable reference for making decision to address the problems in the target case.

“Revise” is to modify the solutions identified from the case base to ensure that the modified solution can be implemented effectively in the target case.

“Retain” is to store the target case in the case base as a new case, which contributes to the development of the cases in the case base.

CBR helps mine the solutions from the cases in the case base, and these cases are usually referred as solution cases which can be used as effective references for making decisions to solve new problems. However, the case base attached to CBR does not include lesson cases. CBR method therefore cannot help understand what triggers the occurrence of the problematic consequences, thus cannot help foresee the potential problems in on-going practices. In other words, CBR focuses on the resolution of existing problems and lacks attention on how to avoid the recurrence of these problems. There is no doubt that the avoidance of problems will save a lot of cost than the resolution of these problems occurred. In this study, the CBR infrastructure will therefore be reshaped into a new architecture, namely, LMS, that can help foresee the potential problems by referring to the lessons mined from previous experiences. A demonstration case will be designated and used to show the application of LMS.

### **3. The architecture of Lessons Mining System (LMS)**

Lessons are often defined as valid knowledge which is learnt from problems or failures (Dash et al., 2016; Pittman et al., 2014). There are two typical approaches to learn lessons from problems: by applying a quantitative lessons learning model (Reason, 2000; Duffield and Whitty, 2015; Labib and Read, 2015), or by creating a lessons-case database (Eken et al., 2015). The effectiveness of employing a quantita-

tive model in generating lessons has been criticized (Pritchett, 1976), whilst learning lessons by creating a lessons-case database is considered feasible and effective (Mannan and Waldram, 2014).

By referring to the CBR infrastructure presented in Figure 1, the LMS architecture is proposed, as shown in Figure 2. LMS includes five elementary components: Lessons-case Representation, Lessons-case Store, Lessons-case Retrieval, Lessons-case Application, and Lessons-case Update.

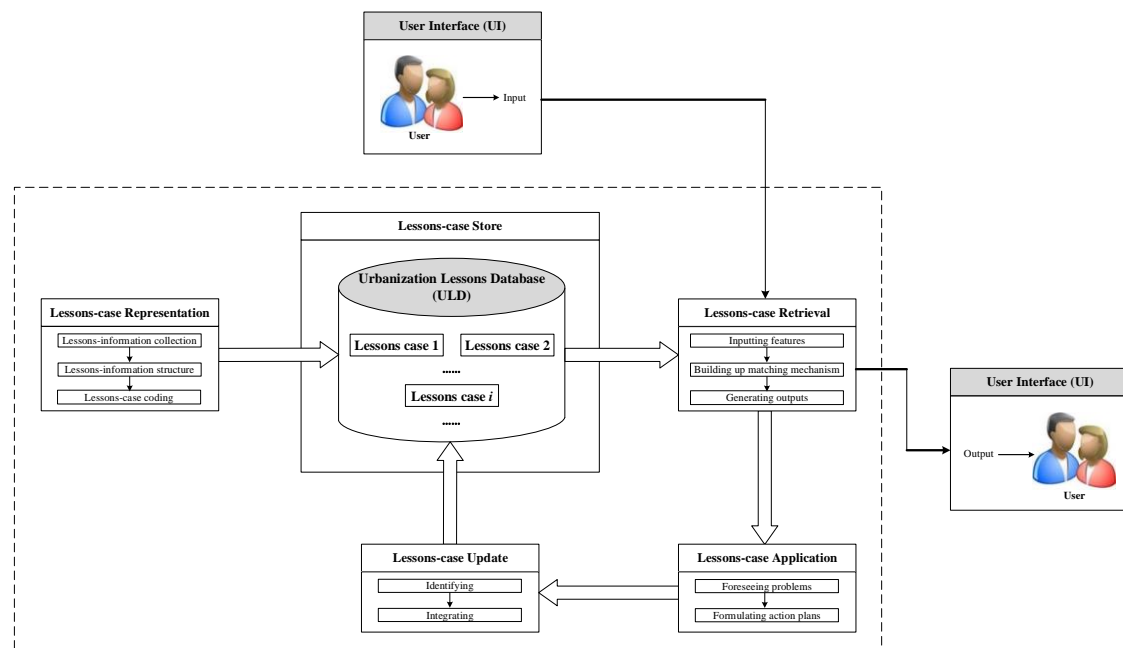


Figure 2. The architecture of Lessons Mining System (LMS)

### 3.1 Lessons-case Representation

Lessons-case Representation in LMS is to define the way in which these lesson cases are organized, including the definition of individual cases and the way the definition information is presented in case base. The specific activities of Lessons-case Representation are elaborated as follows.

#### *Lessons-information collection*

The implementation of LMS is based on the collection of sufficient lesson information associating with various problems in urbanization process. For this purpose,

a comprehensive literature review was conducted on the sources where various lessons of urban problems are reported. The sources used in this study are listed in Table 1.

Table 1. The major sources for collecting lesson information associating with urban problems

Sources	Details of information sources
Books	Ma et al. (2004); Barria and Thajchayapong (2011); Shapiro (2012).
Journal papers	Alam and Ahmed (2013); Black (2003); Brook (2008); Cao et al. (2011); Chen et al. (2016b); Chen et al. (2013); Christidis and Rivas (2012); Contreras and Ferri (2016); Dockery et al. (1993).
Regional reports	European Commission (Revel, 2011).
National reports	The Government of China (Xinhua, 2017), the Central Pollution Control Board in India (Bhawan and Nagar, 2015).
City reports	Xingtai City in China (Zhang et al., 2015), Mexico City (Moreno et al., 2008), Ulaanbaatar City in Mongolia (Amarsaikhan et al., 2014), Tirana City in Albania (Mandija and Zoga, 2012), Delhi City in India (Singh and Peshin, 2014).

### ***Lessons-information structure***

The lesson information collected need to be properly structured. The information for describing a lesson case must be accurate and concise, and the redundant information need to be removed. The information for a lesson case is structured into three parts: (1) a scenario description of the lesson case; (2) a description of the problem category; and (3) a description of the lessons learnt from the problem. The three components of a lesson case are highlighted in Figure 3.

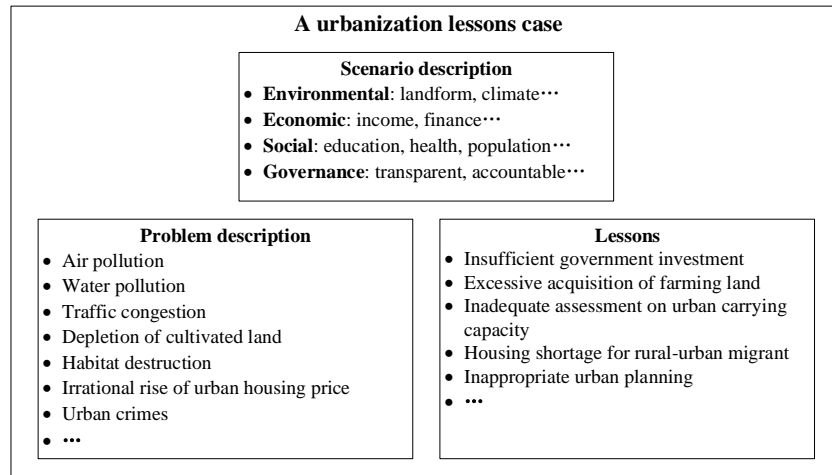


Figure 3. The structure of a lesson case

The description of the lesson scenario is to reveal the real situation in which the problems occurred. The scenario will be expressed by using a set of principal feature variables and their corresponding values. For example, as shown in Figure 3, the feature variables for describing a lesson case include four dimensions, namely, environmental characteristics such as landform and climate; economic performance such as income and finance; social development such as public education and health; and governance performance such as policy transparency and accountability (Shen et al., 2011; Zhang, 2016; Shen et al., 2017c; Shen et al., 2018; de Jong et al., 2015; Shen et al., 2017b). These four-dimension feature variables are described by four types of formats, including crisp symbol, crisp number, interval number, and fuzzy linguistic variable (Shen et al., 2017b). For example, “Crisp number” is used to measure urban population by a definite value of “2000 thousand”.

The description of problem category is to facilitate the access to each individual lesson case. Typical problems associating to urbanization are classified, for example, as shown in Figure 3, air and water pollution, traffic congestion, depletion of cultivated land, habitat destruction, the irrational rise of housing price, and urban crimes (Li et al., 2017; Carrascal Incera et al., 2017; Sun et al., 2018; Wang et al., 2017; Marselis et al., 2017; Malico et al., 2016; Ge, 2017; Alves et al., 2018).

The lessons learnt from the occurrence of problems should be described in the way that can reveal the human errors triggering the problems and alert decision mak-

ers not to commit the errors again. The representation of these lessons should enable effective learning about those human errors committed previously. For example, as shown in Figure 3, these lessons can be insufficient government investment, excessive acquisition of farming land, inadequate assessment on urban carrying capacity, housing shortage for rural-urban migrant, inappropriate urban planning (den Hartog et al., 2018; Chan et al., 2018; Wen et al., 2018; Bayulken and Huisingh, 2015; Puppim de Oliveira, 2013; Hu et al., 2016; Fenton et al., 2015).

### ***Lessons-case coding***

Whilst the representation of a lesson case is guided by the three-components structure shown in Figure 3, each component will be indexed by the corresponding parameters, which are named as scenario parameter, problem parameter, and lessons parameter. Figure 4 demonstrates the parameter-coding structure in LMS.

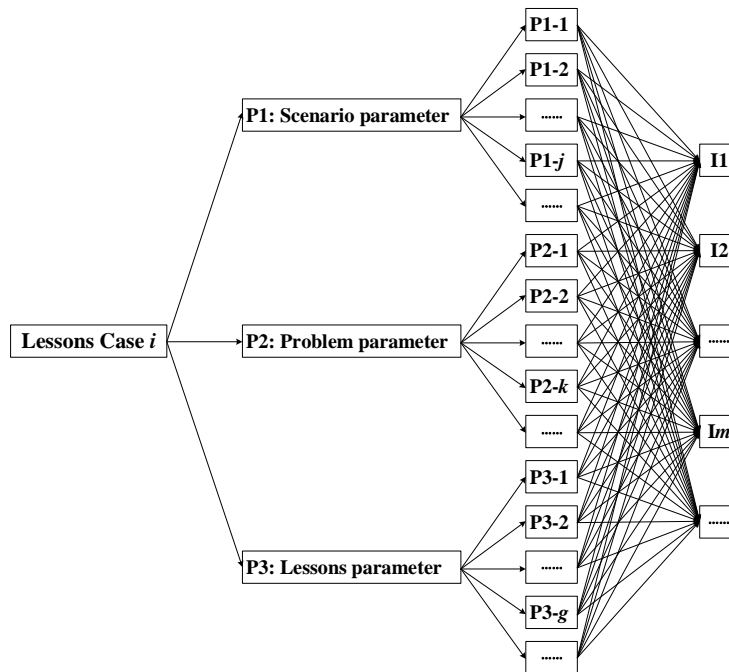


Figure 4. Parameter Coding Structure in LMS

In Figure 4, each type of parameters is defined by a group of feature variables. For example, the scenario parameter (P1) is defined by the feature variables of environmental characteristic, landform and climate, and others. Each feature variable corresponds to a unique code, for example, environmental characteristic is denoted by the

code P1-1, landform and climate is denoted by the code P1-2. The level of performance by each feature variable will be linked to the indicators that measure urbanization performance in different perspectives, as shown in Figure 4. The urbanization performance indicators are coded as I1, I2, ... Im.

### 3.2 Lessons-case Store

In referring to Figure 2, the second elementary component of LMS is to store the represented lesson cases in Urbanization Lessons Database (ULD). These represented lesson cases were standardized by indicators, parameters, semantic relations, and index pointers in ULD, as shown in Figure 5. Specific lesson cases in ULD can be indexed according to the values of indicators and parameters.

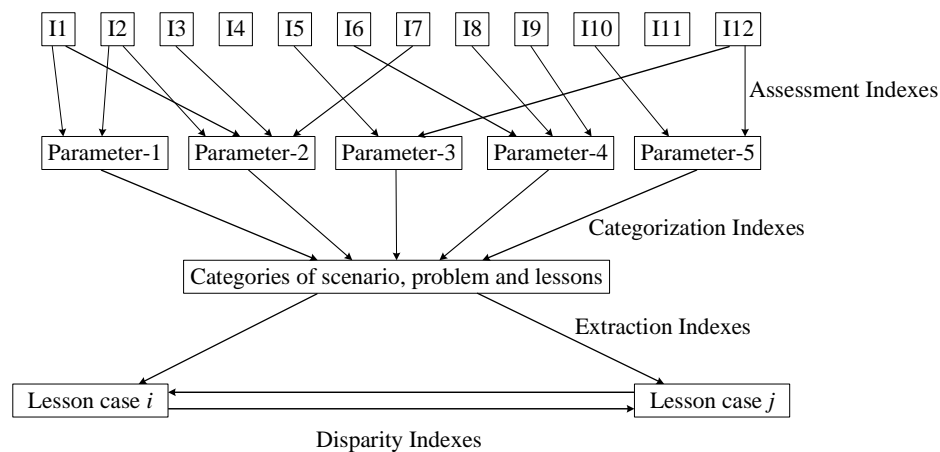


Figure 5. ULD index structure adapted from Shen et al. (2013)

There are four types of indexes in Figure 5:

(1) links from the performance indicators to parameters, named “assessment indexes”;

(2) links from parameters to categories of scenario, problem and lessons, named “categorization indexes”;

(3) links from categories of scenario, problem and lessons to the stored lesson cases, named “extraction indices”;

(4) links between the lesson cases, called “disparity indices”.

---

By using the four types of indexes, similar types of lesson cases can be grouped into one category by a group name. The individual cases within a same group form a hierarchy according to the degree of their memberships in the group. Different groups are inter-linked by a semantic network. The indexed structure of LMS can interpret a lesson case in semantic and pragmatic criteria, instead of purely syntactic ones.

### **3.3 Lessons-case Retrieval**

Lessons-case Retrieval is to mine lesson cases by establishing a matching mechanism between the stored lesson cases in ULD and the target case specified by the LMS user (decision maker). According to Figure 2, Lessons-case Retrieval includes three processes, namely, inputting features, building up matching mechanism, generating outputs.

#### ***Inputting features***

In the process of inputting features, the LMS user will refer to the target scenario and confirm the feature information of the scenario. The target scenario is a specific practice for which the LMS user wants to foresee the possible problems from the practice. The target scenario is described by feature variables and the values of these feature variables are given or defined by the user through the User Interface (UI). By following the UI, the user will be asked to select specifications of the feature variables that can best describe the scenario. For example, the value of landform can be hills, plains, plateaus, basins, etc.

#### ***Building up the matching mechanism***

Matching mechanism is to help match the target scenario with the lesson cases in ULD. In order to establish this matching mechanism, there are two research activities: the calculation of global similarity between the target scenario and all the stored lesson cases, the designation of matching coefficient for filtering out those lesson cases which have lower similarity to the target scenario.



---

Local-Global method (Cheng and Ma, 2015) is used to calculate the global similarity. For this purpose, firstly, the local similarity between target scenario and all individual lesson cases in ULD will be calculated in referring to each feature variables which are described in crisp symbol (Guo et al., 2011; Castro et al., 2009), or interval number format (Slonim and Schneider, 2001), or fuzzy linguistic format (Shiu and Pal, 2004). Secondly, the global similarity between target scenario and each individual lesson cases will be calculated by applying the weighted local similarities (Shen et al., 2017a).

Furthermore, a matching coefficient ( $\alpha$ ) needs to be designated for selecting the useful lesson cases from ULD. The designation of  $\alpha$  is based on the user's preferences and specific circumstance. Those individual lesson cases in ULD whose global similarity values with the target scenario are less than the designated matching coefficient ( $\alpha$ ) will not be shown on the User Interface (UI).

### ***Generating outputs***

By applying the matching coefficient, the lesson cases that the system user wishes to refer to will be mined and generated. There may be a number of mined lesson cases, thus it is necessary to establish the ranks between the cases in order to differentiate the suitability of each mined case for reference. The rank will be established according to the global similarity, and the top cases are considered to be able to provide more valuable references for making decisions in response to the target scenario. On the other hand, the number of mined cases should be sufficient in order to provide effective information, and this can be achieved by adjusting the value of matching coefficient.

### **3.4 Lessons-case Application**

According to the model Lesson-case Application in Figure 2, the mined lesson cases will provide users with two major functions:

#### ***Foreseeing potential problems***

---

Following the Lessons-case Retrieval process in section 3.3, a number of urbanization lesson cases will be mined and ranked in referring to the features of the target scenario. These mined cases have experienced various types of problems, coded by P2 in LMS. The LMS user can foresee whether these problems will occur to the target scenario for which he is addressing. In the context of urbanization practice, the LMS can help decision makers or practitioners in a specific urban environment to foresee the potential problems, if any. This assessment will allow the decision makers implement precautionary measures to avoid the reoccurrence of these problems.

### ***Formulating action plans***

The lessons learnt from each mined lesson case provide important references for the LMS user to make better decisions, thus these problems embodied in the mined cases will not occur in the target scenario. As these mined lessons are different in nature, such as human errors, natural disasters and others, the user can formulate different action plans. As these action plans are designated based on the lessons learnt under similar circumstances, their good effectiveness in application is expected. In other words, the implementation of these action plans can reduce the possibility of the reoccurrence of problems in the future urbanization development.

### **3.5 Lessons-case Update**

According to the model in Figure 2, the function of Lessons-case Update is to integrate new lessons into the ULD. When new problems are encountered in future practices, the lessons from these new problems should be properly examined and described for the inclusion in ULD. In order to avoid the repetition of similar lessons in ULD, it is necessary to filter these new lessons before integrating them in the database by using the established parameter indexes. Only these lessons which have different values in these parameters are stored as new lesson cases. Accordingly, the ULD will be updated.

---

#### **4. Demonstration**

In this section, a demonstration is used to show the procedures of applying the introduced LMS by users (decision makers). In applying the system LMS, the users can foresee the potential problems by referring previous lesson cases and implement proper actions for avoiding the reoccurrence of the related problems in the practice of urbanization. The process of the demonstration is illustrated in Figure 6, in which the target scenario is named Yangwu Town ( $C_0$ ), located in Guizhou province in China. The scenario case has abundant natural resources, such as land resources, and various kinds of agricultural and forestry products. The government of this town is developing the township by infrastructure construction and industrial upgrading and wishes to know the potential problems if any from such development.

##### **The inputs of feature variables**

According to the description of lesson scenario in section 3.1, as shown in Figure 6, there are four dimensions in describing the scenario of a lesson case: environmental, economic, social and governmental dimensions. Each dimension is described by multiple scenario parameters. In this demonstration, eleven feature variables are identified in four dimensions. These feature variables are coded accordingly, as shown in Figure 6. For example, the feature variable “gross output value of primary industry” is coded as P1-1, the feature variable “gross output value of tertiary industry” is coded as P1-2, the feature variable “per capita disposable income of permanent residents” is coded as P1-3, etc. The values of these variables need to be provided by the government of Yangwu Town ( $C_0$ ). For demonstration, the values of these variables are set as follows: P1-1= 19.86, P1-2= 31.50, P1-3= 0.97, as shown in Figure 6.

##### **Lessons-case retrieval**

In Figure 6, the Lessons-case Retrieval process is conducted by matching the target scenario with all the stored lesson cases in ULD. In this demonstration, ten lesson cases are stored in the ULD, including Jianggao Town ( $C_1$ ), Dalucao Town ( $C_2$ ), Yaoshan Town ( $C_3$ ), Boguo Town ( $C_4$ ), Hengdian Town ( $C_5$ ), Shengze Town ( $C_6$ ),

---

Yanjiao Town ( $C_7$ ), Lushi Town ( $C_8$ ), Jichang Town ( $C_9$ ), Nanming Town ( $C_{10}$ ). The values of the variables in each individual scenario of lesson cases are illustrated in Figure 6.

In this demonstration, all the feature variables are measured in the format of crisp number. According to the methods described in Section 3.3, the local similarity between the target scenario and lesson cases in ULD is measured by calculating the distance between the values of the two crisp numbers. The shorter the distance, the more similar the lesson case is considered to the targeted scenario. The local similarity can be obtained through the following formula (Guo et al., 2011; Castro et al., 2009):

$$\text{Sim}(P_{ij}, P_{0j}) = 1 - |P_{ij} - P_{0j}| / (\beta - \alpha), \quad P_{ij}, P_{0j} \in [\alpha, \beta] \quad (1)$$

where  $\text{Sim}(P_{ij}, P_{0j})$  is the similarity between lesson case  $C_i$  and targeted case  $C_0$  in regard to the numerical feature  $P_j$ .  $\alpha$  and  $\beta$  are the lower and upper bounds of the range, respectively.

The global similarity between the target scenario and lesson cases in ULD can be obtained by integrating local similarity of all features:

$$\text{Sim}(C_i, C_0) = \frac{1}{11} \sum_{j=1}^{11} w_j \text{Sim}(Q_{ij}, Q_{0j}) \quad (2)$$

Where  $\text{Sim}(C_i, C_0)$  denotes the global similarity between the lesson case  $C_i$  and target case  $C_0$ .  $\text{Sim}(Q_{ij}, Q_{0j})$  denotes the local similarity between lesson case  $C_i$  and target scenario  $C_0$  in regard to the identified features which can be obtained by using equations (1).  $w_j$  denotes the weighting value of feature  $P_j$ . The weighting of each feature variable in this demonstration will be determined by using the Equal Weight (EW) method, the value of  $w_j$  is equal to 1.0. As a result, the calculations on similarity are listed in Table 2.

For demonstration, the value of the matching coefficient  $\alpha$  is designated as 0.7. By applying this matching coefficient, all the individual lesson cases in ULD whose global similarity to Yangwu Town are less than 0.7 will not be shown on the User In-

---

terface (UI). As a result, the following lesson cases are mined:  $C_2$  ( $\alpha = 0.8734$ ),  $C_3$  ( $\alpha = 0.8849$ ),  $C_4$  ( $\alpha = 0.7008$ ),  $C_8$  ( $\alpha = 0.7897$ ),  $C_9$  ( $\alpha = 0.8306$ ),  $C_{10}$  ( $\alpha = 0.8593$ ).

### **Ranking between the mined lesson cases**

The mined cases are ranked according to their global similarity degree with the target scenario, and the results are shown in Figure 6. For example, Yaoshan Town ( $C_3$ ) ranks first with the global-similarity value of 0.8849.

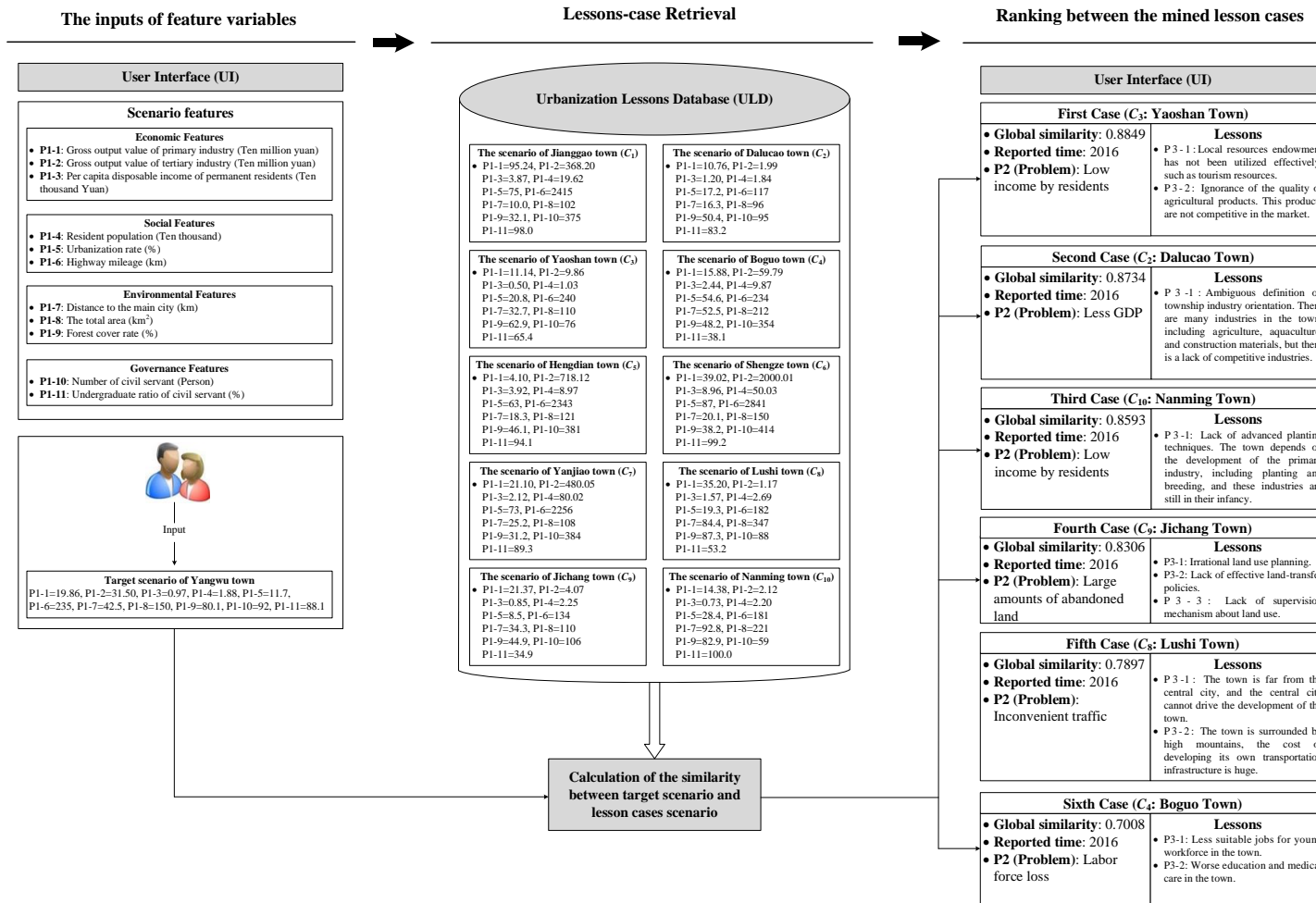


Figure 6. Lessons-mining demonstration

Table 2. The global similarity between target scenario and all the lesson cases in ULD

Lesson cases Similarity		$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$	$C_{10}$
<b>Local similarity</b>	<b>P1-1</b>	0.1729	0.9002	0.9044	0.9563	0.8271	0.7899	0.9874	0.8316	0.9834	0.9399
	<b>P1-2</b>	0.8316	0.9852	0.9892	0.9858	0.6565	0.0152	0.7756	0.9848	0.9863	0.9852
	<b>P1-3</b>	0.6576	0.9730	0.9441	0.8368	0.6513	0.0559	0.8641	0.9286	0.9864	0.9714
	<b>P1-4</b>	0.7754	0.9995	0.9893	0.8998	0.9103	0.3907	0.0108	0.9898	0.9953	0.9959
	<b>P1-5</b>	0.1936	0.9299	0.8841	0.4545	0.3439	0.0408	0.2191	0.9032	0.9592	0.7873
	<b>P1-6</b>	0.1997	0.9567	0.9982	0.9996	0.2263	0.0433	0.2581	0.9805	0.9629	0.9801
	<b>P1-7</b>	0.6075	0.6836	0.8816	0.8792	0.7041	0.7283	0.7886	0.4988	0.8973	0.3925
	<b>P1-8</b>	0.8175	0.7935	0.8470	0.7639	0.8891	1.0000	0.8394	0.2436	0.8472	0.7285
	<b>P1-9</b>	0.1429	0.4714	0.6941	0.4386	0.3929	0.2500	0.1250	0.8750	0.3732	0.9482
	<b>P1-10</b>	0.2028	0.9915	0.9549	0.2620	0.1859	0.0930	0.1775	0.9887	0.9606	0.9071
	<b>P1-11</b>	0.8464	0.9232	0.6467	0.2320	0.9078	0.8310	0.9846	0.4624	0.1843	0.8157
<b>Global similarity</b>		0.4953	0.8734	0.8849	0.7008	0.6087	0.3853	0.5482	0.7897	0.8306	0.8593

---

## 5. Discussion

The above demonstration suggests that the introduced system LMS is applicable and can assist decision-makers in foreseeing the potential problems if any from the current urbanization practice by mining previous lessons. These lessons are important references for designing better decision to promote sustainable urbanization. The key procedures in using the system can be highlighted as follows:

*Building up an adequate database.* The demonstration case highlights the importance of building up an adequate Urbanization Lessons Database (ULD) where there must be sufficient number of lesson cases recorded. This way can ensure that valuable information will be mined for decision references. The lesson cases can be selected from official databases, reports, and journal papers. For example, the lesson cases can be extracted from the Lessons Learnt Database (LLD) designed for construction companies and projects (Ferrada et al., 2016a; Eken et al., 2015).

*Data input for feature variables about the target scenario.* In order to effectively mine lessons, the system users must input adequate information about their own cases (target case). For example, the Yangwu town government in the demonstration case must be clear about the scenario it encounters and input the values of case parameters, such as gross output value of primary industry, gross output value of tertiary industry, per capita disposable income of permanent residents, resident population, urbanization rate, etc.

*Establishing the match mechanism between the target scenario and all the recorded lesson cases in the database ULD.* For establishing the proper matching mechanism, the key issue is to ensure the adequacy of similarity calculation. The calculation in this study is based on the Local-Global method. It is therefore considered that the matching process built in LMS is reliable.

*Foreseeing the potential problems and planning proper actions.* The mined lesson cases are important references for decision makers to foresee the potential problems from the urbanization practices they are addressing. For example, in the



---

above demonstration, the major problems experienced by the top-three mined cases is related to weak performance of economy, such as low income by township residents in Yaoshan Town and Nanming Town, and less GDP in Dalucao Town. Therefore, it is considered that these problems are likely to reoccur in the target scenario of Yangwu. The government of Yangwu Town should be aware that the low-income problem by residents in Yaoshan Town may lead to serious consequences, such as poverty by the residents, and fiscal deficit by the local government. There is no doubt that the cost of solving these serious consequences is huge. On the other hand, the lessons provided by the mined lesson cases would facilitate decision makers to take proper measures for addressing the problems foreseen to target scenario. For the case of the demonstration, the lessons provided by the top-three mined lesson cases include “Local resources endowment has not been utilized effectively, such as tourism resources”, “Ignorance of the quality of agricultural products”, “Ambiguous definition on township industry orientation”, and “Lack of advanced planting techniques”. It is considered that Yangwu Town will be presented with the problem of weak economy if similar kinds of human errors are committed. Therefore, the government of Yangwu Town is advised to implement proper policies in advance, such as enabling the effective utilization of resources endowment, recruiting experts to help famers in the process of agricultural production, attracting talents to assume the role of consults for making better township planning for adjusting the local industry structure, and developing township jointly with external investors.

## **6. Conclusions**

This paper introduces the system LMS as a new mechanism for foreseeing the potential problems in a specific urbanization practice by referring to the previous lessons learnt. The application of the LMS will generate these valuable lessons learnt previously thus can support decision-making process in developing better actions for practicing sustainable urbanization. In other words, the decision has better

---

quality in producing measures to ensure that the previous problems will not happen again in the practice where the decision makers are facing because that the decision is based on sufficient lessons learnt under similar circumstances. Without this lessons mining mechanism, previous lessons cannot be truly learnt, thus many problems have been recurring in practice. The demonstration in the paper tells that the system is effective and applicable.

LMS opens the door to the lessons sharing on the urban problems or failure practices. It represents an innovative mechanism in conducting research in the discipline of sustainable urbanization. This mechanism enables to carry out comparative analysis on a wide range of past lesson practices, thus obtain valuable information to assist decision-makers in formulating better actions for promoting sustainable urbanization. The application of the system LMS enables to foresee potential problems from the current urbanization practice.

The limitations of LMS in its current stage are implicit in its Urbanization Lessons Database (ULD), which included limited number of lesson cases for empirical evidence. It is appreciated that empirical lesson cases, particularly about the outcome from the implementations of lessons-learnt decisions to new scenarios, cannot be easily obtained. It is therefore the intention of this research team to extend the research to empirical study in the future.

### **Acknowledgment**

This paper was supported by the National Social Science Foundation of China (Grant No. “15BJY038” and “17ZDA062”), Social Science Foundation of Chongqing university (Grant No. “2017CDJSK03PT26”), and Fundamental Research Funds for the Central Universities (Grant No. “2018CDXYJG0047”).

### **References**

ALAM M A, AHMED F 2013. Urban Transport Systems and Congestion: A Case Study of Indian Cities. Transport & Communications

---

Bulletin for Asia & the Pacific [J], 82.

ALVES L G A, RIBEIRO H V, RODRIGUES F A 2018. Crime prediction through urban metrics and statistical learning. *Physica A: Statistical Mechanics and its Applications* [J], 505: 435-443.

AMARSAIKHAN D, BATTSENGEL V, NERGUI B, et al. 2014. A Study on Air Pollution in Ulaanbaatar City, Mongolia. *Journal of Geoscience and Environment Protection* [J], 2: 123-128.

BARRIA J A, THAJCHAYAPONG S 2011. Detection and Classification of Traffic Anomalies Using Microscopic Traffic Variables [M]. IEEE Press.

BAYULKEN B, HUISINGH D 2015. Are lessons from eco-towns helping planners make more effective progress in transforming cities into sustainable urban systems: a literature review (part 2 of 2). *Journal of Cleaner Production* [J], 109: 152-165.

BHAWAN P, NAGAR E A 2015. Central Pollution Control Board. Central Pollution Control Board Cpcb Cpcb India [J].

BLACK J 2003. Intussusception and the great smog of London, December 1952. *Archives of Disease in Childhood* [J], 88: 1040.

BROOK R D 2008. Cardiovascular effects of air pollution. *Nature Clinical Practice Cardiovascular Medicine* [J], 115: 175.

CAO J, YANG C, LI J, et al. 2011. Association between long-term exposure to outdoor air pollution and mortality in China: A cohort study. *Journal of Hazardous Materials* [J], 186: 1594.

CARMONA M A, BARBANCHO J, LARIOS D F, et al. 2013. Applying case based reasoning for prioritizing areas of business management. *Expert Systems with Applications* [J], 40: 3450-3458.

CARRASCAL INCERA A, AVELINO A F T, FRANCO SOL S A 2017. Gray water and environmental externalities: International patterns of water pollution through a structural decomposition analysis. *Journal of Cleaner Production* [J], 165: 1174-1187.

CARRILLO P, RUIKAR K, FULLER P 2013. When will we learn? Improving lessons learned practice in construction. *International Journal of Project Management* [J], 31: 567-578.

CASTRO J L, NAVARRO M, SNCHEZ J M, et al. 2009. Loss and gain functions for CBR retrieval. *Information Sciences* [J], 179: 1738-1750.

CHAN F K S, CHUAH C J, ZIEGLER A D, et al. 2018. Towards resilient flood risk management for Asian coastal cities: Lessons learned from Hong Kong and Singapore. *Journal of Cleaner Production* [J], 187: 576-589.

CHEN S, YI J, JIANG H, et al. 2016a. Ontology and CBR based automated decision-making method for the disassembly of mechanical products. *Advanced Engineering Informatics* [J], 30: 564-584.

---

CHEN X, ZHANG L W, HUANG J J, et al. 2016b. Long-term exposure to urban air pollution and lung cancer mortality: A 12-year cohort study in Northern China. *Science of the Total Environment* [J], 571: 855.

CHEN Y, EBENSTEIN A, GREENSTONE M, et al. 2013. Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences of the United States of America* [J], 110: 12936.

CHENG J C P, MA L J 2015. A non-linear case-based reasoning approach for retrieval of similar cases and selection of target credits in LEED projects. *Building and Environment* [J], 93: 349-361.

CHRISTIDIS P, RIVAS J N I 2012. Measuring road congestion. *Jrc Working Papers* [J].

CONTRERAS L, FERRI C 2016. Wind-sensitive Interpolation of Urban Air Pollution Forecasts. *Procedia Computer Science* [J], 80: 313-323.

DASH A K, BHATTCHARJEE R M, PAUL P S 2016. Lessons Learnt from Indian Inundation Disasters: An Analysis of Case Studies. *International Journal of Disaster Risk Reduction* [J], 20.

DE JONG M, JOSS S, SCHRAVEN D, et al. 2015. Sustainable-smart-resilient-low carbon-eco-knowledge cities; making sense of a multitude of concepts promoting sustainable urbanization. *Journal of Cleaner Production* [J], 109: 25-38.

DEN HARTOG H, SENEGERS F, XU Y, et al. 2018. Low-carbon promises and realities: Lessons from three socio-technical experiments in Shanghai. *Journal of Cleaner Production* [J], 181: 692-702.

DOCKERY D W, RD P C, XU X, et al. 1993. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine* [J], 329: 1753.

DRUPSTEEN L, HASLE P 2014. Why do organizations not learn from incidents? Bottlenecks, causes and conditions for a failure to effectively learn. *Accident Analysis & Prevention* [J], 72: 351-358.

DUFFIELD S, WHITTY S J 2015. Developing a systemic lessons learned knowledge model for organisational learning through projects. *International Journal of Project Management* [J], 33: 311-324.

DYE C 2008. Health and urban living. *Science* [J], 319: 766-769.

DYING A C 2009. Are Cities Dying ? *Journal of Economic Perspectives* [J], 12: 139-160.

EKEN G, BILGIN G, DIKMEN I, et al. 2015. A Lessons Learned Database Structure for Construction Companies ☆. *Procedia Engineering* [J], 123: 135-144.

ERIC STEMNA C B, DAVID CLIFF, MAUREEN E. HASSALL 2017. Failure to learn from safety incidents: Status, challenges and opportunities.

---

Safety Science [J].

FENTON P, GUSTAFSSON S, IVNER J, et al. 2015. Sustainable Energy and Climate Strategies: lessons from planning processes in five municipalities. *Journal of Cleaner Production* [J], 98: 213-221.

FERJENCIK M, DECHY N 2016. Three accidents in European dynamite production plants: An attempt to improve the external lessons learning. *Journal of Loss Prevention in the Process Industries* [J], 44: 12-23.

FERRADA X, N EZ D, NEYEM A, et al. 2016a. A Cloud-based Mobile System to Manage Lessons-learned in Construction Projects ☆. *Procedia Engineering* [J], 164: 135-142.

FERRADA X, N EZ D, NEYEM A, et al. 2016b. A Lessons-learned System for Construction Project Management: A Preliminary Application ☆. *Procedia - Social and Behavioral Sciences* [J], 226: 302-309.

FERRADA X, SEP LVEDA M, SERPELL A, et al. 2014. A Lessons-Learned Mobile System for Construction Companies: Motivation and Design [C] //; City. 157-165.

FINNIE G, SUN Z 2003. R5 model for case-based reasoning. *Knowledge-Based Systems* [J], 16: 59-65.

GE J 2017. Endogenous rise and collapse of housing price: An agent-based model of the housing market. *Computers, Environment and Urban Systems* [J], 62: 182-198.

GU D, LIANG C, ZHAO H 2017. A case-based reasoning system based on weighted heterogeneous value distance metric for breast cancer diagnosis. *Artificial Intelligence in Medicine* [J], 77: 31.

GUO Y, HU J, PENG Y 2011. Research on CBR system based on data mining. *Applied Soft Computing* [J], 11: 5006-5014.

HAQUE N 2017. Exploratory analysis of fines for water pollution in Bangladesh. *Water Resources & Industry* [J], 18: 1-8.

HODSON R 2016. Urban health and well-being. *Nature* [J], 531: S49.

HU M-C, WADIN J L, LO H-C, et al. 2016. Transformation toward an eco-city: lessons from three Asian cities. *Journal of Cleaner Production* [J], 123: 77-87.

KAMRIN M A 2014. Donora: Air Pollution Episode. *Encyclopedia of Toxicology* [J]: 222-223.

KIM E, RHEE M 2017. How airlines learn from airline accidents: An empirical study of how attributed errors and performance feedback affect learning from failure. *Journal of Air Transport Management* [J], 58: 135-143.

LABIB A, HARRIS M J 2015. Learning how to learn from failures: The Fukushima nuclear disaster ☆. *Engineering Failure Analysis* [J], 47:

---

117-128.

LABIB A, READ M 2015. A hybrid model for learning from failures [M]. Pergamon Press, Inc.

LI S, FENG K, LI M 2017. Identifying the main contributors of air pollution in Beijing. *Journal of Cleaner Production* [J], 163: S359-S365.

LIU J, LOW S P, WANG L F 2017. Critical success factors for eco-city development in China. *International Journal of Construction Management* [J]: 1-10.

LOVE P E D, TEO P, DAVIDSON M, et al. 2016. Building absorptive capacity in an alliance: Process improvement through lessons learned. *International Journal of Project Management* [J], 34: 1123-1137.

MA J, LIU N Y, SULLIVAN L R 2004. China's water crisis [M]. EastBridge.

MALICO I, CARRAJOLA J, GOMES C P, et al. 2016. Biomass residues for energy production and habitat preservation. Case study in a montado area in Southwestern Europe. *Journal of Cleaner Production* [J], 112: 3676-3683.

MANDIJA F, ZOGA P 2012. Air pollution assessment on city of Tirana [C] //; City. 790.

MANNAN M S, WALDRAM S P 2014. Learning lessons from incidents: A paradigm shift is overdue. *Process Safety & Environmental Protection* [J], 92: 760-765.

MARSELIS S M, FENG K, LIU Y, et al. 2017. Agricultural land displacement and undernourishment. *Journal of Cleaner Production* [J], 161: 619-628.

MORENO T, QUEROL X, PEY J, et al. 2008. Spatial and temporal variations in inhalable CuZnPb aerosols within the Mexico City pollution plume. *Journal of Environmental Monitoring* [J], 10: 370.

NOORHOSSEINI S A, ALLAHYARI M S, DAMALAS C A, et al. 2017. Public environmental awareness of water pollution from urban growth: The case of Zarjub and Goharrud rivers in Rasht, Iran. *Science of the Total Environment* [J], s 599–600: 2019-2025.

OCHOA J J, TAN Y, QIAN Q K, et al. 2018. Learning from best practices in sustainable urbanization. *Habitat International* [J].

PARANAGAMAGE P, CARRILLO P, RUIKAR K, et al. 2012. Lessons learned practices in the UK construction sector: current practice and proposed improvements. *Engineering Project Organization Journal* [J], 2: 216-230.

PITTMAN W, ZHE H, HARDING B, et al. 2014. Lessons to be learned from an analysis of ammonium nitrate disasters in the last 100 years. *Journal of Hazardous Materials* [J], 280: 472-477.

PRITCHETT D V 1976. Econometric policy evaluation: A critique [C]

---

//; City. 63-64.

PUPPIM DE OLIVEIRA J A 2013. Learning how to align climate, environmental and development objectives in cities: lessons from the implementation of climate co-benefits initiatives in urban Asia. *Journal of Cleaner Production* [J], 58: 7-14.

RANI B, SINGH U, CHUHAN A K, et al. 2011. Photochemical smog pollution and its mitigation measures. *Journal of Advanced Scientific Research* [J].

REASON J 2000. Human error: models and management. *Western Journal of Medicine* [J], 172: 393.

REVEL D 2011. A Roadmap for Moving to a Competitive Low Carbon Economy in 2050.

SARTORI F, MAZZUCHELLI A, GREGORIO A D 2016. Bankruptcy Forecasting Using Case-Based Reasoning: the CRePERIE Approach. *Expert Systems with Applications* [J], 64: 400-411.

SERVOS M R, MUNKITTRICK K R, CONSTANTIN G, et al. 2013. Science and management of transboundary lakes: Lessons learned from the global environment facility program. *Environmental Development* [J], 7: 17-31.

SHAPIRO J 2012. *China's environmental challenges* [M]. Polity Press.

SHEN L, REN Y, XIONG N, et al. 2018. Why small towns can not share the benefits of urbanization in China? *Journal of Cleaner Production* [J], 174: 728-738.

SHEN L, YAN H, FAN H, et al. 2017a. An integrated system of text mining technique and case-based reasoning (TM-CBR) for supporting green building design. *Building & Environment* [J], 124.

SHEN L, YAN H, ZHANG X, et al. 2017b. Experience mining based innovative method for promoting urban sustainability. *Journal of Cleaner Production* [J], 156.

SHEN L, ZHANG Z, ZHANG X, et al. 2017c. Measuring incoordination-adjusted sustainability performance during the urbanization process: Spatial-dimensional perspectives. *Journal of Cleaner Production* [J], 143: 731-743.

SHEN L Y, OCHOA J J, SHAH M N, et al. 2011. The application of urban sustainability indicators – A comparison between various practices. *Habitat International* [J], 35: 17-29.

SHIU S, PAL S K 2004. *Foundations of Soft Case-Based Reasoning* [M]. John Wiley & Sons.

SINGH S, PESHIN S K 2014. Air Pollution Scenario over Delhi City. *Environment & Sustainable Development* [J]: 77-85.

SLONIM T Y, SCHNEIDER M 2001. *Design issues in fuzzy case-based reasoning* [M]. Elsevier North-Holland, Inc.

---

STREET P 1997. Scenario workshops: A participatory approach to sustainable urban living? *Futures* [J], 29: 139-158.

SUN C, LUO Y, LI J 2018. Urban traffic infrastructure investment and air pollution: Evidence from the 83 cities in China. *Journal of Cleaner Production* [J], 172: 488-496.

SURAJI A 2003. Construction Accidents and the Lesson Learnt from 1000 Cases. *International Journal of Construction Management* [J], 3: 41-49.

TUNG Y H, TSENG S S, WENG J F, et al. 2010. A rule-based CBR approach for expert finding and problem diagnosis. *Expert Systems with Applications* [J], 37: 2427-2438.

UN 2010. World urbanization prospects: the 2009 Revision.

UNDESA 2015. World urbanization prospects - the 2015 revision. New York: Department of economic and social affairs [J].

WANG Y, CAO J, LI W, et al. 2017. Exploring traffic congestion correlation from multiple data sources. *Pervasive and Mobile Computing* [J], 41: 470-483.

WEN Z, HU Y, LEE J C K, et al. 2018. Approaches and policies for promoting industrial park recycling transformation (IPRT) in China: Practices and lessons. *Journal of Cleaner Production* [J], 172: 1370-1380.

WEN Z, MENG F, DI J, et al. 2016. Technological approaches and policy analysis of integrated water pollution prevention and control for the coal-to-methanol industry based on Best Available Technology. *Journal of Cleaner Production* [J], 113: 231-240.

XINHUA 2017. The notification of Ecological Environment of Qilian Mountain National Nature Reserve in Gansu Province from General Office of the CPC Central Committee and State Council General Circular [http://www.gov.cn/xinwen/2017-07/20/content\\_5212107.htm](http://www.gov.cn/xinwen/2017-07/20/content_5212107.htm) [J].

YANG C J, CHEN J L 2011. Accelerating preliminary eco-innovation design for products that integrates case-based reasoning and TRIZ method. *Journal of Cleaner Production* [J], 19: 998-1006.

YANG H L, WANG C S 2009. Recommender system for software project planning one application of revised CBR algorithm. *Expert Systems with Applications* [J], 36: 8938-8945.

YIN S X, YIN J 2013. Atmospheric Particulates Pollution Control in Europe. *Global Science Technology & Economy Outlook* [J].

YU S C, YONG J L, CHOI S S B 2017. Is there more traffic congestion in larger cities? -Scaling analysis of the 101 largest U.S. urban centers. *Transport Policy* [J], 59: 54-63.

ZHANG, ENZHONG, YANG, et al. 2015. Analysis of Meteorological Causes of Serious Air Pollution in Xingtai City Based on Inversion Layer. *Meteorological & Environmental Research* [J]: 5-9.

ZHANG D, LIU J, LI B 2014. Tackling Air Pollution in China—What



---

do We Learn from the Great Smog of 1950s in LONDON. Sustainability [J], 6: 5322-5338.

ZHANG H, DAI G L 2018. Research on traffic decision making method based on image analysis case based reasoning. Optik - International Journal for Light and Electron Optics [J], 158.

ZHANG X 2016. Sustainable urbanization: a bi-dimensional matrix model. Journal of Cleaner Production [J], 134: 425-433.

ZHAO J, SUIKKANEN J, WOOD M 2014. Lessons learned for process safety management in China. Journal of Loss Prevention in the Process Industries [J], 29: 170-176.

# **A lessons mining system for searching references to support decision making towards sustainable urbanization**

Jinhuan Wang<sup>a,b</sup> Liyin Shen<sup>a,b,\*</sup> Yitian Ren<sup>a,b,c</sup> J.Jorge Ochoa<sup>d</sup> Zhenhua Guo<sup>a,b</sup> Hang Yan<sup>a,b</sup> Zezhou Wu<sup>e</sup>

a School of Construction Management and Real Estate, Chongqing University, Chongqing, China

b International Research Center for Sustainable Built Environment, Chongqing University, Chongqing, China

c Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong

d School of Natural and Built Environments, University of South Australia, Adelaide, South Australia 5001, Australia

e Department of Construction Management and Real Estate, College of Civil Engineering, Shenzhen University, Guangdong, China

\*Corresponding author. Tel: +15823335713

E-mail address: [wangjinhuan99@qq.com](mailto:wangjinhuan99@qq.com) (J. H. Wang), [shenliyincqu@163.com](mailto:shenliyincqu@163.com) (L. Y. Shen)

[rtyijia@163.com](mailto:rtyijia@163.com) (Y. T. Ren), [Jorge.OchoaPaniagua@unisa.edu.au](mailto:Jorge.OchoaPaniagua@unisa.edu.au) (J. Ochoa)

[gzh1000100@163.com](mailto:gzh1000100@163.com) (Z. H. Guo), [yanhang302@sina.com](mailto:yanhang302@sina.com) (H. Yan)

[wuzezhou@szu.edu.cn](mailto:wuzezhou@szu.edu.cn) (Z. Z. Wu)

Dear Editor,

We would like to resubmit our previous rejected manuscript with NO. "JCLEPRO-D-18-02875" for possible publication. Based on the suggestions of Dr. Berge and the comments of three reviewers, we have intensively revised the manuscript. The detailed corrections are listed below in our **Response to Reviewers' Comments**. We hope the revised manuscript now is suitable for your journal.

The new revised manuscript is entitled "**A lessons mining system for searching references to support decision making towards sustainable urbanization**". We believe that the following aspects of this manuscript will make it interesting to general readers of your journal. The recurrence of similar problems caused by human errors in urbanization process present a huge challenge to achieving sustainable urban development. The knowledge learnt from these problems should become lessons and should be applied to avoid the recurrence of these problems. A Lessons Mining System (LMS) is proposed. LMS can facilitate decision-makers to foresee the potential problems thus take proper measures for addressing the problems. And a demonstration of Yangwu Town is used to show the application of the system.

We wish it to be considered for publication in "**Journal of Cleaner Production**". No conflict of interest exists in the submission of this manuscript, and manuscript is approved by all authors for publication. I would like to declare on behalf of my co-authors that the work described was original research that has not been published previously, and not under consideration for elsewhere, in whole or in part.

We deeply appreciate your consideration of our resubmission, and we look forward to receiving comments from the reviewers. If you have any queries, please don't hesitate to contact me at the address below.

Yours sincerely,

Liyin Shen

Address: School of Construction Management and Real Estate, Chongqing University, No.174 Shazheng Street, Shapingba District, Chongqing, PR China

Tel: +86 15823335713

Email: shenliyincqu@163.com

## **Response to Reviewers' Comments**

Dear Editor,

Thank you for considering our research manuscript and giving us the opportunity to resubmit the manuscript in *Journal of Cleaner Production*. Special thanks are to the editor and reviewers for giving us constructive comments and suggestions for revision. These comments are very helpful for us to revise and improve the quality of the paper. We have studied all the reviewers' comments carefully and have made corrections and improvements accordingly. We hope that the revised manuscript meets the requirements.

Many thanks and best regards!

Yours sincerely

Liyin Shen

Professor

School of Construction Management and Real Estate, Chongqing University

Email: shenliyincqu@163.com

---

Please find our responses to reviewers' comments as follows.

**Reviewer 1:**

<b>No</b>	<b>Comment</b>	<b>Response</b>
1	This paper seeks to fill a literature gap by creating an architecture of Lessons Mining System (LeMS) to assist in gaining knowledge from previously failed sustainable urbanization practices.	Thanks for the reviewer's comment. The research aims have been further clarified in the introduction section in the revised manuscript. This paper aims to develop a Lessons Mining System (LMS) which can assist decision makers to retrieve previous lesson cases to help foresee the potential problems if any in their current urbanization process, thus proper measures can be taken for addressing these problems.
2	This paper appears to be the reciprocal of Shen et al. (2013), which was heavily cited and openly referenced throughout. When asked to review this manuscript for the prestigious Journal of Cleaner Production, I was initially happy to do as sustainable urbanization is a research area of mine. Furthermore, machine learning-based decision-making tools are also viewed as important in making progress towards sustainability. Unfortunately, my excitement for this paper was short-lived due to its	Thanks to reviewer's constructive comments. The original idea of this paper is from Experience Mining System (ExMS) introduced by Shen et al. (2013). However, the Lessons Mining System (LMS) of this paper is a new method which is based on the Case-based Reasoning (CBR), this has been further clarified in research method section, details of revision can be found in the revised paper.

	relatively lack of innovation or improvement from Shen et al. (2013) original work in a lower-status journal.	
3	<p>Additionally, I was not impressed with how poorly the paper was written given the prestige of Journal of Cleaner Production. Additionally, although structured okay, the manuscript is not clear or direct, is inconsistent, and lacks the writing quality expected from five corroborating authors.</p>	<p>Thanks to reviewer's comments. Good efforts have been devoted to improving the writing quality by inviting assistance of an English native speaker. The presentation of the paper has been carefully polished. Careful proof-reading work has been done through the full paper. Details of the revision can be found in the revised manuscript.</p>
4	<p>Specifically, the literature should set up the study and provide an overview of urban problems, which suggest the importance of doing the study. Rather, this paper went into detail regarding air and water pollution and skimmed others.</p>	<p>The authors appreciate reviewer's constructive suggestions. The introduction section in the revised paper presents an overview of urban problems as comprehensive as possible, which include air and water pollution, traffic congestion, depletion of cultivated land, habitat destruction, the irrational rise of housing price, urban crimes and so on. All these urban problems are supported by literatures. And it is emphasized that it is not these problems themselves but their recurrences should cause concerns and be prevented.</p>
5	<p>It also suggested its importance by name dropping Shen et al. (2013) in the abstract.</p>	<p>Thanks to reviewer's comment. This expression has been deleted in the abstract in the revised manuscript.</p>

6	I was also shocked that only a “hypothetical demonstration is used to show the application of the system.”	Thanks to reviewer’s constructive comments. We have paid good efforts to revise the quality of demonstration section, the hypothetical lesson cases have been replaced by the township cases in the real practice. In the demonstration section in the revised manuscript, ten lesson cases are stored in Urbanization Lessons Database (ULD), eleven feature variables are identified. And by applying the Lessons Mining System, the way how to input feature variables, how to retrieve lesson cases from Urbanization Lessons Database (ULD), and how to improve decision-making quality have been further clarified. The details of revision can be found in the revised manuscript.
7	In a time when data abound, and solutions are needed, why not employ the study direction on a sustainable urbanization issue (i.e., since you focused on urban air quality, urban heat island)?	Thanks to reviewer’s constructive advices. This paper focuses on the urban problems caused by human errors in previous practices, and applying the lessons learnt from these problems to propose proper measures to prevent possible problems in order to achieve urbanization sustainability.
8	Topically, the manuscript was interesting but it lacks scientific innovation.	Thanks to reviewer’s comment. The authors have paid good efforts to revise the paper. The revised manuscript proposed a new method to facilitate decision makers to foresee the potential problems if any in their current urbanization process thus proper measures can be taken to address these problems.
9	Therefore, even if the paper went through a significant revise and resubmit to fix the	Thanks to reviewer’s comment. The previous manuscript has been greatly modified and carefully polished. The authors hope this revised manuscript will meet the requirements

	<p>aforementioned, I would recommend it for publication in a high-profile journal such as <i>Journal of Cleaner Production</i>.</p>	<p>of <i>Journal of Cleaner Production</i>.</p>
10	<p>My suggestion to the author(s) would be to take a step back from the study to see what contribution(s) to the scientific community are made with this paper.</p>	<p>Thanks for the reviewer's constructive suggestion. The contribution of this study has been further addressed in the revised manuscript. A new method based on Case-based Reasoning (CBR) has been proposed, which is able to facilitate decision-makers to foresee the potential problems if any in their current urbanization process thus proper measures can be taken to address and prevent the reoccurrence of these problems.</p>
11	<p>Some questions that should be asked: Is this work novel?</p>	<p>Thanks to the reviewer's comments. The innovation of this revised manuscript is that a new method based on Case-based Reasoning is proposed, which aims to help the decision-makers to prevent the reoccurrence of the potential problems by applying previous lessons. This has been further addressed in the revised manuscript.</p>
12	<p>Were the methods conducted correctly?</p>	<p>Thanks to the reviewer's comments. The authors have made substantial revisions to further present the research methods in the revised manuscript. The Lessons Mining System (LMS) in our revised manuscript is a method developed from Case-based Reasoning (CBR). The Local-global method was used to conduct the similarity matching.</p>
13	<p>Does the work provide a theoretical contribution?</p>	<p>Thanks to reviewer's comments. Lessons Mining System (LMS) in this revised paper opens a door to the lessons sharing of urban problems. It represents an innovative</p>



		<p>mechanism in conducting research in the discipline of sustainable urbanization. This mechanism enables to carry out comparative analysis on a wide range of past lesson practices, thus obtain valuable information to assist decision-makers in formulating better actions for promoting sustainable urbanization.</p>
14	<p>Does the work provide a methodological contribution?</p>	<p>Thanks to the reviewer's comments. Case-based Reasoning (CBR) is a method that helps extract the effective solutions adopted previously in addressing certain type of problems and applies these solutions as decision-making references to solve a new problem (Yang and Wang, 2009). It is widely used in many fields for decision making (Chen et al., 2016a; Zhang and Dai, 2018), problems diagnosis (Gu et al., 2017; Tung et al., 2010), products design (Yang and Chen, 2011; Shen et al., 2017a), and business management (Carmona et al., 2013; Sartori et al., 2016). In the research method section in the revised paper, the Lessons Mining System (LMS) is proposed by developing and innovating the CBR.</p>
15	<p>Could someone replicate my study?</p>	<p>Thanks for the reviewer's question. The Lessons Mining System (LMS) in our revised manuscript can facilitate decision-makers (LMS user) to foresee the potential problems if any in their current urbanization process thus proper measures can be taken to address these problems. Decision-makers can foresee the potential problems if any in the current urbanization process by inputting feature attributes, the result is replicating.</p>

16	<p>I will come back to some details in the forthcoming sections. To end, it should be noted, that there is a publication within this material; however, in its current form I do not think it is to the caliber of <i>Journal of Cleaner Production</i>. Since this paper is the reciprocal of Shen et al. (2013), and I find the paper of relatively low scientific innovation, I suggest trying to publish this manuscript in the same journal <i>Automation in Construction</i>.</p>	<p>Thanks to reviewer's comment. Good efforts have been paid by the authors to improve the quality of the research paper. The authors hope this revised manuscript will meet the requirements of <i>Journal of Cleaner Production</i>.</p>
17	<p>Might I suggest simplifying the title to: "A lessons mining system for supporting decision-making on sustainable urbanization strategies"</p>	<p>The authors appreciated reviewer's suggestion. By cooperating reviewer's advice, the title has been revised as "A lessons mining system for searching references to support decision making towards sustainable urbanization".</p>
18	<p>Do not name-drop in the abstract and doing so the way you did decrease the scientific value straight-away for your study.</p>	<p>Thanks to the reviewer's constructive comments. The name-drop of Shen et al. (2013) has been deleted in the abstract of the revised manuscript.</p>
19	<p>The writing of this manuscript is poor, with many sentences with too many words, and</p>	<p>Thanks for the reviewer's comments. Good efforts have been devoted to improving the writing quality by inviting assistance of an English native speaker. The sentences have</p>

	<p>other fragmented sentences without enough. I suggest getting a scientific copy editor to review your work first before sending it to review. Write simply, as it can be a beautiful thing if many readers can understand your research story. If not a native English speaker/writer, nor strong at writing in the English language, it is recommended to have an English technical writer review your work before sending it into journal review. For the most part, your article was easy to read but many sentences were confusing, lacked proper syntax, or grammar.</p>	<p>been carefully polished and careful proof-reading work has been done throughout the paper.</p>
20	<p>Many statements or concepts need more literature support. Furthermore, a random diversity of ethnicities cited, from many high profile international journals, is needed to fully grasp motherhood concepts.</p>	<p>The authors appreciated the reviewer's constructive comments. The revised manuscript has been greatly modified. More references from high-profile international journals have been learned and cited to support the concept, theory, and method of this research paper. The details of revision can be found in the revised manuscript.</p>

21	The article is too long in spots and too short in other places. A paper should be balanced and all aspects should support the study, not detract.	The authors appreciated the reviewer's comments. The content of this paper has been revised to be more balanced between different sections. Details of revision can be found in the revised manuscript.
22	Tables and figures need to stand-alone, with their captions providing enough information that the reader does not need to go back into the text.	The authors appreciated the reviewer's constructive advice. The tables and figures have been revised to contain more relevant information in the revised paper, which are stand-alone.
23	This seems foolish, as if you can inform "what not to do" surely you can also provide "what to do." Seems like a poor decision-making tool, and largely academic for gaining more publications. Since the two methods use almost the same architecture, they can now accomplish both ends of the sustainable urbanization spectrum (successes and failures).	Thanks to the reviewer's comments. The original idea of Lessons Mining System (LMS) of our manuscript is from Experience Mining System (ExMS) introduced by Shen et al. (2013). However, the LMS is based on the Case-based Reasoning (CBR) method and aims to foresee and prevent the reoccurrence of the potential problems in the current urbanization process by learning the lessons retrieved from previous problems. The authors appreciate that preventing the recurrence of those possible problems is essential to promote the sustainability of urbanization development.
24	I suggest using some synonym for "lessons" throughout the paper.	Thanks for the reviewer's constructive suggestion. The word "lessons" has been further explained in section 2 and section 3 in the revised manuscript: Lessons are often defined

		as valid knowledge which is learned from previous problems or failures (Dash et al., 2016; Pittman et al., 2014). “knowledge learnt from urban problems caused by human errors” can be regarded as the synonym of “lessons”. The details of revision can be found in the revised manuscript.
25	The second paragraph should summarize all urban environmental issues, not focus primarily on air and water. In example, habitat destruction, invasive species, or anything biological in focus were not talked about at all.	The authors appreciate the reviewer’s constructive suggestions. In the revised manuscript, the authors provide more dimensions of urban problems, which includes air and water pollution, traffic congestion, depletion of cultivated land, habitat destruction, the irrational rise of housing price, and various types of urban crimes. The details of revision can be found in the revised paper.
26	The reviewer should not have to focus on sentence syntax, structure, or grammar; however, there are many issues here with this paper.	Thanks to the reviewer’s comments. Good efforts have been devoted to improving the writing quality. The sentences of the paper have been polished and careful proof-reading work has been done throughout the paper.
27	Double-blind peer-review is a preferred method, as in this case my expectations for excellence was much higher and expected with five authors listed.	The authors appreciate reviewer’s comments. The double-blind peer-review is indeed a perfect method. Your profound knowledge and unique insights have helped the authors greatly improve the quality of this paper. We are very honored to be able to get the guidance of internationally renowned scholars like you.

28	There should be a balance between text and figures/tables, and most journals would like 6-8 (combined) to support a paper such as this.	The authors appreciate the reviewer's advice. The text and figures/tables have been revised to be more balanced in the revised manuscript, which includes two tables and six figures. The details of revision can be found in the revised manuscript.
29	Since this paper is the reciprocal of Shen et al. (2013), and I find the paper of relatively low scientific innovation, I suggest trying to publish this manuscript in the same journal <i>Automation in Construction</i> .	Thanks to the reviewer's comment. Good efforts have been devoted to improve the quality of the paper. The scientific innovation has been further addressed in the revised manuscript, this paper aims to foresee and prevent the reoccurrence of the potential problems in the current urbanization process by learning the lessons retrieved from previous problems. The authors hope this revised manuscript will meet the requirements of <i>Journal of Cleaner Production</i> .
30	Further readings/recommended citations: Strunk, W. (2007). <i>The elements of style</i> . Penguin.	Thanks to the reviewer's advice. After reading carefully the book by Strunk, W. (2007), the authors have benefited a lot on improving the quality of paper writing.

**Reviewer 2:**

<b>No</b>	<b>Comment</b>	<b>Response</b>
1	<p>The authors presented a study using text mining to learn from past failure cases related to sustainable urbanization. This work is timely and urgent due to our increasing capabilities to collect a large quantity of data and cases, and the fast pace of urbanization. The potential of this work is thus very significant. The topic well fits the journal. The discussion on the overall framework has been laid out nicely. This reviewer would like to raise a few points for the authors to consider.</p>	<p>Thanks to the reviewer's very supportive and encouraging comments.</p>
2	<p>1) Focusing on failure in sustainable urbanization is novel, but defining failure is challenging as failure in sustainable urbanization can mean differently in different cases. Thus, it is important to explain how</p>	<p>The authors appreciated the reviewer's suggestion. To avoid confusion, failures will be replaced by problems emerged during the urbanization development in the revised manuscript. The definition and explanation of these problem is further added into section 3.1. Details of revision can be found in the revised manuscript.</p>

	failure is defined in this context.	
3	2) The authors discussed principles for representing failure cases, but it would be great if the authors specifically discuss the actual parameters, categories, and scenarios. Although this information is presented in the demonstration to a certain level (Fig. 5), it is not clear how, for example the parameters are sufficient for describing failure cases.	Thanks to the reviewer's constructive comment and advice. The discussion of the actual scenario features has been added into “/Data input for feature variables about the target scenario/” of section 5 in the revised manuscript. The principles for representing failure cases has also been further clarified in section 3.1. The application of feature variables explaining the target scenario has been clarified and demonstrated in section 4.
4	3) It is great that the work is based on the previous work of the author(s). To facilitate understanding, maybe the authors want to add examples explaining Indicators, Parameters, and Categories. It will help readers better understand how the mining module and coding work.	Thanks to the reviewer's constructive suggestion. The examples explaining indicators and parameters of scenario, problem, and lessons are added in section 3.1 in the revised manuscript.
5	4) Wonder if the authors have more information about the demonstration. Ideally, if some results, even hypothetical, can be	Thanks to the reviewer's constructive comments. The authors have devoted great efforts to improve the quality of demonstration section. The hypothetical lesson cases have been replaced by township cases in the real practice. In demonstration of the revised



	<p>included, they will better support the discussions and conclusions.</p>	<p>manuscript, ten lesson cases are stored in Urbanization Lessons Database (ULD), eleven feature variables are identified. And the way how to input feature variables, how to retrieve lesson cases from Urbanization Lessons Database (ULD), and how to improve decision-making quality by using mined lesson cases have also been further clarified in the demonstration section. Details of revision can be found in the revised manuscript.</p>
--	--	--

**Reviewer 3:**

<b>No</b>	<b>Comment</b>	<b>Response</b>
1	<p>According to the authors: “In order to develop the architecture of Lessons Mining System (LeMS), the Experience Mining System (ExMS) introduced by Shen et al. (2013)* will be used as a reference tool.”</p> <p>It seems to me that the major difference between ExMS and LeMS is that the former extracts valuable experiences from past successful cases to solve new problems in</p>	<p>Thanks to the reviewer’s constructive comments. The original idea of this paper is from Experience Mining System (ExMS) introduced by Shen et al. (2013). However, the Lessons Mining System (LMS) is an innovative method which is based on the Case-based Reasoning (CBR) and aims to foresee as well as prevent the reoccurrence of the potential problems in the current urbanization process by learning the lessons retrieved from previous problems. The scientific innovation of LMS has been further addressed in the revised manuscript. And to better help readers to understand the application of LMS, Section 4 (Demonstration) has been modified with good efforts by showing the way how to input feature variables, how to retrieve lesson cases from</p>

	<p>promoting sustainability of urbanization, whereas the latter does the same with failure cases. Using a framework that has been developed previously as a base for a new research is justified. However, a significant portion of the paper is devoted to explanation of the framework and not too much in its application and use.</p>	<p>Urbanization Lessons Database (ULD), and how to improve decision-making quality by using the mined lesson cases.</p>
2	<p>Again according to the authors: “A hypothetical demonstration is used to show the application of the system.”</p> <p>In the case study the City of VSC has a serious problem of the air pollution and is looking for “not to do lessons” in their plan for infrastructure construction and industrial upgrading. The use of the LeMS system provided them the following failure cases:</p> <p>the Meuse Valley fog incident in Belgium</p>	<p>Thanks to the reviewer’s constructive suggestions. In Section 4 (Demonstration) in the revised manuscript, the hypothetical lesson cases have been replaced with township cases in the real practice. A list of “not to do lessons” has been shown in Figure 6 in Section 4, for example, “Local resources endowment has not been utilized effectively, such as tourism resources” and “Ignorance of the quality of agricultural products, these products are not competitive in the market” in <math>C_3</math> of Yaoshan Town.</p> <p>The “not to do lessons” is stored in a specific lesson case in Urbanization Lessons Database (ULD). The user (decision-maker) can input feature variables of this scenario into LMS, the mined lesson cases will facilitate user to foresee potential problems thus proper measures can be taken to address and prevent the occurrence of the problems.</p>

	<p>the American Donora smog incident  the Great Smog incident of London  the water pollution in Chongqing of China  the Los Angeles photochemical smog episode</p> <p>I agree that this demonstration shows that the system is capable of identifying failure cases related to air pollution but this probably could have been done using key words search in few databases. I was hoping to see a list of “not to do lessons.” The LeMS system should be able to store “not to do lessons” for each failure case and provide them to the user when the user has a similar case.</p>	
3	<p>My suggestion is that the authors elaborate more on application of the LeMS system and its capabilities for the user.</p>	<p>Thanks to the reviewer’s suggestions. The authors have devoted good efforts to improve the quality of Demonstration of Lessons Mining System (LMS) in Section 4 in the revised manuscript, showing the way decision-maker (user) to input feature variables, to retrieve lesson cases from Urbanization Lessons Database (ULD), and to improve decision-making quality by using the mined lesson cases. The capabilities of LMS for</p>

		the user have been further addressed in “/Foreseeing the potential problems and planning proper actions/” of section 5. Details of revision can be found in the revised paper.
4	<p>I also suggest a careful editing of the paper for grammar. For example, see the following examples on how editing the paper will improve its readability:</p> <p>Page 3 – line 3: “LeMS can facilitate decision-makers to engage in effective decision-making process and select strategies for promoting sustainable urbanization by mining lessons from existing failure practices” should be revised as “LeMS can facilitate decision-making process and help decision makers select strategies for promoting sustainable urbanization by mining lessons from existing failure practices”.</p>	<p>Thanks to the reviewer’s kind reminder. The expression of “LeMS can facilitate decision-makers to engage in effective decision-making process and select strategies for promoting sustainable urbanization by mining lessons from existing failure practices” has been deleted in revised manuscript.</p>
5	<p>Page 5 – line 26: “Respiratory diseases, such as pneumonia, are very popular” should be</p>	<p>Thanks to reviewer’s comment. The expression of “Respiratory diseases, such as pneumonia, are very popular” has been deleted in revised manuscript. And the word</p>

	revised as “Respiratory diseases, such as pneumonia, are very common”.	“popular” which was used to describe the serious problems or diseases is replaced by the word “common” in the revised manuscript.
6	Page 5 – line 15: “A typical strategy for avoiding problems recurring is to learn the lessons from past practices which suffered from similar problems” should be revised as “A typical strategy for avoiding problems recurring is to learn from past practices with similar conditions”.	Thanks to the reviewer’s comment. The expression of “A typical strategy for avoiding problems recurring is to learn the lessons from past practices which suffered from similar problems” has been deleted in revised manuscript.
7	Page 5 – line 39: “Zhao et al. (2014) affirmed the importance of lessons learned from past chemical accidents to help industries to reduce the risk of catastrophic chemical accidents that may occur in future” should be revised as “Zhao et al. (2014) affirmed the importance of lessons learned from past chemical accidents to help the chemical industry reduce the risk of catastrophic accidents in future”.	Thanks to the reviewer’s kind reminder. The correction has been made in the revised manuscript.

8	Page 5 – line 50: “could have helped to prevent” should be revised as “could have helped prevent”.	Thanks to the reviewer’s kind reminder. The correction has been made accordingly in the revised manuscript.
9	Page 7 – line 20: “what should not to do” should be revised as “what not to do”.	Thanks to the reviewer’s kind reminder. The correction has been made in the revised manuscript.
10	Page 14 – line 6: “As mentioned early” should be revised as “As mentioned earlier”.	Thanks for your comment. The expression of “As mentioned early” has been deleted in revised manuscript.

**Highlights:**

- A Lessons Mining System (LMS) is proposed.
- LMS is a new method of mining lessons learnt from urbanization problems.
- LMS can facilitate decision-makers to foresee potential problems.
- LMS can facilitate decision-makers to improve decision-making quality.
- A demonstration of Yangwu is used to show the application of LMS.

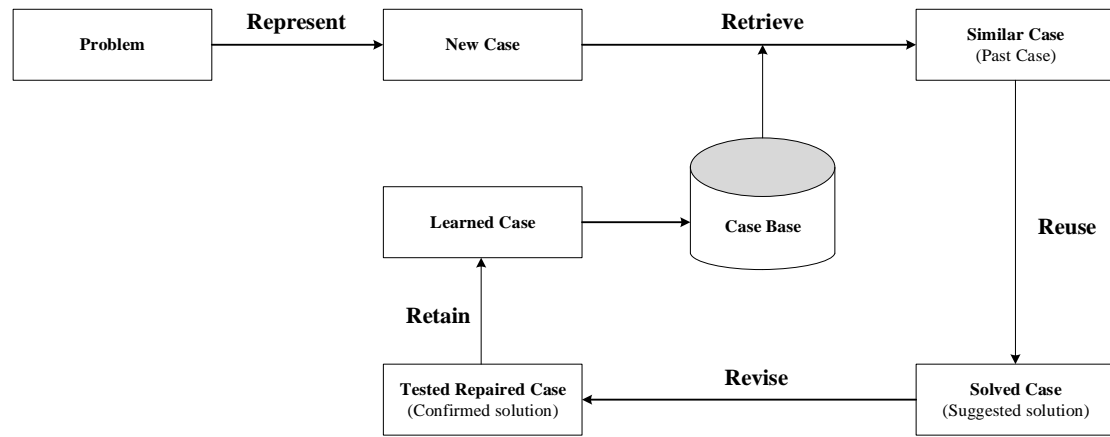


Figure 1. Case-based Reasoning (CBR) cycle

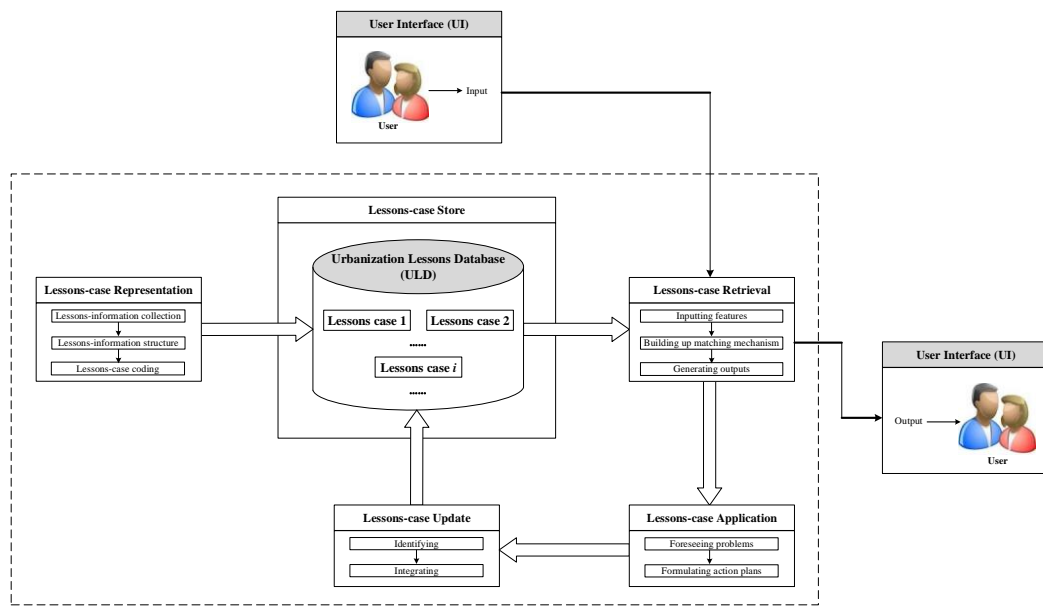


Figure 2. The architecture of Lessons Mining System (LMS)



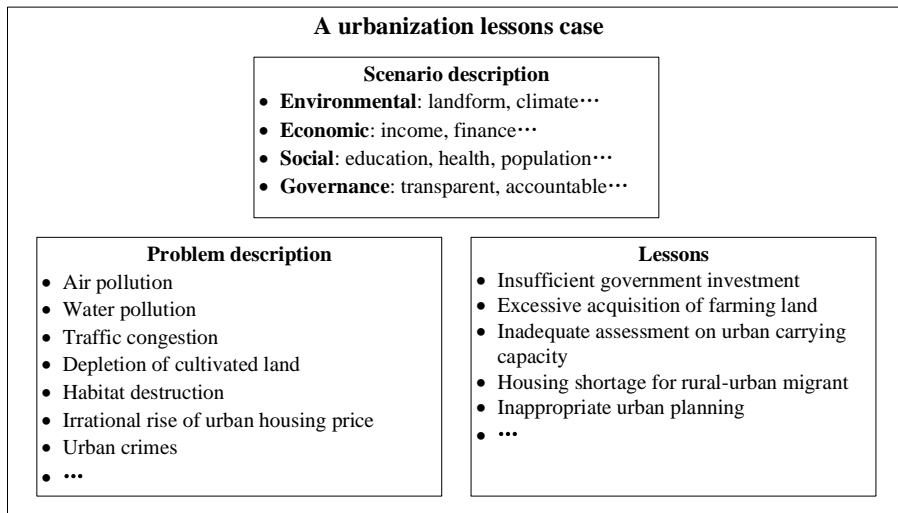


Figure 3. The structure of a lesson case

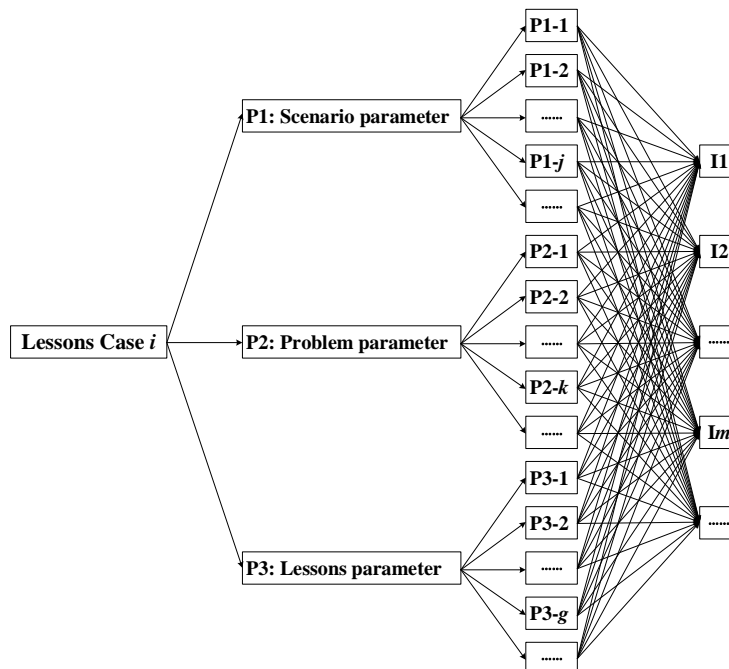


Figure 4. Parameter Coding Structure in LMS

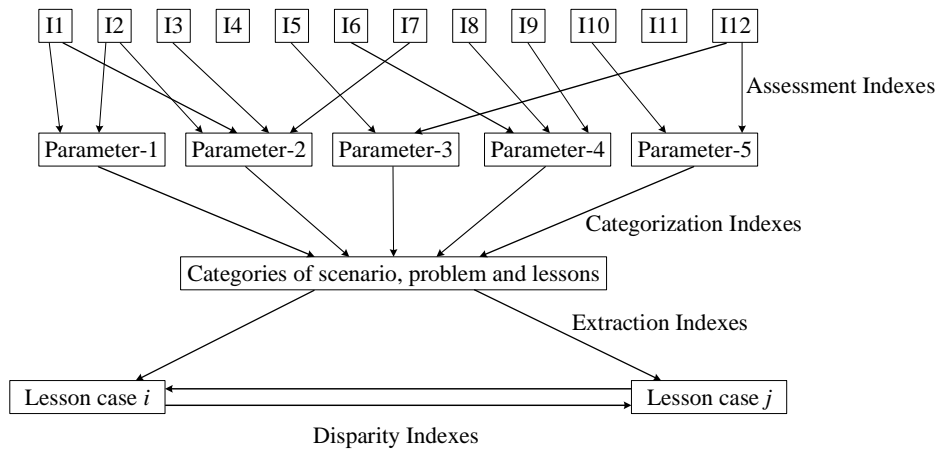


Figure 5. ULD index structure adapted from Shen et al. (2013)

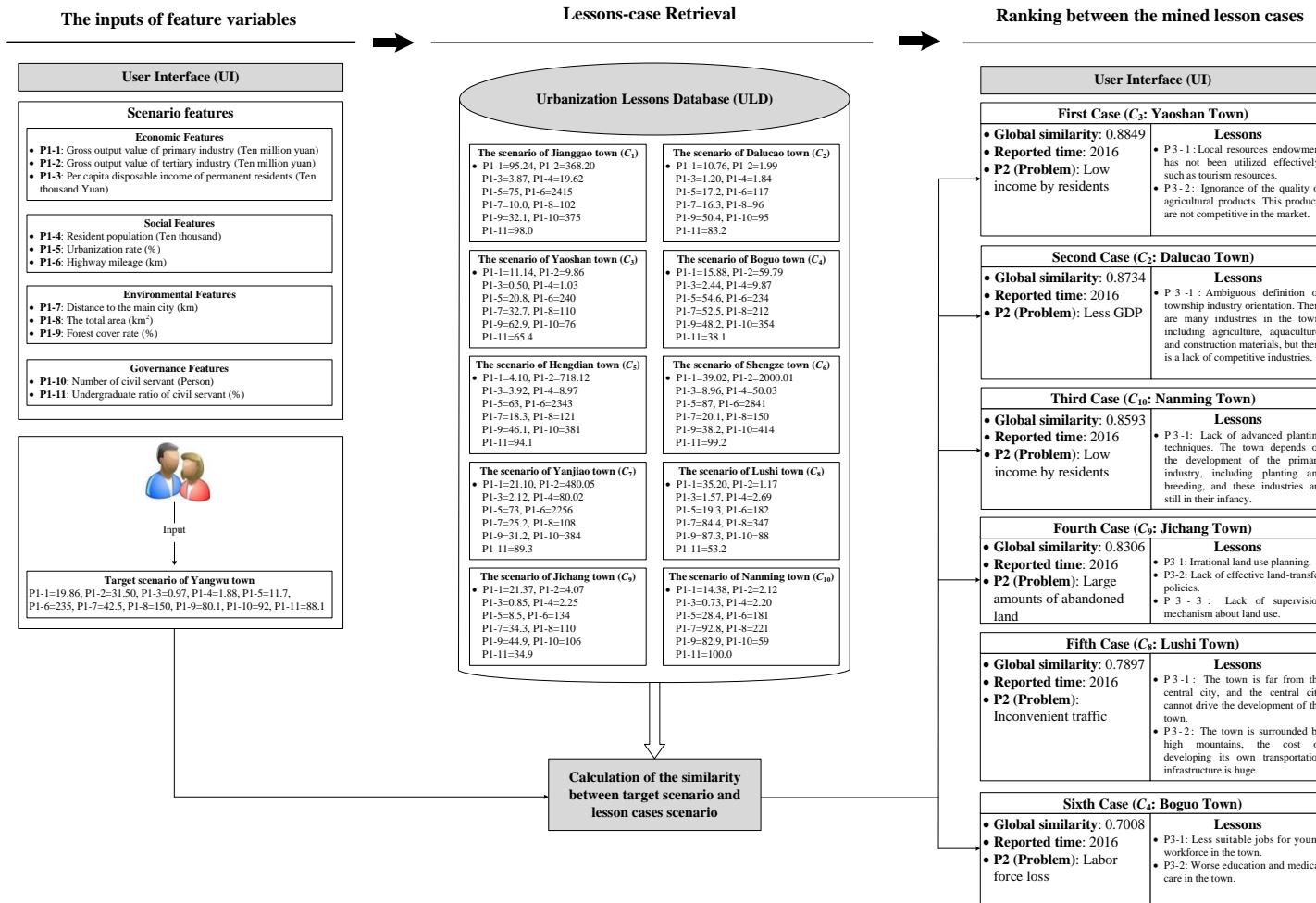


Figure 6. Lessons-mining demonstration

Table 1. The major sources for collecting lesson information associating with urban problems

Sources	Details of information sources
Books	Ma et al. (2004); Barria and Thajchayapong (2011); Shapiro (2012).
Journal papers	Alam and Ahmed (2013); Black (2003); Brook (2008); Cao et al. (2011); Chen et al. (2016b); Chen et al. (2013); Christidis and Rivas (2012); Contreras and Ferri (2016); Dockery et al. (1993).
Regional reports	European Commission (Revel, 2011).
National reports	The Government of China (Xinhua, 2017), the Central Pollution Control Board in India (Bhawan and Nagar, 2015).
City reports	Xingtai City in China (Zhang et al., 2015), Mexico City (Moreno et al., 2008), Ulaanbaatar City in Mongolia (Amarsaikhan et al., 2014), Tirana City in Albania (Mandija and Zoga, 2012), Delhi City in India (Singh and Peshin, 2014).

Table 2. The global similarity between target scenario and all the lesson cases in ULD

Lesson cases Similarity		$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$	$C_{10}$
<b>Local similarity</b>	<b>P1-1</b>	0.1729	0.9002	0.9044	0.9563	0.8271	0.7899	0.9874	0.8316	0.9834	0.9399
	<b>P1-2</b>	0.8316	0.9852	0.9892	0.9858	0.6565	0.0152	0.7756	0.9848	0.9863	0.9852
	<b>P1-3</b>	0.6576	0.9730	0.9441	0.8368	0.6513	0.0559	0.8641	0.9286	0.9864	0.9714
	<b>P1-4</b>	0.7754	0.9995	0.9893	0.8998	0.9103	0.3907	0.0108	0.9898	0.9953	0.9959
	<b>P1-5</b>	0.1936	0.9299	0.8841	0.4545	0.3439	0.0408	0.2191	0.9032	0.9592	0.7873
	<b>P1-6</b>	0.1997	0.9567	0.9982	0.9996	0.2263	0.0433	0.2581	0.9805	0.9629	0.9801
	<b>P1-7</b>	0.6075	0.6836	0.8816	0.8792	0.7041	0.7283	0.7886	0.4988	0.8973	0.3925
	<b>P1-8</b>	0.8175	0.7935	0.8470	0.7639	0.8891	1.0000	0.8394	0.2436	0.8472	0.7285
	<b>P1-9</b>	0.1429	0.4714	0.6941	0.4386	0.3929	0.2500	0.1250	0.8750	0.3732	0.9482
	<b>P1-10</b>	0.2028	0.9915	0.9549	0.2620	0.1859	0.0930	0.1775	0.9887	0.9606	0.9071
	<b>P1-11</b>	0.8464	0.9232	0.6467	0.2320	0.9078	0.8310	0.9846	0.4624	0.1843	0.8157
<b>Global similarity</b>		0.4953	0.8734	0.8849	0.7008	0.6087	0.3853	0.5482	0.7897	0.8306	0.8593