International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:5, 2012

Development of Decision Support System for House Evaluation and Purchasing

Chia-Yu Hsu, Julaimin Goh and Pei-Chann Chang

***Abstract*—**Home is important for Chinese people. Because the information regarding the house attributes and surrounding environments is incomplete in most real estate agency, most house buyers are difficult to consider the overall factors effectively and only can search candidates by sorting-based approach. This study aims to develop a decision support system for housing purchasing, in which surrounding facilities of each house are quantified. Then, all considered house factors and customer preferences are incorporated into Simple Multi-Attribute Ranking Technique (SMART) to support the housing evaluation. To evaluate the validity of proposed approach, an empirical study was conducted from a real estate agency. Based on the customer requirement and preferences, the proposed approach can identify better candidate house with consider the overall house attributes and surrounding facilities.

***Keywords*—**decision support system, real estate, decision analysis, housing evaluation, SMART

I. Introduction

HOUSE purchasing is one of important decision due to the largest expenses in a lifetime for most people. Based on different customer demand and purchasing ability such as income, each house could be satisfied to certain of customer. In order to facilitate house purchasing, real estate agencies play an important role to provide information and coordination between the sellers and customers. With advanced development in information technology and intelligent applications in recent years, real estate market has been driven to adopt innovative e-transformations [8]. Therefore, most of information regarding the housing can be captured in different real estate agencies in the website. Web service for surveying related information from the websites of real estate agencies is critical for customer before they communicate with real estate agencies. However, most websites of house agencies only provide ranking-based results by area, price, and housing size for customers without considering the overall customer preferences. To maintain the competitive advantage and capture the market sharing, house agencies need to match requirement of customers effectively. Housing attributes and surrounding environment of housing are considered for purchase housing.

Chia-Yu Hsu is the assistant Professor with the Department of Information Management, Yuan Ze University, Chungli, 32003, Taiwan (phone: +886-3-4638800 ext.2793; fax: +886-3-435-2077; e-mail: cyhsu@saturn.yzu.edu.tw).

Julaimin Goh is the master student with the Department of Information Management, Yuan Ze University, Chungli, 32003, Taiwan (e-mail: julaimon@hotmail.com).

Pei-Chann Chang is the Professor with the Department of Information Management, Yuan Ze University, Chungli, 32003, Taiwan (e-mail: iepchang@saturn.yzu.edu.tw).

Most information of housing attributes including price, housing size, address and number of room are provided without lacking for quantitative information of surrounding environments for each house. In particular, some facilities can enhance the quality of living such as school, super market, convenience store and some facilities will distract people such like funeral and tomb. Therefore, the spatial information of surrounding area needs to be identified by using Geographic Information System (GIS). However, most information available on the web is only readable to humans through presentation oriented HTML pages [7] [9].

This study aims to develop a decision support system for housing purchasing, in which Simple Multi-Attribute Ranking Technique (SMART) is applied to support the housing evaluation based on customer preference. The proposed methods provide not only housing attributes but also information of surrounding environment for making decisions of housing purchasing. With the developed decision support system of housing purchasing, customers can get required information effectively to enhance decision quality.

II. Literature Review

A. Taiwan Real-Estate Condition

Real-estate agency is to provide service and make communication between buyer and seller. According to Department of Land Administration of Taiwan until 2011, there are around 1685 agencies store which provide service for customer. The types of house purchasing can be classified into two types [1]. The first one is “full-time entrusted” means that one customer only entrusts with one real estate agency. The “generally entrust” means that one customer entrusts with many real estate agencies. However, customers need to use a lot of time for searching and matching their ideal house.

House is one of luxury product with heterogeneity, expensive and low buying frequency that can be waste a lot of time and spend a lot of money on it. Generally, the housing buyers in Taiwan need around six months to find an ideal house without including investor [3].

House interior factors and location place are two mainly considered for buyer [2]. Each buyer also has different preference in ideal house. House factors can be divided into three aspects including environmental, social, and personal as listed in Table I [2]. Commonly real-estate agency website provides housing sequence for buyer by search-based engine, but only binary information for their location of facilities nearby is provided (i.e. yes and no). Customer usually considers not only house attribute but also traffic conditions, surrounding facilities and environmental. Therefore, we consider the distance in between house and facilities and incorporate such information for buyers into match their preference.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:5, 2012

TABLE I
HOUSE FACTORS AND DEFINITION [2]

Factors	Definition	Type
Environmental	can cause pollution or relaxation place	Recycle sites, funeral/tomb, gas station, park and other environmental factors.
Social	making people life much easier,	Convenience store, supermarket, train station and other social factor.
Personal	condition life or wealthy	House price, location work and other personal factors.

A. Simple Multi-Attribute Ranking Technique (SMART)

SMART is one of the Multi-Attribute Decision Making (MADM) techniques. SMART was proposed by Edwards [6] in 1977. SMART has been widely used in various properties independent decision-making problems and used simple calculation weights; that will reduce user on making decision on what house that they will buy [5]. It is popular because its analysis incorporates a wide variety of quantitative and qualitative criteria [13]. SMART uses the simple additive weight (SAW) method to obtain total values for individual alternatives, helping to rank them according to order of preference [10] [11]. It has been successfully applied in many MADM problems.

Most people have different preferences and likeness for purchasing ideal house and usually need a long time in house matching. SMART considers multiple criteria or condition in decision making environments. In dealing with multi-attribute assessment of problems, often with the help of some tools and standardized operating procedures, to help decision-makers to achieve systematic decision-making purposes. Each attribution of weights to criteria should be made by the scaling constants (parameters associated with the degree of importance of the criteria), that indicate the value of tradeoffs between the various pairs of attributes [12]. We focus on considering the two kinds criteria. The first is finite numeric input for housing attribute, for example, unit price, house size, and numbers of room. The second is the select input whether what kind facilities that they want in the surrounding house. SMART can provide the overall evaluation including housing itself and surrounding environment for candidate housing according to the customers' preferences. Ratings of alternatives in SMART are assigned directly in the natural scales of criteria. SMART model is independent of the alternative, so that changing the number of alternatives considered will not in itself change the decision scores of the original alternatives. If new alternatives are likely to be added to the model after its initial construction, and the alternatives are amenable to a direct rating approach, then SMART would be a good choice.

III. PROPOSED APPROACH

The notations and terminologies used in this study are listed as follows.

x_{ij}	Each variable house attributes.
r	Room input.
n	Total elements
b_{ij}	Purchase ability
b_{high}	Upper bound of purchase ability
b_{low}	Lower bound of purchase ability.
p_{high}	Upper bound of price per area input.
p_{low}	Lower bound of price per area input.
R_j	Customer preference.
RS_{ij}	Preference score.
RB	Parameter on purchase.
RH	Parameter on house size.
RP	Parameter on price per area.
WS_j	Smart weight on each house case.
WB_{ij}	Purchase weight on each house case
FS_j	Final score of each house as evaluation.
SA_{ij}	Score of each house attribute.

The proposed framework is shown in Fig. 1. In the first phase, user inputs house factors and preferences. In particular, housing attribute and housing facility are considered as listed in Table II. The positive scores of each attribute represent the positive house factor. The negative scores represent the negative impact for candidate housing in the surrounding environment. Price per area is used rather than total price to measure the cost of purchasing house. Total price represents the purchasing ability. In the second phase, different kindly of weights are calculated for further evaluation. In the third phase, the overall score for each candidate house will be calculated and then the houses with large score are identified.

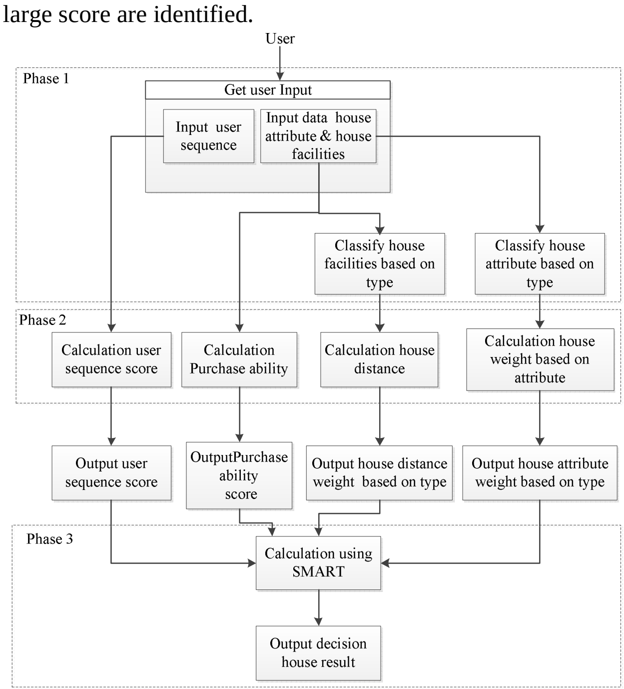

Fig. 1 Research framework

TABLE II
CONSIDERATION OF HOUSE FACTORS

Factors	Type	Sub-type
House attribute	Price per area	-
	House area	-
	Room	-
	Community	Security
		Parking lot
Surrounding Facility	Education	Kindergarten
		Elementary school
		Junior school
		High school
		University
	Social living	Bank
		Convenience store
		Deptartment store
		Supermarket
		Clinic
		Post office
		Entrance science-park
	Transportation	Train station
		Bus station
		High speed rail station
	Negative impact	Life-service
		Tomb
		Recycle
		Gas station

A. Calculation of House Attribute

Both user preference sequences and house attributes are considered that user prefer; where we described in Fig. 1 in phase 1. In the framework we divide into 3 important parts calculation. In this phase we calculate each house attribute based on type each user wants, as we listed in below:

01. House price per area

House price per area score is calculated by using the curve line estimation as shown in Equation (1). To consider the characteristic of price per area, the score within the limit of user-determined decreases slowly. Each minimum and maximum input value will extend 20% limits to consider more candidates.

$$SA_{ij} = \frac{1-\exp\left[-\dfrac{p_{high}-x_{ij}}{RP}\right]}{1-\exp\left[-\dfrac{p_{high}-p_{low}}{RP}\right]} \quad (1)$$

$$RP = p_{low} - (p_{low} \times 0.7) \quad (2)$$

Fig. 2 shows that the calculation of house price per area given interval between 15 until 30 (in thousand)

Step 1: Divide the interval

Step 2: Use (1) for the simultaneous equations, finding SA_{ij} and RP.

Step 3: Use SA_{ij} calculate each house attribute score of house price per area attribute

Fig. 2 House area size interval

02. House size

We count the house size score also using the curve line estimation as shown in Equation (3). The considered hose size is extended 20% based on the upper bound and lower bound. The difference in this attribute calculation is decade until 0.4.

$$SA_{ij} = \frac{1-\exp\left[-\dfrac{(x_{ij}-h_{low})}{RH}\right]}{1-\exp\left[-\dfrac{(h_{high}-h_{low})}{RH}\right]} \quad (3)$$

$$RH = h_{low} - (h_{low} \times 0.9) \quad (4)$$

03. Number of room

To quantity the number of room for each house, it is defined based on user input and the number of rooms. If the number of room (x) is equal to the number of room in house, the room weight will be 1. The score for other number of rooms will decrease.

$$x = \begin{cases} x_{ij}-2 & SA_{ij}=0.1 \\ x_{ij}-1 & SA_{ij}=0.6 \\ x_{ij} & SA_{ij}=1 \\ x_{ij}+1 & SA_{ij}=0.8 \\ x_{ij}+2 & SA_{ij}=0.2 \end{cases} \quad (5)$$

04. Community

For the weight calculation of community, it is depends on whether house included the parking lot or security. If user preferences want parking lot and security, than if there is one house match with the both condition than the community weight will be 1 and if just match in one condition the community weight will be 0.5. The second parts are calculation of house facilities. For house facilities we mainly divided into four types; education, social living, transportation, and negative factors. In this paper, the distance between house and facilities are estimated by using Harvesine formula [4].

B. Calculation of House Facilities

Three different calculations based on their type of facilities are used for quantification of community.

01. Based on nearest distance

Count the nearest distance between house and house factor. m^{i}_{jk} is the nearest distance house and facility, x^{i}_{j} is the threshold that we set based on expert experience and $m^i_{jk}*$ is the weight of each facilities that used for calculation

$$m^i_{jk}* = \begin{cases} 1 & m^i_{jk} \le x^i_j \\ 1 < x < 0 & x^i_j < m^i_{jk} \le x^i_j \\ 0 & m^i_{jk} > x^i_j \end{cases} \quad (6)$$

02. Based on density

It means that how many facilities that surrounding house area. n^i_{jk} is the count of the facilities that is in x^i_j threshold.

$$n^i_{jk} = \begin{cases} 1 & m^i_{jk} \le x^i_j \\ 0 & m^i_{jk} > x^i_j \end{cases} \quad (7)$$

$n^i_{jk}*$ is the weight of the each facilities based on density used for calculation.

$$n^i_{jk}* = \begin{cases} 1 & n^i_{jk} \le x^i_j \\ 1 < x < 0 & x^i_j < n^i_{jk} \le x^i_j \\ 0 & n^i_{jk} > x^i_j \end{cases} \quad (8)$$

03. Based on nearest distance and density

To combine the weight of nearest distance and density. $m^i_{jk}{}'$ is the weight of each facilities based on nearest distance and density.

$$m^i_{jk}{}' = \begin{cases} 1 & \frac{m^i_{jk} + n^i_{jk}*}{2} \ge 1 \\ m^i_{jk} + n^i_{jk}* & \frac{m^i_{jk} + n^i_{jk}*}{2} < 1 \end{cases} \quad (9)$$

C. Weight considered factor and calculation final score.

Beside the attributes we mention; in this thesis also considered user purchase ability, in transformation to score using the estimation of curve adaptive line as shown in Equation (10). Also in each minimum and maximum input value will add 20%. The difference in this attribute calculation is decade until 0.6 .

$$wb_{ij} = \frac{1 - \exp\left[-\frac{(x_{ij} - b_{low})}{RB}\right]}{1 - \exp\left[-\frac{(b_{high} - b_{low})}{RB}\right]} \quad (10)$$

$$RB = b_{low} - (b_{low} \times 0.9) \quad (11)$$

In the last main calculation weight is for the calculation the user sequence, as in below step:

Step 1: Define the consider value of decision-maker.

Step 2: Determine the attribute for decision making and the available option.

Step 3: Re-arrange preferences attribute from the most preferences element to less as $A_1, A_2, ..., A_n$ corresponding to order $R_1, R_2, ..., R_n$.

Step 4: Calculate weight preference by using Rank Sum (RS) to transform to weight score are shown in (12).

$$RS_{ij} = \frac{n + 1 - R_{ij}}{n(n+1)/2} \quad (12)$$

Where w_i a preference weight of each element is, n is the total elements, and R_i is the given preference.

D. Weighted Considered Factors

In the last phase, after each attribute, facilities, purchase ability weight and sequence weight, for the final evaluation recommendation real estate, as in below step:

Step 1: Calculate each attributes weight.

Step 2: Calculate the total score for each preference and attribute in (13).

$$ws_{ij} = \sum_{i=1}^{n} RS_{ij} \times SA_{ij} \quad (13)$$

Step 3: Calculate the final score in Equation (14).

$$FS_{ij} = (0.5 \times ws_{ij}) + (0.5 \times wb_{ij}) \quad (14)$$

Step 4: Order from highest FS_{ij} to lowest in top 30; the highest score is the best selection according on user preference.

IV. Empirical Study

A. Data Collection.

To validate the proposed approach, an empirical study was conducted from a real estate agency in Hsinchu, Taiwan. Hsinchu covers an area around 40.2 square meters with a diversity of environment, residential density and has had a constant increase in property prices. Total 1200 secondhand houses collected around six months were used as candidates in the DSS. The distances between each house and house facilities are calculated through Geographic Information System.

To illustrate the proposed approach, one customer is selected from the historical database. He wants to purchase a house around the Hsinchu city, and look for a price per area in between 15 to 25 thousand per area with affordable price between 500 and 700, the house size is between 50 ping (1ping = 36 square foot.) and 70 ping reason for purchasing a new house is because in their old house room is not enough.

TABLE III
CALCULATION OF USER PREFERENCE WEIGHT

Type	User Preference (n)	Preference Weight (RS_{ij})
Price per area	1	0.25
House area	2	0.214
Room	4	0.143
Community	5	0.107
Education	6	0.071
Social living	3	0.179
Transportation	0	0
Negative impact	7	0.036

The housing is far from their child school also nearby didn't have any convenience and has a gas station nearby. He also needs a place to park his car. His preference for the housing factors includes price per area, house area size, education, social living, parking lot, room and the gas station. The considered factors and preferences are calculated as listed in Table III.

B. Evaluation and Result Discussion.

After the calculation of user preference weight, we can calculate the overall score for each house. Table IV lists the score of first five candidate houses and Table V lists detail house information of each candidate. The condition of house ID 513 matches with customer where they can find the nearest elementary school near house. Although house 1768 has the nearest convenience store but total price house is the highest among these five houses and the score is not highest. To evaluate the effectiveness of proposed approach, sorting-based method which is mainly ranked by price per area and house area is used to examine the validity. In particular, the both first 20 candidates are selected in Table VI, respectively. The average score of candidates provided by proposed approach (92.14) have higher score than consideration based on price per area approach (88.96). Result consideration by price per area shows that most of the evaluation result is out limit of purchasing ability of user.

V. CONCLUSION

House purchasing is critical decision due to the complex considered factors including house attributes and surrounding facilities. Customer would like to purchase cheap house but also to consider the surrounding area and live quality. This study applies GIS to incorporate the information of surrounding facilities and proposes a decision model for house evaluation by SMART for house buyer. According to the empirical study, based on the customer preferences, the proposed approach can identify the better candidate houses than traditional sorting-based approach. Further research can consider more surrounding facilities including subway, work place and select other cities to evaluate the validity of proposed approach.

ACKNOWLEDGMENTS

This research is appreciatively supported by National Science Council, Taiwan (NSC99-2622-E-007-015-CC3; NSC 100-2410-H-155 -048).

REFERENCES

[1] L. Y. Lin, and C. H. Huang, "An Integrated Information System for Real Estate Agency Based on Service-Oriented Architecture," *IEEE International Symposium on Service-Oriented System Engineering*, pp. 184-189, Dec. 2008.

[2] Q. Z. Thomas, and Q. M. Zhou, "Optimal Spatial Decision Making Using GIS: a Prototype of a Real Estate Geographical Information System (REGIS)," *International Journal of Geographical Information Science*, vol. 15, pp. 307-321, Jun. 2001.

[3] M. L. Chou, and C. O. Chang, "Influential Factors of Home Buyers' Search Duration," *Management Rev.*, vol. 24, pp. 133-150, 2004.

[4] R. W. Sinnott, "Virtues of the Haversine," *Sky and Telescope*, vol. 68, no. 2, pp. 159, 1984.

[5] Y. M. Wang, and C. Parkan, "Multiple Attribute Decision Making Based on Fuzzy Preference Information on Alternatives: Ranking and Weighting," *Fuzzy Sets and Syst.*, vol. 153, pp.331-346, Aug. 2005.

[6] F. H. Barron, and B. E. Barrett, "The Efficacy of SMARTER-Simple Multi-Attribute Rating Technique Extended to Ranking," *Acta Psychologica*, vol. 93, pp. 23-36, 1996.

[7] W. Edwards," How to use Multiattribute Utility Theory for Social Decision Making," *IEEE Trans. on Systems, Man, and Cybern.*, vol. 7, pp. 326-340, 1997.

[8] C. F. Chien, *Decision Analysis and Management* (Book style), Taiwan: Yeh Yeh Book, 2005, pp. 179-219.

[9] P. Kim, "Development of Spatial Decision Support Systems for Residential Real Estate," *Journal of Housing Research*, vol.9, pp. 135-136, 1998.

[10] S. K. Chung, and G. Tor, "Selecting DSS Evaluation Methods," *Journal of Information Technology Management*, vol. 3, pp. 29-37, 1992.

[11] E. K. Zavadskas, A. Kaklauskas, and A. Banaitis, "Real Estate's Knowledge and Device- Based Decision Support System," *International Journal of Strategic Property Management*, vol. 14, pp. 271-282, 2010.

[12] G. Shi, and K. Barker, "Thematic Data Extraction from Web for GIS and Applications," *Spatial Data Mining and Geographical Knowledge Services*, pp. 273-278, 2011.

[13] S. R. Ahmed, "Applications of Data Mining in Retail Business," *Int. Conf. Information Technology: Coding and Computing*, vol.2, pp. 455-459, Apr. 2004.

[14] Q. P. Xiao, M. C. Fu, G. J. Wang, and J. J. Zhang, "House Site Selection Decision Support System Based on ArcGIS," *Int. Conf. Management and Service Science*, pp. 1-4, Sept. 2009.

Chia-Yu Hsu is an assistant Professor of Information Management at the Yuan Ze University. He received B.S. in Statistics from the National Cheng Kung University in 2002. He received the M.S. and Ph.D. degrees in Industrial Engineering and Engineering Management from the National Tsing Hua University in 2004 and 2009, respectively. His research interests include decision analysis and decision support system, data mining, manufacturing intelligence, yield enhancement. His research works appear in Computers & Industrial Engineering, Journal of Intelligent Manufacturing, and Journal of Quality.

Julaimin Goh is master student in Information Management at the Yuan Ze University. He received B.S. in Management Information System from the Chung Yuan Christian University in 2008. His research interests include system development, decision support system, and intelligent algorithm.

Pei-Chann Chang is a Professor at the Department of Information Management, Yuan-Ze University. He received the M.S. and Ph.D. degrees from Lehigh University, Bethlehem, PA, in 1985 and 1989, respectively. He is the Senior Editor of the Journal of the Chinese Institute of Industrial Engineering. He is the author or coauthor of more than 60 publications in international journals. His current research interests include financial time-series forecasting, evolutionary computation, fuzzy neural applications, production scheduling, forecasting, case-based reasoning, and applications of soft computing.

TABLE IV
HOUSE EVALUATION RESULT (ONLY FIRST FIVE HOUSE)

Type	Preference Weigh RS_{ij}	House 1 (513)	House 2 (1800)	House 3 (404)	House 4 (1119)	House 5 (1768)
Price per area	0.25	1	1	1	1	1
House area	0.214	1	0.976	0.993	1	1
Room	0.143	1	1	1	1	1
Community	0.107	1	1	1	1	1
Education	0.071	1	1	1	1	1
Social living	0.179	0.905	0.728	0.714	0.637	0.6
Transportation	0	0	0	0	0	0
Negative impact	0.036	1	1	1	1	1
Purchase Ability	X	0.5	0.5	0.499	0.5	0.496
Total score		95.5	93.7	93.7	93.2	93.1

TABLE V
HOUSE INFORMATION DETAIL

Type	House 1 (513)	House 2 (1800)	House 3 (404)	House 4 (1119)	House 5 (1768)
Price (in thousand)	468	550	588	568	650
Price per size (in thousand)	5.85	10	9.8	7.573	9.043
House area (ping)	80	55	60	75	71.88
Room	4	4	4	4	4
Parking Lot	✓	✓	✓	✓	✓
Kindergarten distance(in Kilometer)	0.08	0.3	0.26	0.16	1.04
Convenience store distance(in Kilometer)	0.06	0.11	0.12	0.18	0.02
Gas station distance(in Kilometer)	0.16	0.68	0.44	0.71	1.66

TABLE VI
COMPARISON TABLE

Calculation by Consideration Price per area					Proposed Method				
House ID	Price	Price per area	House area	Score	House ID	Price	Price per area	House area	Score
1630	998	13.307	75	92.68	513	468	5.85	80	95.5
149	836	15.2	55	91.79	1800	550	10	55	93.7
701	1058	17.975	58.86	91.46	404	588	9.8	60	93.7
1803	998	17.787	56.11	91.31	1119	568	7.573	75	93.2
1828	958	17.022	56.28	91.31	1768	650	9.043	71.88	93.1
105	928	17.663	52.54	90.33	318	658	8.773	75	92.9
45	1398	20.938	66.77	90.18	444	458	7.633	60	92.8
2091	1180	20.345	58	89.99	93	598	8.543	70	92.5
498	798	16.556	48.2	89.14	1870	688	13.406	51.32	91.7
1633	1288	17.788	72.41	88.84	1407	698	13.207	52.85	91.6
1206	1158	13.986	82.8	88.63	2063	658	13.514	48.69	91.6
1058	988	17.964	55	88.33	477	678	11.63	58.3	91.5
1834	1290	16.669	77.39	87.16	2038	698	13.601	51.32	91.4
2094	1288	16.555	77.8	87.16	711	590	9.833	60	91.3
1964	1180	19.281	61.2	87.03	124	575	10.21	56.32	91.2
256	730	12.167	60	86.93	2120	658	13.847	47.52	91.2
2131	808	13.871	58.25	86.93	271	598	13.289	45	91.1
2164	808	13.871	58.25	86.93	1018	598	7.475	80	91
749	1568	21.896	71.61	86.64	1096	668	10.277	65	90.9
966	928	16.873	55	86.46	924	598	9.967	60	90.9
Avg. Score				88.96	Avg. Score				92.14