The Impact of Community Safety on House Ranking

Zijun Yao* Yanjie Fu‡ Bin Liu◊ Hui Xiong+†

Abstract

It is well recognized that community safety which affects people's right to live without fear of crime has considerable impacts on housing investments. Housing investors can make more informed decisions if they are fully aware of safety related factors. To this end, we develop a safety-aware house ranking method by incorporating community safety into house assessment. Specifically, we first propose a novel framework to infer community safety level by mining community crime evidences from rich spatio-temporal historical crime data. Then we develop a ranking model which fuses multiply community safety features to rank house value based on the degree of community safety. Finally, we conduct a comprehensive evaluation of the proposed method with real-world crime and house data. The experimental results show that the proposed method substantially outperforms the baseline methods for house ranking.

Keywords: Community safety, House ranking, Spatio-temporal

1 Introduction

Community safety describes the degree that people live without fear of crime, such as the risk of being victimized in burglary, robbery, or assault. It has become a fundamental buying factor of houses nowadays. Community safety issues can severely damage the value of a house by: 1) endangering occupants and properties; 2) degrading living and business environments, e.g., people are less likely to rent for living or business; 3) obstructing the development of the area, e.g., new houses or infrastructure are less likely to be built nearby. Therefore, from the perspective of investors, it is requisite to be aware of potential community safety issues.

Empirical studies have confirmed the importance of community safety factors for house appraisals. For example, a standard deviation increase in the local density of property crime causes a 10% decrease in the price of an average property in London [1]; the move-in of a sex offender leads to a 2.3% fall in nearby housing prices in Hillsborough County, Florida [2]. These evidences show that the value of houses can be significantly influenced by community safety issues.

Motivated by the above, it is appealing to provide a tool for investors to rank house values based on the degree of community safety. For investigating the impacts of community safety on house values, historical crime especially house burglary (also called break-in) is a valuable resource for the following reasons: 1) burglary directly threats houses, 2) burglary is commonly spread in all locations, 3) burglary provides sufficient historical cases for investigation. With rapid advances in positioning technology, data with fine-grained locations such as coordinates of crime records is now available. This allows us to appraise houses via their neighboring crimes which make direct impacts.

Although there are a few studies which have investigated the impacts of crime on house appraisals [1–3], they suffer from two limitations: 1) naive crime statistics (e.g., counting crime cases), which can be improved by sophisticated crime analysis to get in-depth understanding of community safety; 2) traditional appraisal models (e.g., Hedonic regression) which include multiply influencing aspects in house appraisal function, mask the impacts of community safety. Unlike prior studies, we want to comprehensively learn local safety levels by employing factors not limited to naive crime statistics. Moreover, we want to solely focus on the influence of community safety on house appraisal. Therefore, we tackle two research challenges in this paper, **Challenge 1**: what crime analysis can be done to generate in-depth understanding of community safety; **Challenge 2**: how to systematically model the impacts of community safety on house values without effects of other aspects, such as neighborhood income level and rating of nearby schools?

For **Challenge 1**, we identify two categories of discriminative crime evidences that comprehensively describe community safety. The first category is based on crime severity which focuses on property losses led by different burglary crimes. Since available crime information does not provide actual losses explicitly, we derive evidences to infer the severity of crimes implicitly: occurrence address evidence and occurrence time evidence

*‡◊+Rutgers University. Email: {zijun.yao, yanjie.fu, binben.liu, hxiong}@rutgers.edu.

†Contact Author.

of crimes. The second category is based on temporal correlation which considers the correlation of crimes to learn community safety. The crime temporal correlation reflects an important phenomenon called near repeat in criminology, which implies degraded community safety and increased victimization risk [4, 5]. Therefore, we mine near repeat series to discover temporal correlations of crimes around houses. Based on near repeat series, we extract evidences to detect temporal correlations: series size evidence, series length evidence and series intensity evidence.

For **Challenge 2**, we propose a ranking model to understand the impacts of community safety. Since mostly investors want to compare houses rather than knowing the exact value, ranking houses from the perspective of community safety can help investors differentiate low value houses from the other houses. Therefore, we propose a house safety-aware (HSA) ranker which combines all extracted community safety features to rank houses according to house values. In addition, we integrate distinctive house profile such as neighborhood income, nearby school rating and house build year to differentiate the comparability of house pairs in pair-wise ranking objective for enhancing ranking accuracy of community safety features. Last, we optimize the ranking model by jointly preserving the house ranking consistency and maximizing the value prediction accuracy.

In summary, in this paper we strategically leverage rich spatio-temporal crime data for effective house ranking. We highlight our key contributions as follows:

- We present an advanced crime analysis (e.g., crime evidence mining) to comprehensively infer the community safety by exploiting historical crime data.
- We develop a safety-aware ranking model by incorporating the comparability of house pairs into the optimization of pair-wise ranking objective, in a way that we better model the impact of community safety without the effects of other aspects, and thereby enhance the ranking accuracy.
- We validate our method with real-world dataset. Experiments shows that our proposed ranking method not only provides better explanations in safety impacts on house values, but also demonstrates a substantial improvement in ranking accuracy compared with baseline ranking methods.

2 Problem Definition

Usually, when people use house values to indicate the quality and benefits of a houses, they actually mean unit values (e.g., values per square footage) since total values are also affected by floor areas of houses. Moreover, from the perspective of investors, ranking houses in terms of their quality and benefits, instead of predicting absolute appraised values, is more needed for making investment decisions. Therefore, to provide investors with a tool to compare houses, we rank houses based on unit values by taking community safety degrees into account. In the rest of this paper, "house value" will be used to represent the unit value of a house.

(a)

OFFENSE_ID:	2010200143220200
OFFENSE_TYPE_ID:	burglary-residence-by-force
OCCURRENCE_DATE:	2010-04-28 08:29:59
INCIDENT_ADDRESS:	2909 N IVANHOE ST, DENVER
LATITUDE:	39.7586462
LONGITUDE:	-104.9214312

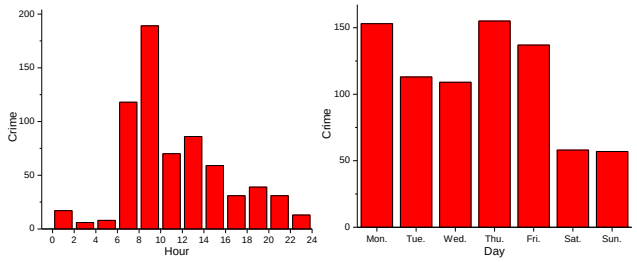
(b)

Figure 1: Example of (a) crime sequence of a house, (b) a crime record.

We are given a set of I houses $H = \{h_1, h_2, ..., h_I\}$ where each house has a location (e.g., latitude and longitude) and corresponding house values $Y = \{y_1, y_2, ..., y_I\}$ where y_i denotes the value per square footage in dollar of house h_i. We are also given profiles of houses in H where each house has several house characteristics such as neighborhood income, nearby school rating and house build year. Last, we are given the complete historical house burglary records of the area, each burglary crime is denoted by $< loc, add, t >$, which has a location loc, an address add and an occurrence timestamp t. The task is training a model to rank a testing set of houses in an ascending order according to their house value by exploiting historical burglary data. We propose to accomplish the task with a two-step framework: 1) extracting and aggregating community crime evidences for learning the community safety of different houses from historical crime data, 2) ranking houses by incorporating the impacts of community safety.

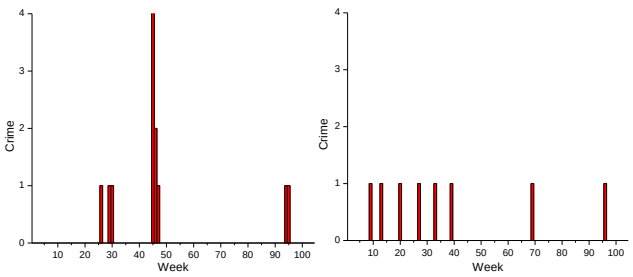
3 Community Crime Evidence Extraction

In this section, we study how to extract and aggregate community crime evidence for learning community safety features of houses. Figure 1a shows an example how we collect five historical crimes around house h_i to form its crime sequence $C_i = \{c_1, c_2, ..., c_5\}$. To get the crime sequence of a house, we collect all the historical crimes which occurred within d meters (e.g., 200 meters) of the house and order the crimes by occurrence time from oldest to newest. We mine crime evidences for house h_i based on its specific crime sequence C_i. Figure 1b shows a sample of crime records. In the following, we will extract crime evidences in two categories: 1) crime severity, and 2) crime temporal correlation. Finally we aggregate crime evidences by each evidence type to generate house-level community safety features.



(a) (b)

Figure 2: Hourly and daily number of burglaries in a residential area during 2009-2014.



(a) Low value house (b) Normal value house

Figure 3: Weekly number of burglaries near two houses during 2009-2010.

3.1 Category Based on Crime Severity. Crime severity indicates the damage level made by a crime. For burglary, it is usually determined by the property losses in a crime. However, we do not have the explicit description about the actual losses of valuables from available crime information. Therefore, we attempt to infer the severity of a burglary implicitly from other crime information. To this end, we propose to mine two evidences for a burglary to assess its severity.

3.1.1 Occurrence Address Evidence. Knowing the detailed occurrence address of burglaries, we can retrieve the appraisal of the victimized houses (e.g., 265900 dollars). Intuitively, the loss led by a burglary is proportional to the appraisal of the victimized house. Higher appraisal usually means the victimized house has more rooms or the house is more luxury. Either possibility gives burglars a higher chance to collect more valuables. Based on this intuition, we can infer the possible loss in a burglary crime by knowing the appraisal of the victimized house via the burglary address. Therefore, we propose the first type of evidence:

$$E_1(c) = Appraisal(add_c), \tag{3.1}$$

where c is a burglary crime. add_c is the occurrence address of burglaries. *Appraisal* denotes the appraisal of the victimized house located at burglary address.

3.1.2 Occurrence Time Evidence. The occurrence time of crimes is another information for inferring the severity of a burglary. House burglary has a unique character that burglars always have to make sure there is no occupant at home to commit crimes. Since the most common cause of people leaving home is for work or school, the confidence for committing burglaries should be strongly correlated to people's working schedule. In Figure 2a which shows the number of burglary by every two hours in workdays, most of the burglaries concentrate in the time range from 6:00 to 16:00 which is the time people are usually out for work. In Figure 2b which shows the number of burglary by every day in weeks, burglaries happened much more frequently in workday (Monday to Friday) compare to in weekend (Saturday, Sunday). Based on this observation, during different time slots, the general confidence of burglars for committing a crime should be different.

We propose a *time entropy* to model the burglary confidence of different time slots by analyzing how many houses were victimized during the time slots. Specifically, we define a time slot in two dimensions: 1) two hours of a day and 2) workday or weekend. For example, a time slot can be 8:00 to 10:00 in every workday. Then let k denotes a time slot, $C_{k,i}$ is the set of burglaries occurred in kth time slot at ith house, and C_k is the set of all burglaries occurred in kth time slot. The probability that a randomly picked burglary occurred in kth time slot belongs to the ith house is $P_{k,i} = |C_{k,i}|/|C_k|$. We define the Shannon entropy of time slot k as follow:

$$Entropy(k) = -\sum_{i:P_{k,i}\neq 0} P_{k,i} \cdot logP_{k,i}. \tag{3.2}$$

A higher time entropy implies a confident time slot during which houses' occupants are more possible to be not at home. On the other hand, a lower time entropy indicate a worse time slot during which occupants are less possible to be not at home. Therefore, we infer the severity of a burglary by the entropy of the time slot it occurred in. If a burglary occurred during a high entropy time slot, we consider it may result more losses since burglars can take longer time for searching valuables and have fewer chance to be discovered. We propose the second type of evidence:

$$E_2(c) = Entropy(ts_c), \tag{3.3}$$

where c is a burglary crime and ts_c is the time slot during which the burglary occurred.

3.2 Category Based on Temporal Correlation. Given a crime sequence of a house, the temporal correlation evidence aims to consider the temporal proximity among burglaries. By analyzing temporal correlation of crimes, we can infer the local community

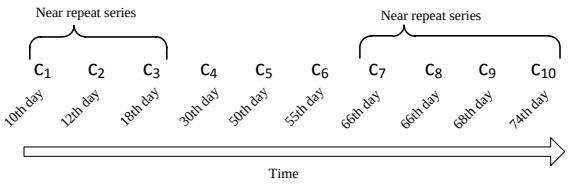


Figure 4: An example of near repeat series mining.

safety. Figure 3 shows the weekly statistics of burglary happened within 400 meters of two houses respectively during 2013-2014. Figure 3a shows the burglary near a low value house, we can see that many burglaries tend to cluster together temporally. On the contrary, Figure 3b shows burglary near a normal value house, we can see that these crimes behave more independently. This temporal correlation can be explained by near repeat phenomenon in criminology research [4, 5]. Usually, the burglaries which repeatedly occurred in the same area within short interval means that they are very likely to be committed by the same burglars. Preference by returning burglars implies worse community safety issues and brings higher crime risks to the area [6]. Therefore, we can expect that a stronger temporal correlation of crimes leads to worse community safety of houses.

For capturing the temporal correlation, we propose to mine *near repeat series* for houses. Basically, a near repeat series is mined from the crime sequence of a house and it is a set of crimes in which every crime happened very shortly after the previous one except the first. A house can have none to multiple near repeat series. Formally, we define a near repeat series as follows:

DEFINITION 1. *A **near repeat series** s of a house h_i consists of N adjacent crimes $s = \{c_1, c_2, ..., c_N\}$ in which every crime c_n belong to the h_i's crime sequence: $c_n \in C_i$ and there is no other near repeat series s' makes that $s \subseteq s'$. Every near repeat series meets two conditions: 1) the number of crimes in s should be not less than a minimum size threshold θ: $|s| \geqslant \theta$, and 2) the interval between any two adjacent crimes should be not longer than a maximum time threshold τ: $\forall n \in [2, N]$, $t_{c_n} - t_{c_{n-1}} \leqslant \tau$ where t_c represents the timestamp of a crime c.*

Figure 4 shows an example of mining near repeat series. Given a crime sequence $C_i = \{c_1, ..., c_{10}\}$ with their occurrence date of a house h_i, c_1 is the oldest crime while c_{10} is the latest one. Suppose we define the maximum time threshold to be 7 days while the minimum size threshold to be 3 crimes. We can find the first near repeat series to be $s_1 = \{c_1, c_2, c_3\}$ since $t_{c_2} - t_{c_1} = 2$ and $t_{c_3} - t_{c_2} = 6$. Because $t_{c_4} - t_{c_3} = 12$, repeat series s_1 stops at c_3. Although the interval between c_5 and c_6 is less than 7 days, they can not find c_4 or c_7 to reach the minimum size threshold. Last, from c_8 to c_{10}, each of them has a less than 7 days interval from its previous one, therefore we find the second near repeat series $s_2 = \{c_7, c_8, c_9, c_{10}\}$ by qualified intervals of adjacent crimes and qualified series size.

We propose three evidences based on near repeat series to assess local community safety.

3.2.1 Series Size Evidence. Series size evidence which measures the number of crimes in a near repeat series, is a significant indicator for community safety situation. For a house with less safety issues, the nearby crimes should act independently and they are unlikely to form large size near repeat series. If a near repeat series has large size, there is high probability that the area of house has serious safety issues. Therefore, we propose the third evidence:

$$E_3(s) = |s|, \tag{3.4}$$

where s is a near repeat series consists of crimes.

3.2.2 Series Length Evidence. Series length evidence measures the length of period a series lasts in order to learn the community safety situation. As the near repeat series lasts longer, it is more difficult for burglars to commit repeat crime because of increasing police attention. If near repeat series lasts long, it means that the area is promising for burglars, therefore it indicates a worse community safety situation for local houses. We propose the forth evidence:

$$E_4(s) = t_{c_N} - t_{c_1} + 1, \tag{3.5}$$

where t_{c_1} and t_{c_N} represent the date of the first crime c_1 and the last crime c_N of the near repeat series s.

3.2.3 Series Intensity Evidence. Series intensity uses the shortest interval occurred in a near repeat series to learn the community safety situation. An intensive near repeat series means that at least two crimes of it have very short interval thus shows a dangerous sign of the area. For example, given two size-3 near repeat series with similar length, the first series has the interval $\{0-day, 7-day\}$ while the second series has the interval $\{3-day, 4-day\}$. The first series has a higher intensity since the first two crimes of it occurred in a the same day with 0-day interval. The second series has a lower intensity since the shortest interval in the second series is 3-day. Therefore, we propose the fifth evidence:

$$E_5(s) = \max_{2 \leq n \leq N} \{\tau - (t_{c_n} - t_{c_{n-1}}) + 1\}, \tag{3.6}$$

where $t_{c_n} - t_{c_{n-1}}$ represents the interval days between crime c_n and its previous adjacent crime c_{n-1} of series s, τ represents the maximum time threshold.

3.3 Evidence Aggregation. We aggregate evidences by types for generating house-level community safety features. As we introduced, for a house h_i, we have a crime sequence $C_i = \{c_1, c_2, ..., c_{K_i}\}$ and a set of near repeat series $S_i = \{s_1, s_2, ..., s_{J_i}\}$. The evidence belongs to crime severity category $(E_1(c), E_2(c))$ is extracted by crime cases while the evidence belongs to temporal correlation category $(E_3(s), E_4(s), E_5(s))$ is extracted by near repeat series. Therefore, for a house h_i, we will have a community safety feature vector $X_i = \{x_{i1}, x_{i2}, ..., x_{i5}\}$ as following:

$$x_{im} = \begin{cases} \sum_{c \in C_i} E_m(c) & \text{if } m \in (1,2) \\ \sum_{s \in S_i} E_m(s) & \text{if } m \in (3,4,5) \end{cases} \tag{3.7}$$

4 A Safety-Aware House Ranking Model

In this section, we propose a House Safety-Aware (HSA) model which ranks houses by incorporating the degrees of community safety.

4.1 Model Specification. Let us define the input of the model to be X_i, y_i and P_i for a house h_i, where X_i denotes M-size vector of community safety features, y_i denotes the ground truth of house value and P_i denotes the L-size vector of house profile characteristics. We want to train a function $f_i = W^T X_i$ which formulates the house value by having $y_i = f_i + \epsilon$, where W denotes the vector of weights for safety features and ϵ denotes the error term which subjects to Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Thus, we have $y_i \sim \mathcal{N}(f_i, \sigma^2)$.

4.2 Objective Function. We propose to jointly model the accuracy of house value prediction and the consistency of house ranking prediction in an objective function. Let the parameter for estimation to be W, model hyperparameter to be $\Phi = \{\sigma^2, b^2\}$, observed ground truth to be $\mathcal{O} = \{Y, R\}$ where Y and R denote the value and rank of houses. Then we have the posterior probability:

$$Pr(W|\mathcal{O}, \Phi) = Pr(\mathcal{O}|W, \Phi) Pr(W|\Phi). \tag{4.8}$$

First, let us address the probability of observed data. We model $Pr(\mathcal{O}|W, \Phi)$ as a joint probability of house value prediction $Pr(Y|W, \Phi)$ and house ranking prediction $Pr(R|W, \Phi)$. **Modeling prediction accuracy**, we use $Pr(Y|W, \Phi) = \prod_{i=1}^{I} \mathcal{N}(y_i|f_i, \sigma^2)$ to ensure value prediction accuracy of houses. **Modeling ranking consistency**, we adopt a pair-wise probability to ensure ranking correctness of all house pairs. Suppose the I houses has already been ranked by house value in ascending order. Given two index i, j which has $i < j$, we should always have $y_i < y_j$ and $h_i \to h_j$ which means the rank of house h_i is higher than the rank of h_j in ground truth. Therefore, we use $Pr(R|W, \Phi) = \prod_{i=1}^{I-1} \prod_{j=i+1}^{I} Pr(h_i \to h_j|W, \Phi)$ to represent the probability that h_i is correctly ranked higher than h_j by model for all house pairs. We adopt Sigmoid function to represent the probability of pair-wise ranking consistency: $Pr(h_i \to h_j) = \frac{1}{1+exp(-(f_j - f_i))}$.

Table 1: Characteristics in house profile.

Characteristics	Description
Neighborhood Characteristics	
Household Income	Average annual household income
High Educated Ratio	Ratio of residents with least bachelors degree to residents who are at least 25 years old
Population Growth	Percentage growth of population from 2000 to 2010
Surrounding Characteristics	
School Rating	Average rating of the nearest public high, middle and primary schools (A school has rating 1 to 5)
Point-of-Interest (POI) Diversity	Number of diverse categorical tags extracted from all the POIs which locate within d meters of the house
Check-in Density	Average number of social network check-in within d meters in every workday after hours 6 PM to 6 AM
Build Characteristics	
Land Area (in sqft), Bedroom Number and Build Year	

Integrating house profile into ranking consistency. House is a kind of distinctive property which has various characteristics, such as the house profile shown in Table 1. If two houses have too large differences in house profile, their value difference does not help the model to learn the impacts of community safety. For example, if two houses have too different build year (e.g., 1950 vs. 2010), then the impact of community safety on house value may be overridden by the impact of build year. Therefore, including the rank observation of dissimilar house pair in optimization objective will jeopardize the ranking prediction capacity of community safety features. Based on this motivation, we propose to weight house pairs differently by the comparability of house pairs by exploiting profile data. Specifically, we use characteristics in Table 1 to compute the similarity between every house pair by $D_{ij} = -\sqrt{\sum_{l=1}^{L} |p_l^i - p_l^j|^2}$, where i and j denote the house pair, p_l denotes the lth characteristic in profile vectors. D_{ij} is normalized as a real number between 0 and 1. Then, we incorporate D_{ij} as the comparability into the pair-wise probability of ranking consistency. We have the new ranking

consistency probability: $Pr(h_i \to h_j|W,\Phi)^{D_{ij}}$, where D_{ij} is assigned as an exponent to corresponding house pair's ranking consistency probability. **Benefit**: when the similarity D_{ij} between h_i and h_j is high such as the extreme 1, the impact of their ranking probability $Pr(h_i \to h_j)$ will be fully preserved in objective function. On the other hand, when D_{ij} is low such as the extreme 0, the impact of $Pr(h_i \to h_j)$ will be fully blocked outside of objective function. In this way, we differentiate the importance of different house pairs for objective function, thus we can better use community safety for house ranking.

We present the final probability of observed data:

$$Pr(\mathcal{O}|W,\Phi) = Pr(Y|W,\Phi)Pr(R|W,\Phi) = \prod_{i=1}^{I}\mathcal{N}(y_i|f_i,\sigma^2)\cdot\prod_{i=1}^{I-1}\prod_{j=i+1}^{I}\left(\frac{1}{1+exp(-(f_j-f_i))}\right)^{D_{ij}}. \tag{4.9}$$

Next, let us address the prior distribution of W which is the last part in posterior distribution. We model $Pr(W|\Phi)$ as Gaussian distribution with 0 mean, where b^2 represents the variance of parameter w_m. We have the prior distribution formally:

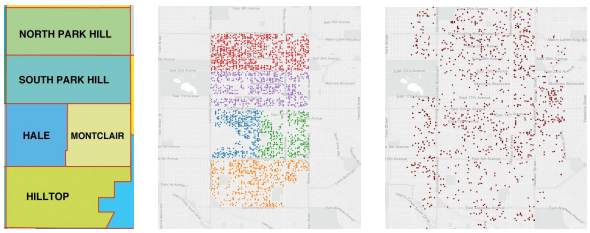
$$Pr(W|\Phi) = \prod_{m=1}^{M}\mathcal{N}(w_m|0,b^2). \tag{4.10}$$

4.3 Parameter Estimation. Given the posterior distribution in Equation 4.8, we want to find the optimal W to maximize the probability. The log posterior distribution is:

$$\mathcal{L}(W|Y,R,\sigma^2,b^2) = -\frac{1}{2\sigma^2}\sum_{i=1}^{I}(y_i-f_i)^2 + \sum_{i=1}^{I-1}\sum_{j=i+1}^{I}D_{ij}ln\frac{1}{1+exp(-(f_j-f_i))} - \frac{1}{2b^2}\sum_{m=1}^{M}w_m^2. \tag{4.11}$$

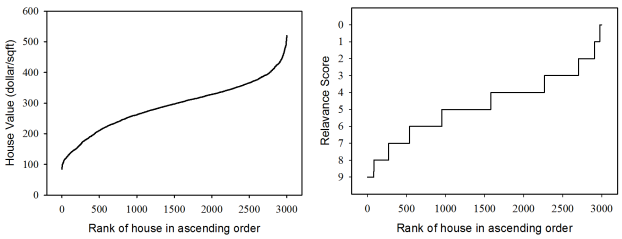
To maximize the log posterior, we utilize gradient ascent method to update parameter w_m by $w_m^{(t+1)} = w_m^{(t)} + \alpha \times \frac{\partial\mathcal{L}}{\partial w_m}$, where α is the learning rate and $\frac{\partial\mathcal{L}}{\partial w_m}$ is the derivatives according to Equation 4.11:

$$\frac{\partial\mathcal{L}}{\partial w_m} = \frac{1}{\sigma^2}\sum_{i=1}^{I}(y_i-f_i)x_{im} - \frac{1}{b^2}w_m + \sum_{i=1}^{I-1}\sum_{j=i+1}^{I}D_{ij}\frac{exp(-(f_j-f_i))}{1+exp(-(f_j-f_i))}(x_{jm}-x_{im}). \tag{4.12}$$



(a) (b) (c)

Figure 5: (a) Official neighborhoods, (b) Houses, (c) Burglary crimes.



(a) (b)

Figure 6: (a) Value of ranked houses, (b) Relevance score of ranked houses.

5 Experimental Results

In this section, we present a comprehensive experiment to evaluate the proposed method on real-world dataset.

5.1 Experimental Data. All the data of houses and crimes are collected from Denver Open Data Catalog [7]. For houses dataset, since house comparisons usually happen in the same type with not very far distance, we restrict the house type to only single family detached home which is the major type in U.S., and collect 3000 houses evenly spread in a major residential region of north Denver which consists of five adjacent official neighborhoods as shown in Figure 5a, 5b. All the house values are appraised in 2015. Figure 6a shows the values of 3000 houses in ascending order. For ranking purpose, we evenly split of range of house value into 10 levels and give the level of lower values the higher relevance score. Therefore, as shown in Figure 6b, the real house value y_i in experiment is the relevance score from 0 to 9 which shows how low a house value is. For crime dataset, we collect residential forcible burglaries happened in Denver during 2009 to 2014. Totally, we find 1131 forcible burglary crimes which are related to our collected houses as shown in Figure 5c. For house profiles, we collect neighborhood data from demographic of 2010 & 2000 US Census, POIs and check-in (9/2010 to 1/2011) data from Foursquare, public school data from official public school rating [8].

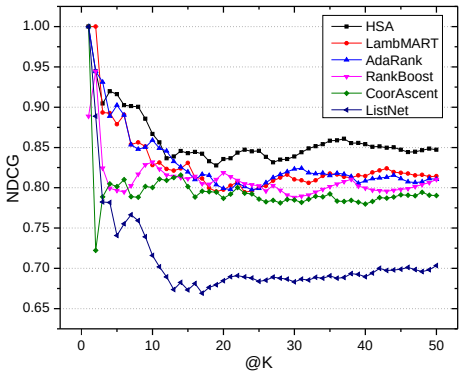


Figure 7: NDCG performance comparison.

5.2 Baseline Algorithms. To validate the effectiveness of our proposed method, we compare it with several traditional ranking algorithm: 1) LambdaMART [9], which employs the Lambda function of LambdaRank as gradients in the learning of Multiple Additive Regression Trees (MART). 2) AdaRank [10], which plugs the evaluation measures into the framework for boosting optimization. 3) RankBoost [11], which adopts AdaBoost [12] for the pair-wise classification. 4) Coordinate Ascent [13], which applies coordinate ascent technique in unconstrained optimization. 5) ListNet [14], which defines the loss function by the probability distribution on permutations.

We adopt RankLib [15] for baseline algorithm implementation. For LambdaMART, we set number of trees = 300, number of leaves = 10. For AdaRank, we set number of round = 500, tolerance = 0.002. For RankBoost, we set number of rounds = 300. For Coordinate Ascent, we set number of random restarts = 5, tolerance = 0.001. We randomly split 3000 houses to be 4:1 where 2400 for training set and 600 for testing set.

5.3 Evaluation Metrics

5.3.1 Normalized Discounted Cumulative Gain (NDCG). NDCG is obtained from Discounted Cumulative Gain (DCG) which measures the ranking quality by calculating the cumulative gain from the top of the result list to the particular rank position K. $DCG@K = rel_1 + \sum_{i=2}^{K} \frac{rel_i}{\log_2(i)}$ where rel_i represents the relevance score of the result at position i. Then we compute Ideal DCG (IDCN) which represents the maximum possible DCG till position K by sorting the result list by relevance. Last we obtain the normalized DCG: $NDCG@K = \frac{DCG@K}{IDCG@K}$.

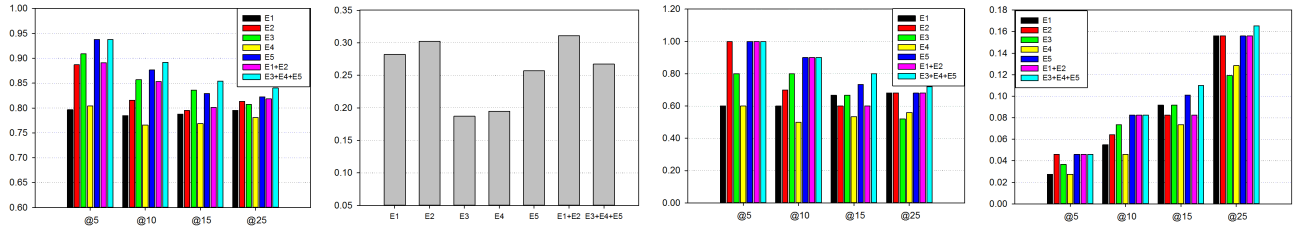
5.3.2 Kendall's tau coefficient (Tau). Kendall's Tau coefficient measures the ranking quality by rank correlation: the similarity of the orderings of houses between predicted ranking list and ground truth ranking list. Let $(\hat{r_i}, r_i)$ be the rank of house h_i in predicted ranking list and ground truth ranking list. Any pair of houses $(\hat{r_i}, r_i)$ and $(\hat{r_j}, r_j)$ are concordant if both $\hat{r_i} > \hat{r_j}$ and $r_i > r_j$ or if both $\hat{r_i} < \hat{r_j}$ and $r_i < r_j$. They are discordant, if $\hat{r_i} > \hat{r_j}$ and $r_i < r_j$ or if $\hat{r_i} < \hat{r_j}$ and $r_i > r_j$. We obtain the Kendall's tau coefficient: $Tau = \frac{\#concordant - \#discordant}{\#concordant + \#discordant}$.

Table 2: Performance of each algorithm.

Metrics	Lamb MART	Ada Rank	Rank Boost	Coor Ascnt	List Net	HSA
NDCG@5	0.8788	0.9023	0.7967	0.8015	0.7408	**0.9160**
NDCG@7	0.8537	0.8532	0.8026	0.7890	0.7664	**0.9013**
NDCG@10	0.8280	0.8590	0.8320	0.8003	0.7159	**0.8666**
NDCG@15	0.8309	0.8197	0.8113	0.8015	0.6733	**0.8429**
NDCG@25	0.8007	0.7993	0.8025	0.7859	0.6839	**0.8457**
Tau	0.2137	0.2733	0.1611	0.2326	0.2471	**0.3146**

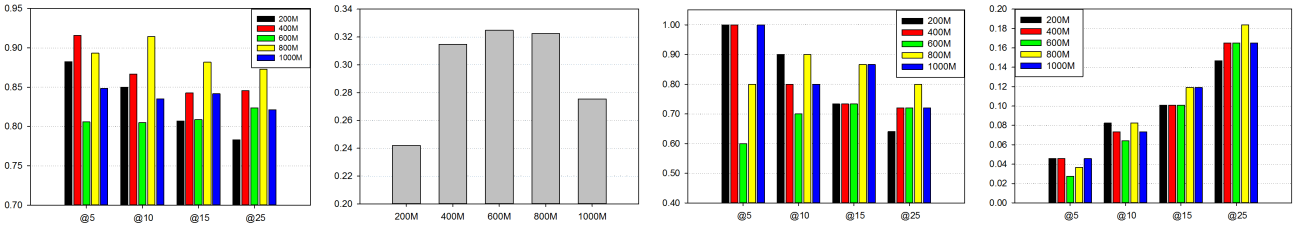
5.3.3 Precision and Recall. Precision measure the fraction of retrieved houses which are relevant. Recall measure the fraction of relevant houses that are retrieved. In our case, we consider the low value house with 7-9 relevance score as relevant, and consider other house with 0-6 relevance score as irrelevant. In retrieved top K ranking houses, we calculate the precision by $Precision@K = \frac{|h_K \cap h_{\geq 7}|}{|h_K|}$ and the recall by $Recall@K = \frac{|h_K \cap h_{\geq 7}|}{|h_{\geq 7}|}$, where h_K and $h_{\geq 7}$ denote the set of retrieved houses and the set of relevant houses.

5.4 Performance Evaluation on House Safety-Aware (HSA) Ranking. We compare the performance of proposed HSA and baseline algorithms with the metrics of NDCG and Kendall's tau coefficient. Figure 7 shows the NDCG@K of each algorithm from K=1 to K=50. Overall, we can see that HSA outperforms all baselines. The improvement start to be obvious since K = 15. Second, we notice that LambdaRank and AdaRank which performance closely reach the second best overall performance. Compared to the rest of baseline algorithms, these two achieve obvious improvement of NDCG performance when K is smaller than 10. Moreover, RankBoost and Coordinate Ascent perform similarly. They both do not perform well when K is smaller than 15. Then their performance returns when K grows large. Last, ListNet makes the lowest performance, which is far behind other baseline algorithms. Table 2 shows the numerical result of NDCG at small K and Kendall's Tau coefficient. On NDCG @5, @7, @10, @15 and @25, HSA shows the consistent advantage. On Kendall's Tau coefficient, HSA achieves the best performance of 0.3146, the second best is AdaRank which gets 0.2733. Overall, HSA shows obvious advance in both NDCG and Kendall's Tau coefficient. In sum-



(a) NDCG@K (b) Kendall's Tau (c) Precision@K (d) Recall@K

Figure 8: Performance of different crime evidences.



(a) NDCG@K (b) Kendall's Tau (c) Precision@K (d) Recall@K

Figure 9: Performance of different crime collection radius.

mary, the results shows that HSA model which incorporates house pairs comparability into ranking objective optimization can effectively increase the performance of house ranking with community safety.

5.5 Performance of Different Crime Evidences. We compare the performance of every single evidence type as well as two evidence combinations in NDCG, Tau, Precision and Recall. In Figure 8, E1 (Occurrence address), E2 (Occurrence time), E3 (Series size), E4 (Series length) and E5 (Series intensity) represent the five evidence types we extract. The combination of E1+E2 denotes the crime severity based category while the combination of E3+E4+E5 denotes the temporal correlation based category. First let us see the performance of single evidences. From the perspective of top K ranking measured by NDCG, E3 and E5 have the best performance and E2 preforms well too. In Precision and Recall, E5 performs the best. E2 and E3 perform well too. For the overall ranking consistency by Kendall's Tau, E1 and E2 show the best performance, E5 also does well. Then let us see the performance of evidence combinations based on two evidence categories. The crime severity based combination E1+E2 outperforms the single evidence E1 or E2 in ranking quality of both NDCG and Tau coefficient. The temporal correlation based combination E3+E4+E5 outperforms the single evidence E3, E4 or E5 in both of NDCG and Tau coefficient as well as Precision and Recall. Comparing E1+E2 with E3+E4+E5, we find that E3+E4+E5 consistently provides better top K ranking quality. Generally, the temporal correlation category evidences perform better than crime severity category evidences in differentiating low value houses via community safety conditions.

5.6 Performance of Different Crime Collection Radius. Since we only consider the crimes which occurred within a certain distance of a house as impactful crimes, we want to explore what the proper distance is for learning a house's community safety. For example, if the radius is 400M, we will learn community safety by the crimes within 400 meter of a house. Figure 9 shows the performance on 5 different radius: 200M, 400M, 600M, 800M and 1000M in metrics of NDCG, Tau, Precision and Recall. From Kendall's Tau coefficient, we can see that radius in 400M, 600M and 800M outperform other distance in the quality of overall ranking. From NDCG, we can observe that 800M and 400M perform the best but the performance of 600M falls. One possible reason is that some houses can not cover more burglaries when the radius increases because it may cover non-residential blocks (e.g., square). When the radius continuously increases to 800M, the circle area overcomes the effects of non-residential blocks and reaches sufficient crimes for safety assessments. Therefore, we can have two insights from the results: 1) The distance from 400M to 600M which is a walking distance provides the most impactful crime for a house. 200M is too short to collect sufficient crimes while 1000M overly collects crime which do not generate real impacts. 2)

Proper radius also depends on specific geographical situation which may disable some radius.

6 Related Work

This work can be grouped into three research categories. The first category is the study of the appraisal and ranking of real property. The works in [2, 3] show that after a registered sex offender moves into a neighborhood, nearby housing prices would be declined in response. The work in [1] reports that property crimes have a significant negative impact on property price in London area. The work in [16] concludes that decreases of perceived security level in victimization survey is associated with decreases of the real property valuation of a district. The works in [17, 18] model the effects of geographical dependencies and function diversities for ranking estate investment values. The work in [19] explores the effects of people's moving behaviors and online reviews on real estate ranking.

The second category belongs to criminology research. The work in [20] finds that there is a dramatically enhanced risk of repeat burglaries for a house immediately after an initial burglary happened. The work in [5] shows that repeat victimization is more likely in high-crime than in low-crime areas, and that the recommitting by same offenders plays a key role in repeat victimization. The works in [4, 21] show that the elevated crime risk after the initial crime not only comes to the victims itself but also spread to nearby areas.

The last category is the research in Learning-to-rank (LTR) algorithm. There are three categories of LTR algorithm, point-wise ranking directly predicts the relevance degree of a document, such as [22] which adopts regression to solve the problem of ranking. Pairwise ranking output the relative order for a pair of two documents. The work in [23] applies the SVM technique to classify orders for document pairs. Last, list-wise ranking model the entire ranking of a whole set of documents. The work in [14] defines the loss function by using the probability distribution on permutations.

7 Conclusion

In this paper, we presented a systematic study on ranking house by leveraging spatio-temporal crime data. Specifically, we first extracted community crime evidences in two categories: crime severity and crime temporal correlation. Moreover we proposed effective approach to ranking houses based on value by incorporating the house specific features of community safety. Also, we integrated the impacts of popular house profile in optimization to enhance the proposed ranking model. Finally, extensive experimental results on real-world crime and house data validated the performance of the proposed method.

Acknowledgement. This research was partially supported by Futurewei Technologies, Inc. Also, it was supported in part by Natural Science Foundation of China (71329201).

References

[1] S. Gibbons, "The costs of urban property crime," *The Economic Journal*, vol. 114, no. 499, pp. F441–F463, 2004.
[2] J. C. Pope, "Fear of crime and housing prices: Household reactions to sex offender registries," *Journal of Urban Economics*, vol. 64, no. 3, pp. 601–614, 2008.
[3] L. Linden and J. E. Rockoff, "Estimates of the impact of crime risk on property values from megan's laws," *The American Economic Review*, pp. 1103–1127, 2008.
[4] J. H. Ratcliffe and G. F. Rengert, "Near-repeat patterns in philadelphia shootings," *Security Journal*, vol. 21, no. 1, pp. 58–76, 2008.
[5] E. R. Kleemans *et al.*, "Repeat burglary victimisation: results of empirical research in the netherlands," *Crime prevention studies*, vol. 12, pp. 53–68, 2001.
[6] I. Hearnden and C. Magill, *Decision-making by house burglars: offenders' perspectives*. Home Office. Research, Development and Statistics Directorate, 2004.
[7] Denver open data catalog. [Online]. Available: http://data.denvergov.org/
[8] School performance framework. [Online]. Available: http://spf.dpsk12.org/
[9] C. J. Burges, "From ranknet to lambdarank to lambdamart: An overview," *Learning*, vol. 11, pp. 23–581, 2010.
[10] J. Xu and H. Li, "Adarank: a boosting algorithm for information retrieval," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2007, pp. 391–398.
[11] Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer, "An efficient boosting algorithm for combining preferences," *The Journal of machine learning research*, vol. 4, pp. 933–969, 2003.
[12] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.
[13] D. Metzler and W. B. Croft, "Linear feature-based models for information retrieval," *Information Retrieval*, vol. 10, no. 3, pp. 257–274, 2007.
[14] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li, "Learning to rank: from pairwise approach to listwise approach," in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, pp. 129–136.
[15] The lemur project/ranklib. [Online]. Available: http://sourceforge.net/p/lemur/wiki/RankLib/
[16] P. Buonanno, D. Montolio, and J. M. Raya-Vílchez, "Housing prices and crime perception," *Empirical Economics*, vol. 45, no. 1, pp. 305–321, 2013.
[17] Y. Fu, H. Xiong, Y. Ge, Z. Yao, Y. Zheng, and Z.-H. Zhou, "Exploiting geographic dependencies for real estate appraisal: A mutual perspective of ranking and clustering," in *KDD 2014*. pp. 1047–1056.
[18] Y. Fu, G. Liu, S. Papadimitriou, H. Xiong, Y. Ge, H. Zhu, and C. Zhu, "Real estate ranking via mixed land-use latent models," in *KDD 2015*. pp. 299–308.
[19] Y. Fu, Y. Ge, Y. Zheng, Z. Yao, Y. Liu, H. Xiong, and N. J. Yuan, "Sparse real estate ranking with online user reviews and offline moving behaviors," in *ICDM 2014*. pp. 120–129.
[20] N. Polvi, T. Looman, C. Humphries, and K. Pease, "The time course of repeat burglary victimization," *British Journal of Criminology*, vol. 31, no. 4, pp. 411–414, 1991.
[21] M. Townsley, R. Homel, and J. Chaseling, "Infectious burglaries. a test of the near repeat hypothesis," *British Journal of Criminology*, vol. 43, no. 3, pp. 615–633, 2003.
[22] W. S. Cooper, F. C. Gey, and D. P. Dabney, "Probabilistic retrieval based on staged logistic regression," in *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 1992, pp. 198–210.
[23] R. Herbrich, T. Graepel, and K. Obermayer, "Large margin rank boundaries for ordinal regression," *Advances in neural information processing systems*, pp. 115–132, 1999.