

# Site Design Case Retrieval and Strategy Generation based on Extension Set

Lemei Li, Guannan Fu\*, Xuechao Liu, Weiwei Hao

<sup>1</sup>College of civil engineering and architecture, Shandong University of Science and Technology, Qingdao, Shandong 266590, China

Email: skd994647@sdust.edu.cn

## Article Info

Volume 83

Page Number: 1849 - 1858

Publication Issue:

July-August 2020

## Abstract

In the big data era, designers face huge difficulty in collecting numerous cases of site design on the Internet. To enhance retrieval efficiency of designers, the present study develops a method based on extension set to retrieve site design cases and approaches to formulate innovative strategies. First, data acquisition software is adopted to acquire site design case data, and the data server, database development and programming software are comprehensively and jointly employed to build a case base of site design. Second, based on the retrieval condition converted from site design contradiction, site design cases are initially screened out by correlation function, and then they achieve dynamic optimization sorting based on extension set theory, as an attempt to satisfy personalized needs of designers. Lastly, innovative strategies for site design are automatically formulated on case base. As suggested from case test, this method is capable of helping designers efficiently retrieve valuable cases of site design on the Internet, thereby presenting some theories and methods for intelligent site design.

**Keywords:** Site Design, Case Base, Extension Set, Correlation Function, Innovation Strategy

## Article History

Article Received: 06 June 2020

Revised: 29 June 2020

Accepted: 14 July 2020

Publication: 25 July 2020

## 1. Introduction

Site design refers to a vital link between the specific construction planning and the general layout design, and it takes up a critical part of the early stage of architectural design. High quality site design can implement the constraint of urban planning into architectural design, as well as offering strategic support to address the contradictions and innovation problems. Existing big data in the Internet is capable of presenting considerable cases for professional designers. However, how to build an efficient query system to screen out and integrate site design cases and yield design ideas by data analysis and mining will be undoubtedly of high significance to boosting the intelligent process of site design.

Through referencing numerous research materials, it is suggested that a certain research gap remains in retrieval of

site design case, whereas many researchers have made some attempt to develop architectural design case retrieval system and to build case database. Viktor, A.<sup>[1-4]</sup> et al. from the Institute of computer science at Hildesheim University verified through a case study that at the early phase of architectural design, the architectural design retrieval system based on rules and cases can more significantly apply to the query of innovative design strategies. In the subsequent study, a multi-agent retrieval method for the case-based semantic search of architectural design was developed. Cekmis, A. [5] proposed a fuzzy logic application for site planning and design. The fuzzy layout planning model employed in the study could help designers generate the optimal allocation scheme by supporting their reasoning mode and decision-making mechanism. Andrés, C. [6] et al. of Georgia Institute of Technology studied the application of CBRR system at the early architectural design phase and pres-

entered the first development phase of a case-based reasoning system that supports the early conceptual design of commercial buildings. Ramon Lopez, D. [7] et al. of the Institute of artificial intelligence solved the problem of introducing originality into the solution proposed by CBR system. The IDIOM system developed by Ian, S. [8] of the Swiss Federal Institute of technology refer to a case-based reasoning system for organizing building layout. Such system allows architects to select their favorite cases to build a case library, through which human-computer interaction and the use of preferences can be realized more effectively. PRECEDENTS computer aided architectural design system developed by Oxman RE and Oxman RM [9] of Haifa Institute of technology, Israel, refer to a case memory model adopted to computerize excellent design cases and store them in a case database, and supports semantic retrieval. Flemming, U. of Carnegie Mellon University [10] has extensively studied the application of SEED software at the early phase of architectural design and highlighted that SEED can automatically store the solutions generated by the system as cases and retrieve them for reuse under similar problems; as a result, the retrieval efficiency of architects is enhanced. Dave, B. [11] et al. of Swiss Federal Institute of technology developed a case-based reasoning system for architectural space design (CADRE), capable of combining parts of multiple design cases to present architects' novel design solutions.

By reviewing previous literature, it is suggested that the existing research exhibits the following defects in site design case retrieval and intelligent scheme design:

- (1) The retrieval research of relevant cases is largely reported in architectural cases, and the special search of site design is lacked. Besides, search for architectural cases primarily focuses on local databases, and the analysis and search of Internet cases have been rarely conducted.
- (2) Insufficient mechanism to translate the contradiction in site design into search conditions, formulate coping strategies and build strategy library.

- (3) Lacked satisfaction sorting analysis of case retrieval. Under a huge number of retrieved cases, several problems will occur (e.g., the uneven quality and manual secondary retrieval). Even because the retrieval condition setting is not sufficiently intelligent, some excellent cases that can trigger the designer's inspiration are sorted backward, which cannot be retrieved.

Extension set theory is capable of quantitatively expressing the transformation of things<sup>[12]</sup>. It can individually assess site design case requirements by analyzing retrieval behavior of site designers. Given the wide distribution, diverse types and dynamic changes of Internet big data, the present study proposes a method to retrieve site design cases and methods to formulate strategy based on extension set. Besides, it attempts to lay theoretical and methodological basis for intelligent site design by enhancing the Internet big data retrieval efficiency of site design cases.

## 2. Methodology

The core content of the present study refers to integrate extension set theory and data analysis mining tools. Under the retrieval condition of site design contradiction, site design cases are initially screened out through correlation function and then sorted in terms of dynamic optimization according to extension set theory to satisfy personalized needs of designers.

The implementation path of the present study consists of three modules, i.e., construction of case base of site design (module 1), optimization of retrieval function for case base of site design (module 2), as well as formulation of site design innovation strategy (module 3). Module 1: Nutch software is adopted to collect site design cases on the Internet, and a case base of site design is established. Module 2: Retrieval function for case base of site design is optimized by the extension set theory and data analysis and mining tools. Module 3: Retrieved case results are analyzed, and site design innovation strategies are formulated. (Fig. 1)

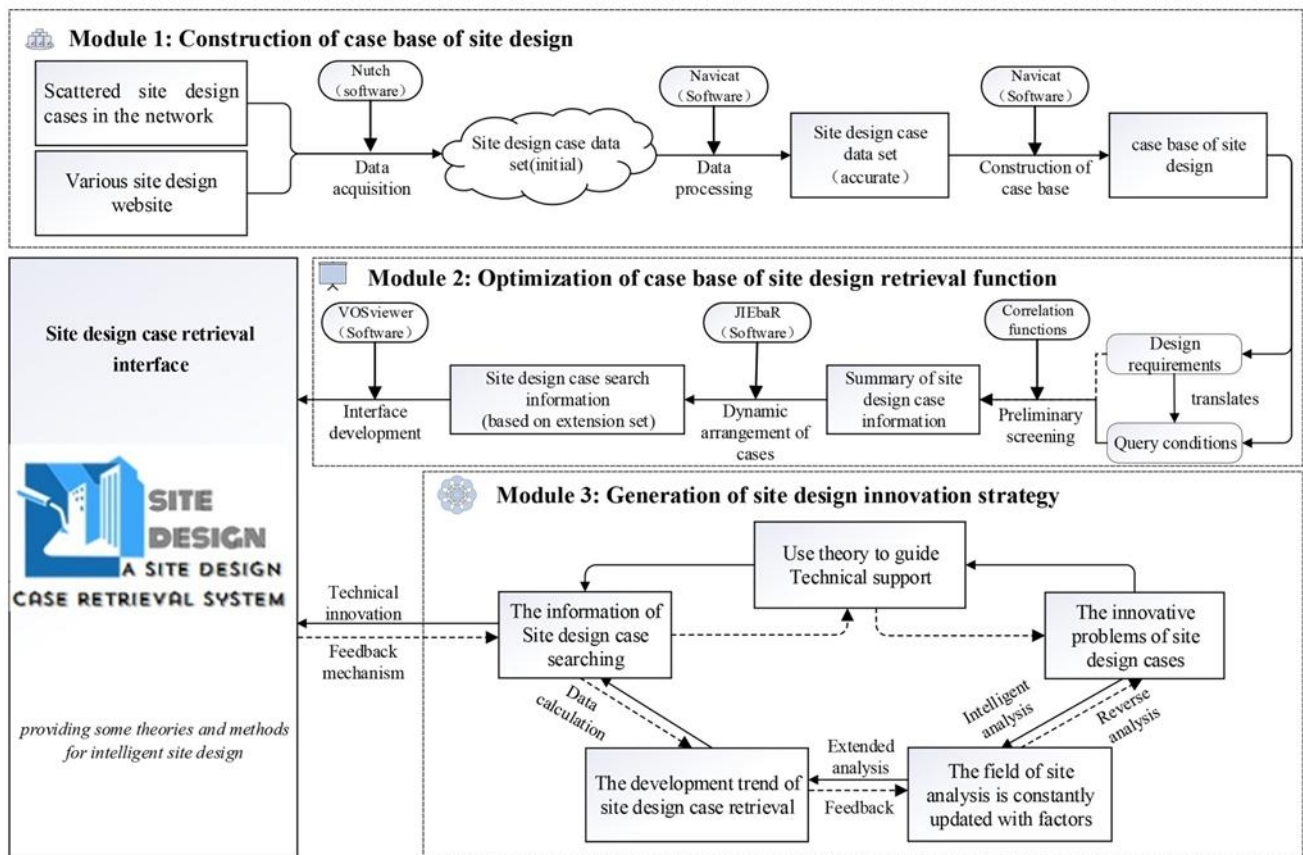


Fig. 1. Site design case retrieval method flow based on extension set

### 2.1. Construction of case base of site design

Site design cases summarize the design objectives, skills and experiences suggested in the design results when the designer performs the task of site design. Achievements are largely stored on the Internet as design taskbook, project planning book, upper planning condition, current situation photo, design text, design drawing, design model, achievements publicity video, etc. The original data of the case base of site design should be acquired from the Internet according to certain collection rules. The major characteristics of data collection are listed in Table 1.

Table 1. Characteristics of building case data

Source of case	Scattered cases in the network, All kinds of construction website, The website of design company
Acquisition mode	Internet collection, real-time collection, offline collection
Basic situation	Project name, project type, site

	location, design time, completion time, design unit, designer and project introduction
Approval of data	Project number, building height, site materials, land type, site area, building area, floor area ratio, green space ratio, building density, number of parking spaces
Review of results	Page views, comments, favorites, comments, retweets, comments
Image data	The name, The category, The theme, The style

Note: the above features are summarized by the author according to the case retrieval habits of several site designers, users can adjust and supplement according to their own needs.

Data collection of site design cases should comply with the principles of pertinence, extensiveness and typicality, and establish collection order from time dimension, space

dimension and emotion dimension. The collection process is elucidated as follows. ① Search for various site design case website via network information retrieval, as well as using Nutch and other software to collect Internet data and update in real time on the premise of permission. Moreover, the possible gap of network data is supplemented by offline collection to achieve the extensiveness and diversity of site design case data collection. ② Extraction of basic profile features, generally located at the end of the webpage (e.g., case names, site types, completion time, and site approval indicators). ③ Network engineers can set the corresponding retrieval mode according to the basic overview and approval data listed in Table 1, as well as combining webpage design with retrieval rules. Boolean logic, truncation retrieval, weighted retrieval and position operator can be adopted to enhance the retrieval efficiency and accuracy of complex cases. Hidden in the text description of the project, considerable basic profile features should be collected as accurately as possible by analyzing designer's language habits and complying with language collection rules. ④ Case review falls into user evaluation and correlation indexes. User evaluation is classified into the comment content and the user rating. The comment content indicates the user's subjective attitude towards the case, and the rating reflects the user's satisfaction degree of retrieval. Emotional analysis software can be imported to automatically determine the emotional category of the case review, which lays a solid basis for data classification. Correlation index is classified into two types, i.e., attention index (e.g., page views and comments) and the recognition of the index (e.g., the amount of collection, thumb-up volume, and forwarding volume). ⑤ Set the data export rule and automatically import the collected data into the MySQL database.

With the application of cloud computing technology, computer data processing ushered into a new era, and computer data processing receives great convenience. Cloud computing can build computing, storage, network and other resources into a unified resource pool via virtualization, as well as effectively reducing the difficulty of resource allocation and management<sup>[13]</sup>. This technology will significantly enhance the dynamic updating ability of data to build case base of site design. Dynamic data collection can be realized by Nutch software on cloud server.

Data processing refers to a vital preparation work before the reconstruction of site design database, mainly aiming to solve the problems of data duplication, data loss and data format inconsistency<sup>[14]</sup>. Considerable repeated data will reduce the spatial efficiency of the case database and slow data analysis. The case database can exploit Navicat's duplicate checking technology to retain the original website of site design cases and delete other reprinted case contents. Insufficient key data acts as a significant bottleneck in site design case retrieval, whereas the data in many cases is not really missing, but not recognized, hidden in the long text introduction. Similar problems can be flexibly solved through keyword extraction, feature item collection, global constants supplement, center measure filling, artificial screening, etc.<sup>[15]</sup>. Overall, unified data format is an important prerequisite for data retrieval. Data inconsistencies involve site design cases primarily existing geographical coordinates, measurement units, language environment, synonyms and differences of professional terms, etc. The unification of the mentioned data can be achieved with SPSS software to establish the conversion rule table, which facilitates designer to obtain more accurate and professional retrieval condition data of site design.

To achieve site design data collection, analysis, integration and storage, the present study will build a case base of site design based on MySQL database management software. The database consists of site condition data table, natural environment data table, social environment data table, site image data table, site designer data table, site design organization data table, user demand data table, administrator data table, etc.

## 2.2. Use extension set theory to optimize the retrieval function of case base of site design

### 2.2.1. Build a mechanism to transform site conflicts into search items

The important difference between the retrieval method based on contradiction problems and the general condition retrieval method is that some unfavorable factors caused by the contradiction in site design are added as the retrieval items. The solution of unfavorable factors is the difficulty and innovation of site design, and then the corresponding design strategy is generated. Besides the basic elements

(e.g., site function, site area and site traffic) that are prone to contradictory problems, these search items also consist of the keywords of specific contradictory problemstermed as problem characteristics. For instance, the design of a burial theme site with historical relicsthat is planned to be protected and restored in the old residential area has serious cultural conflicts between the site function and the surrounding housing need. Such question can determine the site type in the question category as “memorial site”, and the question features as “folk taboo” or related words.

However, the input of such correlation words has some difference in the description of problem characteristics. For this reason, the present study defines the synonym list of site contradiction. During the retrieval, the system expands

$$M = (\text{folk taboo, contradiction, burial area and residential area}) \Leftarrow \begin{cases} M_1 = (\text{national taboo, contradiction, islamic area and Catholic district}) \\ M_2 = (\text{family taboo, contradiction, town area and outside the house area}) \\ M_3 = (\text{personal taboos, contradiction, dynamic and static area}) \end{cases}$$

### 2.2.2. Preliminary sorting of site design cases based on correlation function

In extension logic, relevance function is used to express the relevance of problems, that is to say, there is a certain degree of similarity between things<sup>[15]</sup>. With this tool, the relevance of site design cases and search conditions can be screened. Such type of correlation screening is not consistent with general retrieval in that it can be preliminarily sorted according to the correlation degree.

(1) Correlation degree of site indicators. Interval correlation function can be adopted for calculation. Besides, four relational degree algorithms of extenics<sup>[12]</sup> are adopted respectively to solve four types of typical site design index query.

① Indicators have defined range boundary value (e.g., a proposed transportation hub distribution site) according to the Greenland rate search case. Designers consider site that should reflect the environment quality and the iconic landscape image, the preliminary set of the ideal value is 35%, while the lower limit of the field rate index is superior control rules demands and keep its traffic evacuation function limit, so the index of the acceptable range is set to [20

here retrieval items into synonym stable according to the implication analysis principle of extenics; subsequently, it matches the relevant data of the site design case, as a attempt to achieve the goal of accurately inquiring the related contradictory cases. The principles of implication analysis are as follows:

The problem feature item is regarded as a primitive, which starts from a primitive, and it can contain and extend multiple synonym primitives<sup>[12]</sup>,

$$M = (O, C, V) \Leftarrow \begin{cases} M_1 = (O_1, C_1, V_1) \\ M_2 = (O_2, C_2, V_2) \\ M_3 = (O_3, C_3, V_3) \end{cases}$$

For instance,

%, 60%]; this formula can fit the following  $X = \langle a, b \rangle$   
 $M \in X$

$$k(x) = \begin{cases} \frac{x-a}{M-a}, & x \leq M; \\ \frac{b-x}{b-M}, & x \geq M. \end{cases} \quad (1)$$

② The index has a minimum value requirement but no maximum value requirement (e.g., the site's requirement for the retreat distance of the main road). Thus, the calculation method of correlation function is as follows:

The optimal point is a finite interval, and the optimal correlation function is  $X = \langle a, +\infty \rangle$   $M \in X$

$$k(x) = \begin{cases} \frac{x-a}{M-a}, & x \leq M; \\ \frac{M}{2x-M}, & x \geq M. \end{cases} \quad (2)$$

③ Indexes have maximum requirements but no minimum requirements (e.g., building height control in site design). Thus, the calculation method of correlation function is defined below:

The positive domain is a finite interval, and the optimal correlation function is  $X = \langle -\infty, b \rangle$   $M \in X$

$$k(x) = \begin{cases} \frac{M}{2M-x}, & x \leq M; \\ \frac{x-b}{M-b}, & x \geq M. \end{cases} \quad (3)$$

④ Noscope requirements. This type of situation is special, generally existing in the subjective evaluation index of site design cases. Thus, the calculation method of correlation function is written as follows:

The positive domain is a finite interval, and the optimal correlation function is  $X \in (-\infty, +\infty)$ ,  $M \in X$

$$k(x) = \begin{cases} \frac{1}{1+M-x}, & x \leq M; \\ \frac{1}{x+1-M}, & x \geq M. \end{cases} \quad (4)$$

(2) Regional correlation degree of the site. Overall, large categories of cases are selected by regional climatic characteristic zoning and regional topography and geomorphology characteristic zoning related to the site design. Subsequently, the linear distance between the site design site and the case site is calculated according to the longitude and latitude coordinates. The premise condition refers to determining the longitude and latitude coordinates of all cases (e.g., "Qingdao" into "119°30'-121°00' east longitude and 35°35'-37°09' north latitude"). After obtaining the linear distance, the interval correlation function above can be adopted to sort the correlation degree.

(3) Correlation degree of site serial number. On the whole, the degree of correlation is converted into index-type features according to its serial number order (e.g., industrial and classification in site design,  $M_1, M_2, M_3, M_4$  Numbers 1, 2, 3, 4 can be corresponding, and then calculated according to the relational degree algorithm of index-type features).

(4) Discrete correlation degree of the site. In many cases, the value of key features of the site is discrete; for instance, the landscape design style of the site is modern, classical, compromise, etc. For these non-numerical discrete values, they can also be expressed in the quantitative system of 1,

$$\mathcal{E}(T) = \{(u, y, y') \mid u \in U, y = k(u) \in \mathcal{R}; y' = Tk(u)\} \quad (7)$$

Is an extension set on the domain  $U$ , and the extension function of  $y = k(u)$   $\mathcal{E}(T)$   $y' = Tk(u)$   $\mathcal{E}(T)$  [12].

2 and 3 as:

$$k(x) = \begin{cases} 1, & x = A_i; \\ 2, & x = B_i; \\ 3, & x = C_i. \end{cases} \quad (5)$$

(5) Comprehensive correlation degree of the site. In many cases, when retrieving some complex correlation degree, it is commonly determined by the combination of multiple features, so the weight of each feature should be determined for comprehensive calculation. Weight can be discussed by decision-makers, designers, users and other stakeholders via group focus discussion and lastly calculated with Delphi method.

When the weight coefficients of feature  $c_1, c_2, \dots, L$  and  $c_m$ , are  $\lambda_1, \lambda_2, \dots, L$  and  $\lambda_m$ , and satisfy the conditions of feature  $\sum_{i=1}^m \lambda_i = 1$ , the comprehensive correlation degree of stakeholder is expressed as [15]:

$$K(B) = \sum_{i=1}^m \lambda_i k_i(c_i(B)) = \sum_{i=1}^m \lambda_i k_i(x_i) \quad (6)$$

### 2.2.3. Dynamic sorting of site design cases based on extension set

The conventional planning and architectural design case sorting method only reflects the closeness between case and retrieval condition, but ignores the potential demand degree of case searchers. Therefore, after sorting the basic relational degree, users' search history traces of site design cases can be further analyzed, and the transformation effect of search requirements can be expressed by extension set theory, as an attempt to achieve dynamic sorting of site design cases on demand. The principles are as follows:

The case base of site design is regarded as  $U$ , and the site design case  $U$  is any element in  $U$ ,  $k$  is a mapping from  $U$  to the real domain, and  $T$  is a given transformation, called  $\mathcal{E}$

As indicated from the mentioned definitions, the extension set can be adopted to express the degree to which site desi

gncasesexhibitcertaincharacteristicsandthetransformationeffectof“yes”and“no”withtheNumbersin $(-\infty, +\infty)$ . It can be employed to express the processes of quantitative change (stability domain) and qualitative change (extension domain). The zero bound and the extension bound contain points of qualitative change, beyond which things vary qualitatively<sup>[12]</sup>.

The query traces of site design cases and search users are recorded, the evaluated value of site design cases in the exist in evaluation period and the next evaluation period is defined as correlation degree, and T denotes the transformation that causes the alteration of the evaluated value of site design cases in two cycles<sup>[12]</sup>. Where  $y = k(u)$  is the normalized value of the frequency of the site design case that is consulted within a period, as defined by the standard method of min-max

$$y = k(u) = 2(f - f_{\min}) / (f_{\max} - f_{\min}) - 1 \quad (8)$$

Through calculation, the site design cases that satisfy the search conditions are split into the corresponding five domains (i.e., the positive qualitative change, positive quantitative change, negative qualitative change, negative quan-

titative change, and zero)<sup>[12]</sup>. The grade value and correlation degree of site design cases are summarized according to the weight to form the final case extension evaluation result, which lays a solid basis for dynamic sorting of site design cases. This type of case retrieval method sorted by potential satisfaction can be effectively referenced in a case-based manner to solve contradictory problems in site design.

#### 2.2.4. Case base of site design retrieval function development

To enhance the efficiency of site design case user's retrieval, the vital functions of the case database (e.g., Internet data collection and acquisition, data analysis and processing, data format conversion, data storage and query) are further designed. The case base of site design retrieval function framework is designed by referencing the basic ideas of Guo Qiang and others scholar on building case database system, with BlueGriffon v3.0.1 as a web development tool and Jsp10 web programming language, as well as based on Microsoft IIS 7.0 web servers software. (Fig.2)

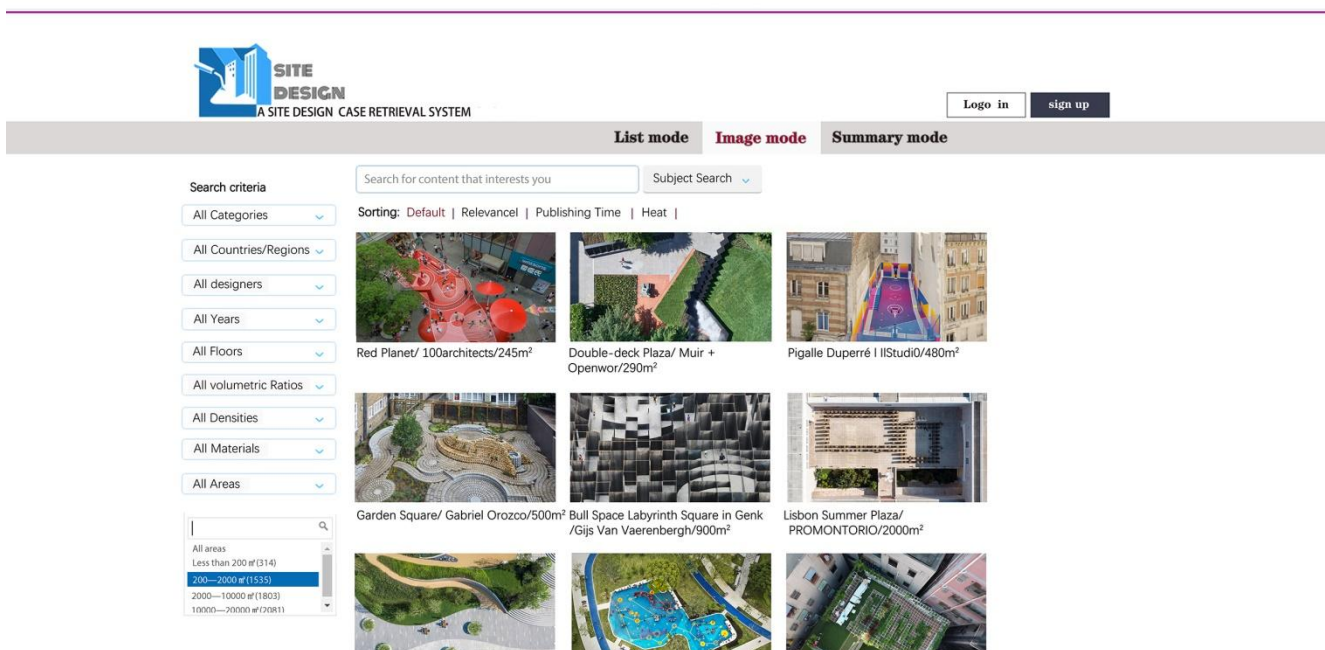


Fig.2. Case base of site design retrieval interface

The retrieval function of the system consists of four areas, i.e., display mode setting area, query mode setting area, qu-

ery condition setting area, as well as query results sorting area. Searchers can choose two ways, i.e., design subject qu-

ery and design contradiction query. In the query condition setting area, search cases scope can be limited (e.g., design type, designer, design place, design year, plot ratio, density and land area). The results are ordered by attention, relevance, release time, or system default. Among them, based on the theory of extension set, the case ordering according to the searcher's needs is realized by analyzing the retrieval behavior. Relevancy cannot only be used to calculate the sorting of the correlation degree between the general features of cases and the search conditions (e.g., the characteristics of release time and construction place), but also be able to sort the processing effect of the contradiction with the specific design according to the scheme, i.e., the significant advantage of this system compared with the conventional case retrieval system.

For query results display, there are three ways, i.e., support information list preview, literature summary preview and result thumbnail preview. In the result thumbnail preview mode, several corresponding site images can be presented according to different drawing modes (e.g., the diagram patterns of the landscape effect, the plan design, and the node construction).

### 2.3. Site design innovation strategy generation

Innovative strategies in site design often emerge with the process of solving design contradictions. This system will draw upon the powerful data analysis and processing ability of a computer to automatically generate considerable innovative design strategies when retrieving the cases that solve the contradictory problems of site design. Distinctive innovative elements, innovative methods and innovative technologies are used to assist site designers to break through present thinking dilemma and obtain design inspiration from numerous ideas.

To formulate innovative design strategies, compared with manual search cases, the advantages of this retrieval system are primarily reflected below: it can efficiently retrieve cases sorted by correlation degree from considerable cases; without considering the basic properties of the site (e.g., land area and land character), it can directly start with contradiction problems and select those valuable cases that are conducive to solving similar contradiction even though the basic properties exhibit great differences, as an attempt

to present ideas from the original starting point of innovation strategy.

### 3. Case test

Site design is considered the design activities that organize the relationships among various elements of the site by investigating the current, planning, standard and user conditions of the base and by delving into the feasibility and suitability of the project. Though some limited conditions of the site objectively caused difficulties for architectural design and landscape design, they also highlighted the direction for the innovation of the site design. Through the site design innovation to solve the mentioned difficulties, it will be transformed into a referenced design strategy for the subsequent design scheme to present novel ideas.

For instance, during site design, the steep slope of mountainous land form often imposes considerable difficulties on the design. A question is raised that how to exploit the height difference to achieve the innovative design according to local conditions? The following provides an example of the site design of mountainous park to conduct case retrieval based on the solution of this contradiction problem.

(1) Set the query mode to "contradictory query" and input the contradictory problem to be solved in its retrieval box. In the description of specific contradiction problems, the fundamental contradiction to be solved should be expressed clearly with professional terms (e.g., "mountain terrain height difference suitability treatment"). However, an amateur case database covers numerous cases, architectural cases involved diverse design factors, and a wide variety of information are intricate, which will waste time and reduce the retrieval efficiency. Accordingly, the search scope can be further narrowed by defining the basic information about the case (e.g., design type, design year and land area) in the search condition bar on the left side of the interface. Then, "design type" is set as "park". Thus, the ultimate condition of the case retrieval of the whole contradiction problem refers to "the suitability treatment of terrain elevation difference in mountainous park".

(2) Set the display mode of query results. Query results can be expressed according to three modes, i.e., the designer's personal need information list preview, result thumbnail preview and literature summary preview. In this study, the



eliteratureabstractpreviewmodeisselectedtodisplayth  
equeryresults.

(3)Setthesortingmethodofqueryresults.Besidesthedef  
aultsorting,threesortingmethods(i.e.,relevance,relas  
etimeandpopularity)areadded.Thecorrelationdegreer  
eferstotheorderofthetreatmenteffectoftheschemeofthe  
specificdesigncontradictoryproblem.This time,thissor  
tisselected.

(4)Afterthementionedstepsareset,the“contradictionqu  
ery”buttonisclickedtoretrieveandfindrelevantsitedesi  
gncases.Withthecasebaseofsitedesignastheplatform,t  
hetextanddrawinginformationofcasescanappear,andt  
hesitedesigninnovationstrategycanbeautomaticallyge  
nerated,asillustratedinthe“solutionstrategy”columnat

theendofeachcase.Itissuggestedthatthesearchresultsa  
ndoneoftheexamplesisthemining siteparkinBelgium.T  
heplatformaimstotreattheheightdifferencewiththemet  
hodofsteps.Thesiteisplannedinto4gradientsbycomply  
ingwiththeappropriategradient,andtheplatformissetatt  
heintersectionoftwoadjacentgradientsfortouriststoaha  
vearest.Inthiscase,thefourslopesgraduallyincrease,an  
dthesidesofthemountainareconsideredanadventurous  
prismatic surface.The designcleverlyindicatethedesig  
nthemeofwin-wincooperation(thehigheryouclimb,the  
moredifficultitwillbe,whichrequirescooperationandm  
tualencouragementuntilyoureachthetopofthemounta  
in).Thiscaseemploysthespecificdesignthemeandtechn  
iquetoskillfullysolvethe designproblemofthelargeslop  
elandinthesitedesignofthemountainpark.(Fig.3)

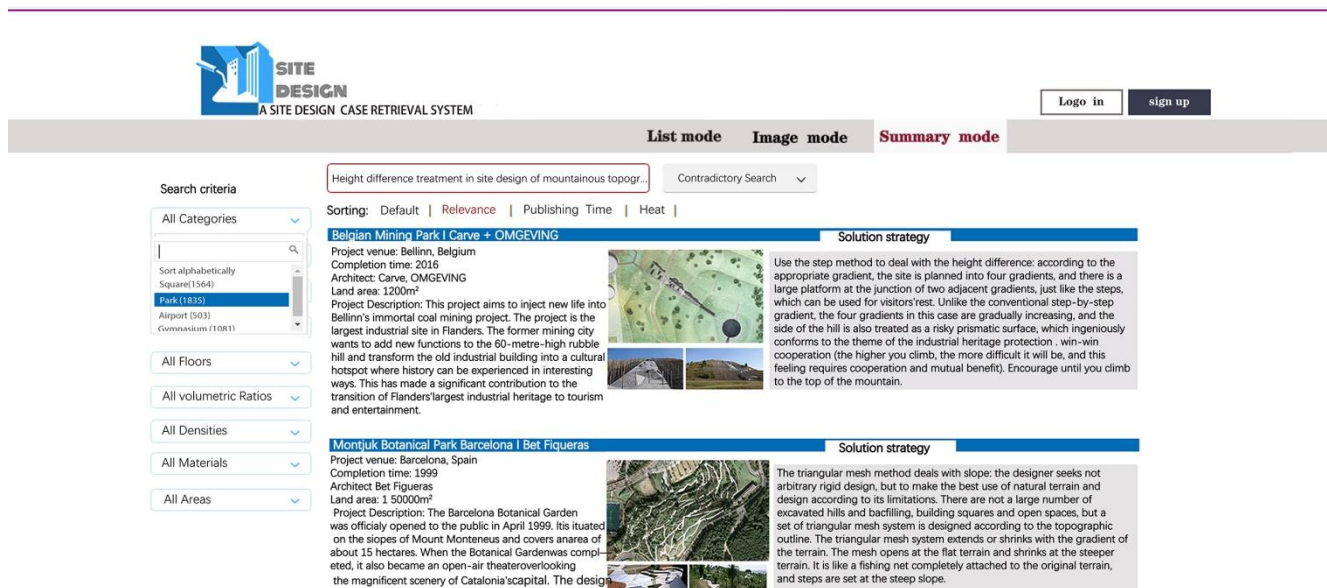


Fig.3.Searchresultsundersitedesignstrategygenerationmode

#### 4. Conclusions

A method of site design case retrieval and methods of strategy generation based on extension set refer to comprehensively using extension set theory and data analysis tools, meet the needs of the case users, convert the site contradictions into the search conditions, apply correlation function to the preliminary screening of site design cases, and exploit extension set to achieve dynamic sorting according to the search requirements, as to achieve the optimized retrieval

val function of the site design case database.

(1) The present study proposes the steps of data collection, data processing, database construction, case sorting optimization and site design innovation strategy generation of site design, and assist designers to efficiently find valuable site design cases from massive Internet data. (2) As revealed from the case test, the database is comprehensive, efficient and convenient, especially suitable for site design beginners. (3) The application of database technology and extension method enhances the ability of designers to adopt

computer technology to collect and integrate data on the Internet, as well as to exploit data to automatically generate design strategies, which speeds up the intelligent process of site design to some extent.

### Acknowledgement

This work is supported by Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (No. 2019RCJJ025).

### Reference

1. Viktor, A., Christoph, L., Johannes, R., Saqib, B., et al. (2016). Comparative evaluation of rule-based and case-based retrieval coordination for search of architectural building designs. ICCBR2016-24th International Conference on Case-Based Reasoning Research and Development, Atlanta, pp. S16-S31.
2. Viktor, A., Christoph, L., Saqib, B., Klaus-Dieter, A., et al. (2016). Thinking with containers: A multi-agent retrieval approach for the case-based semantic search of architectural designs. ICAART2016-Proceedings of the 8th International Conference on Agents and Artificial Intelligence, Saxony, pp. S149-S156.
3. Viktor, A. (2017). Current state and further development of a case-based framework for early phases of architectural conceptual design. CEUR Workshop Proceedings, pp. S224-S228.
4. Qamer Uddin, S., Johannes, B., Viktor, A., et al. (2017). Semantic pattern-based retrieval of architectural floor plans with case-based and graph-based searching techniques and their evaluation and visualization. 6th International Conference on Pattern Recognition Applications and Methods, pp. S50-S60.
5. Cekmis, A. (2016). Fuzzy logic in architectural site planning design. Procedia Computer Science, 102(8): 176-182. <https://doi.org/10.1016/j.procs.2016.09.386>
6. Andrés, C., Preetam, C., Urjit, B., Ashwin, R. (2010). CBArch: A Case-Based Reasoning system for conceptual architectural design. 10th International Conference on Design and Decision Support Systems 4, Eindhoven, pp. S30-S38.
7. Ramon Lopez, D., David, M., Derek, B., David, L., et al. (2005). Retrieval, reuse, revision and retention in case-based reasoning. Knowledge Engineering Review, 20(3): 215-240. [10.1017/S0269888906000646](https://doi.org/10.1017/S0269888906000646)
8. Ian, S., Claudio, L., Boi, F. (1995). Spatial composition using cases: IDIOM. 1st International Conference on Case-Based Reasoning, pp. S88-S97.
9. Oxman RE, Oxman RM. (1993). Rememberance of things in the past: design precedents in libraries, Automation in Construction. 2(1): 21-29. [10.1016/0926-5805\(93\)90032-S](https://doi.org/10.1016/0926-5805(93)90032-S)
10. Flemming, U. (1994). Case-based design in the SED system. Automation in Construction, 3(2-3): 123-133. [10.1016/0926-5805\(94\)90014-0](https://doi.org/10.1016/0926-5805(94)90014-0)
11. Dave, B., Schmitt, G., Shih, S., etc. (1995). Case-Based Spatial Design Reasoning. In European Conference on Case-Based Reasoning, 2nd European Workshop on Case-Based Reasoning, Paris, pp. S198-S210.
12. Yang, C., Wen, C. (2013). Extensics: Theory, Method and Application. Science Press, 7(5): 27-43, 185-191. [10.1016/j.procs.2015.07.045](https://doi.org/10.1016/j.procs.2015.07.045)
13. Pankaj Deep, K., Gitanjali, S. (2015). Scalable Database Management in Cloud Computing. Procedia Computer Science, 70(10): 650-667. [10.1016/j.procs.2015.10.102](https://doi.org/10.1016/j.procs.2015.10.102)
14. Han, J., Micheline, K., Pei, J. (2012). Data Mining Concepts and Techniques. China Machine Press, 58-66.
15. Guo, Q., Zou, G., Sun, T. (2018). Discovery of the Knowledge on the Demands of Building Users Based on Extension Clustering. Revue d'intelligence artificielle, 32(5-6): 705-718.