Int. Journal for Housing Science, Vol.40, No.2 pp. 123-132, 2016
Published in the United States

CRIME RISK ASSESSMENT IN RESIDENTIAL BUILDINGS: OPENINGS

G. Agarbati, M. D'Orazio
Department of Civil and Building Engineering, and Architecture (DICEA)
Polytechnic University of Marche, Ancona
Italy

U. Yazgan
Earthquake Engineering and Disaster Management Institute
Istanbul Technical University, Istanbul
Turkey

ABSTRACT

In recent years the issue of urban security has assumed an increasing importance, showing a transformation whose features help to define specificity of every city. The growth of offenses, the increased crime rates, the growth of the global economic difficulties, lead up to an increase in demand for social security often addressed to the local government aspects having as its object the daily life. Even though security and police resources are the first to be cut back in the shrinking economy, the present turmoil requires even greater concentration of security and law enforcement assets to keep our property, people, and information safe. Given the situation, the need to implement new procedures based on innovative theories requires a different approach to the problem of urban security. In fact, based on a theory called "CPTED" (Crime Prevention Through Environmental Design), an evaluation method capable to estimate crime risk of the urban district in its entirety has been developing. The risk of crime for a building is related to its vulnerability (i.e. presence of security leaks) and the hazard (i.e. crime rate) in the neighborhood. The method proposed in this study, aims at providing a rational measure for evaluating burglary vulnerability of a building. In this article, an analysis of burglary data, collected by observing selected building parameters, is presented. Openings are the basic element of the evaluation. Their characteristics, but especially their relationship with the surroundings, influence the decision of a burglar in performing a criminal act. Through a preliminary qualitative evaluation based on existing literature, a list of basic indicators has been selected.

0146-6518/02/123-132, 2016
Copyright©2016 IAHS

版权所有(c)国家科技图书文献中心 D25062000629ZX00

Statistical approach reveals how some of these indicators have a greater influence than others in the final outcome. An algorithm capable to grade openings first, and buildings then, referring to burglary vulnerability, has been carried out.

Key words: Urban security, Crime risk analysis, CPTED, Sustainability

Introduction

Crime and insecurity feeling can affect city lifestyle, as well as the management and the attractiveness of certain urban areas. When people feel threatened, they modify their behavior and the manner how they use the city. A lot of people do not go out in the evening, they do not use public transport during certain times, they do not hang out at specific urban areas considered dangerous, and they lock into their dwellings or heavily guarded district (as gated communities). Consequently, the quality of life is affected by loss of freedom. Jane Jacobs [1] was the first "key figure" concerned with environmental prevention in order to deter criminal activity. But C.R. Jeffery [2] was the founding father of a theory called "Crime Prevention Through Environmental Design" (CPTED), which main goal is to reduce the opportunities of crime that maybe inherent in the design of structures or in the design of neighborhoods. Then this thesis has been carried on by Newman [3], Coleman, Saville, Crowe [4], and so on, by changing his name over the Countries (Secured by Design in UK, Designing Out Crime in Australia, New Urbanism in America, etc.). Nowadays it is common to call it CPTED, since the basic principles are the same. These principles are often summarized using the following key concepts: Access Control, Territoriality, Surveillance, Target Hardening, Image and Activity Support. Access Control, is about actively keeping certain people out of buildings/enclosures, and the structures, procedures and technologies to achieve this, whilst admitting those people with a right to·be there; Territoriality covers the human motivation to control space, who enters it and what people do within it; Surveillance concerns how people, sometimes aided by design and technology, can act as crime preventers, whether police, employees, owners or general public, by seeing or hearing suspicious behavior, and take some appropriate action; Target Hardening is about making physical structures like walls, windows and doors resistant to attack and penetration by criminals; Image covers the appearance of a building, place or neighborhood, not just aesthetics but relating also to social reputation and stigma of the place and its inhabitants; Maintenance contributes to appearance, obviously, but also to issues like effectiveness of security systems; Activity support is the beneficial effect of having significant numbers of people in, or passing through, a particular place, who are doing routine, honest activities, so that offenders have less opportunities to commit crime [5]. A large quantity of literature has been produced over the years. Most of that is related to social aspect and users' perception, or gives a technical guidance to prevent criminal acts. Often statements are redundant and very few attempts have been made in order to carry out a tool capable to rate crime risk of

版权所有(c)国家科技图书文献中心 D25062000629ZX00

built environment in a rational manner. Only European Standards in crime reduction by urban planning and building design (CEN 14383) seems to represent complete packages containing technical, physical, social and organizational measures. In particular, part 3 [6] describes risk analysis of the vulnerability of dwellings to burglary. On the basis of this method, the following work aims to carry out a rational model to perform a crime risk assessment from a different point of view.

Phases, Materials and Methods

The starting point was collecting urban and architectural features dealing with specific criminal events, leaving out social and criminal aspects of CPTED theory, since not closely related to physical vulnerability. Relying on FEMA 452 [7], three different layers of defense have been established. These levels correspond to: features linked to the neighborhood (environment around the building), characteristics of the entire property (space adjacent to the building), and area within the building (perimeter of internal areas). The main interest of the proposed method has been directed towards the evaluation of the third layer of defense, where windows represent the weakest part of the building. Therefore, all those parameters referred to the peripheral space have been excluded in order to quickly perform field tests. Openings, and characteristics related to, are the object of the evaluation. All the main features that affect offenders' criminal behavior have been picked out and written down in Table 1. It shows 5 indicators: Level of Floor (LoF), Type of Road (ToR) and Building type (BiF) present in Front of the opening, Level of Lighting (LoF) and Visibility (V) from the outside of the property. Each feature is scored using a 3 level scheme. The scores (ordinal type) "1" represents the lowest risk level and "3" the highest one.

Table 1 : Indicators

ITEM	CHARACTERISTIC	SCORE
Level of Floor (LoF)	hard to reach	1
	reachable by climbing aids	2
	very easy to reach	3
Type of Road (ToR)	main road	1
	secondary road	2
	no streets	3
Building in Front (BiF)	mixed-use building	1
	mono-use building	2
	no buildings	3
Lighting Level (LL)	good	1
	poor	2
	no light	3
Visibility (V)	good	1
	limited	2
	absent	3

版权所有(c)国家科技图书文献中心 D250620006292X00

Through Table 1, the field test starts. Since police archives and other useful information were out of reach, direct interviews to acquaintances on site allowed us to obtain needed data to perform the analysis. Only people from burgled buildings have been contacted, so as to have at least one broken-opening for each building. 160 openings have been obtained from the survey. Table 1 describes which kind of features have been checked out during the inspection for each aperture, counting out those with fixed bars or impossible to reach by a potential burglar. Despite subjectivity of the assessment present at this step, the expert evaluator has established objective criteria in order to rigorously define the difference between scores.

Once got data, Bayes' Theorem [8] and Mutual Information [9] have been applied in order to have an idea about findings from the following analysis. Facing with a study of a dichotomous dependent variable and multiple independent quantitative variables, multiple logistic regression model [10] is the one best suited to the analysis. The model could not be defined properly with the linear regression equation since the value of Y, given predictors set as E (Y|X), cannot take on any value from minus infinity to plus infinity, but it must necessarily be within the range [0,1], since it represents the conditional probability P(B|X) of the burglary event. Therefore, the interest is not the expected (or predicted) value, as in the linear regression, but the probability that a subject belongs, at least, to one of the two groups. Then, the function to be solved becomes as follows:

$$\eta = \alpha + x_1\beta_1 + x_2\beta_2 + \cdots x_n\beta_n \quad (1)$$

where $x_1, x_2, \ldots, x_n$ are the predictors, α and β are coefficients and n is the number of predictors. The resolution of the equation consists on the evaluation of model parameters ($\alpha, \beta_1, \beta_2, \ldots$ *etc.*). While these parameters are convenient for testing the usefulness of predictors, $Exp(\beta_1, \beta_2, \ldots)$ represents the ratio-change in the odds of the event of interest for a one-unit change in the predictor. In such estimation for logistic regression analysis, maximum likelihood algorithm (MLE) is used to estimate model parameters so as to maximize the function (log-likelihood function), which reveals how likely it is in order to obtain the expected value of Y given the values of independent variables. The process is repeated (iteration) until the improvement capacity of the function is infinitesimal (it converges). Backward method, used on this test by IBM® SPSS Statistic software [11], fits in one time all the variables into the model and then removes them one by one, depending on the significance of the coefficients obtained by Likelihood Ratio test. This procedure is repeated recursively until only variables with statistically significant coefficients remain within the model. As a further check, we have repeated the analysis using forward stepwise methods, by starting with a model that does not include any of the predictors. At each step, the predictor with the largest score statistic, whose significance value is less than a specified value (0.05), is added to the model. Once obtained predictors, through logit function (2), probability (P) is calculated as follows:

版权所有(c)国家科技图书文献中心 D25062000629ZX00

$$P = \frac{\exp(\eta)}{1+\exp(\eta)} \quad (2)$$

The model is complete, but it is needed to verify its goodness of fit, its predictability, and the significance of variables. In the linear regression model, the coefficient of determination, R^2, summarizes the proportion of variance in the dependent variable associated with the predictor (independent) variables, with larger R^2 values indicating that more of the variation is explained by the model, to a maximum of 1. For regression models with a categorical dependent variable, it is not possible to compute a single R^2 statistic that has all of the characteristics of R^2 in the linear regression model, so approximations are computed instead. Following methods are used to estimate the coefficient of determination. Cox and Snell's R^2 [12] is based on the log likelihood for the model compared to the log likelihood for a baseline model. However, with categorical outcomes, it has a theoretical maximum value of less than 1, even for a "perfect" model. Nagelkerke's R^2 [13] is an adjusted version of the Cox & Snell R-square that adjusts the scale of the statistic to cover the full range from 0 to 1. While these statistics can be suggestive on their own, they are most useful when comparing competing models for the same data.

It has been more useful to save the predicted probabilities, and then construct a Receiver Operating Characteristic (ROC) Curve. For each case, the predicted response is "Yes" if that cases' model-predicted probability is greater than the cutoff value specified in the dialogs (0.5 as default). The ROC Curve procedure provides a useful way to evaluate the performance of classification schemes that categorize cases into one of two groups. The ROC curve is a visual index of the accuracy of the assay. The area under the curve represents the probability that the assay result for a randomly chosen positive case will exceed the result for a randomly chosen negative case. Sensitivity is the probability that a "positive" case is correctly classified, and specificity is the probability that a "negative" case is correctly classified.

Algorithm Definition and Goodness of Fit

After got data from the surveys, by applying Bayes' Theorem and measuring their Mutual Information, we had an idea about expected outcomes. Table 2 shown below has been worked out from Bayes' Theorem: each cell in this table presents conditional probabilities (P) for burglary event (B) given a certain score.

Table 2 : Conditional probabilities obtained using Bayes' Theorem

P(B|LoF=1) = 0.02	P(B|ToR=1) = 0.05	P(B|BiF=1) =	P(B|LL=1) = 0.03	P(B|V=1) = 0.01
P(B|LoF=2) = 0.12	P(B|ToR=2) = 0.17	P(B|BiF=2) =	P(B|LL=2) = 0.06	P(B|V=2) = 0.08
P(B|LoF=3) = 0.25	P(B|ToR=3) = 0.13	P(B|BiF=3) =	P(B|LL=3) = 0.22	P(B|V=3) = 0.42

版权所有(c)国家科技图书文献中心 D25062000629ZX00

The mutual information (I) for pairs of each parameter (LoF, ToR, BiF, LL, V) and occurrence of burglary (B) are presented in Table 2.

Table 3 : Mutual Information values

I(LoF;B) = 0.0755
I(ToR;B) = 0.0102
I(BiF;B) = 0.0124
I(LL;B) = 0.0506
I(V;B) = 0.1551

From this first analysis we have figured out that burglary event (B), given a specific feature, has a quite homogeneous performance (at low likelihood of the event corresponds low score), but Type of Road (ToR) and Building in Front (BiF) have a slightly different trend. Even mutual information shows that ToR has a very low value, it means that knowing this variable does not reduce the uncertainty over B. On the other hand, visibility (V) is the factor which exchanges most information with B, and the correlation values reported in Table 4 confirms this statement.

Table 4 : Correlations

		B	LoF	BiF	LL	V
B	Pearson Correlation	1	0.299**	0.117	0.236**	0.446**
	Sig. (2-tailed)		0.000	0.113	0.001	0.000
	N	185	185	185	185	185

Starred Pearson correlation values mean that correlation is significant at the 0,01 level (2- tailed). BiF has the lowest value, but, anyway, useful to carry on the research. ToR has been excluded because of its negative Pearson value ($\rho_{x,y}$), then negative correlation with B. Usually, $\rho_{x,y}$ greater than 0,3 represents a good correlation.

Logistic regression analysis confirmed about the misleading value of ToR. Both forward stepwise method and backward method stopped iterations leaving out this indicator, pointing out its troubled significance. The two methods have chosen the same variables, so we can be fairly confident that it's a good model. Table 5 shows the last step of the procedure, in which B column represents the variables in equation (1).

Table 5 : Variables in the equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for	
								Lower	Upper
Last Step	LoF	4.164	0.896	21.590	1	0.000	64.308	11.105	372.416
	BiF	1.627	0.948	2.945	1	0.086	5.087	0.794	32.615
	LL	3.554	0.863	16.972	1	0.000	34.953	6.444	189.590
	V	3.854	0.967	15.888	1	0.000	47.164	7.090	313.725
	Constant	-33.653	7.310	21.193	1	0.000	0.000		

版权所有(c)国家科技图书文献中心 D2506200629ZX00

Almost all values of significance level of the Wald statistic are small (less than 0.05); that means these parameters are useful to the model. The equation resulting from the iterative logistic regression process is as follows:

$$\eta = -33.653 + 4.164LoF + 1.627BiF + 3.554LL + 3.854V \quad (3)$$

Hosmer-Lemeshow statistic indicates a poor fit if the significance value is less than 0.05. The results presented in Table 6 confirm that the model adequately fits the data.

Table 6 : Hosmer and Lemeshow test

Step	Chi-square	df	Sig.
1	3.495	7	0.836

This statistic draws together the observations into groups of "similar" cases, and it is then computed based upon these groups, as shown in Table 7.

Table 7 : Contingency table for Hosmer and Lemeshow test

		B = 0.00		B = 1.00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	23	23.000	0	0.000	23
	2	16	16.000	0	0.000	16
	3	19	18.999	0	0.001	19
	4	22	21.985	0	0.015	22
	5	20	19.954	0	0.046	20
	6	22	22.667	1	0.333	23
	7	23	21.349	0	1.651	23
	8	14	14.451	5	4.549	19
	9	1	1.597	19	18.403	20

Table 8 shows that all the variables chosen by backward stepwise method have significant changes in -2 log-likelihood. The change in -2 log-likelihood is generally more reliable than the Wald statistic. As in the other tables, we note that only the BiF value is a bit out on the performance of other coefficients.

Table 8 : Model if term removed

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	LoF	-50.403	57.144	1	0.000
	BiF	-23.749	3.836	1	0.050
	LL	-39.516	35.368	1	0.000
	V	-42.979	42.294	1	0.000

The model proposed in Table 9 (with the largest R^2 statistic) shows an alternative index of goodness of fit related to the R^2 value.

版权所有(c)国家科技图书文献中心 D25062000629ZX00

Table 9 : Model summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
last	43.663[a]	0.427	0.780
a. Estimation terminated at iteration number 9 because parameter estimates changed by less than 0.001.			

Classification Table 10 shows the practical results using the logistic regression model. For cases used to create the model, 157 of the 160 burgled openings are classified correctly, and 20 of the 25 non-burgled are classified correctly. Overall, 95.7% of the cases are classified correctly. This high percentage show a high performance model considering the last iteration step, where ToR is excluded and BiF is present. By choosing to ignore BiF (considering as a poor significance variable), overall percentage decreases, and the predictability of the model as well. For this reason we decided to leave BiF in the model.

Table 10 : Classification table

	Observed		Predicted		
			B		Percentage Correct
			0.00	1.00	
Final step	B	0.00	157	3	98.1
		1.00	5	20	80.0
	Overall Percentage				95.7

Figure 1 shows the receiver operating characteristic (ROC curve) in which AUC (Area Under the Curve) is equal to 0.97725 and the asymptotic significance is less than 0.05. It means that using the proposed model is much better than guessing (B=1 is the positive level).

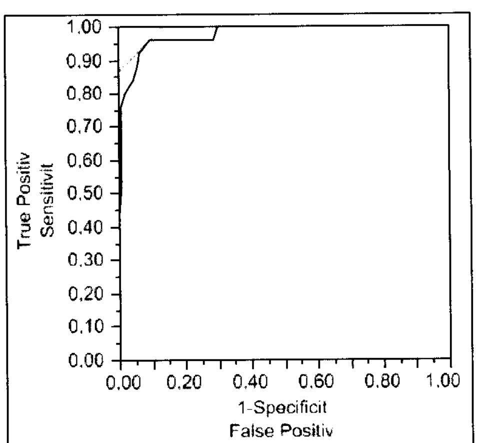

Figure 1 : Receiver Operating Characteristic Curve

版权所有(c)国家科技图书文献中心 D2506200629ZX00

Discussion of Results

Outcomes have shown a good reliability of the algorithm: Hosmer-Lemeshow statistic with significance equal to 0.836, Nagelkerke's R^2 equal to 0.78, and AUC equal to 0.97725. From Table 4 we have figured out that Level of Floor, Visibility and Lighting Level are the most correlated features with a potential criminal act. Figure 2 represent the logit function (2) for collected data from the surveys. It shows not a uniform concentration of data throughout the function because of limited data gathered from the survey, but the typical shape of logit function is easily recognizable. More data of burgled openings could better perform this curve.

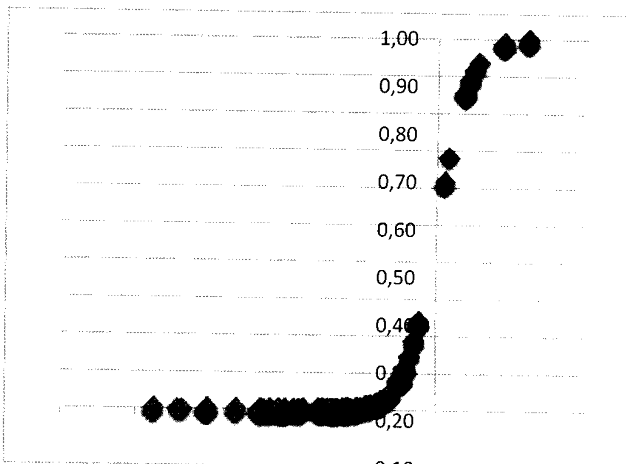

Figure 2 : Logit function

Conclusion

The proposed procedure has been calibrated with field data compiled after some crime events occurred last years in a town near Ancona (Italy). First, a simple screening procedure based on sidewalk survey has been developed in this study [14]. Then, starting from a qualitative analysis of a similar case for Urbino (Italy) [15], and updating it by recent researches, how much some indicators affect potential value for burglary risk in residential buildings has been estimated. Furthermore, the algorithm turned out from the survey is capable to evaluate, in a rational manner, burglary risk of openings first, and of entire building then. The proposed procedure is intended to serve as an initial step both for the treatment of a large-scale spread crime risk, and for detailed treatment of each individual property in the portfolio at risk. In this way you have a specific value about how much a building is vulnerable, according to its physical/architectural configuration and which are the most exposed weak spots.

版权所有(c)国家科技图书文献中心 D2506200629ZX00

References

1. J. Jacobs, "Death and life of great American cities", New York, Vintage Books, 1961.
2. C. R. Jeffery, "Crime prevention through environmental design", Beverly Hills, Sage Publications, 1971.
3. O. Newman, "Defensible space: Crime prevention through urban design", New York, Macmillan, 1972.
4. T. D. Crowe, "Crime prevention through environmental design", Woburn (MA), Butterworth-Heinemann, 1991.
5. P. Ekblom, "Redesigning the Language and Concepts of Crime Prevention Through Environmental Design", in Proceedings of 6th Ajman International Urban Planning Conference: City and Security, Ajman, United Arab Emirates, March 2013.
6. CEN/TS 14383/3, Prevention of Crime – Urban Planning and Building Design: Dwellings, October 2005.
7. FEMA 452, Risk Assessment: A how-to guide to mitigate potential terrorist attacks against buildings, Buildings and Infrastructures Protection Series, July 2010.
8. E.T. Jaynes, "Bayesian Methods: General Background." In Maximum-Entropy and Bayesian Methods in Applied Statistics, by J. H. Justice (ed.). Cambridge: Cambridge Univ. Press, 1986
9. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, "Section 14.7.3. Conditional Entropy and Mutual Information". Numerical Recipes: The Art of Scientific Computing (3rd ed.), New York: Cambridge University Press, 2007.
10. W. David, Hosmer, Lemeshow, Stanley, "Applied Logistic Regression" (2nd ed.), Wiley, 2000.
11. IBM, "IBM SPSS User's Guide", International Business Machines (IBM) Corp., New York, 2010.
12. D. R. Cox, E. J. Snell, "Analysis of binary data", 2nd edition, Chapman and Hall / CRC Press, 1989.
13. N. J. D. Nagelkerke, "A note on a general definition of the coefficient of determination", Biometrika, Vol. 78, No. 3., pp. 691-692, 1991.
14. H. Sucuoglu, U. Yazgan, A. Yakut, "A Screening Procedure for Seismic Risk Assessment in Urban Building Stocks", Earthquake Spectra, Vol. 23, No. 2, p. 441-458, 2007.
15. R. Montagna, E. Pandolfi, G. Agarbati, "Rehabilitation of Urbino-2 slum-2 – Criteria to Upgrade the Safety Level in Building Design", in Proceedings of 39th IAHS World Congress on Housing, Milan, Italy, September 17-20, 2013.

版权所有(c)国家科技图书文献中心 D2506200629ZX00