Methodology, Modeling and Case Studies of Break-in Offence Patterns Using Fuzzy Logic and Markov Chains

by

Simon Xin Li
B.A.Sc., Shanghai Jiao Tong University, 1993

A PROJECT REPORT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING
in the
School of Engineering Science

© Simon Xin Li 2003
SIMON FRASER UNIVERSITY
November 2003

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.
APPROVAL

Name: Simon Xin Li
Degree: Master of Engineering
Title of Thesis: Methodology, Modeling, and Case-Studies of Break-in Offence Patterns using Fuzzy Logic and Markov Chains

Examinng Committee.

Dr. William A. Gruver
Professor, Engineering Science, SFU
Senior Supervisor

Dr. Patricia L. Brantingham
Professor, Criminology Department, SFU
External Supervisor

Wesley Speerbrecker
Manager, Research & Development, CPAL Inc.
External Supervisor

Date Approved: November 26, 2003
PARTIAL COPYRIGHT LICENCE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay:

Author: 

(Signature) 

Simon Xin Li

(Name) 

Nov. 26, 2003

(Date)
Abstract

This exploratory study incorporates the methodology of fuzzy logic and Markov chains for environmental burglary pattern analysis. We treat fuzzy models based on a variety of fuzzy sets, operators, hedges, and reasoning methods, and specifically create a characterization of the models in the suspect's *modus operandi* ranking. Markov chain models are then used to manipulate the *fuzzy numbers* and *possibility values*. Finally, *fuzzy arithmetic* methods are explored to evaluate the quality of the risk analysis.

The fuzzy logic models are constructed and simulated in MATLAB. Fuzzy sets are represented by membership functions with triangular curves, trapezoidal curves, Gaussian curves, and Pi-curves combined with fuzzy operators and inference methods. *If-Then* rules formulate the conditional statements that comprise the fuzzy logic. Fuzzy inference is based on *Mamdani's method*, where the *center of area* and *mean of maximum* strategies are adopted for defuzzification.

In addition to MATLAB, models are extended with C++. These extensions which implement fuzzy sets, operators, hedges, are packaged into dynamic linked libraries. Case studies are analyzed using a multiple-document interface, and created using Visual C++ and *Microsoft Foundation Class (MFC)* to examine burglary pattern cases. These MATLAB models and C++ extensions then enable the development of case studies which model pattern cases.

The case studies focus on the burglary models and examples from the literature. Variables are chosen from the residential entry/exit points, property taken, behaviors on the scene, overt security, and these are interpreted as fuzzy sets. The expert-knowledge derived from criminological theory (specifically environmental criminology) is analyzed and characterized with *linguistic variables, fuzzy hedges, If-Then rules, transition matrices*, and *reward matrices*.
Dedication

To Kristy, Mom, Dad, and Vincent
Acknowledgments

I would like to give thanks to Dr. William A. Gruver who has been supervising my study and research for the past three years. Without his patience and flexibility, this project would not have been possible. I would also like to thank Crime Prevention Analysis Lab Inc. (CPAL) and Forensic Logic Inc. (FLI) which helped find the general direction of the research, provided GIS maps and graphics, and assisted me to practice and rehearse presentations. I wish to thank Chris Giles and Dr. Patricia L. Brantingham for providing research guidelines on burglary pattern models and editing the project report. Finally, I want to give tribute to my family and faithful friends for their encouragement and support.
Contents

Approval .............................................................................................................................. ii
Abstract .......................................................................................................................... iii
Dedication ......................................................................................................................... iv
Acknowledgments ............................................................................................................. v
Contents .......................................................................................................................... vi
List of Tables ..................................................................................................................... ix
List of Figures ................................................................................................................... xi
List of Examples ............................................................................................................... xiv

1 Introduction ...................................................................................................................... 1
  1.1 Crime Patterns ........................................................................................................... 1
  1.2 Fuzzy Logic .................................................................................................................. 2
    1.2.1 Fuzziness and Contestability ................................................................................. 3
    1.2.2 Certainty and Possibility ....................................................................................... 4
  1.3 Project Objective ......................................................................................................... 5
  1.4 Research Overview ..................................................................................................... 5
  1.5 Research Tools ............................................................................................................ 6
  1.6 Organization of the Report .......................................................................................... 7
  1.7 Research Beneficiary .................................................................................................. 7

2 Environmental Criminology Theory ............................................................................. 8
  2.1 Introduction ................................................................................................................ 8
    2.1.1 Routine Activities Theory ...................................................................................... 8
    2.1.2 Rational Choice Theory ....................................................................................... 9
    2.1.3 Crime Pattern Theory .......................................................................................... 10
  2.2 Criminological Models .............................................................................................. 11
    2.2.1 Offender Behavior Cycle ..................................................................................... 12
    2.2.2 Target Selection Model ....................................................................................... 13
    2.2.3 Repeat Offender and Hot Place ........................................................................... 15
  2.3 Hypothesis and Assumption ...................................................................................... 16
    2.3.1 Individual Behavior and Pattern Aggregation ...................................................... 16
    2.3.2 Risk Distribution ................................................................................................ 17
  2.4 Applicability of Intelligent Systems to Burglary Patterns ....................................... 18
    2.4.1 Fuzziness and Linguistics .................................................................................... 18
    2.4.2 Offender, Target, and Cue .................................................................................. 20
    2.4.3 Spatial Risk Analysis ............................................................................................ 21
    2.4.4 Possibility/Probability and Prediction ................................................................... 22
  2.5 Artificial Intelligent Models for Burglary Patterns ................................................... 22
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.1 Fuzzy Environmental Factor Analysis</td>
<td>23</td>
</tr>
<tr>
<td>2.5.2 Spatial Risk Analysis</td>
<td>24</td>
</tr>
<tr>
<td>2.5.3 Markov Chain Analysis</td>
<td>25</td>
</tr>
<tr>
<td>2.6 Research Data</td>
<td>26</td>
</tr>
<tr>
<td>2.6.1 Newlands Study</td>
<td>27</td>
</tr>
<tr>
<td>2.6.2 Overview of Data Collection</td>
<td>27</td>
</tr>
<tr>
<td>2.6.3 Target District</td>
<td>27</td>
</tr>
<tr>
<td>2.6.4 District Overview</td>
<td>28</td>
</tr>
<tr>
<td>2.6.5 Environmental Factors</td>
<td>29</td>
</tr>
<tr>
<td>2.6.6 Research Concerns</td>
<td>31</td>
</tr>
<tr>
<td>2.7 Summary</td>
<td>31</td>
</tr>
<tr>
<td>3 Fuzzy Set Theory</td>
<td>33</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>33</td>
</tr>
<tr>
<td>3.2 Basics of Fuzzy Set</td>
<td>33</td>
</tr>
<tr>
<td>3.2.1 Membership Functions</td>
<td>33</td>
</tr>
<tr>
<td>3.2.2 Linguistic Variables</td>
<td>35</td>
</tr>
<tr>
<td>3.2.3 Instances and Attributes</td>
<td>36</td>
</tr>
<tr>
<td>3.3 Fuzzy Sets</td>
<td>41</td>
</tr>
<tr>
<td>3.3.1 Statistical Variables</td>
<td>42</td>
</tr>
<tr>
<td>3.3.2 Physical Variables</td>
<td>47</td>
</tr>
<tr>
<td>3.3.3 Dichotomous Variables</td>
<td>52</td>
</tr>
<tr>
<td>3.3.4 Interval Variables</td>
<td>53</td>
</tr>
<tr>
<td>3.3.5 Nominal Variables</td>
<td>56</td>
</tr>
<tr>
<td>3.3.6 Psychometric Variables</td>
<td>58</td>
</tr>
<tr>
<td>3.4 Summary</td>
<td>58</td>
</tr>
<tr>
<td>4 Fuzzy Logic and Reasoning</td>
<td>60</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>60</td>
</tr>
<tr>
<td>4.2 Fuzzy Logic Operations</td>
<td>60</td>
</tr>
<tr>
<td>4.2.1 Basic Zadeh-Type Fuzzy Set Operations</td>
<td>60</td>
</tr>
<tr>
<td>4.2.2 C++ Implementation</td>
<td>62</td>
</tr>
<tr>
<td>4.3 Fuzzy Reasoning</td>
<td>63</td>
</tr>
<tr>
<td>4.3.1 Linguistic Hedges</td>
<td>63</td>
</tr>
<tr>
<td>4.3.2 Approximation Reasoning</td>
<td>65</td>
</tr>
<tr>
<td>4.3.3 Fuzzy Propositions</td>
<td>66</td>
</tr>
<tr>
<td>4.3.4 Monotonic Reasoning (Tsukamoto’s Method)</td>
<td>69</td>
</tr>
<tr>
<td>4.3.5 Complex Monotonic Reasoning</td>
<td>76</td>
</tr>
<tr>
<td>4.4 Fuzzy Compositional Rules of Inference</td>
<td>78</td>
</tr>
<tr>
<td>4.5 Methods of Decomposition and Defuzzification</td>
<td>79</td>
</tr>
<tr>
<td>4.5.1 Composite Moments (Centroid)</td>
<td>80</td>
</tr>
<tr>
<td>4.5.2 Composite Maximum</td>
<td>82</td>
</tr>
<tr>
<td>4.5.3 Singleton Geometry Representations</td>
<td>83</td>
</tr>
<tr>
<td>4.6 Summary</td>
<td>89</td>
</tr>
</tbody>
</table>
5 Possibility and Fuzzy Arithmetic................................................................................. 90
5.1 Fuzzy Number........................................................................................................ 90
5.2 Fuzzy Arithmetic .................................................................................................. 92
5.3 Ordering/Ranking of Fuzzy Numbers...................................................................... 92
  5.3.1 Possibility measure on Ordering of Fuzzy Numbers ...................................... 92
  5.3.2 Linear Ordering of Fuzzy Numbers ................................................................. 94
  5.3.3 Approximation of Fuzzy Number Ranking ...................................................... 95

6 Case Studies ................................................................................................................. 99
6.1 Design Methodology................................................................................................. 99
6.2 Case Study: Environmental Factor Analysis............................................................. 102
  6.2.1 Hypothesis and Assumptions ........................................................................... 102
  6.2.2 Scalable Monotonic Chaining ........................................................................ 102
  6.2.3 Single Attribute Analysis ................................................................................ 105
  6.2.4 Multi-Attribute Analysis .................................................................................. 118
  6.2.5 Discussion......................................................................................................... 125
6.3 Case Study: Spatial Risk Analysis .......................................................................... 126
  6.3.1 Hypothesis and Assumptions ........................................................................... 126
  6.3.2 Risk Analysis: Intrinsic Attributes .................................................................. 126
  6.3.3 Risk Analysis: Spatial Distance Decay ........................................................... 130
  6.3.4 Hybrid Risk Analysis ...................................................................................... 135
  6.3.5 Discussion......................................................................................................... 139
6.4 Case Study: Markov Chain Model ......................................................................... 140
  6.4.1 Hypothesis and Assumptions ........................................................................... 140
  6.4.2 Transition Matrix and Dynamic Pattern Characteristics ............................... 141
  6.4.3 Reward Matrix and Pattern Prediction ............................................................ 146
  6.4.4 Decision Quality Analysis ............................................................................... 156
6.5 Summary...................................................................................................................... 161

7 Conclusions ..................................................................................................................... 163
7.1 Problems in Newlands’ Study and Improvements .................................................... 163
7.2 Strengths and Limitations ......................................................................................... 165
7.3 Fuzzy Logic Model Versus Probabilistic Model ....................................................... 167
7.4 Contributions ............................................................................................................. 167
7.5 Future Research ......................................................................................................... 168

Bibliography ..................................................................................................................... 169
# List of Tables

Table 2.1 Environmental factors and attributes ................................................................. 30
Table 3.1 Fuzzy set definition for residential break-in ratio ............................................. 34
Table 3.2 Fuzzy sets with nominal quantities .................................................................... 37
Table 3.3 Fuzzy sets with ordinal quantities ...................................................................... 38
Table 3.4 Fuzzy sets with dichotomy quantities ................................................................. 39
Table 3.5 Fuzzy sets with interval quantities .................................................................... 40
Table 3.6 Fuzzy sets with ratio quantities ........................................................................ 41
Table 3.7 Fuzzy set definition for residential break-in ratio ............................................ 42
Table 3.8 Fuzzy Set definition for Break-In-Frequency ................................................... 45
Table 3.9 Fence Height in high crime blocks ................................................................... 49
Table 3.10 Fuzzy set definition for residential break-in ratio ......................................... 49
Table 3.11 Railing Design in high crime blocks ............................................................... 51
Table 3.12 Fuzzy set definition for residential break-in ratio: Railing Design ................ 51
Table 3.13 Crime Prevention Decals (all crime blocks) ..................................................... 52
Table 3.14 Fuzzy set definition for residential break-in ratio: Crime Prevention Decals ................................................................................................................. 52
Table 3.15 Building Type in high crime blocks ............................................................. 54
Table 3.16 Building Type in all crime blocks .................................................................. 54
Table 3.17 Building Type in All Crime Blocks ................................................................ 55
Table 3.18 Fuzzy set definition for residential break-in ratio: Building Type ................. 55
Table 3.19 Fuzzy set definition for residential break-in ratio: Fence Type ..................... 57
Table 3.20 Fuzzy set definition for residential break-in ratio: Fence Type ..................... 57
Table 4.1 Fuzzy set definition for residential break-in ratio: Fence Height .................. 68
Table 4.2 Crime rate and basement design in high crime blocks .................................... 70
Table 4.3 Crime rate and basement design in all blocks .................................................. 71
Table 4.4 High crime blocks ......................................................................................... 85
Table 4.5 High crime blocks ......................................................................................... 86
Table 4.6 High Crime Blocks ......................................................................................... 87
Table 4.7 All Blocks ........................................................................................................ 87
Table 6.1 Crime rate and basement design in all blocks ................................................... 104
Table 6.2 Building Age in all crime blocks ..................................................................... 106
Table 6.3 Fuzzy set definition for residential break-in ratio ........................................... 106
Table 6.4 COA decomposition: Residential Characteristic in high crime blocks ........... 123
Table 6.5 COA decomposition: Barriers and Deterrents in high crime blocks ............. 123
Table 6.6 COA decomposition: Entry and Access in high crime blocks ....................... 124
Table 6.7 COA Decomposition: activity node risk assessment in high crime blocks......128
Table 6.8 Spatial Factors: membership functions, operations, results.........................134
Table 6.9 Hybrid Factors: membership functions, operations, and results...................138
Table 6.10 Fuzzy domain and set: visiting volume, police patrol, and block-watching.................................................................150
Table 6.11 Comparison: visiting volume, police patrol, and block-watching................150
Table 6.12 Comparison: state vector of break-ins in with/without decision...............156
Table 6.13 Kaufmann and Gupta’s Criteria: Linear ordering of fuzzy number..........158
Table 6.14 Linear Ordering of Fuzzy Number: left and right dominance...............160
List of Figures

Figure 1.1 Spatial restraints and crime density .......................................................... 1
Figure 1.2 Street block patterns ................................................................................ 2
Figure 1.3 Crisp Sets and Fuzzy Sets ......................................................................... 3
Figure 1.4 Fuzzy C++ Analysis Tools: Burglar Risk Assessment GUI ..................... 6
Figure 2.2 Awareness Space (Individual Offender) .................................................. 11
Figure 2.1: A Specific Behavior Cycle: Burglary ....................................................... 12
Figure 2.4 Decrease Research Intensity and Target Selection ................................. 17
Figure 2.5 Risk Distribution ..................................................................................... 18
Figure 2.7 Spatial Risk Interactions .......................................................................... 21
Figure 2.8 Illustration of activity nodes, pathways, and edges .................................. 23
Figure 2.9 Illustration of Spatial Risk Analysis ....................................................... 24
Figure 2.10 Illustration of Markov Chain Analysis .................................................. 26
Figure 2.8 Road Map ............................................................................................... 28
Figure 3.1 Venn diagram and membership function: crisp set and fuzzy set ............... 34
Figure 3.2 Linguistic variables: Building Age .......................................................... 36
Figure 3.2 Discrete and continuous membership functions ....................................... 40
Figure 3.3 Membership function for output variable: break-in ratio ....................... 43
Figure 3.4 Membership function for statistical input variable: Break-In-Frequency .... 46
Figure 3.5 A physical input variable ......................................................................... 48
Figure 3.6 Membership function: Fence Height ........................................................ 50
Figure 3.7 Membership function: Railing Design ..................................................... 51
Figure 3.8 Membership function: Crime Prevention Decals ...................................... 53
Figure 3.9 Membership function: Building Type ...................................................... 55
Figure 3.10 Membership function: Fence Type ........................................................ 58
Figure 4.1 Fuzzy Operations .................................................................................... 61
Figure 4.2 Illustration: break-in ratio hedges ............................................................ 66
Figure 4.3 Fuzzy Inference: Proposition, Composition, Decomposition ................... 69
Figure 4.4 Fuzzy reasoning: Monotonic Reasoning in high crime blocks ................ 72
Figure 4.5 Basement Design: input fuzzy set, S-curve ............................................ 74
Figure 4.6 Basement Design: output fuzzy set, S-curve ........................................... 74
Figure 4.7 Basement Design: Input and output ........................................................ 75
Figure 4.8 Two-dimensional Monotonic Reasoning inference space ....................... 77
Figure 4.9 Defuzzifying the fuzzy set of break-in ratio (COA) ................................. 81
Figure 4.10 Defuzzifying the fuzzy set of break-in ratio (MOM) .............................. 82
Figure 4.1 Singleton geometry models vs. region-based models: membership functions ......................................................... 83
Figure 4.12 Singleton geometry models vs. region-based models: Proposition, (Composition) and Decomposition ......................... 84
Figure 4.13 Singleton geometry decomposition: fence type at front (high crime blocks) .......................................................... 85
Figure 4.14 Singleton geometry decomposition: fence type at rear (high crime/all blocks) .......................................................... 88
Figure 5.1 Evolution of Fuzzy Number ......................................................................................................................... 90
Figure 5.2 Triangular Fuzzy Number (TFN) ................................................................................................................... 91
Figure 5.3 Maximizing and minimizing set ..................................................................................................................... 95
Figure 5.4 Ranking two triangular fuzzy number ............................................................................................................... 97
Figure 6.1 Surface Comparison ........................................................................................................................................ 100
Figure 6.2 Comparisons of Decomposition Methods ........................................................................................................ 101
Figure 6.3 Scalable Monotonic Chaining Scheme: Break-in Risk Assessment of Residential characteristics ......................... 103
Figure 6.4 Membership function: Building Age ................................................................................................................ 107
Figure 6.5 Fuzzy rules: Building Age ................................................................................................................................ 108
Figure 6.6 Inferencing: Building Age .................................................................................................................................. 108
Figure 6.7 Output surface: Building Age .......................................................................................................................... 109
Figure 6.8 Risk assessment: Building Age at 15 years old ................................................................................................. 111
Figure 6.9 Risk assessment: Building Age at 35 years old ................................................................................................. 112
Figure 6.10 Input shapes: basement design (Residential Characteristics), Preventive Decals (Barriers and Deterrents), and Windows Location on first floor at rear (Entry and Access) ................................................................. 127
Figure 6.11 Inference and result: applying AND operator (minimum) ...................................................................................... 128
Figure 6.12 Comparison: geometry implication vs. fuzzy composition ..................................................................................... 129
Figure 6.13 Spatial influence: adjacent areas ...................................................................................................................... 131
Figure 6.14 Regional Park: before and after above operator ................................................................................................. 132
Figure 6.15 Sky Train Station: before and after below operator .............................................................................................. 132
Figure 6.16 Park and Sky Train Station: before and after above, below, and must (min) operators ...................................................... 132
Figure 6.17 Public school zone: before and after VERY and CLOSE operators ...................................................................... 133
Figure 6.18 Public school zone: comparison of original shape, VERY, and CLOSE operators .................................................. 133
Figure 6.19 Spatial Risk Analysis: shapes of interim and final results ...................................................................................... 133
Figure 6.20 Hybrid factors: shapes of spatial factors and residential factors .............................................................................. 136
Figure 6.21 Residential Factors: before and after applying OR operator ................................................................................. 136
Figure 6.22 Spatial Factors: before and after applying fuzzy operators ................................................................................ 137
Figure 6.23 Hybrid Factors: before and after applying OR operator to the fuzzy sets of residential and spatial.................................137
Figure 6.24 Burglary distribution and diffusion during PNE fair.................................142
Figure 6.25 Burglary distribution and diffusion during PNE fair with crime preventions.................................147
Figure 6.26 Markov Chain Model: diagram and pseudo code.................................149
Figure 7.1 Clustering in high crime blocks and low crime blocks.................................163
Figure 7.2 Clustering based on preempted concept of hot spot.................................164
Figure 7.3 Knowledge bases and the fuzzy reasoning process.................................166
List of Examples

Example 3.1: Fence Height ................................................................. 48
Example 3.2: Railing Design ............................................................... 50
Example 3.3: Crime Prevention Decals ............................................... 52
Example 3.4: Building Type ............................................................... 54
Example 3.5: Fence Type ................................................................. 57
Example 4.1: C++ Code for complement of fuzzy sets ....................... 62
Example 4.2: C++ code for intersection of fuzzy sets .......................... 62
Example 4.3: C++ code for union of fuzzy sets .................................... 63
Example 4.4: C++ implementation of Concentration ......................... 64
Example 4.5: C++ Implementation of Dilation ..................................... 64
Example 4.6: C++ Implementation of Intensification ........................... 64
Example 4.7: Fence Height ............................................................... 68
Example 4.8: Basement Design ......................................................... 70
Example 4.9: C++ Implementation of min operation ............................. 78
Example 4.10: C++ Implementation of max operation ......................... 79
Example 4.11: C++ Implementation of Fuzzy Additive operation .......... 79
Example 4.12: C++ Implementation of Centroid method ..................... 81
Example 4.13: C++ Implementation of MOM method ......................... 83
Example 4.14: Fence Type (Triangular Membership Function) ............. 85
Example 6.1: Building Age ............................................................. 105
Example 6.2: C++ code for Building Age factor .................................. 110
Example 6.3: C++ code for creating fuzzy set for Building Age factor .... 112
Example 6.4: C++ code for creating fuzzy rules for Building Age factor ... 113
Example 6.5: C++ code for defuzzificating fuzzy rules for Building Age factor .... 118
Example 6.6: C++ code for simple summarizing the Barriers and Deterrent factors: ... 119
Example 6.7: C++ code for summarizing the barriers and deterrent factors in COA method ...................................................................... 122
Example 6.8: C++ code for summarizing decision support of Synthetical Risk Analysis ................................................................. 130
Example 6.9: C++ code for summarizing decision support of spatial risk analysis .... 134
Example 6.10: C++ code for summarizing decision support of hybrid Risk Analysis ... 139
Scenario 1: Less Effective ............................................................... 146
Scenario 2: More Effective .............................................................. 146
1 Introduction

This research studies the mathematical modeling of an environmental criminological pattern, and illustrates the use of artificial intelligence (AI) methods to compose case studies based on the work of Malcolm Newlands [2]. We begin with an overview of crime pattern theory with respect to burglary, and then briefly outline the research project.

1.1 Crime Patterns

Crime patterns can be viewed at the macro level and the micro level. The patterns at the macro level usually involve characteristics of an area or district. At this level, the structures of a region result in spatial patterns, for instance, clusters of the crime activities. Crime Pattern Theory developed by Brantingham and Brantingham [8] uncovered a strong connection between structure of the urban landscape and spatial patterning of criminal incidents. Figure 1.1 shows an example of crime activities (arsons) clustered along railways, rivers, and major city streets. The density of these activities is strongly affected by physical structures such as streets, bridges, and residential and commercial premises.

Figure 1.1 Spatial restraints and crime density
(By Permission ©CPAL Inc.)
The patterns at a micro level usually involve the aggregation of individual offenders’ selection of signals and cues from private and public premises (i.e., houses, shops, parks, schools, transits, and airports). Aggregations of criminal events expose spatial patterns, and the decision-making process of individuals indicates motivation and behavior. At macro level, the attributes of premises result in statistical crime aggregate patterns, for instance, different break-in rates at street blocks. Figure 1.2 shows the structural differences between interior street blocks and border street blocks. Brantingham and Brantingham found that “when the crime rates by block were compared with the location of the blocks in the neighborhood sets, the hypothesized structures emerged: blocks which were in the border areas had higher break-in ratios than blocks which were in the interior of the neighborhood sets.” [7, pp276] Therefore, crime patterns exist at macro and micro levels. The patterns at micro levels are accumulated, which results in the macro level patterns of an area.

![Figure 1.2 Street block patterns](By Permission ©Patricia Brantingham) [7]

### 1.2 Fuzzy Logic

In 1965, Lotfi Zadeh [23] introduced fuzzy set theory as a mathematical way to represent vagueness in linguistics. Initially, fuzzy logic was applied in engineering systems, especially in control systems and pattern recognition, and enjoyed a broad application in Japan during the 1980s. Fuzzy logic has become a standard technology in many
industries: computer science, engineering, socio-economic and management sciences, business and finance.

1.2.1 Fuzziness and Contestability

*Fuzzy set theory* can be considered a generalization of classical set theory. Classical set theory works well under simple and isolated natural circumstances, but is insufficient to resolve contemporary problems with complexity, interactions, and human subjectivity. In a classical (non-fuzzy) set, an element of the universe either belongs to or does not belong to the set, which means the membership of an element is crisp — it is either *yes* (in the set) or *no* (not in the set). A fuzzy set is a generalization of an ordinary set in that it allows the degree of membership for each element to range over the unit interval \([0, 1]\). Thus, the membership function of a fuzzy set maps each element of the universe of discourse to its range space which, in most cases, is set to the unit interval. [5, pp3]

![Diagram showing crisp and fuzzy sets](image)

*Figure 1.3 Crisp Sets and Fuzzy Sets*

Compared with experimental tools in natural or physical sciences, the methods of quantitative data collection for sociology consist of observations of physical traces, naturally occurring behaviors, or other phenomena. The data are often measured in the form of checklists, or coding schemes, or interview. Error and incompleteness are incurred in sampling in measurement and result in how users interpret, where subjectivity arises. The observations may be particularly informative but vague about the activities
because of ignorance or inabilities or nonverbal behaviors. Fuzzy logic theory has an ability to describe naturally occurring events in natural settings. [4, pp299]

When investigators or criminologists assess residential break and enter risk, they evaluate the past break-in ratio in the area or nearby areas, and estimate the vulnerability of the residence. Some variables are vague and ambiguous of measurement, called fuzziness; those that are verbally vague in description are called linguistics. If both are not taken into account during the assessment, erroneous decisions may occur, potentially jeopardizing the investigation or insurance policy. Fuzzy logic addresses issues associated with intrinsic imprecision and vagueness rather than those directly concerned with the failure of our measuring devices or the inaccuracy of measurements. Intrinsic imprecision is associated with a description of the properties of a phenomenon, not with our measurement of the properties using some external device. [6, pp49]

1.2.2 Certainty and Possibility

Initially risk assessment statistically summarizes intrinsic factors, (i.e., residential character, barrier and deterrent, access and entry) in Newlands' study [2], and spatial factors such as the risk influences from neighboring areas, plus other empirical knowledge. These influences are evaluated in linguistic forms, which is an approximation to phenomena. The approximation contains a degree of imprecision and uncertainty. Fuzzy numbers and fuzzy arithmetic are valuable tools for handling imprecise or uncertain data.

Furthermore, possibility analysis using a Markov Chain model is explored to determine the possibilities in consecutive events, called risk prediction. This research first studies the possibility and probability of an event, then evaluates the quality of risk assessment by comparing risks in the form of fuzzy numbers before and after a decision. The prediction process for assessment is a key component in strategic planning and resource allocation.
1.3 Project Objective

The objective of this project is to build risk analysis models, grounded in environmental criminology through the use of fuzzy logic and Markov chains. The risk analysis models are intended to discover characteristics of environmental factors, recognize spatial patterns, and predict reoccurrence. *Fuzzy Arithmetic* methods are applied to determine the quality of the assessment. The research helps improve the quality of human decision-making in investigation or insurance risk.

The models are created and simulated using the *MATLAB Fuzzy Toolbox* [10]. We design, develop, and implement a fuzzy logic analysis tool, combining the interests of academic research with industrial application. Fuzzy logic functions are designed and extended into C++.

1.4 Research Overview

The project incorporates fuzzy logic and Markov chains for burglary pattern analysis. The foundation of the research is fuzzy set theory. First, the original data and burglary offence models were taken from the honors thesis of Newlands entitled *An Analysis of Burglary Offence Patterns in a Vancouver Neighborhood* [2]. In order to convert the original data to fuzzy sets, certain assumptions are required. His research compared high-crime blocks with all blocks in the target area to identify significant crime patterns. Additionally, he classified original data into three categories: *Residence Characteristics, Barriers and Deterrents, Entry and Access*.

By demonstrating that fuzzy set theory and Markov chains provide workable models for criminological risk analysis, I will show the advantage of adopting these models over traditional analysis. To achieve the latter objective, I will establish several case studies at the following three levels:

Environmental factors: analysis of physical signals and cues in fuzzy set theory;
Spatial Risk Analysis: analysis of aggregations of offence events in fuzzy logic theory;
Time-series Prediction: analysis of trends and evaluation of the quality of the decisions. These AI modeling and methodology are illustrated in the case studies and summarized in the conclusions.

1.5 Research Tools

The MATLAB Fuzzy Logic Toolbox allows users to create and edit fuzzy inference systems by using graphical tools, command-line functions, and C code. The integrated nature of the MATLAB environment provides users with the ability to create tools to customize the Fuzzy Logic Toolbox [10]. A Multiple-Document Interface (MDI) application is created using Visual C++ 6.0 and Microsoft Foundation Class (MFC). The design of the graphical user interface facilitates the analysis of the case studies. All fuzzy models are hard-coded in the program. The results are graphically displayed and listed in the grid on the right-hand side. C++ is a general-purpose programming language with intrinsic facilities for bundling code packages into objects. This project uses fuzzy sets, fuzzy set operators, fuzzy set hedges, and fuzzy reasoning to build a fuzzy decision-support system. I do not discuss how to write the source code in detail for coding fuzzy sets, fuzzy operators, fuzzy set hedges, and fuzzy reasoning. All basic functionalities can be obtained from The Fuzzy Systems Handbook [6]. The basic fuzzy functions are unchanged. For this application, the C++ code from the latter reference has been modified, especially for advanced features such as implications and defuzzification.

![Figure 1.4 Fuzzy C++ Analysis Tools: Burglar Risk Assessment GUI](image)
1.6 Organization of the Report

The project will first discuss, in Chapter 2, the basic concepts of crime pattern theory and introduce some models and examples of burglary offence analysis. In Chapters 3, 4 and 5, the principles and techniques of fuzzy set theory and Markov chains are described, and examples and scripts are presented for MATLAB and C++. Chapter 6 presents the analysis of the fuzzy model using case studies where the models, solutions, and applications are discussed extensively for each case.

The case studies are categorized in three levels. First level analysis is introduced to describe the characteristics through fuzzy set theory. The methods cover both a single factor and multi-factors. The basic reasoning methods are discussed including monotonic and center of area (COA, centroid). Second level analysis provides decision support analyses including linguistic, operations, environmental factor analysis, spatial risk assessment, and hybrid risk analysis. These models associate the relationship between cues and burglary offences. The correlation discloses burglary patterns and indicates vulnerability. In the third level, an analysis of Markov chain models and fuzzy arithmetic is provided to predict risks and to evaluate burglary-preventive precautions. Chapter 7 provides a conclusion and discussion highlighting the strengths and limitations of this study, and suggestions for future research.

1.7 Research Beneficiary

Researchers, developers, and institutions from academics and industry benefit from the application of artificial intelligence to jurisdiction projects. These potential users or groups are found in several industries: insurance, police, government, public safety agencies, and academic institutions.
2 Environmental Criminology Theory

Research on burglary offence patterns is grounded in the logic, theory, and research findings of environmental criminology and uses these elements as guides for crime analysis. In this chapter, I introduce research on burglary offence patterns, the crime site selection model, and then discuss the relationships between the target selection model and the artificial intelligence model with respect on modeling, methodology, and implementation.

2.1 Introduction

Three theoretical perspectives have inspired continued development in the field of environmental criminology. These three theoretical perspectives are: (1) *Routine Activities Theory* by Cohen and Felson [26]; (2) *Rational Choice Theory* by Cornish and Clarke [28]; and (3) *Crime Pattern Theory* by Brantingham and Brantingham [24]. This review will focus on research examining target/victim selection cues of offenders, research devoted to the criminogenic nature of certain places, and research examining the prevalence and patterns of repeat offending.

2.1.1 Routine Activities Theory

In 1979, Cohen and Felson developed a ground-breaking theory to explain the changing nature of the burglary offence in the United States. They believed three elements are necessary for a crime to occur: a motivated offender, a suitable target, and the absence of a capable guardian(s). In this theory, these three conditions must converge “in time and space” for a predatory offence to occur. It assumes that the motivated burglar is present and a suitable target is one which the burglar is attracted to, based on the offender’s assessment of target attractiveness cues. Additionally, there must be an absence of a guardian (or guardians) who could stop the burglary from occurring. Moreover, Cohen and Felson argue that crime is dependant on the changing routine activities of victims and targets (i.e. changes to the durability and manufacturing of products). *Routine activities*
theory discovered the connections between burglary offence patterns and the presence of offenders, guardians, targets and victims. [27]

2.1.2 Rational Choice Theory

Rational choice theory is based on the assumption that offending is purposive behavior and designed to satisfy the needs of the burglar in some sense. These theorists emphasize the need to consider the rational decision process of the offender, but note that in most cases the decision should be considered one based on the notion of limited rationality. In 1978, Brantingham and Brantingham proposed a model of target selection that is based on a multi-staged decision process of offender [22]. They asserted that a motivated offender:

Goes through a “multi-staged decision process which seeks out and identifies, within the general environment, a target or victim positioned in time and space” [22, pp107]. These cues are derived from the environment in which an offender is operating. An offender uses these cues for target selection and learns these cues “through experience” or “social transmission” from other experienced offenders. Through experience, an offender develops “individual cues, clusters of cues, and sequences of cues associated with ‘good’ targets” [22, pp108] and these develop into crime templates. Crime targets are accepted based on their correspondence to a crime template.

Although a single offender may possess a multitude of crime selection templates, “once the template is established, it becomes relatively fixed and influences future searching behavior” [22, pp108].

Much of the research dedicated to the understanding of cue and template use in target selection has focused on robbery offenders and burglary offenders. However, Burglary research shows that burglars use similar cues when selecting a residence to victimize [17]. Although ethnographic studies indicate that offenders use cues and develop crime templates for victim/target selection, this research concludes that offenders rarely incorporated all preferred cues. Instead, due to drug use and cash-intensive lifestyles,
offenders choose victims that “appeared to meet their minimal subjective criteria for an acceptable victim” [32, pp87-88].

2.1.3 Crime Pattern Theory

Brantingham and Brantingham note that crime of different types is not evenly distributed across the urban landscape in space and time, nor are offenders and victims/targets evenly distributed. In 1984, they developed *Crime Pattern Theory* which takes account of the importance of zoning and population flows in a city. They believe that offenders are people with rationality and follow a sequential, hierarchical course in selection of the actual target [7, 17].

Brantingham and Brantingham introduced *awareness space* to explain the pattern of personal and property crimes by focusing on the movement patterns of both potential victims and offenders [7, 17]. An *Awareness Space Model* more depends on the offender’s perception (*knowledge*) of the environment and emphasizes three inter-related concepts to explain the movement pattern of the people (both offenders and potential victims): activity nodes, pathways, and edges. The first concept is an activity node. According to Brantingham and Brantingham, activity nodes are centers of high activity where individuals “spend the majority of their time” [25, pp36] and include the home, school, work, places of entertainment, shopping areas, etc. The second concept is pathways. Pathways are the routes that connect a person’s activity nodes. Pathways would include, but are not limited to, streets, major streets, sidewalks, footpaths, etc. These paths may be traveled by foot, public transit or automobile. As people travel these paths to and from activity nodes with some regularity, the “paths and narrow areas surrounding them become known spaces to the people who travel them” [25, pp35].

Third, Brantingham and Brantingham introduced the concept of edges [31]. An edge is a boundary that cannot easily be traversed. These boundaries can be both physical and perceptual. A physical edge would include rivers, forests, bridges, etc., whereas a perceptual edge would include areas that people are afraid of, such as a rival gang territory or areas with a large discrepancy in socioeconomic status, e.g. a middle class person may experience discomfort when entering high crime inner city areas.
Conversely, an inner city offender may wish to victimize lucrative targets in affluent areas, but generally would not venture into the affluent areas because anonymity is absent.

![Figure 2.2 Awareness Space (Individual Offender) (By Permission ©Patricia Brantingham) [8, pp27-54. Adapted]](image)

These three concepts form a person's\/offender's awareness space [24, pp358]. The theory predicts offenders will commit crimes in their awareness spaces, either along a path, along an edge or around an activity node. This theory assumes that crimes are less likely to occur, all else being equal, outside the awareness spaces of offenders, because they are not aware of the location of attractive targets, they would have to devote much time and attention to the routines of guardians, and they are less cognizant of escape routes. For a conceptual illustration of an awareness space refer to Figure 2.2.

2.2 Criminological Models

The aforementioned theories provide several models to study crime patterns from the micro (e.g., highly variable decision-making of individual offenders) to the macro (e.g., aggregations of criminal events). An Offender Behavior Cycle explains the motivations behind offenders; and the Target Selection Model analyzes offender cues and crime templates used in the selection of targets. The Awareness Space Model explains how an offender/victim moves in an urban landscape under the structures of activity nodes, pathways, edges and understands the characteristics of the environment.
2.2.1 Offender Behavior Cycle

Studies of the offender behavior cycle helps researchers understand the causes behind the burglary offences, (e.g. the motivations to start, the techniques to execute, and the illicit markets to convert stolen goods into cash). According Scarr [3]:

"Burglary is a crime against a place, or against property, not against people and itself is behavior. Like all behavior, it involves needs to be met, opportunities to meet them, perception of these opportunities, means to take advantage of such opportunities, satisfactions when needs are met, decisions about alternate routes to need-meeting, and the existence of outside interference in the process. Thus, schematically, the following elements are necessary in any approach to burglary, in the hypothetical cyclical order, as indeed they are a form of motivated behavior:

- Needs that may be met through successful burglarizing.
- Knowledge of burglary technology.
- Opportunities to burglarize.
- Path to meet needs.
- Choice of burglary over other paths.
- Attempt, which succeeds in the complete cycle.
- Conversion of the burglarized goods into a useful form.
- Satisfactions for the act.
- Reinforcement of the whole series of steps in the cycle, thus increasing its probability of reoccurrence."

![Figure 2.1: A Specific Behavior Cycle: Burglary. [3, pp117]](image)

The cycle of the elements enables us to organize burglary information, acquire the patterns of behavior, similarities, and trends. This study relates these elements in the
cycle such as the *opportunities perceived to burglarize, path to meet needs*, and *choice of burglary over other paths*. The model provides a platform to better understand and interpret the analytical results of each case study. [3, pp3-5]

### 2.2.2 Target Selection Model

During the offender behavior cycle, offenders are driven by motivations from different sources with diverse explanations: *etiological models or theories, economics, sociology,* and *culture*. However, the *Offender Behavior Cycle* fails to account for the process of selecting certain targets over others. This decision process is accounted for through the work of Brantingham and Brantingham [22, pp107-108] who developed a theoretical model of target selection that posits:

"II. Given the motivation of an individual to commit an offence, the actual commission of an offence is the end result of a multi-staged decision process which seeks out and identifies, within the general environment, a target or victim positioned in time and space.

a) In the case of high-affect motivation, the decision process will probably involve a minimal number of stages.

b) In the case of high-instrumental motivation, the decision process locating a target or victim may include many stages and much careful searching.

III. The environment emits many signals, or cues, about its physical, spatial, cultural, legal and psychological characteristics.

   (a) These cues vary from generalized to detailed.

IV. An individual motivated to commit a crime uses cues (either learned through experience or learned through social transmission) from the environment to locate and identify targets or victims.

V. As experience grows knowledge grows, an individual motivated to commit an offence learns which individual cues, clusters of cues, and sequences of cues are associated with "good" victims or targets. These cues, cue clusters, and cue sequences (spatial, physical, social, temporal, and so on) can be considered a template which is used in victim or target selection. Potential victims or targets are compared to the template and either rejected or accepted, depending on the congruence.

   (a) The process of template construction and the search process may be consciously conducted, or these processes may occur in an unconscious, cybernetic fashion so that the individual cannot articulate how they are done.

VI. Once the template is established, it becomes relatively fixed and influences future searching behavior, thereby becoming self-reinforcing.

VII. Because of the multiplicity of targets and victims, many potential crime selection templates could be constructed. But because the spatial and temporal distribution of targets and victims is not regular, but clustered and patterned, and because human environmental perception has some universal properties, individual templates have similarities which can be identified."
These theorists accept that individuals exist who are motivated to commit specific offences and that the sources of motivation are diverse. Different etiological models or theories may appropriately be invoked to explain the motivation of different individuals or groups. Additionally, the strength of such motivation varies and the character of such motivation varies from affective to instrumental. From this viewpoint, we can treat the research as a decision-making and knowledge-based system and develop a series of subject-centric decision support and expert systems.

According to Brantingham and Brantingham, a burglary offense target is, all else being equal, located within an awareness space of an offender. An offender enters an area, searches generally for a target, then gradually aims at a particular opportunity, and finally plans to commit the burglary offense. This search process is partially structured. Therefore, Brantingham and Brantingham [7, 22] speculated that as part of this structured search process:

An offender engages in a partially structured search process in which he locates first an attractive target or neighborhood

- Then an attractive sub-area, e.g., a block
- Then a building within this sub-area
- and finally, a unit within the building.

There has been much support for the Brantingham and Brantingham model of target selection in robbery and burglary research [1, 17]. An offender may follow this process completely or partially, and in most cases target sites are selected based on opportunity and are not carefully planned. For example, Cromwell, et al. [17, pp41] described three common opportunistic patterns:

- The burglar happened by the potential burglary site at an opportune moment when the occupants were clearly absent and the target was perceived as vulnerable (open garage door, windows, etc.).
The site was one that had been previously visited by the burglar for a legitimate purpose (as a guest, delivery person, maintenance worker, or other such activity).

The site was chosen after "cruising" neighborhoods searching for a criminal opportunity and detecting some overt or subtle cue as to vulnerability or potential for material gain.

To summarize, offenders looking for a target search for cues and acquire knowledge of the environment. Criminological research has provided evidence for this opportunistic process of target selection:

- Offenders indicate they generally keep their eyes open for targets, and if they see an attractive target, they try to remember it for future reference. [17]
- Offenders selected their targets from within a narrow "activity space", that is, areas they passed through in daily journeys from home to work or school, and to the main social and shopping locations they frequent. [19]
- Similar observations were found with juvenile delinquents. [20]

2.2.3 Repeat Offender and Hot Place

An offender seeks targets in a territory in which they are familiar and are reluctant to cross boundaries. As experience grows knowledge grows, and a template or templates of cues, cue clusters, and cue sequences (spatial, physical, social, and so on) are formed and used in target selection [22, pp107-108]. Recent studies have found that there exists a small proportion of offenders who commit a disproportionate amount of crime [17]. Some research indicates that some offenders develop a preference for robbery and burglary offences. _Target Selection Theory_ explains that these offenders would have the most developed and fixed crime templates for target selection. This finding is important for this research to apply AI to analyze burglary offence patterns.

Furthermore, environmental criminologists focus on the importance of place and its relationship to crime. Researchers have found that certain places experience a disproportionate amount of crime [24]. Brantingham and Brantingham conceptualize this disproportionate level of criminal activity attributed to places by referring to them as
crime generators and crime attractors [31]. Crime generators are places or areas “to which large numbers of people are attracted for reasons unrelated to any particular level of criminal motivation they might have or to any particular crime they might end up committing” [30, pp7]. These areas or places “produce crime by creating particular times and places that provide appropriate concentrations of people and other targets in settings that are conducive to particular types of criminal acts” [30, pp7]. Examples of crime generators include the development of new rapid transit routes and stations, opening a new bar, opening new shopping centers, etc. These places generate crime that was not prevalent in an area. Crime attractors “are particular places, areas, neighborhoods, districts which present well-known criminal opportunities to which strongly motivated intending criminal offenders are attracted because of the known opportunities for particular types of crime” [30, pp8]. These places or areas include inner city ghettos, opening a needle exchange clinic in a high crime area, etc. Often the difference between these two concepts is theoretical. However, the major differentiator here is that a crime generator produces crime in an area that was absent prior to the establishment of the place, while a crime attractor intensifies criminal activities already present in a particular area. This finding provides a basis for the analysis burglary patterns on or around public sites (e.g., Sky Train stations, schools, parks, P.N.E.).

2.3 Hypothesis and Assumption

2.3.1 Individual Behavior and Pattern Aggregation

Although the motivations and behaviors of individuals are highly variable, Crime Pattern Theory predicts that all else being equal, attractive victims or targets that fall within the awareness space of an offender are chosen over other possible targets. Individuals follow fairly fixed patterns when moving about in an area. Spatial clusterings are repeatedly found around criminal residence or in paths that criminal used to go to and way from home. The shape of the overall pattern depends on the concentration of housing, shopping, entertainment, and work locations. In this research, the statistical data are the aggregations of the individuals’ offence events, and the crime patterns can be recognized in their target selections within Northern Hastings, and hypothetical areas.
(such as P.N.E., sky train stations, public school, regional parks, etc.). Environmental criminology theories support this assumption. [24, pp355-357]

2.3.2 Risk Distribution

The interaction of the locations of potential targets and the awareness and activity spaces of potential criminals produces the patterns of crime occurrence [24, pp362]. Environmental criminologists have found that offenders travel very short distances from activity nodes to crime sites. To illustrate, Figure 2.4 shows that, as an offender travels away from a sky train station, the probability of target selection decreases as the distance increases from that activity node. This means more crimes are committed directly adjacent to the sky train station, where they are familiar and conduct their target search in higher intensity. Burglary offence risk closely correlates to target search intensity by distance. As the searching process is increasingly further away from the central point of activity node, where offenders are less and less familiar, the intensity of searching falls.

Figure 2.4 Decrease Research Intensity and Target Selection
(By Permission ©Patricia Brantingham) [8, pp27-54. Adapted]
The risk distribution is obviously not uniform, but can be described by mathematical models. In this research, fuzzy logic is applied to approximate the pattern of the risk distribution. Assume that public sites (e.g., sky train stations, schools, parks, P.N.E.) are hypothetically considered as aggregations of activity nodes of individual offenders in Northern Hastings area, and are treated as hot places. The distance can be defined as either travel distance or travel time between a nearby spot to the center of the activity space or hot place. The rate of distance decay can be linear or non-linear, and the risk near a spot can then be estimated by a distance decay function. This assumption provides a basis to decision-support studies in the research.

2.4 Applicability of Intelligent Systems to Burglary Patterns

The AI analysis of the environmental constraints and the decision process of offenders leads to the discovery of burglary patterns. In this section, we discuss the applicability of AI to analyze crime patterns.

2.4.1 Fuzziness and Linguistics

It can be argued that social science data are either fuzzy or linguistic. This is evident in the data used as the basis of this research. For example, Newlands described fence height in terms of grouped height categories \{under 3', 3' - 4', 4' - 6', 6'\}, but this can be
operational in a fuzzy terms as fence height \{low, medial, medial high, high\}. The reason is the offenders are unlikely to measure fence height, instead, they have the concept of high or low, seeking for a match between the concept and their mental perception, and finally describe fence height verbally for plotting offences. They actually approximate the characterization of phenomena that might be too complex or too ill-defined to be amenable to description in conventional quantitative terms [6, pp114]. The concept of low and high is natural and more applicable when observers describe the attribute of fence height. Therefore, the attribute is linguistic in description. Routinely, the linguistic concept is introduced when a researcher begins to analyze an offence pattern in form of survey questionnaires, checklists or coding schemes or when a person (such as investigator, a victim, or witness, etc.) describes a crime scene. In addition, more linguistic words are widely used such as some, near, about, very, etc.

On the other hand, the limitation of Newlands’ category design causes ambiguity and vagueness, for example, fence heights at 3’ or 4’ or 6’ could be classified into two categories because of observers’ choice and inaccurate measurement. Observer choice causes ambiguity, which is the uncertainty associated with choice and difficulty in making a choice between two or more alternatives. For example, the observer could place fence heights at 4’ in crisp sets 3’ – 4’ and 4’ – 6’ because of the category design. Inaccurate measurement causes vagueness, which is the uncertainty associated with the difficulty of delimitating sharp or precise boundary in classifying measures. For example, fence heights at 3.9’ and 4.1’ are classified in two distinct categories but contribute no significant difference on output space. In addition, the observation is purposely designed for measurement and the measurement demands accuracy. However, any measurement is not free from instrument error and human intervention. Observation (measurement) of fence height could be affected by viewpoint, landscape, color, shape. Therefore, the category design of fence height is viewed as fuzziness of criminological data.

Both fuzziness and probability concepts are related to uncertainty but quite different. Probability is concerned with whether an event occurs or not. Offender Decision Model
shows the behaviors of individual offenders exhibit the *uncertain* and offenders are driven by motivations from different sources with diverse explanations: *etiological models or theories, economics, sociology,* and *culture.* *Fuzziness* measures the degree to which an event occurs, other than whether it occurs or not. The statistical data of Newlands' are aggregations of offense events and reflect the physical characteristics of cues and signals rather than individual behaviors. The *uncertainty* largely arises from the methods of quantitative measurement in crime scene investigation, witness interview, and criminological survey and questionnaire. The design of fuzzy sets depends on the nature of *uncertainty* that lies in criminological data.

Therefore, both *fuzziness* and *linguistic* are applicable in environmental criminology, but sometimes they overlap and sometimes one prevails over the other.

2.4.2 Offender, Target, and Cue

*Rational choice theory* states that criminals make decisions, and that is they choose among alternative courses of action according to their perceptions of the risks and gains associated. Additionally, a degree of rationality can be attributed to offenders in the planning and executing their burglary offences. The related models such as *Offender Behavior Cycle* and *Target Selection Model* can be applied to analyze the decision-making process of an offender through physical representations of cues. Therefore, the risks on a target site can be understood through cues and signals. Although this research is based on statistical data that are not focusing on individual behaviors, these models assist with the analysis of the aggregation of the offence events in order to recognize the crime patterns.

A crime selection template is virtually a knowledge-based *database* system and a rule-based inference system. Offenders develop a set of rules and logics to reflect their rationality spontaneously or intentionally. The *database* contains *data* — motivations, instruments, alcohol and drugs, targets, markets, and risks (fears of punishment). The *inference system* contains *logical rules* — searches, selection, execution, cashing out, and prosecution.
On the other hand, people usually describe their observations and perceptions in a linguistic and vague way, so that the offender-based data are usually imprecise, incomplete or uncertain—all these attributes are considered to exhibit a strong sense of fuzziness. The logical rules offenders think and narrate in their minds are linguistic, thus they may use words to describe and decide, such as about, near, around, possible, must, not, very, much, extreme, and somehow. All these words are considered to exhibit a strong sense of linguistics.

By studying imitation of decision-making process and characteristics of cues and signals (e.g., in fuzzy inference methods), the burglary offence patterns are recognizable in a geographic area. The patterns are an aggregation of individual offenders’ target selections: cues/templates within their awareness spaces.

2.4.3 Spatial Risk Analysis

The search process is influenced by the objective site characteristics and by the spatial risk interaction of neighborhood sites. The objective site characteristics can be described in mathematical models; the risk interactions can be estimated by a risk decay function. Figure 2.7 depicts the uneven spatial distribution of stationary targets (such as residences, shops, sky train stations, etc.) superimposed over the physical characteristics in the target area.

![Figure 2.7 Spatial Risk Interactions](image)
The patterns of individuals are then aggregated in a geographic area. Therefore, the target attractiveness can be understood mathematically and can be represented by AI models. The decision support theory and knowledge-based system can simulate the process of perception under such a space. The data mining technologies can find the patterns of the signals and cues in terms of similarity or differentiation [1, 7]. The Case Study Part II (Section 6.3) is dedicated to the topic and illustrates the methodology.

2.4.4 Possibility/Probability and Prediction

The risk of an event, namely the risk at a specific point of time is frequently estimated. Offence behaviors exhibit a strong sense of uncertainty and the patterns of the behaviors have to be derived statistically out of chaos, but if the risk aggregation is known at the previous time, the chance of possible risks can be estimated. For this reason, the current research adopts the techniques of possibility/probability theory to determine the chance/likelihood about whether an event (risk) occurs or not. Fuzzy number and fuzzy arithmetic are essential to the study of possibility problems in fuzzy set theory, and Markov chain models extend the capability to predict the possible trends in the time line. With the aid of fuzzy logic and Markov chains, this research studies the signals and cues emitted from the physical and social environments, and analyzes how to discern the patterns of signals and cues in the mist of the uncertainty and chaos. The Case Study Part III (Section 6.4) is dedicated to the topic and illustrates the methodology.

2.5 Artificial Intelligent Models for Burglary Patterns

Based on the assumptions of the previously discussed models developed in the field of environmental criminology and applicability of AI, intelligent burglary pattern models can be established.
2.5.1 Fuzzy Environmental Factor Analysis

In this research, Fuzzy Environmental factor analysis studies characteristics of environmental factors in cues and signals, and determines relationship between the characteristics and the corresponding burglary offence risks. The environmental factors are physical and psychological signals and cues of targets within a given area. Figure 2.8 shows different types of sites and routes that have various characteristics in constructions, traffic, residence, business, police, population, geological terrain, etc. This analysis turns to study the relationship between burglary risk and characteristics of signals and cues.

![Figure 2.8 Illustration of activity nodes, pathways, and edges](image)

Newlands correlated statistical data in Northern Hastings and pointed out the risks associated to attributes of the environmental factors using bivariate analyses (i.e. the burglary rate/ratio was correlated with different characteristics of residence) [2]. In this research, intelligent models are applied to handle fuzziness of the environmental factors, determine the risk for an attribute, and consolidate individual risks among multiple attributes of those factors into a potential estimation of overall risk. Therefore, the victimization risk of a site is assessable by inputting values of attributes into fuzzy logic models.
2.5.2 Spatial Risk Analysis

The assumption of risk distribution indicates that the risk of a site is influenced by neighborhoods. The risk could vary either by travel distance or travel time. This variation could be proportional or un-proportional and is out of scope in the study. However, decision-makers could include a measure of the variation in their rule-based estimation. Fuzzy reasoning can approximate the decision-making process.

Figure 2.9 illustrates the analysis mechanism. Suppose for Spot D, the risks are known in surrounding neighborhoods Spot A, B, C, then the risk at Spot D is in part a result of the external spatial influence from Spot A, B, C and the internal characteristics of the environmental factors of Spot D. The internal characteristics are assessable according to environmental factor analysis. The external influence is assessable for an experienced decision-maker in rule-based estimation. For example, the rule may resemble the following:

If risk A is very high, risk D is around 0.5 time of risk A;
If risk B is middle, risk D is around the risk B;
If risk C is quite low, risk D is around 1.5 time of risk C.

![Figure 2.9 Illustration of Spatial Risk Analysis](By Permission ©CPAL Inc.)
Fuzzy operation has a mathematical interpretation for very, around, and quite. Fuzzy reasoning is rule-based and has capacity to "understand" if-then rules. Therefore, the risk at node D can be computed in fuzzy logic.

2.5.3 Markov Chain Analysis

When both internal characteristics and external spatial influences are known, a timeline dimension can be added to explore the trend of ever-changing risks on a site and neighborhoods. Markov chain model can compute a state-based analysis of characteristics of signals and cues on a site. A decision-maker can bet for the possibility of an event and an array can be constructed for the possible trend for either a risk or a characteristic at a state (past, at present, future). The risks at past and current states contribute to future risk. Therefore, all possible risks can be predicted in Markov chain, compared and evaluated in fuzzy arithmetic, and finally the best prediction can be derived from those risks. Figure 2.10 illustrates that characteristics and influence vary from state 1 (past), to state 2 (at present), to state 3 (Future). At the final state, the risk at Spot D is the result of three sources of data:

- Internal characteristics and spatial influence at future state;
- Internal characteristics and spatial influence at current state, and time-line influence;
- Internal characteristics and spatial influence at past state, and time-line influence.

The later two contributions have an additional time-related influence in addition to internal characteristics and spatial influence.
2.6 Research Data

This research uses Newlands’ honors thesis as a data source, and follows his terminologies and format [2]. However, this study has limitations, including:

- 1980s’ data and questionnaire design might not reflect the present situation in the target area;
- Social science data might have reliability problems (particularly undergraduate data), for example, fence height \( \{ \text{under } 3', 3' - 4', 4' - 6', 6' \} \), the definition is not as natural as the concepts of high and low, that is not easy to understand; and the definition is binary, for example, two close values \( \{3.9' \} \) and \( \{4.1' \} \) are classified to two distinct categories, consequently accuracy is lost in measurement.
- Baseline data collection was not designed for fuzzy set, and the transformation costs accuracy and correctness.

Therefore, we treat the data hypothetically and extend the definitions of data type for convenience and simplicity. Thus, the results and conclusions are hypothetical and used for demonstration purposes. Newlands categorized the data into high crime blocks and all crime blocks. The initial MATLAB simulation indicated the data in high crime blocks more readily displayed the offence patterns, consequently the data from high crime blocks have been chosen for case studies unless the data was absent in high crime blocks.
2.6.1 Newlands Study

Newlands [2] studied the burglary offence patterns in the Northern Hastings district of Vancouver, B.C. He targeted the environmental factors, gathered data, tabulated the data, and examined the relationships between the burglary offences and individual environmental characteristics. In this section, I will briefly discuss his study. In my study, all data were obtained from his thesis *An Analysis of Burglary Offence Patterns in A Vancouver Neighborhood*.

2.6.2 Overview of Data Collection

Newlands collected burglary data from the Vancouver Police Department from January 1, 1979 to October 31, 1982. A total of 341 burglary offences occurred within the time frame in Northern Hasting area. A total of 1056 premises and forty street blocks were located in the target area. Newlands input the data into SPSS and performed individual cross-tabulations on each variables in Table 2.1.

2.6.3 Target District

The target area is Northern Hastings located in Vancouver East, whose southern side borders Hastings street, the east side borders Pandora Park, and the west side neighbors Pacific National Exhibit (PNE). Its northern side ends at the Burrard inlet and has no exits. The area has a high-density, primarily low-income population, and serval major travel and transportation routes, public schools, parks and commercial premises. The area is a typical Vancouver East District and is well-known for its high rate of burglary.
2.6.4 District Overview

1056 premises and forty street blocks were located in the target area. The residential blocks were studied and had reported a high rate of burglary. More than 20% of the housing units had property stolen or attempted burglaries had occurred. Of the 82 block faces in the study area, 28 block faces were identified as having a high break-in ratio.

2.6.4.1 Border Block and Interior Block

Of the 47 border block faces, 19 of these faces (40.4%) were identified as possessing a high break-in ratio. In addition, border block faces comprised 51.1% of the total 92 possible block faces and the high break-in ratio border block faces accounted for 67.8% of all the high break-in ratio faces within the entire study area. Other high break-in ratio faces were located opposite Callister Park and Hastings Elementary School.

The areas identified as containing the most high break-in ratio faces were areas bordering major pathways, both vehicular and pedestrian. Newlands’ findings suggested that burglars traveling along a pathway would be more likely to recall a building depending on the attributes of movement, contour, size, shape and surface. Six of the nineteen border faces with high break-in ratios (31.6%), were opposite the PNE fair grounds. This indicated that the burglary offence, in part, was influenced by the PNE fair grounds including the various PNE events (shows, exhibitions) and its facilities serving
2.6.4.2 Elementary School, Playground, Park

Two high crime block faces were adjacent to the Elementary School and two high crime block faces were also adjacent to the park within the study area. Combined, these blocks accounted for 20% of all the high crime block faces.

2.6.4.3 Commercial Premises

Four border faces with a high break-in ratio (21%) were located near a major turn in one of the paths bordering the study area and two residential high crime faces (10%) were located along the Hastings St. shopping strip. The entire shopping strip had a break-in ratio far exceeding that of the residential area. Of the entire 341 burglaries committed, 117 (34%) of these were committed against these commercial premises. The fifty-eight stores located along the shopping strip showed that commercial premises compared to residences appear to have a higher chance of being burglarized.

2.6.5 Environmental Factors

Newlands operationalized the environmental risk factors of the premises into three categories: *Characteristics of Residence, Barriers and Deterrents, and Entry and Access.* Each factor consists of attributes, which are listed in the following table.
<table>
<thead>
<tr>
<th>Environmental Factors</th>
<th>List of Sub-types or Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Residence</td>
<td></td>
</tr>
<tr>
<td>1 Building Type</td>
<td><em>Single Family Dwelling, Multi-Family Dwelling, Apartment Dwelling</em></td>
</tr>
<tr>
<td>2 Basement Design</td>
<td><em>None, Below Ground, Ground Level, Above Ground</em></td>
</tr>
<tr>
<td>3 Building Age</td>
<td><em>New, Average, Old, Under Construction.</em></td>
</tr>
<tr>
<td>4 Number of Balconies</td>
<td><em>None, One, Two to Three, Three or more</em></td>
</tr>
<tr>
<td>5 Building Location</td>
<td><em>Corner, Between Residences</em></td>
</tr>
<tr>
<td>6 Rear Lane Presence</td>
<td><em>Yes, No</em></td>
</tr>
<tr>
<td>Barriers and Deterrents</td>
<td></td>
</tr>
<tr>
<td>1 Garage Type</td>
<td><em>None, Garage Attached, Garage Separate, Carport Attached</em></td>
</tr>
<tr>
<td>2 Crime Prevention Decals</td>
<td><em>Visible, Not Visible</em></td>
</tr>
<tr>
<td>3 Step Design</td>
<td><em>None, Facing Street, Perpendicular to Street</em></td>
</tr>
<tr>
<td>4 Building Condition</td>
<td><em>Good, Average, Poor</em></td>
</tr>
<tr>
<td>5 Porch Landing Area</td>
<td><em>None, No Roof, Roof</em></td>
</tr>
<tr>
<td>6 Railing Design</td>
<td><em>None, Facing Street, Perpendicular to Street</em></td>
</tr>
<tr>
<td>7 Fence Type</td>
<td><em>None, Open Wooden, Closed Wooden, Chain Link, Metal, Cement-Rock</em></td>
</tr>
<tr>
<td>(Not Fuzzy)</td>
<td></td>
</tr>
<tr>
<td>8 Fence Height</td>
<td><em>Under 3’, 3’ - 4’, 4’ - 6’, 6’ +</em></td>
</tr>
<tr>
<td>Entry and Access</td>
<td></td>
</tr>
<tr>
<td>1 Number of Doors</td>
<td><em>None, One, Two or more</em></td>
</tr>
<tr>
<td>2 Door Level</td>
<td><em>Basement, First Floor, Second Floor, Both</em></td>
</tr>
<tr>
<td>3 Door Position</td>
<td><em>Facing Street, Perpendicular to Street, Split</em></td>
</tr>
<tr>
<td>(Both Facing and Perpendicular)</td>
<td></td>
</tr>
<tr>
<td>4 Window Location</td>
<td><em>2 -, 3 - 5, 6 +</em></td>
</tr>
</tbody>
</table>

Table 2.1 Environmental factors and attributes [2]
Newlands calculated break-in ratios from the statistical data, and correlated them with each attribute of environmental factors, and uncovered the contribution of each factor to the break-in ratio of the district. He demonstrated that the attributes of each factor are a measure of the break-in risk and can be described and evaluated by crime pattern theory and the quantitative methods.

2.6.6 Research Concerns

The research methods employed by Newlands contain some limitations. The first concern is the data collected in this study. Some methodological problems occurred when Newlands collected the data. The raw data were collected in surveys and interviews, which may introduce sources of error including sampling problems, crisp classification, subjectivity of responses, forgetting, and the honesty of those being interviewed. These problems make the offence patterns are more difficult to recognize. The second concern is the quality of the assumptions by which I converted the criminological model to the mathematical model. However, this research mainly focuses on methodology, modeling and case study, so these problems do not limit the generalizability of fuzzy logic and Markov chain models.

2.7 Summary

Brantingham and Brantingham’s target selection theory [8] explains how signals and cues play a role in the discovery of burglary offence patterns. Two AI research approaches can be studied to explore the crime patterns which include:

- **Subjective**: Explicitly imitate the offenders’ decision-making process: from perception to knowledge, and from inference to decision. From this approach, we know how an offender makes a decision and adapts the environment in the awareness space. The inference process would be found directly in choices that the offender is supposed to make.

- **Objective**: Analyze and find the patterns of the cues and signals, then infer the offenders’ knowledge, logics and consequences. An offender’s decision can be indirectly deduced from signals and cues.
Both approaches need to incorporate the techniques of AI into crime pattern models. Particularly in this study, we utilize risk analysis and pattern recognition systems case by case based on the findings and sample of Newlands [2]. His study on aggregations of signals and cues support the existence of burglary offence patterns. In this research, these signals and cues are converted to mathematical measures (fuzzy sets and fuzzy numbers) in Chapter 3. The methods to explain the influence of signals and cues are discussed in Chapter 4 and Chapter 5. The risk analysis and pattern recognition systems are illustrated through case studies in Chapter 6.
3 Fuzzy Set Theory

3.1 Introduction

In his seminal paper, "Fuzzy Sets", Lotfi Zadeh [23] founded fuzzy set theory, as a mathematical way to represent vagueness in linguistics. A fuzzy set is a generalization of an ordinary set and improves the traditional system by providing "greater generality, higher expressive power, an enhanced ability to model real-world problems, and a methodology for exploiting the tolerance for imprecision." [5, pp3] At present, fuzzy logic has enjoyed success in engineering systems, especially in control systems and pattern recognition, and accomplished broad application in Japan during the 1980s, and in Europe and North America in 1990s. Fuzzy logic became a standard technology in many industries: computer science, engineering, socio-economic and management sciences, business and finance.

3.2 Basics of Fuzzy Set

3.2.1 Membership Functions

A fuzzy set eliminates the sharp boundary that divides members from non-members in the group and gains the capability to describe vagueness in the real world. Thus, a gradual transition replaces the abrupt crisp edge between full membership and non-membership. Hence, Lin and Lee [5] defined a fuzzy set as follows:

A fuzzy set $\tilde{A}$ in the universe of discourse $U$ can be defined as a set of ordered pairs,

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) | x \in U\},$$

(3.1)

where $\mu_{\tilde{A}}(\cdot)$ is called the membership function (or characteristic function) of $\tilde{A}$ and $\mu_{\tilde{A}}$ is the grade (or degree) of membership of $x$ in $\tilde{A}$, which indicates the degree that $x$ belongs to $\tilde{A}$. The membership function $\mu_{\tilde{A}}(\cdot)$ maps $U$ to the membership space $M$, that is, $\mu_{\tilde{A}}(\cdot): U \rightarrow M$. when $M=\{0,1\}$, set $A$ is non-fuzzy, and $\mu_{\tilde{A}}(\cdot)$ is the characteristic function of
the crisp set $A$. For fuzzy sets, the *range* of the membership function $M$ is a subset of the nonnegative real numbers whose supremum is finite. In most general cases, $M$ is set to the unit interval $[0, 1]$. [5, pp11]

![Figure 3.1 Venn diagram and membership function: crisp set and fuzzy set [24]](image)

Fuzzy set membership functions describe the curve surface, and the nature of the fuzzy set members. The selection of the functions is *subjective* in nature; however, it cannot be assigned arbitrarily. Membership functions can be chosen to represent the different nature of variables. Both the MATLAB Fuzzy Toolbox and C++ fuzzy components provide basic membership functions as follows:

<table>
<thead>
<tr>
<th>Membership Functions</th>
<th>MATLAB Fuzzy Toolbox [23]</th>
<th>C++ Fuzzy DLLs [6]</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-curve (Sigmoid/Logistic)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Z-curve</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Triangular</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Trapezoidal</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>PI curve</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Beta curve</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Bell curve</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gaussian curve</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 3.1 Fuzzy set definition for residential break-in ratio
3.2.2 Linguistic Variables

Zadeh (1975) introduced the concept of linguistic variables. This is a means to approximate the characterization of phenomena that are too complex or too ill-defined to be amenable to description in conventional quantitative terms. Linguistic variables, is a basic concept in fuzzy logic and approximate reasoning, has values expressed in words or sentences in a natural or artificial language. For example, residential break-in ratio is a linguistic variable if it consists of the values such as \{low, sort of low, mid, rather mid, sort of high, high, very high\} instead of crisp set value \{0\%, 1\%, 2\%, ..., etc.\}. A linguistic variable takes words or sentences in natural or artificial languages. A generalization of the linguistic variable is as follows:

“A linguistic variable is a variable of a higher order than a fuzzy variable, and it takes fuzzy variables as its values. A linguistic variable is characterized by a quintuple \((x, T(x), U, G, M)\) in which \(x\) is the name of the variable; \(T(x)\) is the term set of \(x\), that is, the set of names of linguistic values of \(x\) with each value being a fuzzy variable defined on \(U\); \(G\) is a syntactic rule for generating the names of values of \(x\); and \(M\) is a semantic rule for associating each value of \(x\) with its meaning.” [5, pp115]

A linguistic variable defines a concept of our everyday language. For example, the attribute Building Age can be defined as the verbal concept of age as New, Average, Old. These values are the verbal word in our daily lives. A set of linguistic variables is the vocabulary of fuzzy logic systems.
3.2.3 Instances and Attributes

When generating a fuzzy set, a fuzzy number, or a linguistic variable as an input to a decision support system, the nature of the data needs to be reflected. Considering the basic data mining concept, the input data to a decision-support scheme is a set of instances. These instances are the things that are to be classified, or associated, or clustered. An instance is interpreted as a fuzzy set, which is a collection of data characterized by its values on a fixed, predefined set of features or attributes. The value of an attribute for a particular instance is a measurement of the quantity that the attribute refers to. There is a broad distinction between quantities that are numeric and ones that are nominal. Numeric attributes, sometimes called continuous attributes, measure numbers — either real or integer value. (Note that the term continuous is routinely abused in this context: integer-valued attributes are certainly not continuous in the mathematical sense.)

Nominal attributes take on values in a pre-specified, finite set of possibilities and are sometimes called categorical. But there are other possibilities. Nominal attributes are sometimes called categorical, enumerated, or discrete. Statistics texts often introduce “levels of measurement” such as nominal, ordinal, interval, and ratio. [11]
3.2.3.1 Nominal quantities

Nominal quantities have values that are distinct symbols. The values themselves serve only as labels or names — hence the term nominal. For example, in the Fence Type fuzzy set \{None, Open Wooden, Closed Wooden, Chain Link, Metal, Cement-Rock\}, no relation is implied among these attributes — no ordering or distance measure. It certainly does not make sense to add the values together, or multiply them, or even compare their size. A rule using such an attribute can only test for equality or inequality.

<table>
<thead>
<tr>
<th>Physical Input Variables</th>
<th>List of Sub-types or Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Residence</td>
<td></td>
</tr>
<tr>
<td>Barriers and Deterrents</td>
<td></td>
</tr>
<tr>
<td>1 Garage Type</td>
<td>None, Garage Attached, Garage Separate, Carport Attached</td>
</tr>
<tr>
<td>2 Fence Type</td>
<td>None, Open Wooden, Closed Wooden, Chain Link, Metal, Cement-Rock</td>
</tr>
<tr>
<td>3 Porch Landing Area</td>
<td>None, No Roof, Roof</td>
</tr>
<tr>
<td>Entry and Access</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2 Fuzzy sets with nominal quantities

3.2.3.2 Ordinal Quantities

Ordinal quantities are ones that make it possible to rank order the categories. However, although there is a notion of ordering, there is no notion of distance between the intervals. For example, in the Building Age fuzzy set \{New, Average, Old, Under Construction\}, the values are ordered. Whether Under Construction $\rightarrow$ New $\rightarrow$ Average $\rightarrow$ Old or Old $\rightarrow$ Average $\rightarrow$ New $\rightarrow$ Under Construction, is a matter of convention — it does not matter which is used so long as consistency is maintained. What is important is that New $\rightarrow$ Average lies between the other two. Ordinal attributes are generally called numeric, or perhaps continuous, but without the implication of mathematical continuity.
Note that the better term for fence height \{under 3', 3' - 4', 4' - 6', 6'+\} might be fence height \{low, medial, medial high, high\}. The term is conserved for consistency of original format of Newlands' study. The problem in the crisp set is that the values of measurements \{4'\} could be classified to both \{3' - 4'\} and \{4' - 6'\}. Errors are usually expected in this crisp classification of social science study. Fuzzy set theory can better solve the classification problem. The vagueness is included in the fuzzy set definition and the values of measurements \{4'\} can be tolerated to be classified to either fuzzy set \{media\} or \{medial high\}, errors can be minimized to an acceptable level. Same explanation is applicable to Window Location.

<table>
<thead>
<tr>
<th>Physical Input Variables</th>
<th>List of Sub-types or Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristics of Residence</strong></td>
<td></td>
</tr>
<tr>
<td>1  Basement Design</td>
<td>None, Below Ground, Ground Level, Above Ground</td>
</tr>
<tr>
<td>2  Building Age</td>
<td>New, Average, Old, Under Const.</td>
</tr>
<tr>
<td>3  Building Location</td>
<td>Corner, Between Residences</td>
</tr>
<tr>
<td><strong>Barriers and Deterrents</strong></td>
<td></td>
</tr>
<tr>
<td>1  Step Design</td>
<td>None, Facing Street, Perpendicular to Street</td>
</tr>
<tr>
<td>2  Building Condition</td>
<td>Good, Average, Poor</td>
</tr>
<tr>
<td>3  Railing Design</td>
<td>None, Facing Street, Perpendicular to Street</td>
</tr>
<tr>
<td>4  Fence Height</td>
<td>Under 3', 3' - 4', 4' - 6', 6'+*</td>
</tr>
<tr>
<td>5  Window Location</td>
<td>2 - , 3 - 5, 6 +</td>
</tr>
<tr>
<td><strong>Entry and Access</strong></td>
<td></td>
</tr>
<tr>
<td>1  Door Position</td>
<td>Facing Street, Perpendicular to Street, Split (Both Facing and Perpendicular)</td>
</tr>
</tbody>
</table>

Table 3.3 Fuzzy sets with ordinal quantities
3.2.3.3 Dichotomous Quantities

Dichotomous quantities are a special case of the nominal scale, which has only two members—often designed as true and false, or yes and no. Such attributes are sometimes called Boolean. For example, Crime Prevention Decals fuzzy set \{Visible, Not Visible\}, Rear Lane fuzzy set \{Yes, No\} are dichotomous variables.

Notice that the distinction between dichotomous and ordinal quantities is not always straightforward and obvious.

<table>
<thead>
<tr>
<th>Physical Input Variables</th>
<th>List of Sub-types or Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Residence</td>
<td></td>
</tr>
<tr>
<td>1 Rear Lane Presence</td>
<td>Yes, No</td>
</tr>
<tr>
<td>Barriers and Deterrents</td>
<td></td>
</tr>
<tr>
<td>1 Crime Prevention Decals</td>
<td>Visible, Not Visible</td>
</tr>
<tr>
<td>Entry and Access</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.4 Fuzzy sets with dichotomy quantities

3.2.3.4 Interval Quantities

Interval quantities have values that are not only ordered but measured in fixed and equal units. A good example is Door Level fuzzy sets \{Basement, First Floor, Second Floor, Second Floor above\}, expressed in floors rather than on the non-numeric scale implied by low or high. It makes sense to talk about the difference between a basement and second floor, and compare that with the difference between another two floors, say First Floor and Second Floor. The variables are known as discrete, and possess connotations of ordering when being discretized to continuous, numeric quantities.
<table>
<thead>
<tr>
<th>Physical Input Variables</th>
<th>List of Sub-types or Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Residence</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Building Type</td>
</tr>
<tr>
<td>2</td>
<td>Number of Balconies</td>
</tr>
<tr>
<td>Barriers and Deterrents</td>
<td></td>
</tr>
<tr>
<td>Entry and Access</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Number of Doors</td>
</tr>
<tr>
<td>2</td>
<td>Door Level</td>
</tr>
<tr>
<td>Statistical Input Variables</td>
<td></td>
</tr>
<tr>
<td>Break-In-Frequency</td>
<td>1, 2, 3 Break-In events in a period</td>
</tr>
</tbody>
</table>

Table 3.5 Fuzzy sets with interval quantities

![Discrete and continuous membership functions](image-url)

Figure 3.2 Discrete and continuous membership functions
3.2.3.5 Ratio Quantities

*Ratio* quantities are those for which the measurement scheme *inherently* defines a zero point. For example, when measuring distance from one object to others, the distance between the object and itself forms a natural zero. Ratio quantities are treated as real numbers: any mathematical operations are allowed. Nonetheless, these quantities *inherently* depend on our scientific knowledge—it is culture-relative, therefore fuzzy.

<table>
<thead>
<tr>
<th>Physical Input Variables</th>
<th>List of Sub-types or Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output Variables</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><em>residential break-in ratio</em></td>
</tr>
<tr>
<td></td>
<td>0~100%*</td>
</tr>
</tbody>
</table>

Table 3.6 Fuzzy sets with ratio quantities

*Note: Newslands defined break-in ratio at 20% as high risk. We found the scope of break-in ratio varies from 0%~35% in most cases of the research. The rates in 60%~100% only have mathematical meanings and are not applicable in any crimp pattern models.

3.3 Fuzzy Sets

In this section, a burglary model, namely the burglary pattern analysis of the northern Hasting-Rupert community, is interpreted and the statistical data is converted into fuzzy sets. This step tests the quality of the hypothesis assumptions and explores these mathematical nature of the burglary patterns.

Every input variable or output variable will be transformed to a fuzzy set by a membership function, therefore it will be represented as a *triangular*, a *trapezoidal*, a *shouldered*, a *Gaussian*, or a *PI* curve. Both MATLAB and C++ code provide the membership functions. The input variables of a burglary pattern consist of a physical characteristic and its burglary *Break-In-Frequency*. Only one variable is used as output variable — residential *break-in ratio*. 

41
3.3.1 Statistical Variables

3.3.1.1 Output Variable: Residential Break-in Ratio

In a linguistic variable, linguistic terms representing approximate values of a base variable, the values of which are real numbers within a specific range, are captured by appropriate fuzzy numbers. Residential break-in ratio is a statistical value and equal to the total crime number divided by the total number of residences in a bounded period of time. It is converted to a typical linguistic variable and is assigned to approximate the potential risk contributed by every physical residential input in a uniform measure. The linguistic variable contains \{low, sort of low, mid, rather mid, sort of high, high, very high\} by a semantic rule, as shown in Figure 3.3, and is defined as the ratio of the number of buildings with break-ins and the total number of buildings within a category. The value is presented as a percentage. The fuzzy numbers, whose membership functions have the triangular or trapezoidal shapes, are defined on the interval \([0,100]\), the domain of the variable, as listed in Table 3.7. The domain is 0 to 100, but here a 20% victimization rate will be considered as high break-in ratio, while more than 40% is denoted as very high.

![Table 3.7 Fuzzy set definition for residential break-in ratio](image)

<table>
<thead>
<tr>
<th>Fuzzy Sets</th>
<th>Membership Function</th>
<th>Break-in ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Start</td>
</tr>
<tr>
<td>Low</td>
<td>Triangular</td>
<td>0</td>
</tr>
<tr>
<td>Sort Of Low</td>
<td>Triangular</td>
<td>0</td>
</tr>
<tr>
<td>Mid</td>
<td>Triangular</td>
<td>5</td>
</tr>
<tr>
<td>Rather Mid</td>
<td>Triangular</td>
<td>10</td>
</tr>
<tr>
<td>Sort Of High</td>
<td>Triangular</td>
<td>15</td>
</tr>
<tr>
<td>High</td>
<td>Triangular</td>
<td>20</td>
</tr>
<tr>
<td>Very High</td>
<td>Trapezoidal</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 3.7 Fuzzy set definition for residential break-in ratio
Figure 3.3 Membership function for output variable: break-in ratio

MATLAB Implementation

```matlab
a=newfis('Burglary_Risk_COA_trim');

% Normalized Risk
a=addvar(a, 'output', 'Posibility', [0 100]);

a=addmf(a, 'output', 1, 'Posibility.Low', 'trimf', [0 0 5]);
a=addmf(a, 'output', 1, 'Posibility.SortOfLow', 'trimf', [0 5 10]);
a=addmf(a, 'output', 1, 'Posibility.Mid', 'trimf', [5 10 15]);
a=addmf(a, 'output', 1, 'Posibility.RatherMid', 'trimf', [10 15 20]);

a=addmf(a, 'output', 1, 'Posibility.SortOfHigh', 'trimf', [15 20 25]);
a=addmf(a, 'output', 1, 'Posibility.High', 'trimf', [20 30 40]);
a=addmf(a, 'output', 1, 'Posibility.VeryHigh', 'trimf', [30 60 100]);
```

C++ Implementation

```cpp
void CFuzzyResidenceRisk::CreateFuzzySetsRisk()
{
    // domain
    dDomain[0]=0; dDomain[1]=100;

    // Risk: Low
    dParms[0]=0;
    dParms[1]=5;
    dParms[2]=0;

    lpFDBRisk[0]=FzyCreateSet("Risk.Low", DECREASE, dDomain, dParms, 2, &
                             iStatus);
    strcpy(lpFDBRisk[0]->FDBdesc, "Risk: Low (in probability)");
```
Risk: Sort of Low

\[ d_{\text{Parms}}[0] = 0; \]
\[ d_{\text{Parms}}[1] = 5; \]
\[ d_{\text{Parms}}[2] = 10; \]

\[ \text{lpFDBRisk}[1] = \text{FzyCreateSet("Risk.SortOfLow", \text{TRIANGLE }, \text{dDomain}, \text{dParms}, 3, \& \text{iStatus});} \]
\[ \text{strcpy(} \text{lpFDBRisk}[1] \text{->FDBdesc,"Risk: Sort of Low (in probability)";} \]

Risk: Mid

\[ d_{\text{Parms}}[0] = 5; \]
\[ d_{\text{Parms}}[1] = 10; \]
\[ d_{\text{Parms}}[2] = 15; \]

\[ \text{lpFDBRisk}[2] = \text{FzyCreateSet("Risk.Mid", \text{TRIANGLE }, \text{dDomain}, \text{dParms}, 3, \& \text{iStatus});} \]
\[ \text{strcpy(} \text{lpFDBRisk}[2] \text{->FDBdesc,"Risk: Mid (in probability)";} \]

Risk: Rather Mid

\[ d_{\text{Parms}}[0] = 10; \]
\[ d_{\text{Parms}}[1] = 15; \]
\[ d_{\text{Parms}}[2] = 20; \]

\[ \text{lpFDBRisk}[3] = \text{FzyCreateSet("Risk.RatherMid", \text{TRIANGLE }, \text{dDomain}, \text{dParms}, 3, \& \text{iStatus});} \]
\[ \text{strcpy(} \text{lpFDBRisk}[3] \text{->FDBdesc,"Risk: Rather Mid (in probability)";} \]

Risk: Sort of High

\[ d_{\text{Parms}}[0] = 15; \]
\[ d_{\text{Parms}}[1] = 20; \]
\[ d_{\text{Parms}}[2] = 25; \]

\[ \text{lpFDBRisk}[4] = \text{FzyCreateSet("Risk.SortOfHigh", \text{TRIANGLE }, \text{dDomain}, \text{dParms}, 3, \& \text{iStatus});} \]
\[ \text{strcpy(} \text{lpFDBRisk}[4] \text{->FDBdesc,"Risk: Sort of High (in probability)";} \]

Risk: High

\[ d_{\text{Parms}}[0] = 20; \]
\[ d_{\text{Parms}}[1] = 30; \]
\[ d_{\text{Parms}}[2] = 40; \]

\[ \text{lpFDBRisk}[5] = \text{FzyCreateSet("Risk.High", \text{TRIANGLE }, \text{dDomain}, \text{dParms}, 3, \& \text{iStatus});} \]
\[ \text{strcpy(} \text{lpFDBRisk}[5] \text{->FDBdesc,"Risk: High (in probability)";} \]

Risk: Very High

\[ d_{\text{Parms}}[0] = 30; \]
dParms[1]=60;
dParms[2]=100;

lpFDBRisk[5]=FzyCreateSet("Risk.VeryHigh", RITESHOULDER , Domain, dParms,3,&iStatus);
strcpy(lpFDBRisk[5]->FDBdesc,"Risk: Very High (in probability)");
  // MdlLinkFDB(lpFDBRisk[5],lpPDBRisk,&iStatus);

return;

}  

3.3.1.2 Input Variables: *Break-In-Frequency*

An input variable to a decision-support scheme is a set of instances. These instances are the things that are to be classified, or associated, or clustered ([11] p. 41). The two types of input variables divided the inputs into two categories: physical input variables and a statistical input variable.

<table>
<thead>
<tr>
<th>Fuzzy Sets</th>
<th>Membership Function</th>
<th>Break-In-Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Start</td>
</tr>
<tr>
<td><em>Low</em></td>
<td>Triangular</td>
<td>0</td>
</tr>
<tr>
<td><em>Mid</em></td>
<td>Triangular</td>
<td>1</td>
</tr>
<tr>
<td><em>High</em></td>
<td>Trapezoidal</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.8 Fuzzy Set definition for *Break-In-Frequency*

In Newlands’ study, *Break-In counts* was denoted for measuring the frequency of burglary offence events in a period and the values were correlated to the categorized attributes. The values were divided into three categories: one Break-In, two Break-Ins, three or more Break-Ins. In this study, we hypothetically denote it as *Break-In-Frequency* and extended the definition to the average frequency of burglary offence events in a bounded period of time. The values are correlated for the categorized attributes as well and are statistical. The domain ranges from 0 to 10, but since the most frequent *break-ins* of a building within a studying period (i.e., six years) is under 3 incidents, the values more than 3 are considered to only have mathematical meanings. The variable will be used as a measure of every physical characteristic in the burglary pattern model. In this study, this variable is denoted as *Break-In-Frequency* (in some
MatLab or C application, it is called *Category*). Three fuzzy sets are *Low, Mid, High,*
which represents *one* (or more), *two* (or more), and *three* (or more) break-ins.

![Membership function for statistical input variable: Break-In-Frequency](image)

**Figure 3.4 Membership function for statistical input variable: Break-In-Frequency**

**MATLAB Implementation**

```matlab
a=newfis('Burglary_Categories_COA_trmf');

% Category
a=addvar(a,'input','Break-In-Frequency',[0 10]);

a=addmf(a,'input',2,'Break-In-Frequency.Low', 'trimf', [0 1 3]);
a=addmf(a,'input',2,'Break-In-Frequency.Mid', 'trimf', [1 2 4]);
a=addmf(a,'input',2,'Break-In-Frequency.Hight', 'trimf', [2 3 10]);
```

**C++ Implementation**

```c++
void CFuzzyResidenceRisk::CreateFuzzySetsCategory(){
    // domain
dDomain[0]=0;dDomain[1]=10;

    // Category: One+
dParms[0]=1;
dParms[1]=3;
dParms[2]=0;

    lpFDBCategory[0]=FzyCreateSet("BIF.Low",DECREASE ,dDomain,dParms, 3,&iStatus);
    strcpy(lpFDBCategory[0]->FDBdesc,"BIF: Low ( no. of burglary case in past three years)");
    // MdlLinkFDB(lpFDBCategory[0],lpPDBRisk,&iStatus);

    // Category: Two+
```

46
dParms[0]=1;
dParms[1]=2;
dParms[2]=4;

lpFDBCategory[1]=FzyCreateSet("BIF.Mid",TRIANGLE ,dDomain,dParms, 3,&istatus);
strcpy(lpFDBCategory[1]->FDBdesc,"BIF: Mid+ ( no. of burglary case in past three years"));
    //    MdlLinkFDB(lpFDBCategory[1],lpPDBRisk,&iStatus);

//    Category:    Three+
dParms[0]=2;
dParms[1]=3;
dParms[2]=10;

lpFDBCategory[2]=FzyCreateSet("BIF.High",RITESHoulder ,dDomain,dParms,3,&istatus);
strcpy(lpFDBCategory[2]->FDBdesc,"BIF: High ( no. of burglary case in past three years"));
    //    MdlLinkFDB(lpFDBCategory[2],lpPDBRisk,&iStatus);

return;
}

3.3.2 Physical Variables

Each physical variable represents a potential environmental cue or signal that exists physically and can be described mathematically. Some input variables physically exist in a residential area or are able to be measured as a characteristic of a building. They affect the burglary pattern in strong or weak relationships. Each physical characteristic contains a few sub-types or attributes, which were converted to fuzzy sets (see Table 1.1). Of those variables, some attributes can be explicitly converted to fuzzy sets such as Fence Height, Window Location, and Building Age because of the analogue property, while others need to have assumptions such as Building Location, Railing Design, Door Level, and Door Position. Some are binary such as Rear Lane Presence, Crime Prevention Decals. Some are discrete such as Building Type, Number of Doors, and Number of Balconies. Some are not related to each other in any form such as Fence Type. One example of the variables is illustrated in Figure 3.5.
In this section, methods used to convert physical variables into fuzzy sets are described. The physical variables are nominal, ordinal, interval, dichotomous, and ratio level, and are fuzzy in nature. When creating a fuzzy set, sufficient data needs to be collected to determine the domain and to validate the assumption.

3.3.2.1 Explicit Ordinal Variables

Of those variables, some attributes can be explicitly converted to fuzzy sets such as Fence Height, Building Location, and Building Age because of the analogue property. In most cases the assumption is not necessary as is the case with Fence Height and Building Age. In other cases, this assumption is still necessary for the variable Building Location, and it is assumed that all residential buildings are lined up along the street within a block, thus, the building location is measured by the distance to the nearest street corner or block side.

Example 3.1: Fence Height

When fence height was studied, significant relationships were found in the rear of homes, but not in the front. Even in the high crime blocks, fence height seemed to have a deterrent effect. There was a trend evident, suggesting that as fences became higher the more likely a home would be burglarized. The type of fence potentially hinders surveillance of a person’s backyard and thereby allows the burglar the cover he/she needs to commit his/her crime. As a result, a conflict may arise between design features needed to give privacy or beauty to urban environment and crime prevention features. [2]
Fence Height

<table>
<thead>
<tr>
<th>Fence Height</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>Under 3'</td>
<td>62</td>
<td>79.5</td>
<td>14</td>
<td>17.9</td>
</tr>
<tr>
<td>3' – 4’</td>
<td>62</td>
<td>74.7</td>
<td>14</td>
<td>16.9</td>
</tr>
<tr>
<td>4’ – 6’</td>
<td>15</td>
<td>62.5</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td>6’ +</td>
<td>3</td>
<td>42.9</td>
<td>2</td>
<td>28.6</td>
</tr>
</tbody>
</table>

Due to the analogue nature (Table 3.9), a fuzzy set is defined by using smooth curve (i.e., Gaussian in the middle and PI at both ends). Other smooth curves can be adopted as well as triangles in the middle and trapezoidals at both ends. The reason for using PI or trapezoidal curves at both ends is that the fence height is rarely lower than three feet and rarely higher than 6 feet, therefore, PI or trapezoidal curves explicitly indicate the assumption that if the fence height is near either end, then the risk curve considerably converges to a limit point.

<table>
<thead>
<tr>
<th>Fuzzy Sets</th>
<th>Membership Function</th>
<th>Start</th>
<th>Peak</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (3-)</td>
<td>Trapezoidal</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Mid (4-)</td>
<td>Triangle</td>
<td>1</td>
<td>3.5</td>
<td>6</td>
</tr>
<tr>
<td>Mid-High (6-)</td>
<td>Triangle</td>
<td>2</td>
<td>5.5</td>
<td>8</td>
</tr>
<tr>
<td>High (6+)</td>
<td>Trapezoidal</td>
<td>5</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 3.10 Fuzzy set definition for residential break-in ratio
3.3.2.2 Implicit Ordinal Variables

While some variables are explicitly ordinal, others need assumptions to measure the difference, such as *Window Location*, *Railing Design*, and *Door Position*. The assumption is subjective; however, a good assumption describes the characteristic of the variable with less error.

Example 3.2: Railing Design

Railings in the front of the home were also significantly related to crime occurrence. Because the railings faced the same way as the steps, it is not clear whether the railings themselves or the step design influenced the results. There is the possibility that both work in combination to produce the results obtained. Nevertheless, railings and steps may act as symbolic barriers that further define an individual’s territory.

Again, during the analysis of all blocks according to no crime and crime a significant relationship existed for homes with railings located in the rear. As earlier mentioned, it was this form of analysis that produced a significant value for homes with steps in the rear. Once again homes with perpendicular railings in the rear had a higher proportion within the no crime category. But in the high crime blocks, railing and step design had no effect on crime. [2]
The railing direction can be considered as an angle either to the building front side or the street. *Facing Street* is 90°, *Perpendicular to Street* is either 0° or 180°, thus, the presence of railings is an angle between 0° to 180°, therefore the problem becomes a study of the break-in risk distribution in relation to the angle from 0° to 180°. *None Railings* also needs to be represented in the fuzzy sets. Since the universe of an angle is 0° to 360°, then we arbitrarily choose 181° to 360° as *None Railings* in a trapezoidal curve, so that the risk does not change while the angle changes from 181° to 360°, and is considered as very safe virtually.

<table>
<thead>
<tr>
<th>Railings Front</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>None</td>
<td>25</td>
<td>54.3</td>
<td>17</td>
<td>37.0</td>
</tr>
<tr>
<td>Facing Street</td>
<td>144</td>
<td>74.6</td>
<td>37</td>
<td>19.2</td>
</tr>
<tr>
<td>Perpendicular to Street</td>
<td>22</td>
<td>84.6</td>
<td>3</td>
<td>11.5</td>
</tr>
</tbody>
</table>

Table 3.11 *Railing Design* in high crime blocks

![Figure 3.7 Membership function: Railing Design](image)

![Table 3.12 Fuzzy set definition for residential break-in ratio: Railing Design](image)
3.3.3 Dichotomous Variables

Some variables are binary such as Rear Lane Presence, Crime Prevention Decals and are defined in such a way as to dichotomize the elements of a given universe of discourse into two groups: Yes or No, Visible or Not Visible. The boundary of the fuzzy set is rigid and sharp and performs a two-class dichotomization. With this type of variable it seems there is not much difference between a crisp set and a fuzzy set. However, the research on such a fuzzy set still demonstrates the potentials of fuzzy logic in solving subjective problems.

Example 3.3: Crime Prevention Decals

Another deterrent to crime was the presence of crime prevention decals. In the category of One (break-in count), a significant relationship existed where homes were equipped with crime prevention decals on the front of the home, but no significant relationships existed for the decals on the rear of homes. When homes in all blocks were analyzed according to no crime and crime categories, a very significant relationship existed (p=0.0062).

<table>
<thead>
<tr>
<th>Decals Front</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>Visible</td>
<td>389</td>
<td>87.2</td>
<td>44</td>
<td>9.9</td>
</tr>
<tr>
<td>Not Visible</td>
<td>404</td>
<td>80.6</td>
<td>82</td>
<td>16.4</td>
</tr>
</tbody>
</table>

Table 3.13 Crime Prevention Decals (all crime blocks)*

* Note: Data for Crime Prevention Decals in high crime blocks are not available. Data from all crime blocks are used instead.

<table>
<thead>
<tr>
<th>Fuzzy Sets</th>
<th>Membership Function</th>
<th>Break-in Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Start</td>
</tr>
<tr>
<td>Visible</td>
<td>Trapezoidal</td>
<td>0</td>
</tr>
<tr>
<td>Not Visible</td>
<td>Trapezoidal</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Table 3.14 Fuzzy set definition for residential break-in ratio: Crime Prevention Decals
The domain runs from 0 to 1 indicating the degree to which the concept Acceptance is Yes or No on visibility of crime prevention decals. The range of the psychometric scale is determined by the level of granularity and fine detail in the model. The degree of visibility is reflected in the psychometric scale underlying the Acceptance parameter. Many factors may affect the Acceptance of the visibility: color, size, position, and surrounding plants, etc. [5, pp.91]

3.3.4 Interval Variables

Interval variables are discrete and ordered but are measured in fixed and equal units. In this burglary pattern model, the discrete variables are Building Type, Number of Doors, and Number of Balconies. Newlands’ research indicated that certain restraints affect the applicability of the assumption. The cross-tabulation analysis of the Building Type is listed in Table 3.15. The building type attribute has three values: single family dwelling, multi-family dwelling, and apartment dwelling. There are only three apartment buildings in high crime blocks, five in all blocks; 258 single family buildings in high crime blocks, and 905 in all blocks. Even in the multi-family dwelling category, there are only 8 buildings in high crime blocks and 40 in all blocks [2]. Thus, questions raised here are: Are data sufficient to describe the burglary pattern in the multi-family dwellings and apartment dwellings?
- Is the burglary pattern model sufficient to compare the relations between single family dwellings, multi-family dwellings, and apartment dwellings?
- Is the burglary pattern model applicable to compare the patterns of multi-family and apartment dwellings with those of other areas?

<table>
<thead>
<tr>
<th>Attributes</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Family Dwelling</td>
<td>189</td>
<td>53</td>
<td>11</td>
<td>5</td>
<td>258</td>
</tr>
<tr>
<td>Multi-Family Dwelling</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Apartment Dwelling</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>192</td>
<td>57</td>
<td>11</td>
<td>9</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.15 Building Type in high crime blocks

<table>
<thead>
<tr>
<th>Attributes</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Family Dwelling</td>
<td>762</td>
<td>119</td>
<td>18</td>
<td>8</td>
<td>907</td>
</tr>
<tr>
<td>Multi-Family Dwelling</td>
<td>31</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>Apartment Dwelling</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>794</td>
<td>127</td>
<td>19</td>
<td>12</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.16 Building Type in all crime blocks

Example 3.4: Building Type

The cross-tabulation of physical design characteristics of buildings yielded significant results when building type was analyzed in high crime blocks. Of the total 258 single family buildings analyzed, 26.7% were found to have crime, while 73.3% of these houses had no crime. However, of the total 8 multi-family buildings, 62.5% were found to have experienced crime compared to 37.5% that did not. But building type was not found to be significant when the same analysis was performed on houses within all blocks while others remained the significance level. [2]
A single family dwelling usually consists of one family per building, but in reality it is quite common for two or three families or roommates to live in the same building and the house may be separated into several living units. Multi-family buildings may have multiple entrances and living units. In these cases, a triangle membership function intuitively describes the data set of the building type and gives a sense that the number of the living units and entrances has a relationship with the building type. Apartment dwellings possess 6 suites to 50 suites. High-rise buildings are treated in a similar fashion.
to apartment buildings. In *Northern Hastings* area, a typical apartment building holds 20–30 suites.

The five apartments in the area have a very high rate of burglary. This observation discloses the characteristic of the *break-in risk* for this type of building and strengthens the conclusion that *apartment buildings* are relatively vulnerable other building types. The local apartments are out-dated in security system and building structure. This makes the situation worsen in the *Northern Hastings* area. For the other building types, the *Multi-family dwelling* type has a higher *break-in ratio* than the *single family dwelling* type. The research in some other districts supports this result. Therefore, the burglary pattern model is applicable. However, both apartments and multi-family buildings are only a very small portion of the all-types buildings—five apartments and eight multi-family buildings. That means any statistical error can be enlarged and degrade the credibility. Thus we must be cautious when comparing the risk between the two building types. [1]

This variable implicitly affects the results of other variables such as *Number of Doors*, and *Number of Balconies*, because apartments and multi-family buildings have more doors and balconies. This suggests that the statistical data might not be mutually exclusive. Weighting might be required to eliminate the mutual inclusion effects when performing aggregation on these variables.

### 3.3.5 Nominal Variables

Distinct attributes are found in some variables that are not related to each other in any form. In this study, only fuzzy set *Fence Type* {None, Open Wooden, Closed Wooden, Chain Link, Metal, Cement-Rock} has this property. Consequently, any of these attributes has no association with any of the rest. Hence their fuzzy sets are mapped onto $\mathbb{R}$ (real line) in any arbitrary order. However, an appropriate and intuitive definition is still required.
Example 3.5: Fence Type

A weak relationship was found in the analysis of fence type (front) according to no crime and crime categories within high crime. But stronger relationships did exist in the analysis of fence type in the rear of homes. [2]

<table>
<thead>
<tr>
<th>Fence Type Rear</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>None</td>
<td>43</td>
<td>67.2</td>
<td>18</td>
<td>28.1</td>
</tr>
<tr>
<td>Open Wooden</td>
<td>69</td>
<td>83.1</td>
<td>12</td>
<td>14.5</td>
</tr>
<tr>
<td>Closed Wooden</td>
<td>25</td>
<td>64.1</td>
<td>11</td>
<td>28.2</td>
</tr>
<tr>
<td>Chain Link</td>
<td>38</td>
<td>70.4</td>
<td>12</td>
<td>22.2</td>
</tr>
<tr>
<td>Metal</td>
<td>5</td>
<td>62.5</td>
<td>1</td>
<td>12.5</td>
</tr>
<tr>
<td>Cement-Rock</td>
<td>8</td>
<td>72.7</td>
<td>2</td>
<td>18.2</td>
</tr>
</tbody>
</table>

Table 3.19 Fuzzy set definition for residential break-in ratio: Fence Type

Notice that each nominal attribute is counted in a number of buildings with the corresponding attribute. The total number of the building for each attribute is able to be mapped onto $R$(real line), so that the area of the membership function is proportional to number of buildings with the same type attribute. Therefore, the building number plays a key role to intuitively define a fuzzy set. The fuzzy sets of fence type are defined as follows:

<table>
<thead>
<tr>
<th>Fuzzy Sets</th>
<th>Membership Function</th>
<th>Break-in Ratio (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Start</td>
<td>Peak</td>
<td>End</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Triangle</td>
<td>0</td>
<td>32</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Open Wooden</td>
<td>Triangle</td>
<td>64</td>
<td>104</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>Closed Wooden</td>
<td>Triangle</td>
<td>147</td>
<td>166</td>
<td>186</td>
<td></td>
</tr>
<tr>
<td>Chain Link</td>
<td>Triangle</td>
<td>186</td>
<td>213</td>
<td>240</td>
<td></td>
</tr>
<tr>
<td>Metal</td>
<td>Triangle</td>
<td>240</td>
<td>243</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>Cement-Rock</td>
<td>Triangle</td>
<td>248</td>
<td>253</td>
<td>259</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.20 Fuzzy set definition for residential break-in ratio: Fence Type
Nominal quantities have no ordering or distance measure, so each attribute (fuzzy set value) on the real line is only an intuitive reflection of its portion of all buildings in the area. When fuzzy logic is applied to this design of no overlapping and no ordering, the result is as same

![Membership function: Fence Type](image)

Figure 3.10 Membership function: Fence Type

3.3.6 Psychometric Variables

Traditionally objective domains are relate to physical objects and units of measure such as meters, liters, dollars, ratios, item counts, and so forth. But psychometric domains are subjective and unitless. Many sociology problems are associated with concepts rather than things, whose fuzzy set domains reflect a subjective, if not quite arbitrary, scale against which the range of the fuzzy model is mapped. These domains are devised by the model builder, and had no unit to measure, and have no corresponding parallel in the real world. This kind of scale, since is comes from the designer's mind, is called a psychometric scale. [4; 5, pp90] In this research, all environmental factors are not psychometric but objective.

3.4 Summary

The basic concepts of fuzzy set theory were discussed in this chapter, namely membership functions and linguistic variables. Some basic concepts of data mining theory were evaluated to aid in the classification of the various burglary environmental cues, with regard to the nature of the variables. The modeling of the variables was conducted in MATLAB simulation environment. The tabular data in the examples were
mapped to input spaces and output spaces in MATLAB fuzzy toolbox. The parameters of
the membership functions were adjusted in a trial-and-error method to maximize the
approximation to the variables. Both theories and simulation techniques are the key
components that incorporate the sociological data in the form of artificial intelligence,
namely converting the burglary data into appropriate fuzzy sets.
4 Fuzzy Logic and Reasoning

4.1 Introduction

As fuzzy sets are an extension of classical crisp sets, fuzzy logic is an extension of classical two-valued logic. As there is a correspondence between classical crisp sets and classical logic, so is there a correspondence between fuzzy set theory and fuzzy logic. Furthermore, the degree of an element in a fuzzy set may correspond to the truth value of a proposition in fuzzy logic. In this chapter, fuzzy operators are first discussed, then hedges and reasoning, and then how much the degree to which the input of a physical cue contributes to the overall burglary offence risk.

4.2 Fuzzy Logic Operations

4.2.1 Basic Zadeh-Type Fuzzy Set Operations

Similar to the cardinality of a crisp set, which is defined as the number of elements in the crisp set, the *cardinality* of fuzzy set $A$ is the summation of the membership grades of all the elements of $x$ in $A$, denoted as $|A|$. When a fuzzy set $A$ has finite support, its cardinality can be defined as a fuzzy set. This *fuzzy cardinality* is denoted as $|A|_f$ and defined by Zadeh as

$$|A| = \sum_{x \in U} \mu_A(x) \quad (4.1)$$

$$|A|_f = \sum_{\alpha \in A} \frac{\alpha}{|A|_a} \quad (4.2)$$

With the basic notations and definitions for fuzzy sets, the basic fuzzy logic operations are:
- Complement: \( \bar{A} \)

\[
\mu_{\bar{A}}(x) \triangleq 1 - \mu_{A}(x) \quad \forall x \in U
\]  

(4.3)

- Intersection:

\[
\mu_{A \cap B}(x) \triangleq \min[\mu_A(x), \mu_B(x)] = \mu_A(x) \wedge \mu_B(x) \quad \forall x \in U
\]  

(4.4)

- Union:

\[
\mu_{A \cup B}(x) \triangleq \max[\mu_A(x), \mu_B(x)] = \mu_A(x) \vee \mu_B(x) \quad \forall x \in U
\]  

(4.5)

![Figure 4.1 Fuzzy Operations](image)

Figure 4.1 Fuzzy Operations
4.2.2 C++ Implementation

As fuzzy logic is an extension of classical two-valued logic, a correspondence exists between fuzzy set theory and fuzzy logic. For example, the union operator corresponds to the logic OR, intersection to AND, and complement to NOT. In this research, these operators are implemented as functions in C++ code. Since fuzzy sets are not crisply partitioned in the same sense as Boolean sets, these operators are applied at the truth membership level. [6]

- **Complement of fuzzy sets**

  A complement of set A contains the all elements that are not in set B. In fuzzy logic, it is produced by inverting the truth function along each point of the fuzzy set and registers the degree to which an element is complementary to the underlying fuzzy concept. [6]

Example 4.1: C++ Code for complement of fuzzy sets

```c
switch(iNotClass)
{
    case ZADEHNOT:
        for(i=0; i<VECMAX; i++)
            pFDBptr->FDBvector[i]=1-pEDBptr->FDBvector[i];
        break;
}
```

- **Intersection of fuzzy sets**

  The intersection of two sets contains the elements that are common to both sets, which is equivalent to the arithmetic or logical AND operation. In conventional fuzzy logic, the AND operator is supported by taking the minimum of the truth membership grades [6].

Example 4.2: C++ code for intersection of fuzzy sets

```c
float PE FzyAND(float fTruth1, float fTruth2)
{
    if(fTruth1<0||fTruth1>1) {return((float)-1);}
The union of two sets contains the elements that belong to both sets, which is equivalent to the arithmetic or logical OR operation. In conventional fuzzy logic, the OR operator is supported by taking the maximum of the truth membership grades [6].

Example 4.3: C++ code for union of fuzzy sets

```c++
float PE_FzyOR (float fTruth1, float fTruth2)
{
    if(fTruth1<0||fTruth1>1) {return((float)-1);} 
    if(fTruth2<0||fTruth2>1) {return((float)-2);} 
    return(max(fTruth1,fTruth2)); 
}
```

4.3 Fuzzy Reasoning

4.3.1 Linguistic Hedges

Hedges play the same role in fuzzy production rules that adjectives and adverbs play in English sentences, and they are processed in a similar manner. Hedges are operators, which are independently created to modify the meaning of the operators or, more generally, of a fuzzy set $A$ to create a new fuzzy set $h(A)$. Fuzzy operators are linguistic, so are the hedges. Hedge can be equated to English words. For example, the linguistic hedge "VERY" could be equated to $m(A(x))^2$, or "SOMewhat" to $m(A(x))^{0.5}$. The mechanics underlying a hedge operation are generally heuristic in nature. That is, the degree to which a fuzzy surface is transformed but the nature of the transformation is not based on a mathematical theory of fuzzy surface topology operations. Instead, they are associated with the perceived fit of the transformation and the psychological confidence.
of the resulting fuzzy region. [5, pp219] The following fuzzy operations are frequently used in defining a linguistic hedge [6]:

- Concentration: \( \text{CON}(A), \text{VERY}(A) \)

\[
\mu_{\text{CON}(A)}(u) = \left( \mu_A(u) \right)^2
\]

(4.6)

Example 4.4: C++ implementation of Concentration

```cpp
for(i=0;i<VECMAX;++i)
    pOutFDBptr->FDBvector[i]=(float)pow(pInFDBptr->FDBvector[i],TWO);
break;
```

- Dilation: \( \text{DIL}(A), \text{FAIRLY}(\text{MORE \ OR \ LESS}), \text{SOMETHING} \)

\[
\mu_{\text{DIL}(A)}(u) = \left( \mu_A(u) \right)^{\frac{1}{2}}
\]

(4.7)

Example 4.5: C++ Implementation of Dilation

```cpp
for(i=0;i<VECMAX;++i)
    pOutFDBptr->FDBvector[i]=(float)pow(pInFDBptr->FDBvector[i],POINTFIVE);
break;
```

- Intensification: \( \text{INT}(A), \text{INDEED} \)

\[
\mu_{\text{INT}(A)}(u) = \begin{cases} 
2 \left( \mu_A(u)^2 \right), & \mu_A(u) \in [0, 0.5] \\
1 - 2(1 - \mu_A(u))^2, & \text{otherwise}
\end{cases}
\]

(4.8)

Example 4.6: C++ Implementation of Intensification

```cpp
if(pInFDBptr->FDBvector[i] >= x)
{
    hv=pInFDBptr->FDBvector[i];
    hv=(float)2*pow(hv,TWO);
    if(hv>1.0) hv=1.0;
    pOutFDBptr->FDBvector[i]=(float)hv;
}
else
{
    hv=pInFDBptr->FDBvector[i];
    ```
hv=(float)1-(2*(pow(1-hv,TWO)));
if(hv<0) hv=0;
pOutFDBptr->FDBvector[i]=(float)hv;
}

4.3.2 Approximation Reasoning

*Linguistic approximation* is a procedure for determining a term from the term set of a linguistic variable such that the meaning of this term is closest to a given fuzzy set. The inferred results usually must be translated into meaningful terms (fuzzy sets) by using *linguistic approximation*, and a final conclusion is drawn from a set of premises \( \{p_1, p_2, \ldots, p_n\} \). To provide the capability of approximate reasoning, fuzzy logic allows the use of *fuzzy predicates*, *fuzzy predicate modifiers*, *fuzzy quantifiers*, and *fuzzy qualifiers* in the propositions.

- Fuzzy predicates: fuzzy predicates are Low, High, Mid.
- Fuzzy Predicate Modifiers: a variety of predicate modifiers are VERY, RATHER, MORE OR LESS, SLIGHTLY, NEAR.
- *Fuzzy quantifiers*: fuzzy logic allows the use of fuzzy quantifiers exemplified by Most, Many, Several, Few.
- Fuzzy qualifiers: True, False, Likely.

Moreover, a fuzzy logic proposition forms a rule expressed in the way of a natural language. Behind the linguistics, a fuzzy proposition consists of fuzzy operators and linguistic hedges. A hedge modifies the shape of the surface of a fuzzy set and causes a change in the membership truth function. Thus, a hedge transforms one fuzzy set into a new fuzzy set. The hedges either intensify the characteristic of a fuzzy set (*very, extremely*), or dilute the membership curve (*somewhat, rather, quite*). On the other hand, the hedges either form the complement of a set (*not*), or approximate a fuzzy region by converting a scalar to a fuzzy set (*about, near, close to, approximately*). For example, the *VERY* hedge intensifies the break-in ratio for the *High* and *Low* membership curves of a burglary offence fuzzy set.
By means of fuzzy operators and fuzzy linguistics, the burglary model rules can be made through an assertion about the relation between break-in ratio and the characteristics of residence, deterrents and barriers, and access and entry. For example, the relationship between the building age and the break-in ratio can be stated as:

If the building age is New and the Break-In-Frequency is Mid, then the break-in ratio is Very High.

4.3.3 Fuzzy Propositions

Fuzzy reasoning consists of an inference engine and a fuzzy rule base. The inference engine is the core of the fuzzy expert system in modeling human decision making within the conceptual framework of fuzzy logic and approximate reasoning. The Fuzzy rule base is a collection of fuzzy IF-THEN rules in which the preconditions and consequents contain linguistic variables. This collection of fuzzy rules is a bridge connecting the inputs and the outputs of the system. A proposition is a statement of the relationships between model variables and one or more fuzzy regions. A series of propositions is evaluated for its degree of truth, and all those that have some truth contribute to the final output state of the solution variable set. Fuzzy propositions are expressed in the form:
$X$ is $Y$ \hspace{1cm} (4.9)

where $X$ is a scalar from the domain, and $Y$ is a linguistic variable. The effect of evaluating a fuzzy proposition is a degree or grade of membership derived from the transfer function:

$$\mu_x \leftarrow (x \in Y)$$ \hspace{1cm} (4.10)

A conditional fuzzy proposition is analogous to a rule in conventional symbolic expert system, that is expressed as $IF-THEN$ rules. The collection of fuzzy rules contains the preconditions and consequents contained in fuzzy or linguistic variables. The general form is: [5, pp145]

$R_i$: If $(x$ is $A_i) \cdot \ldots \cdot (y$ is $B_i)$, then $z$ is $C_i$, $i=1, 2, \ldots , n$, \hspace{1cm} (4.11)

where $x$, $\ldots$, $y$, and $z$ are linguistic variables representing the process state variables. The symbol $[\cdot]$ represents operators $\text{and}$ or $\text{or}$. $A_i$, $B_i$, and $C_i$ are the linguistic values of the linguistic variables $x$, $\ldots$, $y$, and $z$ in the universes of discourse $U$, $\ldots$, $V$, and $W$, respectively. In the case of multiple antecedent propositions, the position of $z$ within $C_i$ is determined by the composite truth of the complete antecedent. For multi-variables or multi-criteria inference, weighing is relative important. Here all variables are equally weighed for simplicity. [5, pp145; 6, pp274]

The fuzzy burglary offence model consists of a series of fuzzy propositions, which established the relationships between a value in the underlying domain of physical characteristic and the burglary offence fuzzy space. For example, the fuzzy reasoning mechanism was explained and discussed for the $\text{Fence Height}$ variable of the fuzzy burglary offence model.
Example 4.7: Fence Height

When fence height was studied, significant relationships were found in the rear of homes, but not in the front. Even in the high crime blocks, fence height seemed to have a deterrent effect. There was a trend evident, suggesting that as fences became higher the more likely a home would be burglarized. The type of fence potentially hinders surveillance of a person’s backyard and thereby allows the burglar the cover he/she needs to commit his/her crime. As a result, a conflict may arise between design features needed to give privacy or beauty to urban environment and crime prevention features. [2]

<table>
<thead>
<tr>
<th>Fence Height Rear</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>Under 3’</td>
<td>62</td>
<td>79.5</td>
<td>14</td>
<td>17.9</td>
</tr>
<tr>
<td>3’ – 4’</td>
<td>62</td>
<td>74.7</td>
<td>14</td>
<td>16.9</td>
</tr>
<tr>
<td>4’ – 6’</td>
<td>15</td>
<td>62.5</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td>6’ +</td>
<td>3</td>
<td>42.9</td>
<td>2</td>
<td>28.6</td>
</tr>
</tbody>
</table>

Table 4.1 Fuzzy set definition for residential break-in ratio: Fence Height

Rules:

If Fence Height is 3’ AND Break-In-Freq is Low, then break-in ratio is Rather Mid;
If Fence Height is 4’-6’ AND Break-In-Freq is Mid, then break-in ratio is Sort of High;
If Fence Height is 6’+ AND Break-In-Freq is Low, then break-in ratio is High.
4.3.4 Monotonic Reasoning (Tsukamoto’s Method)

Monotonic reasoning is also called proportional reasoning, a basic fuzzy implication technique. This is a simplified method based on fuzzy reasoning in which the consequence $C_i$ (Equation 4.11) is required to be monotonic; that is Equation 4.11 was functionally represented by the fuzzy monotonic logic transfer function as follows:

$$z = f((x, Y), w)$$  \hspace{1cm} (4.12)

The composition and decomposition were eliminated from the process of fuzzy reasoning. The transfer function was bounded by two parameters, under a restricted set of circumstances. Therefore, the output value is estimated directly from a corresponding truth membership grade in the antecedent fuzzy regions.
Example 4.8: Basement Design

*Basement design* revealed an association in result that there was a trend suggesting that the higher the basement was above ground, the more likely the home would be broken into. The association was obvious but not significant. In both *high crime blocks* and *all blocks*, the difference of the *break-in ratio* was less than 3% among *Below Ground, Ground Level, and Above Ground*, but the *break-in ratio* was 4–9% higher in *None category* than in other categories. The explanation might lie in the fact that above ground basements potentially provided easy access for the burglar and the types of homes with these basements were usually older multi-family dwellings can expect to be burglarized more often. The trend certainly existed in all categories but not distinct among *Below Ground, Ground Level, and Above Ground*. [2]

If the trend exists in all categories but is not distinct in the attributes (Table 4.2 and 4.3), a fuzzy model cannot be accurately defined in the usual way (i.e., *Mamdani or Sugeno*). It is difficult to understand the exact association of the *break-in risk* in such a rough model (lack of details); therefore, it is difficult to predict the exact *break-in risk* value. However, the relationship is known between the input space of *residential attributes* and solution space of *break-in risk*, e.g., a *proportional monotonic* relationship, thus the trend is predictable.

<table>
<thead>
<tr>
<th>Basement Design</th>
<th>No Crime</th>
<th>Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>None</td>
<td>23</td>
<td>57.5</td>
</tr>
<tr>
<td>Below Ground</td>
<td>39</td>
<td>76.5</td>
</tr>
<tr>
<td>Ground Level</td>
<td>78</td>
<td>75.0</td>
</tr>
<tr>
<td>Above Ground</td>
<td>52</td>
<td>73.2</td>
</tr>
</tbody>
</table>

Table 4.2 Crime rate and basement design in high crime blocks
Table 4.3 Crime rate and basement design in all blocks

<table>
<thead>
<tr>
<th>Basement Design</th>
<th>No Crime</th>
<th>Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>None</td>
<td>116</td>
<td>78.9</td>
</tr>
<tr>
<td>Below Ground</td>
<td>134</td>
<td>83.8</td>
</tr>
<tr>
<td>Ground Level</td>
<td>305</td>
<td>83.3</td>
</tr>
<tr>
<td>Above Ground</td>
<td>238</td>
<td>87.2</td>
</tr>
</tbody>
</table>

Apparently in both categories (high crime blocks and all blocks), the results consistently indicated the relationship between the break-in ratio and the basement design. This can be considered as a simple risk access model. The underlying fuzzy vocabulary for this model consists of two fuzzy sets: *basement design* in terms of the basement elevation and *break-in ratio* in term of percentage. The model is based on the perceived relationship between basement elevation and its break-in risk. The rule is simplified and stated as follows:

*If Basement Design increases then break-in ratio increases.*

Assume that when *Basement Design* increases from *below ground, ground level, above ground*, to *none*, the *break-in risk* ranges from 23.5 to 42.5 or 16.3 to 21.1, consecutively in high crime blocks and all crime blocks. The exception is that the risk is 12.8 for the above ground in all crime blocks.

For the fuzzy set of *Basement Design*, the criteria can be set based on the position of basement ceiling as follows:

*below ground level* means the position of the ceiling is under 1 meter;
*ground level* means the position of the ceiling is under 2 meters;
*above ground level* means the position of the ceiling is under 3 meters;
*none* means the position of the ceiling of the first floor is above 3 meters.

Along an S-curve lie the different Basement Designs of below ground, ground level, above ground level and none (Figure 4.4).
By the same token, the break-in ratios are derived from the corresponding values of Basement Design types along an S-curve and the break-in ratios are 23.5%, 25.0%, 26.8%, and 42.5% on the solution space, respectively.

Figure 4.4 Fuzzy reasoning: Monotonic Reasoning in high crime blocks

For every element x in the domain of Y, its membership is \( \mu_Y [x] \) in the fuzzy region Y. Corresponding to \( \mu_Y [x] \), on the fuzzy region W, the value on the domain axis, z, is the solution to the implication function as follows:

\[
Z_w = f(\mu_Y [x], D_w)
\]  \hspace{1cm} (4.13)

The implication method is characterized by a lack of high-level orthogonality in the consequence fuzzy space. That is, while the antecedent fuzzy expression might be complex, the solution is not produced by any formal method of defuzzification, but by a direct slicing of the consequence fuzzy set at the antecedent’s truth level. The monotonicity of the implication function is associated with the nature of fuzzy domains, not the fuzzy set topology. When the values are selected from the domain in the Basement Design fuzzy set (input space), a corresponding or proportional value is selected from the consequent break-in ratio fuzzy set (solution space).[6, pp277-279]
The function initializes the input values and rule description.

```cpp
void CFuzzyResidenceRisk::InitializeModelData(void){
    RiskRules[1].dblTargetFactor=2.3;
    RiskRules[1].Description = "Basement Design";
    RiskRules[1].Rule = "If Below GND, then Rather Mid";
    RiskRules[1].UOM="Metres";
}
```

This function initializes the monotonic input and output fuzzy sets by certain domains.

```cpp
void CFuzzyResidenceRisk::CreateRiskSMC(void){
    // Factor: Basement Design
    //dblParams[0] = 0;
    //dblParams[1] = 3;
    // domain
dblDomain[0] = 0; dblDomain[1] = 6;

    RiskRules[1].lpFDBTargetFactor=FzyCreateSet("SMC.BasementDesign", GROWTH, dblDomain, dblParams, 0, &iStatus);
    strcpy(RiskRules[1].lpFDBTargetFactor->FDBdesc,"SMC.BasementDesign");
    MdlLinkFDB(RiskRules[1].lpFDBTargetFactor, lpPDBRisk, &iStatus);
    
    // BasementDesign: Risk
    //dblParams[0] = 0;
    //dblParams[1] = 23.5*10;
    //dblParams[2] = 42.5*10;
    // domain
dblDomain[0] = 0; dblDomain[1] = 42.5*10;

    RiskRules[1].lpFDBBreakInRate=FzyCreateSet("SMC.Risk2", GROWTH, dblDomain, dblParams, 0, &iStatus);
    strcpy(RiskRules[1].lpFDBBreakInRate->FDBdesc,"SMC.Risk2: a normalized risk index");
    MdlLinkFDB(RiskRules[1].lpFDBBreakInRate, lpPDBRisk, &iStatus);
}
```

The membership for the fuzzy set is a *S-curve* and shown in Figure 4.5.
This function in Metus dll correlates the input and the output fuzzy sets.

```c
#include "FDB.hpp"     // The Fuzzy Set descriptor
#include "sysdef.hpp"   // System constants and symbolics
/
double PE FzyMcnotonicLogic(
    FDB   *pFromFDBptr,
    FDB   *pToFDBptr,
    double dFromValue,
    float *fPremiseTruth,
    int   *ipStatus)
{
    int iIdxPos;
    float fPTruth;
    double dScalar;
```
The function processes the model. It calls the FzyMonotonicLogic(void) function and correlates the input value with an output value using scalable monotonic chaining method.

float CFuzzyResidenceRisk::ModelProcessSMC(void){
    RiskRules[2].dblNormRisk = FzyMonotonicLogic(RiskRules[2].lpFDBTargetFactor,
                                                  RiskRules[2].lpFDBBreakInRate,
                                                  RiskRules[2].dblTargetFactor,
                                                  &RiskRules[2].fPremiseTruth,
                                                  &iStatus);
    return RiskRules[2].fPremiseTruth;
}

![Diagram of Basement Design: Input and Output](image)

Figure 4.7 Basement Design: Input and Output

Output break-in ratio: 16.27% [0.293]
4.3.5 Complex Monotonic Reasoning

From Equation (4.11), a general form of monotonic reasoning, i.e., the antecedent fuzzy true function can be generated from an arbitrarily complex approximate expression following the general form:

If \((x \text{ is } Y) \cdot (k \text{ is } U) \cdot (s \text{ is } M)\) then \(z \text{ is } W\) \hspace{1cm} (4.14)

The operator \(\cdot\) represents either the AND or OR operation in any of the operator classes. Then equation (4.13) is converted into the following general form: [6, pp282]

\[
Z_w = f\left(\{ (x, Y), (k, U), (s, M) \}, W \right)
\]  \hspace{1cm} (4.15)

Or

\[
Z_w = f\left( \sum_{i=1}^{N} v_i, F_i \right), W \right) \hspace{1cm} (4.16)
\]

Where \(\sum\) is the general aggregation operator acting on the variable and fuzzy set tuples to produce the fuzzy predicate truth value. Geometrically, the aggregate truth of the predicates can be viewed as a point in a fuzzy region bounded by the composite fuzzy sets (see Figure 4.5). For example, both Building Age and Building Condition have their own contributions to the break-in ratio, and aggregation is necessary to determine the truth value for the break-in ratio. For simplicity, all variables are equally weighed here.

The rule is:

Rule1: If Building Age is Old AND Break-In-Freq is Low, the break-in ratio is Mid
Rule2: If Building Condition is Poor AND Break-In-Freq is Low, the break-in ratio is High

Conclusion: break-in ratio is Sort of Mid

The two-dimensional space bounded by the Building Age fuzzy set and the Building Condition fuzzy set gives a compensatory mean (average). The AND operator is used to combine the truth values from the both fuzzy sets. Thus, monotonic reasoning acts as a
proportional correlating function between two general fuzzy regions. An important restriction on monotonic reasoning is its requirement that the output for the model be a single fuzzy variable controlled by a single fuzzy rule (with an arbitrarily complex predicate).

![Figure 4.8 Two-dimensional Monotonic Reasoning inference space](image)

Monotonic reasoning provides a powerful method of linking the truth of fuzzy regions. However, it is less general and more restrictive than the conventional min-max or min-product rules of implication. This is the result of the following two limitations:

- The requirement that the output for the model be a single fuzzy variable controlled by a single fuzzy rule (with an arbitrarily complex predicate)
- The implication function between the two fuzzy regions is expressed as a correlated surface topology.

As the complexity of the predicate proposition increases, the degree to which monotonic reasoning will produce consistently valid results tend to decline. This is a consequence of basic complexity and information theories. [6, pp284]
4.4 Fuzzy Compositional Rules of Inference

The implication space generated by the general compositional rules of inference is derived from the aggregated and correlated fuzzy spaces produced by the interaction of several statements. All the propositions are run in parallel to create an output space that contains information from all the propositions. Each conditional proposition whose evaluated predicate truth is above the current $a$-cut threshold contributes to the shape of the fuzzy representation of the variable on the output solution. There are four types of fuzzy composition operators that can be used in the compositional rule of inference. These correspond to the four operations associated with the $t$-norms [6]:

- **Max-min operation**

The consequent fuzzy region is restricted to the minimum of the predicate truth.

$$\mu_{ch} [X_i] \leftarrow \min(\mu_{pi}, \mu_{ch} [X_i]) \quad (4.17)$$

**Example 4.9: C++ Implementation of min operation**

```c
void PE_FzyCorrMinimum(float fVector[], const float fPredTruth, int *ipStatus)
{
    int i;
    char *sPgmId="mtfzcor";

    *ipStatus=0;
    if(fPredTruth<0||fPredTruth>1)
        { *ipStatus=1;
          MtsSendError(26,sPgmId,MtsFormatFlt(fPredTruth,4));
          return;
        }
    for(i=0;i<VECMAX;i++) fVector[i]=min(fVector[i],fPredTruth);
    return;
}
```

The output fuzzy region is updated by taking the maximum of these minimized fuzzy sets.

$$\mu_{sfs} [X_i] \leftarrow \max (\mu_{sfs}, \mu_{cfs} [X_i]) \quad (4.18)$$
Example 4.10: C++ Implementation of max operation

```cpp
void PE FzyCorrMaximum(float fVector[], const float fPredTruth, int *ipStatus)
{
    int i;
    char *sPgmId="mtfzc";

    *ipStatus=0;
    if(fPredTruth<0||fPredTruth>l)
    {
        *ipStatus=1;
        MtsSendError(26,sPgmId,MtsFormatFlt(fPredTruth,4));
        return;
    }
    for(i=0; i<VECMAX; i++) fVector[i]=max(fVector[i],fPredTruth);
    return;
}
```

- Fuzzy Additive operation

The consequent fuzzy region is reduced by the maximum truth value of the predicate.

$$\mu_{sf}(x) \leftarrow (\mu_{sf}(x) + \mu_{fs}(x))$$  \hspace{1cm} (4.19)

Example 4.11: C++ Implementation of Fuzzy Additive operation

```cpp
case BOUNDEDMEAN:
    for(i=0; i<VECMAX; i++)
        pOutFDBptr->FDBvector[i]=
            min(1,(pOutFDBptr->FDBvector[i]+fMemVector[i])/2);
    break;
```

4.5 Methods of Decomposition and Defuzzification

As previously discussed, the following propositions, when evaluated, will correlate the consequence fuzzy sets Rather Mid and Sort of High, and High to produce a fuzzy set representing the solution variable break-in ratio:

If Fence Height is 3'- AND Break-In-Freq is Low, then break-in ratio is Rather Mid;
If Fence Height is 4'-6' AND Break-In-Freq is Mid, then break-in ratio is Sort of High;
If Fence Height is 6'+ AND Break-In-Freq is Low, then break-in ratio is High.
The evaluation of a proposition produces one fuzzy set associated with the solution variables of the fuzzy burglary offence model. To obtain the actual risk assessment scalar, a value must be found that best represents the content in the fuzzy set break-in ratio. This is a decomposition process, called defuzzification.

Defuzzification is a mapping from a space of expert decisions, defined over an output universe of discourse, into a space of non-fuzzy (crisp) expert decisions. The strategy is aimed at producing a non-fuzzy decision action that best represents the possibility distribution of an inferred decision action. Two commonly used techniques of defuzzification are the Composite Moments method and the Composite Maximum method. [6, pp303]

4.5.1 Composite Moments (Centroid)

The centroid or Center of Area (COA) is the mostly widely used strategy. COA generates the center of area (gravity) of the possibility distribution of a decision. In the case of a discrete universe, this method yields:

\[
Z_{\text{COA}} = \frac{\sum_{j=1}^{n} \mu_c(z_j)z_j}{\sum_{j=1}^{n} \mu_c(z_j)},
\]

(4.20)

where \( n \) is the number of quantization levels of the output, \( z_j \) is the amount of decision output at the quantization level \( j \), namely the \( j \)th domain value, and \( \mu_c(z_j) \) represents its membership value in the output fuzzy set \( C \). If the universe of discourse is continuous, then the COA strategy generates an output decision:

\[
Z_{\text{COA}} = \frac{\int \mu_c(z)zdz}{\int \mu_c(z)dz},
\]

(4.21)
Example 4.12: C++ Implementation of Centroid method

case CENTROID:
    fX=(float)0.0;
    fY=(float)0.0;
    for(i=0;i<VECMAX;i++)
    { 
        fX+=pFDBptr->FDBvector[i]*(float)i;
        fY+=pFDBptr->FDBvector[i];
    }
    if(fY==0)
    {
        *fpGrade=(float)0.0;
        dScalar=0.0;
        *ipStatus=99;
        return(dScalar);
    }
    iTVpos=(int)(fX/fY);
    if(iTVpos<0||iTVpos>VECMAX)
    {
        *fpGrade=(float)0.0;
        dScalar=0.0;
        *ipStatus=77;
        return(dScalar);
    }
    dScalar=FzyGetScalar(pFDBptr,iTVpos,ipStatus);
    *fpGrade=pFDBptr->FDBvector[iTVpos];
    CompleteDefuzz("CENTROID",pFDBptr,dScalar,*fpGrade);
    return(dScalar);

This decomposition technique, called centroid or composite moments defuzzification, finds the balance point of the solution fuzzy region — the center of gravity, and weights the mean of the fuzzy region. The most desirable advantages of centroid defuzzification lie in the following properties [6, pp307]:

Figure 4.9 Defuzzifying the fuzzy set of break-in ratio (COA)
Defuzzified values tend to move smoothly around the output fuzzy region/space, which results in smooth changes in the expected value;
- Simple calculation;
- Applicable to both fuzzy and singleton output set geometries.

4.5.2 Composite Maximum

Some closely related types of composite maximum techniques are the *Mean of Maximum (MOM)*, the *center of maximums*, and the *simple composite maximum*. The MOM strategy generates a decision action that represents the mean value of all local decision actions whose membership functions reach the maximum. In the case of a discrete universe, the decision action may be expressed as:

\[
Z_{\text{MOM}}^{*} = \frac{\sum_{j=1}^{m} \frac{z_j}{m}}{m},
\]

(4.22)

where \(z_j\) is the support value at which the membership function reaches the maximum value \(\mu_c(z_j)\) and \(m\) is the number of such support values [6].

![Figure 4.10 Defuzzifying the fuzzy set of break-in ratio (MOM)](image-url)
Example 4.13: C++ Implementation of MOM method

```c
case AVGMAXIMUM:
    fTVal=(float)0;
    dSumofScalars=0;
    n=0;
    for(i=0;i<VECMAX;i++)
    {
        if(pFDBptr->FDBvector[i]>0)
        {
            n++;
            fTVal=fTVal+pFDBptr->FDBvector[i];
            dScalar=FzyGetScalar(pFDBptr,i,ipStatus);
            dSumofScalars=dSumofScalars+dScalar;
        }
    }
    if(n==0)
    {
        *ipStatus=3;
        return((double)0);
    }
    *fpGrade=(fTVal/n);
    dScalar=dSumofScalars/n;
    CompleteDefuzz("AVGMAXIMUM",pFDBptr,dScalar,*fpGrade);
    return(dScalar);
```

The MOM is also called the average maximum. It finds the mean maximum value of the fuzzy region. An average value often produces a better approximation for a composite maximum approach in cases where the plateaus tend to be clipped at the left or right edges. [6, pp315]

4.5.3 Singleton Geometry Representations

![Singleton geometry models vs. region-based models: membership functions](image)

Figure 4.11 Singleton geometry models vs. region-based models: membership functions
The COA method is used mainly for discrete or continuous values of a fuzzy set. The variables of sociology are usually *dichotomous* or *interval* or *Nominal*. Actually, an interval value is not sufficiently discrete, because a discrete data set is usually discretized from continuous or analogue values. The COA defuzzification method, when slightly simplified, can be represented in a singleton geometric output space. The membership functions of a fuzzy set are replaced by single vertical points.

\[ z_{\text{SingletonCOA}} = \frac{\sum_{j=1}^{n} d_j \mu_C(z_j) z_j}{\sum_{j=1}^{n} \mu_C(z_j)}, \quad (4.23) \]

Where \( d_j \) is the domain value at each singleton. The logic processing in a singleton fuzzy model is shown in Figure 4.8, and numerical integration of the entire fuzzy set surface is eliminated. Instead, the domain value at each singleton is multiplied by the height (true membership value) of the singleton. Generally, it is equivalent to the normal COA. [6, pp328]

![Figure 4.12 Singleton geometry models vs. region-based models: Proposition, (Composition) and Decomposition](image-url)
Example 4.14: Fence Type (Triangular Membership Function):

A weak relationship was found in the analysis of fence type (front) according to no crime and crime categories within high crime. But stronger relationships did exist in the analysis of fence type in the rear of homes. Since no real data for multiplicative rules, the rules are hypothetically generated. [2]

<table>
<thead>
<tr>
<th>Fence Type Front</th>
<th>No Crime</th>
<th></th>
<th>Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
</tr>
<tr>
<td>None</td>
<td>144</td>
<td>70.2</td>
<td>61</td>
</tr>
<tr>
<td>Open Wooden</td>
<td>16</td>
<td>88.9</td>
<td>2</td>
</tr>
<tr>
<td>Closed Wooden</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Chain Link</td>
<td>13</td>
<td>68.4</td>
<td>6</td>
</tr>
<tr>
<td>Metal</td>
<td>4</td>
<td>66.7</td>
<td>2</td>
</tr>
<tr>
<td>Cement-Rock</td>
<td>15</td>
<td>88.2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.4 High crime blocks

Rules (Simple):

1. if fence type at front is cement-rock, open wooden, none, chain link, metal, closed wooden consecutively, then break-in risk is increased to high;

Or

1. if fence type at front is Open Wooden, then break-in ratio is increased to mid;
2. if fence type at front is Close Wooden, then break-in ratio is increased to very high;
3. if fence type at front is Chain Link, then break-in ratio is increased to high;
4. if fence type at front is Metal, then break-in ratio is increased to high;
5. if fence type at front is Cement-Rock, then break-in ratio is increased to mid;
6. if fence type at front is None, then break-in ratio is increased to high;

Figure 4.13 Singleton geometry decomposition: fence type at front (high crime blocks)
Fence types consist of Open Wooden, Closed Wooden, Chain Link, Metal, Cement-Rock and None. Each fence type is independent of others. Because of the nominal nature, no membership function is applicable to the fence type except a unit function, and no overlaps exist between any two curve shapes. The membership function for the input fuzzy set is meaningless and not necessary, though the method of implication remains generally the same as that using fuzzy sets except the transfer is direct. For example, each fence type input is directly transferred to the solution space. The solution space is a singleton geometric output space. Therefore, the singleton geometry representations are used for output defuzzification space. However, the direct method of implication does provide a convenient way to translate the rules to output space—usually an ill-defined space.

<table>
<thead>
<tr>
<th>Fence Type Front</th>
<th></th>
<th>No Crime</th>
<th>Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
</tr>
<tr>
<td>None</td>
<td>144</td>
<td>70.2</td>
<td>61</td>
</tr>
<tr>
<td>Open Wooden</td>
<td>16</td>
<td>88.9</td>
<td>2</td>
</tr>
<tr>
<td>Closed Wooden</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Chain Link</td>
<td>13</td>
<td>68.4</td>
<td>6</td>
</tr>
<tr>
<td>Metal</td>
<td>4</td>
<td>66.7</td>
<td>2</td>
</tr>
<tr>
<td>Cement-Rock</td>
<td>15</td>
<td>88.2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.5 High crime blocks

Rules (Simple):

1. if fence type at front is cement-rock, open wooden, none, chain link, metal, closed wooden consecutively, then break-in risk is increased to high;

Or

1. if fence type at front is Open Wooden, then break-in ratio is increased to mid;
2. if fence type at front is Close Wooden, then break-in ratio is increased to very high;
3. if fence type at front is Chain Link, then break-in ratio is increased to high;
4. if fence type at front is Metal, then break-in ratio is increased to high;
5. if fence type at front is Cement-Rock, then break-in ratio is increased to high;
6. if fence type at front is None, then break-in ratio is increased to mid.
Table 4.6 High Crime Blocks

Rules (Simple):

1. if fence type at rear is open wooden, cement-rock, chain link, none, closed wooden, metal consecutively, then break-in risk is increased to high;

Or

1. if fence type at rear is Open Wooden, then break-in ratio is increased to very mid;
2. if fence type at rear is Close Wooden, then break-in ratio is increased to high;
3. if fence type at rear is Chain Link, then break-in ratio is increased to high;
4. if fence type at rear is Metal, then break-in ratio is increased to high;
5. if fence type at rear is Cement-Rock, then break-in ratio is increased to high;
6. if fence type at rear is None, then break-in ratio is increased to high;

Table 4.7 All Blocks

Rules (Simple):

1. if fence type at rear is open wooden, chain link, cement-rock metal, none, closed wooden, consecutively, then break-in risk is increased to high;
1. if fence type at rear is *Open Wooden*, then break-in ratio is increased to *mid*;
2. if fence type at rear is *Close Wooden*, then break-in ratio is increased to *sort of high*;
3. if fence type at rear is *Chain Link*, then break-in ratio is increased to *very mid*;
4. if fence type at rear is *Metal*, then break-in ratio is increased to *very mid*;
5. if fence type at rear is *Cement-Rock*, then break-in ratio is increased to *very mid*;
6. if fence type at rear is *None*, then break-in ratio is increased to *sort of high*.

![Truth Value Graph](image)

Figure 4.14 Singleton geometry decomposition: fence type at rear (high crime/all blocks)

By eliminating the step to build the input spaces, the rules are *directly* transferred to the output space—the singleton geometry spaces (Figure 4.14). For the attribute of fence types at front, the break-in ratios are resided with nearly-full truth value around mid, high and 100%. For those of fence types at rear (all blocks), the break-in ratios are resided with nearly-full truth value around mid, very mid, sort of high.

One apparent problem is that, for those of fence types at front, only one building is *Closed Wooden* contributes a large portion of the decomposition. Weighing is necessary to reduce the effect which is a result of the size of the overlapping areas. When the effect is insignificant, this step can be skipped. For example, for the attribute of fence types at rear, the value of *metal* type lies between the *Closed Wooden* and *none* types. Both *Closed Wooden* and *none* have sufficient samples and nearly-full truth value, thus their implications are confident and the overlapping areas eliminate the less-confident effect from *metal* type.
To summarize, the technique of the singleton geometry decomposition is suitable to assess the offence risk pattern the nominal residential attributes in a region or a residence. Usually the input characteristic is difficult to understand fully, but this method saves time and efforts to build the input spaces and can be utilized for a quick pre-assessment for burglary investigations.

4.6 Summary

The composite maximum and moments represent the most widely used methods of defuzzification. Particularly, the moments or the centroid method is preferred in a large majority of fuzzy models because it tends to assimilate all the input information in the output fuzzy set and selects a value supported by the knowledge accumulated from each executed proposition. In nearly all burglary offence cases, these methods will provide the best overall expected value. [6, pp314]
5 Possibility and Fuzzy Arithmetic

Possibility theory is associated with fuzzy set theory by defining a possibility distribution as a fuzzy restriction that acts as an elastic constraint on values that may be assigned to a fuzzy variable. This chapter continues to explore the relationship between possibility and fuzzy set, fuzzy arithmetic, and the ordering/ranking problems. Fuzzy arithmetic is a basic tool for computations of the sociologically imprecise or uncertain data.

5.1 Fuzzy Number

The concept of fuzzy number inherits the property of fuzzy set and crisp set. The evolution can be seen in Figure 5.1. The concept of fuzzy number is important to the possibility/probability theory and it represents the possibility/probability values. Fuzzy arithmetic manipulates the fuzzy numbers and measures the different possibilities/probabilities that fuzzy numbers represent.

![Figure 5.1 Evolution of Fuzzy Number](https://example.com/figure5.1.png)

Lin and Lee defined fuzzy number as “A convex, normalized fuzzy set, defined on the real line \( \mathbb{R} \), whose membership function is piecewise continuous or, equivalently, each \( \alpha \)-cut is a closed interval, is called a fuzzy number.” [6, pp17] Typical fuzzy numbers are the triangular function \( T(x; c,a,b) \), trapezoidal function \( \text{Tr.F.N} \), \( S(x;a,b) \), \( c(x;a,b) \), and Gaussian function, which are denoted as \( T(x; c,a,b) \), \( S(x;a,b) \), \( c(x;a,b) \) and \( G(x; \kappa, \gamma) \),
respectively. A fuzzy number $A$ in $\mathbb{R}$ (real line) is one of these fuzzy numbers, if its membership function $\mu_a(x) : \mathbb{R} \rightarrow [0,1]$ is defined by

$$
\text{Triangular } \mu_A(x) = \begin{cases} 
\frac{(x-c)}{(a-c)} & c \leq x \leq a, \\
\frac{(x-b)}{(a-b)} & a \leq x \leq b, \\
0 & \text{otherwise}
\end{cases}
$$

(5.1)

Figure 5.2 Triangular Fuzzy Number (TFN)

$$
S(x; a, b) = \begin{cases} 
0 & x < a \\
2 \left( \frac{x-a}{b-a} \right)^2 & a \leq x < \frac{a+b}{2} \\
1 - 2 \left( \frac{x-b}{b-a} \right)^2 & \frac{a+b}{2} \leq x < b \\
1 & x \geq b
\end{cases}
$$

(5.2)

$$
\pi(x; a, b) = \begin{cases} 
S(x; b-a, b) & x < b \\
1 - S(x; b, b+a) & x \geq b
\end{cases}
$$

(5.3)

$$
G(x, \kappa, \gamma) = e^{-\kappa(y-x)^2} \quad 0 < \kappa < \infty
$$

(5.4)

The triangular function makes it intuitively easy for decision-makers to perform an evaluation. By default, the triangular function is considered, but others are chosen to describe different variable types. The triangular function is denoted as $A=T(x;c,a,b)$ where, $c<a<b$. The parameter $a$ gives the maximal grade of $\mu_A(a)=1$. The triangular fuzzy number is used as a normal and convex fuzzy set of $\mathbb{R}$, (i.e., $\mu_A(a)=\Pi$), and it is the most possible value of the evaluation data. The lower and upper bounds of the domain are $c$ and $b$, which reflect the fuzziness of the data. [5, pp17; 14, 26]
5.2 Fuzzy Arithmetic

Similar to those from a scientific instrument or an engineering device, the values or data from the social sciences are usually imprecise or uncertain and are possibly located inside a closed interval on a real line \( R \); that is, this uncertain value is inside an interval of confidence of \( R, x \in [a_1, a_2] \), where \( a_1 \leq a_2 \). Symbolically, a fuzzy number can be denoted by \( A = [c, a, b] \), where a fuzzy number with a triangular membership function has the maximal grade of \( \mu_A(a) = 1 \), where it is the most possible value of the evaluation data. The parameters \( c \) and \( b \) are the lower and upper bounds of the area of the evaluation data, indicating the fuzziness of the data. The operations and computations on fuzzy numbers are called fuzzy arithmetic, which is one of the major components of possibility theory. Fuzzy arithmetic is a basic tool in computing with fuzzy quantifiers in approximate reasoning. Zadeh (1965) introduced the mathematical operations on intervals of confidence. These operators are addition (+), subtraction (-), multiplication (\( \cdot \)), and division (:\( \div \)), which can be defined on triangular fuzzy number by the extension principle:

Let \( A = [c_1, a_1, b_1] \) and \( B = [c_2, a_2, b_2] \) be two intervals of confidence in \( R \). If \( x \in [c_1, b_1] \), where \( c_1 \leq b_1 \), \( y \in [c_2, b_2] \) where \( c_2 \leq b_2 \), then we write:

- Addition (+), subtraction (-):
  \[
  A(+)B = [c_1, a_1, b_1] (+) [c_2, a_2, b_2] = [c_1 + c_2, a_1 + a_2, b_1 + b_2]
  \] (5.5)
  \[
  A(-)B = [c_1, a_1, b_1] (-) [c_2, a_2, b_2] = [c_1 - c_2, a_1 - a_2, b_1 - c_2]
  \] (5.6)
- Multiplication (\( \cdot \)), division (\( : \)):
  \[
  A(\cdot)B = [c_1, a_1, b_1] (\cdot) [c_2, a_2, b_2] = [c_1 \cdot c_2, a_1 \cdot a_2, b_1 \cdot b_2], c_1, c_2 > 0
  \] (5.7)
  \[
  A(:)B = [c_1, a_1, b_1] (:) [c_2, a_2, b_2] = [c_1/b_2, a_1/a_2, b_1/c_2], c_1, c_2 > 0
  \] (5.8)

5.3 Ordering/Ranking of Fuzzy Numbers

5.3.1 Possibility measure on Ordering of Fuzzy Numbers

A fuzzy number \( A \) is associated with two fuzzy sets, denoted as \( A_p \) and \( A_n \). The set of numbers that are possibly greater than or equal to \( A \) is denoted as \( A_p \) and is defined as:
\[ \mu_{A_p}(w) = \Pi_A((\infty, w)) = \sup_{u \leq w} \mu_A(u) \quad (5.9) \]

Similarly, the set of numbers that are necessarily greater than \( A \) is denoted as \( A_n \) and is defined as:

\[ \mu_{A_n}(w) = N_A((\infty, w)) = \inf_{u \leq w} (1 - \mu_A(u)) \quad (5.10) \]

where \( \Pi_A \) and \( N_A \) are the possibility and necessity measures defined as:

\[ \pi : x \to [0,1] \quad (5.11) \]

\[ \Pi(A) = \max_{x \in A} \pi(x), \quad \forall A \in B \quad (5.12) \]

\[ N(A) = (1 - \Pi(\overline{A})) = 1 - \max_{x \in A} \pi(x) = \min_{x \in A} (1 - \pi(x)) = \min_{x \in A} (1 - \pi(x)), \quad \forall A \in B \quad (5.13) \]

\( \mu_{A_p} \) and \( \mu_{A_n} \) can be viewed as the upper and lower possibility distribution functions of \( A \), respectively. Then the two fuzzy numbers \( A \) and \( B \) can be compared to see which has a greater possibility.

The methods of the optimistic comparison and the pessimistic comparison are widely used to compare \( \mu_M(A) \) and \( \mu_G(A) \). Suppose \( X, Y \) are variables whose domains are constrained by \( \mu_A \) and \( \mu_B \).

- Optimistic: the possibility that the largest value \( X \) can take is at least equal to the smallest value that \( Y \) can take.
  \[ \Pi_A(B_N) = \sup_u \min(\mu_A(u), \sup_{v \geq u} \mu_B(v)) \quad (5.14) \]

- Pessimistic: the necessity that the smallest value \( X \) can take is greater than the largest value that \( Y \) can take.
  \[ N_A(B_N) = 1 - \sup_{u \leq v} \min(\mu_A(u), \inf_{v \geq u} [1 - \mu_B(v)]) \quad (5.15) \]
When the membership function of the fuzzy number $A$ is a linear function, for example, a triangular membership function, then the next step is to calculate the $\Pi_A$ and $N_A$ by fuzzy arithmetic. Kaufmann and Gupta introduced a set of fuzzy arithmetic models to manipulate the fuzzy number for approximation and ranking. Their research made it possible to estimate the possibility and the necessity, and further to evaluate the estimation analysis. [14]

5.3.2 Linear Ordering of Fuzzy Numbers

Kaufmann and Gupta [14, pp37] set three criteria to compare two fuzzy numbers or to rank a set of $n$-fuzzy numbers. They studied the methods for computing the distance between two fuzzy numbers and for obtaining a dissemblance relation of $n$-fuzzy numbers. If the first criterion does not result in a unique linear order, then the second and the third criteria should be used.

- First criterion: the removal
  In case of a triangular fuzzy number (T.F.N.), an ordinary representation is
  \[ A = \frac{c + 2a + b}{4}, \text{where, } A = (c, a, b), \text{ a T.F.N.} \]  
  (5.9)

- Second criterion: mode
  In each class of fuzzy numbers one should look for the mode, and these modes will generate sub-classes.

- Third criterion: divergence
  In the divergence around the mode, we obtain sub-sub-classes, and this criterion may be sufficient to obtain the final linear ordering of fuzzy numbers.

Kaufmann and Gupta's criteria are simple and easy to implement. The method can be utilized for a quick pre-assessment of burglary offences. In Chapter 6, the ranking criteria are fully described and illustrated in case studies.
5.3.3 Approximation of Fuzzy Number Ranking

Let $f$ be the function having real values in $X$ and the highest and the lowest value of $f$ are $\text{sup}(f)$ and $\text{inf}(f)$ respectively (Figure 5.3). At this time, the maximizing set $M=\{(x, \mu_M(x)) \mid x \in A\}$ is defined as a fuzzy set:

$$\mu_M(x) = \frac{f(x) - \text{inf}(f)}{\text{sup}(f) - \text{inf}(f)}, \forall x \in A$$ (5.16)

That is, the maximizing set $M$ is a fuzzy set and is defined by the possibility of $x$ to make the maximum value $\text{sup}(f)$. The possibility of $x$ to be in the range of $M$ is defined from the relative normalized position in the interval $[\text{inf}(f), \text{sup}(f)]$. Here the interval $[\text{inf}(f), \text{sup}(f)]$ denotes the possible range of $f(x)$ to have some values. Similarly, the minimizing set $G=\{(x, \mu_G(x)) \mid x \in A\}$ is defined as a fuzzy set:

$$\mu_G(x) = \frac{f(x) - \text{sup}(f)}{\text{inf}(f) - \text{sup}(f)}, \forall x \in A$$ (5.17)

Moreover, Kaufmann and Gupta introduced a right side variable and a left side variable. These variables can be further transformed into a right utility and a left utility [15]. Suppose for $n$ number of fuzzy numbers, a right utility value of each fuzzy number $A_i$ is defined as follows:
Similarly, a left utility value of each fuzzy number $A_i$ is defined as follows

$$U_L(i) = \sup_x (\mu_{A_i}(x) \land \mu_{A_i}(x)), i = 1,2,..., n$$  \hspace{1cm} (5.18)

Campos and Gonzalez (1989) introduced an evaluation function on the set of real intervals, in which a parameter, called degree of risk, reflects the risk-bearing attitude (tendency or aversion) of a decision maker [16]. Then a total utility of each fuzzy number $A_i$ is defined as:

$$U_T(i) = \lambda U_M(i) + (1 - \lambda) U_G(i), \text{ where } \lambda \in [0,1], i = 1,2,..., n,$$  \hspace{1cm} (5.19)

The function $U_T(i)$ is a convex combination of the interval extremes, where:

$$A = [\mu_M, \mu_G], \mu_M, \mu_G \in R, \mu_M > \mu_G,$$ and the coefficient $\lambda$, a constant, is called degree of risk that a decision maker is bearing. If $\lambda > 0.5$, it implies a risk tendency of the decision maker; if $\lambda < 0.5$, it implies an aversion of the decision maker; if $\lambda = 0.5$, the attitude is neutral. In this study, we arbitrarily choose neutral for weighting the attitude of a decision-maker.

Specifically, if $A_i$ is a fuzzy number with the triangular membership function, then these utilities can be approximated and rewritten by:

$$U_M(i) = \frac{[c_i - \inf(x)]}{[\sup(x) - \inf(x)] - (b_i - c_i)}, \text{ where } A_i = (a_i, b_i, c_i), i = 1,2,..., n$$  \hspace{1cm} (5.20)

$$U_G(i) = \frac{[\sup(x) - a_i]}{[\sup(x) - \inf(x)] + (b_i - a_i)}, \text{ where } A_i = (a_i, b_i, c_i), i = 1,2,..., n$$  \hspace{1cm} (5.21)
where \( A_i = (a_i, b_i, c_i), i = 1, 2, \ldots, n \)

A decision maker can compare a pair of fuzzy numbers \( A_i \) and \( A_j \) according to following criteria:

- If \( U_T(i) = U_T(j) \), then \( A_i = A_j \),
- If \( U_T(i) > U_T(j) \), then \( A_i > A_j \),
- If \( U_T(i) < U_T(j) \), then \( A_i < A_j \).

By comparing the utility \( U_T \), while considering degrees of risk, the decision maker can evaluate the quality of his or her decisions. Further, we introduce a Quality Index to measure the quality of a decision:

\[
Q = \frac{U_T(x) - U_T(x)'}{U_T(x)} \%
\]  
(5.23)

The following criteria are set to qualify a decision:

- If \( Q \geq 50\% \), then the quality of a decision is Good;
- If \( 50\% > Q \geq 0\% \), then the quality of a decision is Fair;
- If \( Q < 0\% \), then the quality of a decision is Poor.
This method can solve sophisticated ranking problems and utilize the tendency of the decision maker. The full illustration appears in Chapter 6.
6 Case Studies

An investigation organization is often faced with a wide spectrum of environmental factors: residential characteristics, barriers and deterrents, access and security. An investigator must decide how to allocate limited financial, staffing, and equipment resources to investigation in order to satisfy jurisdictional, ethical, social, administrative, and political constraints. The techniques of fuzzy logic and Markov chains can be employed into different stages of an investigation to accurately measure the required resource. In this chapter, the applications are studied on a case-basis and techniques are illustrated in fuzzy logic models and Markov chain models.

6.1 Design Methodology

Over the previous discussions in Chapter 5, principals for designing a decision support system are summarized as follows:

- Define input and output variables
- Decide the fuzzy partition of the input and output spaces
- Choose the membership functions for the input and output linguistic variables
- Decide the types and the derivation of fuzzy decision rules
- Design the inference mechanism: a fuzzy implication and a compositional operator, etc.
- Choose a defuzzification operator. [5, pp159]

When different membership functions or defuzzification methods are chosen, the surfaces differ. More or less, these factors affect the expected values on the output space. The possible output surfaces of the membership functions are listed in Figure 6.1.

Mizumoto compared approximately thirty defuzzification methods [21]. Two of the more common techniques are the centroid and maximum methods. In the centroid method, the crisp value of the output variable is computed by finding the variable value of the center of gravity of the membership function for the fuzzy value. In the maximum method, one of the variable values at which the fuzzy subset has its maximum truth value is chosen as
the crisp value for the output variable. There are several variations of the maximum method that differ only in what they do when there is more than one variable value at which this maximum truth value occurs. One of these, the MOM method, returns the average of the variable values at which the maximum truth value occurs. Figure 6.2 indicates the differences in decomposition methods in the COA method, the MOM method, the monotonic method, and the singleton geometry method.

![Figure 6.1 Surface Comparison](image)

**Figure 6.1 Surface Comparison**

a) COA and MOM methods
In this chapter, we explore the application of fuzzy set theory to various decision-support analyses of burglary offences. The case studies start with the methods of fuzzy logic to analyze the patterns of the environmental factors then move on to examples carrying complex fuzzy rules, and finally discuss Markov chains for complex decision-support models. The analysis capability of fuzzy logic and Markov chain model is demonstrated and tested.
6.2 Case Study: Environmental Factor Analysis

In this section, basic fuzzy logic models are applied to analyze the characteristics of the environmental factors and applicable to understand potential risks of cues and signals in activity node, pathways, and edge. These models are Scalable Monotonic Chaining, single attribute analysis, and multi-attribute analysis.

6.2.1 Hypothesis and Assumptions

Awareness Space Model consists of three elements: activity node, pathways, and edges. These elements include the premises such as residences, shops, street blocks, community, parks, schools, roads, bridges, train stations, airports, rivers, and forests. The cues and signals of these structures have influence on target selection of offenders. In this case study, we choose the residences in Northern Hastings as a example of these elements. Newlands described the environmental signals and cues of these residences in three environmental factors: Residential Characteristics, Barriers and Deterrents, and Access and Security. In this case study, we assume that both the environmental criminology theories (Routine Activities, Rational Choice, and Crime Pattern Theory) and the models (Offender Behavior Cycle, Offender Decision Model, Target Selection Model, and Awareness Space Model) are applicable to the residences, then discuss fuzzy logic modeling, methodology, and application in assessing risks of the residences by these factors.

6.2.2 Scalable Monotonic Chaining

Scalable monotonic chaining does not create and then defuzzify a solution fuzzy set. Instead, it manipulates the assessment by mapping the break-in ratio in individual rules to an intermediate break-in measuring the fuzzy set break-in ratio. The result of the mapping is a scalar value from the domain of the break-in ratio indicating the degree of risk for the particular attribute input. To convert the break-in ratio to break-in risk, we need to accumulate, as much possible, the attributes of a residence, because all factors contribute to the final break-in risk on output solution space, and the cumulative effect of each factor should influence the assessment of the break-in risk if each factor has a
significant truth. It is also true among the residential characteristic factors, that to determine a statistically significant factor is sometimes very difficult. Therefore, the relationship can be expressed as

Figure 6.3 Scalable Monotonic Chaining Scheme: Break-in Risk Assessment of Residential characteristics

\[ Z_{SMC} = k \times \sum_{j=1}^{m} z_j \]  (6.1)

Rules:
1. if building type is multi-family dwelling, then break-in risk is increased to high;
2. if basement design is above ground, then break-in risk is increased to high;
3. if building age is new, then break-in risk is increased to high;
4. if # of balconies is two to three, then break-in risk is increased to high;
5. if building location is near street corner, then break-in risk is increased to high;
6. if rear lane presence is invisible, then break-in risk is increased to high.

The fuzzy rules are suggestive, not definitive. The rules accumulate evidence for or against the ultimate value of a solution variable. The degree to which the predicate of a rule is true (or false) indicates the degree to which the solution variable takes the shape of another fuzzy region. For example, the rule says "to the degree that basement design is
representative of the concept above ground, make the fuzzy solution region for break-in risk look more like the shape of the fuzzy set increased to high’’.

The membership value of $\mu = 0.293$ for basement design represents the compatibility between the data value and the concept associated with the above ground fuzzy set (Table 6.1). The value can be correlated with the value of increased to high, and can update the break-in ratio solution variable (Figure 6.3).

<table>
<thead>
<tr>
<th>Residential Characteristics</th>
<th>Input Domain</th>
<th>Output Domain</th>
<th>Scalar Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Input Factor Value</td>
</tr>
<tr>
<td>Building Type (#)</td>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Basement Design (m)</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Building Age (year)</td>
<td>6</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td># of Balconies (#)</td>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Building Location (m)</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Rear Lane Presence(%)</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total Risk</td>
<td></td>
<td></td>
<td>Risk=Sum(break-in ratio)*10</td>
</tr>
</tbody>
</table>

Table 6.1 Crime rate and basement design in all blocks

Multiple rules fire simultaneously, each adding to the shape of break-in ratio. In the end, when all the rules have fired, the shape of the break-in ratio fuzzy set on the output space indicates the preponderance of evidence.

6.2.2.1 Discussion

Some of the highlights of the monotonic chaining model applied to an assessment of break in risk are:

- The assessed risk associated with each factor is based on an associated fuzzy region and its mapping to a risk measuring fuzzy set;
- Each fuzzy set can have different topologies (i.e., different membership functions), and the method allows for a highly nonlinear relationship on a factor-by-factor basis;
• For instances with large number of rules that map a consequent to fuzzy solution set, and many of these rules are true and the solution effect is arithmetically cumulative rather than morphological, and chained monotonic scaling reflects the effect.;

• Avoid rule saturation. In the min-max inference technique, the solution fuzzy set is updated for conditional if-then rules by taking the maximum of the consequent fuzzy set and the solution fuzzy set. This means that after a few rules, the “high water mark” in the solution set will quickly move toward [1.0]. This dictates that rules whose truths are less than the current truth level of the solution fuzzy region will not contribute to the solution. This is not the case in scalable monotonic chaining because of its cumulative nature;

• Although a simple additive accumulation is adopted in this study, the assessment risk value can be weighted by its grade of membership in the consequent fuzzy set or by finding this number in another fuzzy set that calibrates this number to a final risk assessment.

The last attribute also introduces a high degree of nonlinearity into the model. The fuzzy model can handle any number of rules and still maintain the important relationships between the underlying rules and the risk assessment.

6.2.3 Single Attribute Analysis

Building Age is one of the attributes in the environmental factor Residential Characteristic. In this section, the fuzzy logic COA method is discussed to discover the pattern of an attribute. The method is implemented in MATLAB fuzzy toolbox, M-File, and C++.

Example 6.1: Building Age

The physical element of building age also was significant in the high crime blocks. The trend appears that the older the home is the less likely it is to get broken into. The more modern in appearance, the home was hit harder maybe because this type of home was a more attractive target. [2]

Because of the analogue nature (Table 6.1), a fuzzy set is defined by using a smooth curve (i.e., triangle in the middle and trapezoidal at the right end). Other smooth curves
can be adopted as well, such as Gaussian in the middle and \( PI \) at the right end. The reason for using trapezoidal and \( PI \) curves at the right end is that the building age is rarely higher than 40 years, which is usually called an overtime house (OT). Therefore trapezoidal and \( PI \) curves explicitly contain the assumption that if the building age is near overtime, namely 40 years old, then the risk curve considerably converges to a limit point.

<table>
<thead>
<tr>
<th>Building Age</th>
<th>No Crime</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>New</td>
<td>11</td>
<td>44.0</td>
<td>10</td>
<td>40.0</td>
</tr>
<tr>
<td>Average</td>
<td>123</td>
<td>74.5</td>
<td>32</td>
<td>19.4</td>
</tr>
<tr>
<td>Old</td>
<td>57</td>
<td>77.0</td>
<td>14</td>
<td>18.9</td>
</tr>
<tr>
<td>Under Const.</td>
<td>1</td>
<td>50.0</td>
<td>1</td>
<td>50.0</td>
</tr>
</tbody>
</table>

Table 6.2 Building Age in all crime blocks

<table>
<thead>
<tr>
<th>Fuzzy Sets (Year)</th>
<th>Membership Function</th>
<th>Break-in Ratio (%)</th>
<th>Start</th>
<th>Peak</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under Construction</td>
<td>Triangle</td>
<td></td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>New</td>
<td>Triangle</td>
<td></td>
<td>1</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>Average</td>
<td>Triangle</td>
<td></td>
<td>10</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>Old</td>
<td>Trapezoidal</td>
<td></td>
<td>25</td>
<td>35, 50</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 6.3 Fuzzy set definition for residential break-in ratio
As Figure 6.4 shows, the fuzzy model can directly represent super-positional states. The Building Age fuzzy set represents the gradual shift in membership or representation from a point that is completely representative of the set concept (e.g., building age at 20 years old represented by the Average fuzzy set) to a point that is completely unrepresentative of the set concept (e.g., building age at 20 years old represented by the Old fuzzy set). Due to the fact that points in between the Average fuzzy set and the Old fuzzy set can have partial degrees of membership, a fuzzy set can overlap with other fuzzy sets that have complementary or even contradictory meanings.

The rules are listed as follows:

If Break-In-Freq is Low, AND Building Age is Under Construction, then break-in ratio is VeryHigh;
If Break-In-Freq is Mid, AND Building Age is Under Construction, then break-in ratio is Low;
If Break-In-Freq is High, AND Building Age is Under Construction, then break-in ratio is Low;

If Break-In-Freq is Low, AND Building Age is New, then break-in ratio is Mid;
If Break-In-Freq is Mid, AND Building Age is New, then break-in ratio is VeryHigh,
If Break-In-Freq is High, AND Building Age is New, then break-in ratio is SortOfLow;

If Break-In-Freq is Low, AND Building Age is Average, then break-in ratio is SortOfHigh,
If Break-In-Freq is Mid, AND Building Age is Average, then break-in ratio is SortOfLow,
If Break-In-Freq is High, AND Building Age is Average, then break-in ratio is Low;
If Break-In-Freq is Low, AND Building Age is Old, then break-in ratio is SortOfHigh.
If Break-In-Freq is Mid, AND Building Age is Old, then break-in ratio is SortOfLow.
If Break-In-Freq is High, AND Building Age is Old, then break-in ratio is Low.

Figure 6.5 Fuzzy rules: Building Age

Figure 6.6 Inferencing: Building Age
When the inputs are a 25 year-old building with five break-in cases, the break-in risk is quite low at 1.53 for this type of building.

![Figure 6.7 Output surface: Building Age](image)

To conclude, the output surface indicates that the risk decreases when the case count increases, except in areas where the building age is low (Under Construction and New), the possibility is high where five break-in cases occurred; in areas where the building age is low (Under Construction and New), the possibility is Sort of Low to have more than five break-in cases; for those with building age at Average or Old, the possibility is Low to have more than five break-in cases.

### 6.2.3.1 MATLAB Implementation

The first approach is demonstrated in a MATLAB M-File by a manually designed fuzzy logic model. The risk assessment is programmed as an M-File function. The two inputs are $x$ and $y$, which respectively represent the building age in years and the burglary case count. The only output is the break-in ratio of the residential area. The input of the target attribute is building age and accounts for the contribution on the solution space. The rules are defined in Example 6.1. The script is listed as follows.
Example 6.2: C++ code for Building Age factor

```c
% File name: Fun_BuildingAge_COA_trim.m

% A Simple Fuzzy Logic Controller

% Problem Statement: This demonstration uses the Fuzzy Logic Toolbox to design a Fuzzy Inference System (FIS)
% using different defuzzification methods

% Create a new FIS with filename "Fun_BuildingAge_COA_trim.fis". function z= Fun_BuildingAge_COA_trim(x,y)
a=newfis('Durglary_BuildingAge_COA_trim');

% Add input variables x, y and output variable z to the FIS.
a=addvar(a,'input','BuildingAge',[0 50]);

% BuildingAge
a=addmf(a,'input',1,'BuildingAge.UnderConstruction','trimf', [0 0 2]);
a=addmf(a,'input',1,'BuildingAge.New','trimf', [1 7 15]);
a=addmf(a,'input',1,'BuildingAge.Average','trimf', [10 20 35]);
a=addmf(a,'input',1,'BuildingAge.Old','trapmf', [25 35 50 50]);

% Category
a=addvar(a,'input','Categories',[0 10]);

a=addmf(a,'input',2,'Break-In-Frequency.Low','trimf', [0 1 3]);
a=addmf(a,'input',2,'Break-In-Frequency.Md','trimf', [1 2 4]);
a=addmf(a,'input',2,'Break-In-Frequency.High','trimf', [2 3 10]);

% Normalized Risk
a=addvar(a,'output','Posibility',[0 100]);

a=addmf(a,'output',1,'Posibility.Low','trimf', [0 0 5]);
a=addmf(a,'output',1,'Posibility.SortOfLow','trimf', [0 5 10]);
a=addmf(a,'output',1,'Posibility.Mid','trimf', [5 10 15]);
a=addmf(a,'output',1,'Posibility.RatherMid','trimf', [10 15 20]);
a=addmf(a,'output',1,'Posibility.SortOfHigh','trimf', [15 20 25]);
a=addmf(a,'output',1,'Posibility.High','trimf', [20 30 40]);
a=addmf(a,'output',1,'Posibility.VeryHigh','trimf', [30 60 100]);

% Add rules to the FIS.
rule=[...
1 1 1 2 2 2 3 3 3 4 4 4
1 2 3 1 2 3 1 2 3 1 2 3
```
Add rules to the FIS.
\[
a = \text{addrule}(a, \text{rule}');
\]
\[
\%
\text{Use COA defuzzification method to find output Z}
\]
\[
z = \text{evalfis}([x, y], a);
\]

The script was run in MATLAB Command Window as a M-File function with inputs of the Building Age at 15 and the Break-In-Freq categories at 1, 2, and 3 respectively, then the correspondent outputs are 20, 12.5, and 4.1304, which represent the categories of risk (Sort of high, Mid, and Sort of Low).

![MATLAB Command Window](image)

Figure 6.8 Risk assessment: Building Age at 15 years old

Then the script was run in MATLAB Command Window as an M-File function with inputs of the Building Age at 35 and the Break-In-Freq categories at Low (1), Mid (2), and High (3) respectively. The correspondent outputs are 20, 11.3793, 3.5455, which represent the categories of risk (Sort of high, Mid, and Sort of Low).
The results of the M-file script are confirmed by the Rule Viewer in Fuzzy Toolbox as a FIS workspace.

6.2.3.2 C++ Implementation

This section demonstrates a method of implementation in C++ programming. This is an implementation of fuzzy logic reasoning in the COA method for the residential factor Building Age. The procedures are described as below:

1. Create fuzzy set for Building Age factor.

Example 6.3: C++ code for creating fuzzy set for Building Age factor

```c++
void CFuzzyResidenceRisk::CreateBuildingAge(void)
{
    // domain
    dblDomain[0]=0; dblDomain[1]=50;

    // BuildingAge: UnderConstruction
    dblParams[0]=0;
    dblParams[1]=2;
}
```
lpFDBBuildingAge[0]=FzyCreateSet("BuildingAge.UnderConstruction", DECREASE, dblDomain, dblParams, 2, &istatus);
strncpy(lpFDBBuildingAge[0]->FDBdesc,"Building Age: Under Construction (no. of burglary case in past three years)");
MdlLinkFDB(lpFDBBuildingAge[0],lpPDBRisk,&iStatus);

// BuildingAge: New
dblParams[0]=1;
dblParams[1]=7;
dblParams[2]=15;

lpFDBBuildingAge[1]=FzyCreateSet("BuildingAge.New",TRIANGLE,dblDomain,dblParams,3,&istatus);
strncpy(lpFDBBuildingAge[1]->FDBdesc,"Building Age: New Building (no. of burglary case in past three years)");
MdlLinkFDB(lpFDBBuildingAge[1],lpPDBRisk,&iStatus);

// BuildingAge: Average
dblParams[0]=10;
dblParams[1]=20;
dblParams[2]=30;

lpFDBBuildingAge[2]=FzyCreateSet("BuildingAge.Average",TRIANGLE,dblDomain,dblParams,3,&istatus);
strncpy(lpFDBBuildingAge[2]->FDBdesc,"Building Age: Average Age Building (no. of burglary case in past three years)");
MdlLinkFDB(lpFDBBuildingAge[2],lpPDBRisk,&iStatus);

// BuildingAge: Old
dblParams[0]=25;
// dblParams[1]=35;

lpFDBBuildingAge[3]=FzyCreateSet("BuildingAge.Old",RITESHOULDER,dblDomain,dblParams,2,&istatus);
strncpy(lpFDBBuildingAge[3]->FDBdesc,"Building Age: Old Building (no. of burglary case in past three years)");
MdlLinkFDB(lpFDBBuildingAge[3],lpPDBRisk,&iStatus);

return;
}

2. Generate rules and set the method of decomposition

Example 6.4: C++ code for creating fuzzy rules for Building Age factor

void CFuzzyResidenceRisk::CreateRulesBuildingAge(double dblInputA, double dblInputB, int iCorrMethod){
double dblPremiseTruthA,dblPremiseTruthB,dblPremiseTruth;
int iIndexPos;

//final result shape
FzyAddFZYblk(lpVDBRisk,&lpFDBBuildingAge_COA,&lpFSVRisk,&iStatus);
//Apply the rule:
strcpy(lpFDBBuildingAge_COA->FDBId,"BasementAge");
strcpy(lpFDBBuildingAge_COA->FDBdesc,"Basement Age Shape");

//If Break-In-Freq is One+, AND Building Age is Under Construction, then break-in ratio is VeryHigh;
FRAR_BuildingAge[0].Rule ="If Break-In-Freq is One+, AND Building Age is Under Construction, then break-in ratio is VeryHigh";
FRAR_BuildingAge[0].lpFDBCaseCount =lpFDBCaseCountCategory[0];
FRAR_BuildingAge[0].lpFDBTargetFactor =lpFDBBuildingAge[0];
FRAR_BuildingAge[0].lpFDBBreakInRate=lpFDBBreakInRate[6];

dblPremiseTruthA=FzyGetMembership(FRAR_BuildingAge[0].lpFDBCaseCount,dblInputA,&iIndexPos,&iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_BuildingAge[0].lpFDBTargetFactor,dblInputB,&iIndexPos,&iStatus);

dblPremiseTruth=min(dblPremiseTruthA,dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[0].lpFDBBreakInRate,lpFSVRisk ,iCorrMethod,dblPremiseTruth,&iStatus);
FRAR_BuildingAge[0].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is Two+, AND Building Age is Under Construction, then break-in ratio is Low;
FRAR_BuildingAge[1].Rule ="If Break-In-Freq is Two+, AND Building Age is Under Construction, then break-in ratio is Low";
FRAR_BuildingAge[1].lpFDBCaseCount =lpFDBCaseCountCategory[1];
FRAR_BuildingAge[1].lpFDBTargetFactor =lpFDBBuildingAge[0];
FRAR_BuildingAge[1].lpFDBBreakInRate=lpFDBBreakInRate[0];

dblPremiseTruthA=FzyGetMembership(FRAR_BuildingAge[1].lpFDBCaseCount,dblInputA,&iIndexPos,&iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_BuildingAge[1].lpFDBTargetFactor,dblInputB,&iIndexPos,&iStatus);

dblPremiseTruth=min(dblPremiseTruthA,dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[1].lpFDBBreakInRate,lpFSVRisk ,iCorrMethod,dblPremiseTruth,&iStatus);
FRAR_BuildingAge[1].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is Three+, AND Building Age is Under Construction, then break-in ratio is Low;
FRAR_BuildingAge[2].Rule ="If Break-In-Freq is Three+, AND Building Age is Under Construction, then break-in ratio is Low";
FRAR_BuildingAge[2].lpFDBCaseCount =lpFDBCaseCountCategory[2];
FRAR_BuildingAge[2].lpFDBTargetFactor =lpFDBBuildingAge[0];
FRAR_BuildingAge[2].lpFDBBreakInRate=lpFDBBreakInRate[0];

dblPremiseTruthA=FzyGetMembership(FRAR_BuildingAge[2].lpFDBCaseCount,dblInputA,&iIndexPos,&iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_BuildingAge[2].lpFDBTargetFactor,dblInputB,&iIndexPos,&iStatus);

dblPremiseTruth=min(dblPremiseTruthA,dblPremiseTruthB);
FzyCond Proposition(FRAR_Bu ildingAge[2].lpFDBBreak InRate, lpFSVRisk ,ICorrMethod,dblPremiseTruth, &iStatus);

FRAR_Bui ildingAge[2].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is One+, AND Building Age is New, then break-
in ratio is Mid;
FRAR_B ui ildingAge[3].Rule ="If Break-In-Freq is One+, AND
Building Age is New, then break-in ratio is Mid";
FRAR_B ui ildingAge[3].lpFDBCaseCount =lpFDBCaseCountCategory[0];
FRAR_B ui ildingAge[3].lpFDBTargetFactor =lpFDBBuildingAge[1];
FRAR_B ui ildingAge[3].lpFDBBreakInRate=lpFDBBreakInRate[2];

dblPremiseTruthA=FzyGetMembership(FRAR_Bui ildingAge[3].lpFDBCaseCo unt,dblInputA,&iIndexPos,&iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_Bui ildingAge[3].lpFDBTargetFactor,dblInputB,&iIndexPos,&iStatus);

dblPremiseTruth=min(dbIPremiseTruthA,dblPremiseTruthB);
FzyCond Proposition(FRAR_Bui ildingAge[3].lpFDBBreakInRate,lpFSVRisk ,ICorrMethod,dblPremiseTruth, &iStatus);

FRAR_Bui ildingAge[3].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is Two+, AND Building Age is New, then break-
in ratio is VeryHigh;
FRAR_B ui ildingAge[4].Rule ="If Break-In-Freq is Two+, AND
Building Age is New, then break-in ratio is VeryHigh";
FRAR_B ui ildingAge[4].lpFDBCaseCount =lpFDBCaseCountCategory[1];
FRAR_B ui ildingAge[4].lpFDBTargetFactor =lpFDBBuildingAge[1];
FRAR_B ui ildingAge[4].lpFDBBreakInRate=lpFDBBreakInRate[6];

dblPremiseTruthA=FzyGetMembership(FRAR_Bui ildingAge[4].lpFDBCaseCo unt,dblInputA,&iIndexPos,&iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_Bui ildingAge[4].lpFDBTargetFactor,dblInputB,&iIndexPos,&iStatus);

dblPremiseTruth=min(dbIPremiseTruthA,dblPremiseTruthB);
FzyCond Proposition(FRAR_Bui ildingAge[4].lpFDBBreakInRate,lpFSVRisk ,ICorrMethod,dblPremiseTruth, &iStatus);

FRAR_Bui ildingAge[4].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is Three+, AND Building Age is New, then break-
in ratio is SortOfLow;
FRAR_B ui ildingAge[5].Rule ="If Break-In-Freq is Three+, AND
Building Age is New, then break-in ratio is SortOfLow";
FRAR_B ui ildingAge[5].lpFDBCaseCount =lpFDBCaseCountCategory[2];
FRAR_B ui ildingAge[5].lpFDBTargetFactor =lpFDBBuildingAge[1];
FRAR_B ui ildingAge[5].lpFDBBreakInRate=lpFDBBreakInRate[1];

dblPremiseTruthA=FzyGetMembership(FRAR_Bui ildingAge[5].lpFDBCaseCo unt,dblInputA,&iIndexPos,&iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_Bui ildingAge[5].lpFDBTargetFactor,dblInputB,&iIndexPos,&iStatus);

dblPremiseTruth=min(dbIPremiseTruthA,dblPremiseTruthB);
// If Break-In-Freq is One+, AND Building Age is Average, then break-in ratio is SortOfHigh;
FRAR_BuildingAge[6].Rule = "If Break-In-Freq is One+, AND Building Age is Average, then break-in ratio is SortOfHigh";
FRAR_BuildingAge[6].lpFDBCaseCount = lpFDBCaseCountCategory[0];
FRAR_BuildingAge[6].lpFDBTargetFactor = lpFDBBuildingAge[2];
FRAR_BuildingAge[6].lpFDBBreakInRate = lpFDBBreakInRate[4];

dblPremiseTruthA = FzyGetMembership(FRAR_BuildingAge[6].lpFDBCaseCount, dblInputA, &iIndexPos, &iStatus);
dblPremiseTruthB = FzyGetMembership(FRAR_BuildingAge[6].lpFDBTargetFactor, dblInputB, &iIndexPos, &iStatus);

dblPremiseTruth = min(dblPremiseTruthA, dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[6].lpFDBBreakInRate, lpFSVRisk, iCorrMethod, dblPremiseTruth, &iStatus);
FRAR_BuildingAge[6].fPremiseTruth = dblPremiseTruth;

// If Break-In-Freq is Two+, AND Building Age is Average, then break-in ratio is SortOfLow;
FRAR_BuildingAge[7].Rule = "If Break-In-Freq is Two+, AND Building Age is Average, then break-in ratio is SortOfLow";
FRAR_BuildingAge[7].lpFDBCaseCount = lpFDBCaseCountCategory[1];
FRAR_BuildingAge[7].lpFDBTargetFactor = lpFDBBuildingAge[2];
FRAR_BuildingAge[7].lpFDBBreakInRate = lpFDBBreakInRate[1];

dblPremiseTruthA = FzyGetMembership(FRAR_BuildingAge[7].lpFDBCaseCount, dblInputA, &iIndexPos, &iStatus);
dblPremiseTruthB = FzyGetMembership(FRAR_BuildingAge[7].lpFDBTargetFactor, dblInputB, &iIndexPos, &iStatus);

dblPremiseTruth = min(dblPremiseTruthA, dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[7].lpFDBBreakInRate, lpFSVRisk, iCorrMethod, dblPremiseTruth, &iStatus);
FRAR_BuildingAge[7].fPremiseTruth = dblPremiseTruth;

// If Break-In-Freq is Three+, AND Building Age is Average, then break-in ratio is Low;
FRAR_BuildingAge[8].Rule = "If Break-In-Freq is Three+, AND Building Age is Average, then break-in ratio is Low";
FRAR_BuildingAge[8].lpFDBCaseCount = lpFDBCaseCountCategory[2];
FRAR_BuildingAge[8].lpFDBTargetFactor = lpFDBBuildingAge[2];
FRAR_BuildingAge[8].lpFDBBreakInRate = lpFDBBreakInRate[0];

dblPremiseTruthA = FzyGetMembership(FRAR_BuildingAge[8].lpFDBCaseCount, dblInputA, &iIndexPos, &iStatus);
dblPremiseTruthB = FzyGetMembership(FRAR_BuildingAge[8].lpFDBTargetFactor, dblInputB, &iIndexPos, &iStatus);

dblPremiseTruth = min(dblPremiseTruthA, dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[8].lpFDBBreakInRate, lpFSVRisk, iCorrMethod, dblPremiseTruth, &iStatus);

FRAR_BuildingAge[8].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is One+, AND Building Age is Old, then break-in ratio is SortOfHigh;
FRAR_BuildingAge[9].Rule ="If Break-In-Freq is One+, AND Building Age is Old, then break-in ratio is SortOfHigh";
FRAR_BuildingAge[9].lpFDBCaseCount =lpFDBCaseCountCategory[0];
FRAR_BuildingAge[9].lpFDBTargetFactor =lpFDBBuildingAge[3];
FRAR_BuildingAge[9].lpFDBBreakInRate=lpFDBBreakInRate[4];

dblPremiseTruthA=FzyGetMembership(FRAR_BuildingAge[9].lpFDBCaseCount, dblInputA, &iIndexPos, &iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_BuildingAge[9].lpFDBTargetFactor, dblInputB, &iIndexPos, &iStatus);

dblPremiseTruth=min(dblPremiseTruthA,dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[9].lpFDBBreakInRate, lpFSVRisk, iCorrMethod, dblPremiseTruth, &iStatus);

FRAR_BuildingAge[9].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is Two+, AND Building Age is Old, then break-in ratio is SortOfLow;
FRAR_BuildingAge[10].Rule ="If Break-In-Freq is Two+, AND Building Age is Old, then break-in ratio is SortOfLow";
FRAR_BuildingAge[10].lpFDBCaseCount =lpFDBCaseCountCategory[1];
FRAR_BuildingAge[10].lpFDBTargetFactor =lpFDBBuildingAge[3];
FRAR_BuildingAge[10].lpFDBBreakInRate=lpFDBBreakInRate[1];

dblPremiseTruthA=FzyGetMembership(FRAR_BuildingAge[10].lpFDBCaseCount, dblInputA, &iIndexPos, &iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_BuildingAge[10].lpFDBTargetFactor, dblInputB, &iIndexPos, &iStatus);

dblPremiseTruth=min(dblPremiseTruthA,dblPremiseTruthB);
FzyCondProposition(FRAR_BuildingAge[10].lpFDBBreakInRate, lpFSVRisk, iCorrMethod, dblPremiseTruth, &iStatus);

FRAR_BuildingAge[10].fPremiseTruth=dblPremiseTruth;

//If Break-In-Freq is Three+, AND Building Age is Old, then break-in ratio is Low;
FRAR_BuildingAge[11].Rule ="If Break-In-Freq is Three+, AND Building Age is Old, then break-in ratio is Low";
FRAR_BuildingAge[11].lpFDBCaseCount =lpFDBCaseCountCategory[2];
FRAR_BuildingAge[11].lpFDBTargetFactor =lpFDBBuildingAge[3];
FRAR_BuildingAge[11].lpFDBBreakInRate=lpFDBBreakInRate[0];

dblPremiseTruthA=FzyGetMembership(FRAR_BuildingAge[11].lpFDBCaseCount, dblInputA, &iIndexPos, &iStatus);
dblPremiseTruthB=FzyGetMembership(FRAR_BuildingAge[11].lpFDBTargetFactor, dblInputB, &iIndexPos, &iStatus);

dblPremiseTruth=min(dblPremiseTruthA,dblPremiseTruthB);
3. Defuzzify the output space.

Example 6.5: C++ code for defuzzifying fuzzy rules for Building Age factor

```cpp
float CFuzzyResidenceRisk::DefuzzyBuildingAge(int iCorrMethod){
    float fBreakInRate, fRisk, fsetheight;
    fBreakInRate=FzyDefuzzify(lpFDBBuildingAge_COA,lpFSVRisk->FSVdefuzzMethod,&fRisk,&iStatus);
    fsetheight=FzyGetHeight(lpFDBBuildingAge_COA);
    lpFDBBuildingAge_COA->FDBDefuzzy=fBreakInRate;
    lpFDBBuildingAge_COA->FDBDefuzzySup=fsetheight;
    return fsetheight;
}
```

When we design the input space and the output space, we choose the domains, the membership functions, and methods of inference and defuzzification. For example, if Building Age is 13.5 years with one break-in in the past three years, then the resultant possible break-in ratio is estimated at 15.63. (see Table 6.4) Combination of different designs for fuzzy set and fuzzy inference is critical for a good estimation. Scope of domains, types of membership functions, types of fuzzy operators, and methods of defuzzification will affect the estimation. But this effort is rewarding for establishing a good model.

6.2.4 Multi-Attribute Analysis

Based on the single-attribute analysis, multiple attributes of an environmental factor can be further analyzed. This analysis reveals the patterns of the factor with complexity. A systematic design is necessary to build a model incorporating domains, membership functions (shape, overlaps, etc.), methods of implication, and methods of defuzzification. In this section, two software packages (MATLAB and C++) are implemented for the risk assessment model. The MATLAB tool conveniently allows a user to set up a simulation.
In contrast, the C++ program is sophisticated but has strong functionality. The results can be evaluated by the equation (6.1)

\[ R_x = \sum_{n=1}^{k} f(r_n) \]  

(6.1)

\( r_n \) is one of the residential attributes (physical or psychological factors). The break-in ratio of the residence characteristics \( r_x \) is strongly related to \( r_n \), either the physical factors such as the intrinsic properties of the building, nearby neighborhood location (street block, community, district), or cultural and psychological impact such as traditions or neighborhood relationships. For simplicity, equal weighing is opted.

### 6.2.4.1 MATLAB Implementation

In this section, we create several MATLAB M-file scripts to evaluate the characteristics of the factors and sum the results. Inside these scripts, we call the MATLAB functions that we generated for evaluating each factor using the method of COA. The following samples demonstrate the methods to evaluate the complex attributes of a residence that contribute to the output risk.

Example 6.6: C++ code for simple summarizing the Barriers and Deterrent factors:

1. Residential characteristics

```matlab
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% %                   Fuzzy Logic Toolbox FIS Evaluation
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% % %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 
% % Problem Statement: This demonstration uses the Fuzzy Logic Toolbox to design a Fuzzy Inference System (FIS) evaluating multiple factors
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% % %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 
% % Create a new FIS with filename "Fun_Residential_Sum.fis":
% function z=Fun_Residential_Sum(x1,x2,x3,x4,x5,x6)
```
% Coefficient
% Add input variables x, y and output variable z to the FIS.

a=Fun_BuildingTypes_COA_trmf(x1,1);
b=Fun_BuildingAge_COA_trmf(x2,1);
c=Fun_BasementDesign_COA_trmf(x3,1);
d=Fun_BuildingLocation_COA_trmf(x4,1);
e=Fun_RearLanePresence_COA_trmf(x5,1);
f=Fun_BalconyNo_COA_trmf(x6,1);

z=a+b+c+d+e+f;

%%%%%%%%%%%%%%%%
% Filename : Fun_Barrier_Sum.m
%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%
% A Multi-factor Fuzzy Logic Controller
%%%%%%%%%%%%%%%%

2. Barriers and deterrents

%%%%%%%%%%%%%%%%
% Problem Statement: This demonstration uses the Fuzzy Logic Toolbox to design a Fuzzy Inference System (FIS) evaluating multiple factors
%%%%%%%%%%%%%%%%

% Create a new FIS with filename "Fun_Barrier_Sum.fis".
function z=Fun_Barrier_Sum(x1,x2,x3,x4,x5,x6,x7)

% Coefficient
% Add input variables x, y and output variable z to the FIS.

a=Fun_BuildingCondition_COA_trapmf(x1,1);
b=Fun_FenceHeight_COA_trapmf(x2,1);
c=Fun_FenceType_COA_trmf(x3,1);
d=Fun_PorchLandingArea_COA_trmf(x4,1);
e=Fun_PrevDecals_COA_trmf(x5,1);
f=Fun_RailingDesign_COA_trapmf(x6,1);
g=Fun_StepDesign_COA_trapmf(x7,1);

z=a+b+c+d+e+f+g;

3. Entry and access

%%%%%%%%%%%%%%%%
% Filename : Fun_EntryAndAccess_Sum.m
%%%%%%%%%%%%%%%%
% Create a new FIS with filename "Fun_EntryAndAccess_Sum.fis".
function z=Fun_EntryAndAccess_Sum(x1,x2,x3,x4,x5,x6,x7,x8)

% Coefficient
% Add input variables x, y and output variable z to the FIS.

a=Fun_DoorPosition_COA_trapmf(x1,1);
b=Fun_NoOfDoors_COA_trapmf(x2,1);
c=Fun_WLFBsmt_COA_trapmf(x3,1);
d=Fun_WLFLstF_COA_trapmf(x4,1);
e=Fun_WLF2ndF_COA_trapmf(x5,1);
f=Fun_WLR1stF_COA_trapmf(x6,1);
g=Fun_WLR2ndF_COA_trapmf(x7,1);
h=Fun_DoorLevel_COA_trimf(x8,1);

z=a+b+c+d+e+f+g+h;

Further, we can use formulas, such as equation 3.1, or we can obtain the average break-in ratio to evaluate the target residence.

6.2.4.2 C++ Implementation

In this section, a C++ application program is developed to demonstrate the same functionality as previously shown in MATLAB. This is an implementation of a fuzzy logic model. The procedures are described as follows:

1. Initialize the fuzzy model system including templates of domains, membership functions, rules, hedges, and decompositions. For example, the fuzzy control block consists of log files, alpha cuts, and dictionaries.
2. Create fuzzy sets: inputs and outputs (i.e., break-in count, break-in ratio, residential characteristic factors).
3. Load input variables, rules, hedge types, and methods of composition and decomposition.
4. Create solution fuzzy regions, incorporate hedges, and execute propositions. This step evaluates the model proposition through an aggregation process, and generates the final fuzzy regions for each solution variable.
5. Choose the method of the decomposition and defuzzification, which generates the ultimate output values.

The rules are defined in Example 6.2. The script is list as follows.

Example 6.7: C++ code for summarizing the barriers and deterrent factors in COA method

```c++
float CFuzzyResidenceRisk::ReasoningCOA(void)
{
    dblDomain[0]=0; dblDomain[1]=100;
    lpVDBRisk=VarCreateScalar("RiskAccessmentRes", REAL, dblDomain, "0", &istatus);
    MdlLinkVDB(lpVDBRisk, lpPDBRisk, &istatus);
    FzyInitFZYblk(&istatus);

    //lpFDBRisk_COA is final result shape
    FzyAddFZYblk(lpVDBRisk, &lpFDBRisk_COA, &lpFSVRisk, &istatus);

    CreateAllSets();
    InitializeModelData();
    CreateRules(lpFSVRisk->FSVcorrMethod);
    return DefuzzyCOA(lpFSVRisk->FSVdefuzzMethod);
}
float CFuzzyResidenceRisk::DefuzzyCOA(int iDefuzzMethod)
{
    float fRisk;
    mfRisk[0]=DefuzzyBuildingType(iDefuzzMethod);
    fRisk=mfRisk[0];

    mfRisk[1]=DefuzzyBuildingAge(iDefuzzMethod);
    fRisk=fRisk+mfRisk[1];
    mfRisk[2]=DefuzzyBasementDesign(iDefuzzMethod);
    fRisk=fRisk+mfRisk[2];
    mfRisk[3]=DefuzzyNoOfBalconies(iDefuzzMethod);
    fRisk=fRisk+mfRisk[3];
    mfRisk[4]=DefuzzyBuildingLocation(iDefuzzMethod);
    fRisk=fRisk+mfRisk[4];

    mfRisk[5]=DefuzzyRearLane(iDefuzzMethod);
    fRisk=fRisk+mfRisk[5];

    return fRisk;
}
float CFuzzyResidenceRisk::DefuzzyBuildingType(int iCorrMethod)
{
    float fBreakInRate, fRisk, fsetheight;
```
fBreakInRate=FzyDefuzzify(lpFDBBuildingType_COA,lpFSVRisk->FSVdefuzzMethod,&fRisk,&iStatus);
fsetheight=FzyGetHeight(lpFDBBuildingType_COA);
lpFDBBuildingType_COA->FDBDefuzzySup=fsetheight;
return fsetheight;
}

6.2.4.3 Results

<table>
<thead>
<tr>
<th>Residential Characteristics (Output space domain: break-in ratio 0~100%)</th>
<th>Input Domain</th>
<th>Input Value</th>
<th>Scalar Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Break- In-Freq.</td>
</tr>
<tr>
<td>Building Type (#)</td>
<td>0</td>
<td>50</td>
<td>1</td>
</tr>
<tr>
<td>Basement Design (m)</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Building Age (year)</td>
<td>0</td>
<td>50</td>
<td>1</td>
</tr>
<tr>
<td># of Balconies (#)</td>
<td>0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Building Location (m)</td>
<td>0</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>Rear Lane Presence (%)</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Average break-in ratio</strong></td>
<td>Sum(break-in ratio)/6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Risk</strong></td>
<td>Risk=Sum(break-in ratio)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.4 COA decomposition: Residential Characteristic in high crime blocks

<table>
<thead>
<tr>
<th>Barriers and Deterrents (Output space domain: break-in ratio 0~100%)</th>
<th>Input Domain</th>
<th>Input Value</th>
<th>Scalar Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Break- In-Freq.</td>
</tr>
<tr>
<td>Garage Type (nominal)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Preventive Decals(binary)</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fence Type (nominal)</td>
<td>0</td>
<td>259</td>
<td>1</td>
</tr>
<tr>
<td>Fence Height (feet)</td>
<td>0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Step Design (degree)</td>
<td>0</td>
<td>360</td>
<td>1</td>
</tr>
<tr>
<td>Building Condition (%)</td>
<td>0</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>Porch Landing Area (degree)</td>
<td>0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Railing Design (degree)</td>
<td>0</td>
<td>360</td>
<td>1</td>
</tr>
<tr>
<td><strong>Average Break-in ratio</strong></td>
<td>Sum(break-in ratio)/7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Risk</strong></td>
<td>Risk=Sum(break-in ratio)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.5 COA decomposition: Barriers and Deterrents in high crime blocks
The average break-in ratio indicates the possible break-in ratio of a residence based on its intrinsic characteristics in each category. The sum of the break-in ratios is considered a benchmark of the break-in risk of the residence. An investigator or insurer can evaluate the break-in ratio or break-in risk and determine the Achilles' heel. Weighting can be added to the averaging process and unreliable results (e.g., lack of supportive statistical data) can be also filtered out, which increases confidence in the final results. However, the flaw of this arithmetic mean is that the relationship among different factors is not clearly explored on the output space.

On the other hand, arithmetic mean is a good measure of central tendency for roughly symmetric distributions but can be misleading in skewed distributions since it can be
greatly influenced by extreme scores. Therefore, other statistics such as the median may be more informative for skewed distributions.

6.2.5 Discussion

These models showed advantages in handling various situations. *Scalable Monotonic Chaining* is a simple model and intuitive to understand. It provides a quick way to pre-estimate the risk of environmental factors before establishing a larger scale research or application. Single attribute analysis can incorporate various fuzzy inference methods to handle the complexity of an attribute in various data types. The main difficulty of this model is to design appropriate fuzzy sets and inference methods to overcome imprecision and vagueness, which can be a time-consuming process. Multi-attribute analysis provides an example to understand the aggregation of individual attributes, and various arithmetic and statistical algorithms can be used for better estimation.
6.3 Case Study: Spatial Risk Analysis

In the previous section, a problem was examined because of the ambiguous relationship among different factors on the output space. This problem can be resolved with the assistance of fuzzy reasoning. On a geographic area, it is commonly to estimate the risk of a spot. The estimation depends on risks contributed by cues and signals within the spot and spatial neighbourhood influences outside. Assume that these risks are known to investigators or criminologists, and then their knowledge is resembled in fuzzy linguistic variables, fuzzy hedges, fuzzy operators, and if-then rules, and all known risks are synthesized to an estimation of the overall risk on the spot. The strengths of fuzzy reasoning are examined and demonstrated in this section.

6.3.1 Hypothesis and Assumptions

Inside a site, risks of attributes interact; outside a site, risks of neighborhood sites spatially interact. Both affect the risk on the side. If the risks are known in given residential attributes, overall risk can be estimated and weighed based on knowledge of decision-makers. If the risks are known in adjacent sites, then based on the assumption of distance decay, risk can be estimated by the distance (either travel distance or travel time). In this case study, we assume the concept of weighing and distance decay is included in the knowledge of a decision-maker either implicitly or explicitly. All data are hypothetical.

6.3.2 Risk Analysis: Intrinsic Attributes

One activity node, for example, a residential building consists of multiple environmental factors (Table 1). Instead of simply adding up all the break-in ratios of the factors, the data on output spaces can be treated as new inputs, and synthesized into a new output space. Arbitrarily we choose the output data set of basement design (Residential Characteristics), Preventive Decals (Barriers and Deterrents), and Windows Location on first floor at rear (Entry and Access) as input data sets (Figure 6.10).
Furthermore, three rules can be set as:

1. The basement design risk must be Pattern A.
2. The Preventive Decals risk must be Pattern B.
3. The Windows Location on first floor at rear risk must be Pattern C.

The linguistic word must can be interpreted as the fuzzy logic operator AND (Min). Zadeh (1965) initially defined the conventional fuzzy logic operations as intersection (Equation 5.4) as follows:

\[ \mu_{A \cap B}(x) \triangleq \min \{ \mu_A(x), \mu_B(x) \} = \mu_A(x) \land \mu_B(x) \quad \forall x \in U \]

To see the intersection operation, we apply the rules with AND operator, the partial membership organization of the input fuzzy sets creates an area (shaded) where the values are considered to belong to Shape A, B, and C. Figure 6.11 illustrates the process of inference on output space.
Further, we apply the COA method to a fuzzy set on the output space and obtain the output value.

<table>
<thead>
<tr>
<th>Input Factors</th>
<th>Input Domain</th>
<th>Input Value</th>
<th>Scalar Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Output space domain: break-in ratio 0~100%)</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Break-In-Freq</td>
</tr>
<tr>
<td>Basement Design (m) — Residential Characteristics</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Preventive Decals (binary) — Barriers and Deterrent</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1st floor Windows Location at rear (meter) — Entry and Access</td>
<td>0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Break-in ratio (%)</td>
<td>Simple Averaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advanced Inferencing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truth Value</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.7 COA Decomposition: activity node risk assessment in high crime blocks
Compared with simple averaging method, inferencing of risks on activity node reflects the knowledge of the experts. The input fuzzy sets are first conformed to the rules of knowledge base, then the mathematical methods can be applied to obtain the final results. More methods of implication and difuzzifications are optionally used to represent the knowledge of the experts.

6.3.2.1 C++ Implementation

In this section, the same functionality is demonstrated in a C++ application of implementation of a fuzzy logic model. The procedures are described as follows:

1. Initialize the C++ fuzzy model system including templates of domains, membership functions, rules, hedges, and decompositions. For example, the fuzzy control block consists of log files, alpha cuts, and dictionaries.
2. Create fuzzy sets: inputs and outputs (i.e., break-in count, break-in ratio, residential characteristic factors).
3. Load input variables, rules, hedge types, and methods of composition and decomposition.
4. Create solution fuzzy regions, incorporate hedges, and execute propositions. This step evaluates the model proposition through an aggregation process, and generates the final fuzzy regions for each solution variable.
5. Choose the method of the decomposition and defuzzification, generating the ultimate output values.

The rules are defined in Example 6.4. The script is list as follows.
Example 6.8: C++ code for summarizing decision support of Synthetical Risk Analysis

```cpp
float CFuzzyDecisionSupport::ReasoningDS(void)
{
    dblDomain[0]=0;
    dblDomain[1]=100;
    lpVDBRisk=VarCreateScalar("DecisionSupport",REAL,dblDomain,"0",&iStatus);
    MdlLinkVDB(lpVDBRisk,lpPDBRisk,&iStatus);
    FzyInitFZYblk(&iStatus);
    //lpFDBRisk_COA is final result shape
    FzyAddFZYblk(lpVDBRisk,&lpFDBRisk_DS,&lpFSVRisk,&iStatus);
    CreateAllSets();
    InitializeModelData();
    ImplicationMethod(MINMAX);
    CreateRules(lpFSVRisk->FSVcorrMethod);
    return DefuzzyDS(lpFSVRisk->FSVdefuzzMethod);
}
```

6.3.3 Risk Analysis: Spatial Distance Decay

Previously, we discussed the influences of the residential factors on *burglary* risk. These factors are static and intrinsic properties of a residence. According to crime pattern theory, these influences come from nearby public areas (parks and schools), commercials (shops and clubs), and transportation networks (sky train stations, main highways, and highway entrances or exits). The relationship certainly exists between the *break-in ratio* and these area characteristics, but usually they are not clearly defined and are described verbally—such a decision-support process is linguistic and relies on the knowledge of experts. Fuzzy logic presents a set of methods that incorporates human interferences—*Linguistic variables, hedges, operators*. In this section, the spatial influence from the adjacent areas is analyzed based on *If-Then* rules. These influences are external to a residence. Several situations can be accounted for, for example, at a micro level, a spot is adjacent to other spots such as a *school, park, sky train station* (Figure 6.13); at a macro level, a community or district is adjacent to other communities or districts, the target area
of Northern Hastings is the neighbor of the areas of Downtown Eastside, Renfrew heights, and Northern Burnaby.

This case study evaluates the break-in risk influences from neighborhood—adjacent spots such as a public school, a regional park, and a sky train station. The neighborhood must have been investigated and the risk is known as an input to the decision-support system. The neighborhood can be a community in a larger scale or spots in a smaller scale such as the residences of a street block, a few street blocks, or nearby points of interests (public sites, or commercial shops). Based on the known risks in the neighborhood, a prediction can be estimated based on human knowledge and judgments in form of the following If-Then pseudo code:

1. The risk must be above the risk of regional park,
2. The risk must be below the risk of sky train station,
3. If the risk at public school zone is very Low, then the risk is possibly close to the risk at the public school zone.

The rules consist of fuzzy operators, hedges, and linguistic variables such as must, possibly, above, below, very, if-then, etc. The rules (1, 2, 3) only reflect the contribution of risk from the neighboring spots.
Figure 6.14 Regional Park: before and after *above* operator

Figure 6.15 Sky Train Station: before and after *below* operator

Figure 6.16 Park and Sky Train Station: before and after *above*, *below*, and *must* (*min*) operators
Further we apply the COA method to the fuzzy set on the output space and obtain the output values listed in Table 6.8.
### Table 6.8 Spatial Factors: membership functions, operations, results

<table>
<thead>
<tr>
<th>Input Factors</th>
<th>MF</th>
<th>Parameter</th>
<th>Fuzzy Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(break-in ratio 0~100%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="#">Regional park</a></td>
<td>PI</td>
<td>Start: 13</td>
<td>Applied operators</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peak: 13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>End: 13+6.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>above: Min</td>
<td></td>
</tr>
<tr>
<td><a href="#">Sky train station</a></td>
<td>PI</td>
<td>Start: 43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peak: 43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>End: 43+21.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>below: Max</td>
<td></td>
</tr>
<tr>
<td><a href="#">Public school zone</a></td>
<td>triangular</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Break-in ratio (%)]</td>
<td></td>
<td>29.68</td>
<td></td>
</tr>
</tbody>
</table>

| Truth Value                                        | 1       |

### 6.3.3.1 C++ Implementation

The analysis is implemented in a C++ application. The outline of procedures is described as follows:

1. Initialize the fuzzy model system including templates of domains, membership functions, rules, hedges, and decompositions. For example, the fuzzy control block consists of log files, alpha cuts, and dictionaries.
2. Create fuzzy sets: inputs and outputs (i.e., break-in count, break-in ratio, residential characteristic attributes).
3. Load input variables, rules, hedge types, and methods of composition and decomposition.
4. Create solution fuzzy regions, incorporate hedges, and execute propositions. This step evaluates the model proposition through an aggregation process, and generates the final fuzzy regions for each solution variable.
5. Choose the method of the decomposition and defuzzification, generating the ultimate output values.

The rules are defined in Example 6.5. The script is listed as follows.

**Example 6.9: C++ code for summarizing decision support of spatial risk analysis**

```c++
float CFuzzyDecisionSupport::ReasoningDS2(void){

dblDomain[0]=0;
dblDomain[1]=100;
lpVDBRisk=VarCreateScalar("DecisionSupportCase2",REAL,dblDomain,"0",&iStatus);
MdlLinkVDB(lpVDBRisk,lpPDBRisk,&iStatus);
FzyInitFZYblk(&iStatus);
```

134
6.3.4 Hybrid Risk Analysis

A spot is adjacent to public sites, the risk on the spot is partially affected by the intrinsic environmental factors on site (e.g., building age etc) and affected by the spatial risk distribution from neighborhoods. In this section, both the external risks and the internal risks are evaluated. The external risks are the influences from the neighborhood (regional park, sky train station, school zone), and the internal risks are the influence from intrinsic properties (residential character, barrier and deterrent, and entry and access). Based on the known risks in the neighborhood and the residence, a prediction can be estimated based on human knowledge and judgments in form of the following If-Then pseudo code:

1. The risk must below 1.2 times the risk of regional Park,
2. The risk must above 0.7 times the risk of sky train station,
3. The residential character risk is possibly Pattern A,
4. The barrier and deterrent risk is possibly Pattern B,
5. The entry and access risk is possibly Pattern C,
6. If the risk at the school zone is very Low, then the risk is possibly around the risk at school zone.

The rules consist of the fuzzy operators, hedges, and linguistic variables such as around, possibly, very, if-then, etc. The rules (1, 2, 6) reflect the contribution of risks from the neighboring spots, and the other rules reflect the risks from intrinsic attributes of the residence. Pattern A, B, C are inputs of known characteristics, here we arbitrarily pass in the shape of a residential attribute from a category. In addition, rules 1 and 2 have a coefficient, which can enlarge or reduce the regions of shapes, and shift the centers of curves to the lower end or the upper end. The shape of the inputs is displayed in Figure 6.20. The left-hand part of the figure contains the shapes of the spatial attributes;
right-hand side contains the shapes of the residential factors, which were generated in the previous section.

The linguistic variable possible is interpreted as a fuzzy OR operator. After applying the operator, three shapes of the fuzzy sets are composed into one set (shown in Figure 6.21). The shape on the right-hand side is the result of the union operation. The region area of the intermediate result is enlarged. Both intrinsic attributes and external factors from neighborhood regions have possible influences on the final output space.

The operations are displayed in Figure 6.21. The original shapes of the spatial factors are shown on the left graph, then the below and the above operators are applied and the intersection of the shapes is shown in the middle graph. In addition, the area of the region is lessened, the shape is unionized with the shape of the fuzzy set of public school zone, and the final result is shown on the right-hand side graph. The linguistic variables
possible and must are interpreted as a fuzzy OR operator and a AND operator respectively, which changes the area of a region.

Figure 6.22 Spatial Factors: before and after applying fuzzy operators

The union operation is applied because both the region of the residential factors and the spatial factors contributes to the final result.

Figure 6.23 Hybrid Factors: before and after applying OR operator to the fuzzy sets of residential and spatial.

Further we apply the COA method to the fuzzy set on the output space and obtain the output values listed in Table 6.9.
## Table 6.9 Hybrid Factors: membership functions, operations, and results

### Environmental Factors

<table>
<thead>
<tr>
<th>Environmental Factors</th>
<th>Input Domain</th>
<th>Input Value</th>
<th>Scalar Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Break-In Ratio</td>
</tr>
<tr>
<td></td>
<td>0-1 00%</td>
<td></td>
<td>Value (%)</td>
</tr>
<tr>
<td>Basement Design (m)</td>
<td>0 6</td>
<td>1 2.3</td>
<td>0.346</td>
</tr>
<tr>
<td>Residential Characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preventive Decals (binary)</td>
<td>0 1</td>
<td>1 0.7</td>
<td>0.696</td>
</tr>
<tr>
<td>Barriers and Deterrent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st floor Windows Location at rear (meter)</td>
<td>0 10</td>
<td>1 2.5</td>
<td>0.75</td>
</tr>
<tr>
<td>Entry and Access</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Spatial Factors

<table>
<thead>
<tr>
<th>Spatial Factors</th>
<th>Membership Function</th>
<th>Parameter</th>
<th>Fuzzy Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(break-in ratio 0-100%)</td>
<td></td>
<td></td>
<td>Applied operators</td>
</tr>
<tr>
<td>Regional park</td>
<td>PI</td>
<td>15.6-7.8</td>
<td>15.6+7.8</td>
</tr>
<tr>
<td>Sky train station</td>
<td>PI</td>
<td>30.1-15.0</td>
<td>30.1+15.0</td>
</tr>
<tr>
<td>Public school zone</td>
<td>triangular</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>Break-in ratio (8)</td>
<td></td>
<td></td>
<td>23.44</td>
</tr>
<tr>
<td>Truth Value</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

6.3.4.1 C++ Implementation

The analysis is implemented in a C++ application. The outline of procedures is described as follows:

1. Initialize the fuzzy model system including templates of domains, membership functions, rules, hedges, and decompositions. For example, the fuzzy control block consists of log files, alpha cuts, dictionaries, etc.
2. Create fuzzy sets: inputs and outputs (i.e., break-in count, break-in ratio, residential characteristic factors).
3. Load input variables, rules, hedge types, and methods of composition and decomposition.
4. Create solution fuzzy regions, incorporate hedges, and execute propositions. This step evaluates the model proposition through an aggregation process, and generates the final fuzzy regions for each solution variable.

5. Choose the method of the decomposition and defuzzification, generating the ultimate output values.

The rules are defined in Example 6.5. The script is listed below.

Example 6.10: C++ code for summarizing decision support of hybrid Risk Analysis

```cpp
float CFuzzyDecisionSupport::ReasoningDS1(void){
    dblDomain[0]=0;
    dblDomain[1]=100;
    lpVDBRisk=VarCreateScalar("DecisionSupportCase2",REAL,dblDomain,"0",&iStatus);
    MdlLinkVDB(lpVDBRisk,lpFDBRisk,&iStatus);
    FzyInitFZYblk(&iStatus);
    //lpFDBRisk_COA is final result shape
    FzyAddFZYblk(lpVDBRisk,&lpFDBRisk_DS,&lpFSVRisk,&iStatus);
    CreateAllSets1();
    InitializeModelData1();
    ImplicationMethod(MINMAX);
    CreateRules1(lpFSVRisk->FSVcorrMethod);
    return DefuzzyDS1(lpFSVRisk->FSVdefuzzMethod);
}
```

6.3.5 Discussion

Fuzzy logic shows strengths in handling linguistic decision-making. Fuzzy reasoning manipulates inputs and outputs in if-then rules, fuzzy operators, fuzzy approximation, and fuzzy hedges. The analyses utilize fuzzy reasoning methods and extend crime pattern recognition from characteristics of attributes to the aggregation of burglary offences. Therefore, not only the characteristics of attributes on a site are assessable but the spatial burglary patterns are recognizable from neighborhood risk impacts. Especially, Hybrid Risk Analysis provides a balanced weighing in risky factors from external influence from neighborhoods and internal characteristics within a site over a geographical space.
6.4 Case Study: Markov Chain Model

On a target spot, break-in risk varies in attributes and space. However, the attributes and spatial risk distributions are ever-changing, thus break-in risk fluctuates with time. These time-related factors usually reflect changes in weather, transportations, development and construction (lack of security), market and fair (accessible to cash out), event and program, economy, and policy etc. By adding a dimension of time to Target Selection Model, the trend of risks is assessable by analyzing timeline variations in characteristics of attributes and spatial distributions of break-in risks. Markov chain model can track changes from a state to another. Variations of the trends can be betted by an array of possibility/probability values. This study of Markov chain model can be applied to achieve three purposes:

- Analyze the dynamic characteristics of the burglary offence pattern from one time period to another in the aid of the transition matrices;
- Predict the trend of the burglary offence pattern based on the reward matrices of an empirical model;
- Evaluate the quality of the analysis and prediction.

A Markov chain analyses the behavior of the burglary risk analysis over time, whose probability of going from one state to any other state is independent of how the chain reached its current state. Previously, all residential attributes are considered static, or under the same state. In this section, the attributes vary from state to state. With a state, the attributes are treated as the fuzzy number of incremental interval during an on-going event. The operations and computations on fuzzy numbers are called fuzzy arithmetic, discussed in Chapter 5.

6.4.1 Hypothesis and Assumptions

In Vancouver East side, Pacific National Exhibition (PNE) is surrounded by three districts: the Northern Hastings community in the west, 1st Ave in the south, and Boundary St. in the east. The target areas are PNE neighborhoods. Geographically, around PNE fair, the major east-west bound travel routes are Hastings St. and 1st Ave.
The major north-south bound travel routes are Boundary St., Rupert St., and Renfrew St. The most traffic usually occurred in Hastings St., 1st Ave and Boundary St., which was considered as major travel and transportation routes around PNE. During June, July, and August every year, PNE is hosted an annual event — PNE summer fair. Thousands of people from the Lower Mainland visit the PNE fair and incur a large amount increase in traffic, retail, advertisement, and community activities. An increase in burglary offence is observed in the surrounding regions during the event period. In this section, the dynamic characteristics of the burglary offence pattern are studied based on possibility/probability theory, fuzzy logic and Markov chains model. Assumptions need to set up for these models as:

- The physical and sociological characteristics are relatively stable at micro level (i.e., residences) in the region. The capacity for the opportunity is certain and can be obtained in the historical records of burglary offences.

- PNE summer fair causes the physical and sociological characteristics changed at macro level (i.e., traffic flows, visitors, commercial premises, and community venues). The opportunity increases during the PNE fair event but cannot exceed a certain limit. The capacity for the opportunity saturates at a point. The increase can be estimated based on the burglary records in the first month of the event—June, 2002.

- The distribution of the opportunities is geographically even in the PNE neighborhoods. The possibility/probability is virtually equal in any part of the PNE neighborhoods. The observations support this assumption: the travel distance within the PNE neighborhoods are less than 10 minutes; comparing the communities in Vancouver West side, all the communities are rather similar in the regions.

According to empirical knowledge, the offenders are reluctant to revisit the same place within three months*. They target a few houses in an area. Hence at each attempt the offender could only be active in one district or another, the possibility for the next attempt is virtually equal anywhere else in PNE neighborhoods.

* Note: it is debatable for a revisit in three month, or within a week, or even next day, but we hypothetically assume three months is an average period for a revisit.

6.4.2 Transition Matrix and Dynamic Pattern Characteristics

Every year, PNE hosts a summer fair in June, July, and August. Suppose, in 2002 during the PNE fair event, initially all three areas observed a normal Break-in ratio in May but
were disturbed by the annual PNE fair event during summer, the *burglary* rates are recorded respectively from higher to lower in the Northern Hastings, 1st Ave, and Boundary St. In June, the rates hiked in all three districts, but most sharply in the Northern Hastings area and reached its peak, moderately in the 1st Ave district, and lowest in Boundary Street. district. In July, the rate changed moderately in Northern Hastings, while the rate rose faster in other areas, especially in the 1st Ave district the rate reached its peak. In August, the rate declined in the Northern Hastings area and in 1st Ave district moderately, but rapidly increased in Boundary St. district. The burglary distribution and diffusion for these months is illustrated in Figure 6.24.

![Figure 6.24 Burglary distribution and diffusion during PNE fair](image)

The observations above can be summarized as follows:

- PNE fair produced opportunities that attracted *burglary* offenders; the opportunities are incentives to *burglary* activities in neighborhood districts;
- The *burglary* activities occurred along the major travel routes;
- The originally-attracted offenders in the PNE east side gradually lost their interest and traveled from east side to south, then to the west side. Moreover, the originally-attracted offenders in the PNE west side gradually lost their interest and traveled from west side to south, then to the east side; the originally-attracted
offenders in the PNE south side gradually lost their interest and were free to travel
either to the west or to the east.

These observations can be explained as follows:

- A certain amount of offenders (individuals or groups) were attracted to these areas
during PNE fair, and they were opportunists in that they tended to target PNE
neighborhoods near the east side entrances, and traveled from block to block to
seek the most attractive target.
- Around PNE fair, the major east-west bound travel routes are Hastings St. and 1st
Ave. The major north-south bound travel routes are Boundary St., Rupert St., and
Renfrew St. The most traffic usually is observed on Hastings St., 1st Ave and
Boundary St., which are considered major travel and transportation routes around
PNE.
- The offenders were free to travel around PNE fair to identify targets, thus the
density of activities diffused and finally the density reached its equilibrium.

Consequently, based on the observation and explanation, the PNE fair event can be
divided into three monthly states as:

- State 1: the rate increases sharply in Northern Hastings;
- State 2: the rate rises faster in 1st Ave district;
- State 3: the rate soars higher in Boundary St. district.

Markov Chain model is applied to determine the possibility/probability that the burglars
or burglar groups will be active in one district in any trial. The risk changes can be traced
by the transition matrices from state to state. The transition matrices can be set up by
surveys of the investigators or historical records. In this case study, the transition matrix
is assumed unchanged.

Suppose data were collected and it was found that, originally the break-in ratios are 19%,
11%, 9% in Northern Hastings district, 1st Ave district, and Boundary St. district
respectively in June. In July, for the burglars, who usually were active in Northern
Hastings in June, 42% of them stayed in Northern Hastings district, 30% moved to 1st
Ave district and 28% to Boundary St. district. For those who usually were active in 1st
Ave district in June, the corresponding figures were 35%, 35%, and 30%. For those who usually were active in Boundary St. district, the figures were 32%, 32%, and 36%. The transition probability matrix from these observations is:

\[ P = \begin{bmatrix} 0.42 & 0.30 & 0.28 \\ 0.35 & 0.35 & 0.30 \\ 0.32 & 0.32 & 0.36 \end{bmatrix} \]  

(6.4.1)

Note that the sum of each row is equal to 1, because we considered the PNE neighborhood as a whole and each district accounted for a portion of the total break-in cases. The state probability vector \( x_0 \) is applied to determine the state of the system in subsequent periods and represents the break-in ratios in the initial time period. Hence:

\[ x_0' = [0.19, 0.11, 0.09] \]  

(6.4.2)

identifies the state in the initial period for burglary break-in cases in the Northern Hastings district, 1st Ave district, and Boundary St. district. The state probabilities for the next time period are realized by post-multiplying the state vector of the initial period by the transition probability matrix \( P \). Hence:

\[ x_1 = x_0'P = \begin{bmatrix} 0.19 & 0.11 & 0.09 \end{bmatrix} \begin{bmatrix} 0.42 & 0.30 & 0.28 \\ 0.35 & 0.35 & 0.30 \\ 0.32 & 0.32 & 0.36 \end{bmatrix} = \begin{bmatrix} 0.1471 & 0.1243 & 0.1186 \end{bmatrix} \]  

(6.4.3)

For the second period, the month of August, assume that the transition matrix unchanged. The burglars kept traveling to neighboring districts. Hence:
If the PNE event would continued for the third period, the month of September assume that the transition matrix unchanged. Hence:

\[
x_2' = x_1'P = x_0'P^2 = \begin{bmatrix} 0.19 & 0.11 & 0.09 \\ 0.35 & 0.35 & 0.30 \\ 0.32 & 0.32 & 0.36 \end{bmatrix}^2 \begin{bmatrix} 0.42 & 0.30 & 0.28 \\ 0.35 & 0.35 & 0.30 \\ 0.32 & 0.32 & 0.36 \end{bmatrix} = \begin{bmatrix} 0.3710 & 0.3206 & 0.3084 \\ 0.3655 & 0.3235 & 0.3110 \\ 0.3616 & 0.3232 & 0.3152 \end{bmatrix}
\]

\[
x_2' = x_1'P = x_0'P^2 = \begin{bmatrix} 0.1432 & 0.1256 & 0.1212 \end{bmatrix}
\]

(6.4.4)

A general form of the state probability for period \( n \) is represented by:

\[
x_n' = x_0'P^n
\]

(6.4.6)

when \( n=4 \), \( x' = x_n' = x_{n+1} \),

\[
P^4 = \begin{bmatrix} 0.3663 & 0.3223 & 0.3114 \\ 0.3663 & 0.3223 & 0.3114 \\ 0.3663 & 0.3223 & 0.3114 \end{bmatrix}
\]

\[
x_4' = x_5' = x_6' = \ldots = x_n' = \begin{bmatrix} 0.1429 & 0.1257 & 0.1214 \end{bmatrix}
\]

(6.4.7)

the state vector \( x' \) is equal to the state vector \( n+I \), where \( x_n' \) is independent of the initial state vector \( x_0' \), and the \( P \) is called steady-state vector for the Markov chain. The elements of \( x' \) are called the steady-state probabilities, given by:

\[
x' = x'P
\]

(6.4.8)
This model explains that, the impact of the PNE fair first waned in the Northern Hastings district. The offenders were more likely to seek new targets in 1st Ave district and Boundary St. district. The number of burglary offences were foreseen to rise in July and August, and in the long run the break-in risk would have converged to \[ [0.1429 \ 0.1257 \ 0.1214] \] regardless of the initial state. However, preventative measures, such as block watching and street patrolling, can reduce the impact of the event.

Since the transition matrix is assumed unchanged in this case study, this dynamic analysis relies mathematically on the initial transition probability matrix $P$ and eliminates any other factor in the real environment. Consequently, the results are never manifested in the real world. To avoid this limitation, the rewords state matrices before and after PNE summer fair event are introduced and discussed in the next section.

### 6.4.3 Reward Matrix and Pattern Prediction

Suppose there are two scenarios described the burglary risks during the period from May to July, 2002 as follows:

**Scenario 1: Less Effective**

- State 1: the rate was *about Low* in May before the PNE fair, and the rate was normal compared with the same period in the past. This state is marked as $P^1$;

- State 2: the rate rose to *High* in June during the PNE fair, and less effective burglary prevention actions had been taken. The state is marked as $P^2$.

**Scenario 2: More Effective**

- State 2: the rate rose *near to High* in June during the PNE fair, and less effective actions had been taken to prevent the break-in crime. The state is marked as $P^2$.

- State 3: the rate was *about Mid* in July during the PNE fair. More effective actions had been taken to prevent the break-ins. The state is marked as $P^3$. 
Based on the two scenarios, *burglary* risks are predicted for the period of August, 2002, and the confidence of the prediction is evaluated to determine its quality. (See Figure 6.25)

In this section, the reward matrices are introduced into the Markov Chain model which allow user inputs that reflect the changes in environment and precautions of human decisions. In the reward matrices, the data reflect the imprecise and linguistic nature of human decisions, the reward matrix values are considered as fuzzy numbers, and the operations of the numbers are adopted from fuzzy arithmetic. The algorithm is proposed under an empirical study of *break-in* risk analysis combined with fuzzy set theory and Markov chain. In this case study, the reward matrices are incorporated a hypothetical empirical model for illustration purposes. The reward matrices are derived from statistical data from surveys or historical records.

The empirical model contains the residential attributes or sociological factors. These attributes and factors are individually accounted for the value on the final solution space.
The data of the attributes and factors are formulated into an equation based on the historical statistical data, or the knowledge of experts.

6.4.3.1 Markov Chain Reward Matrix Model

Suppose that the system moves to State \( j \) at the first state, whose reward is \( r_{ij} \). The expected reward over all \( n \) transition can be expressed as \( r_{ij} + v_i(n-1) \), where \( v_i(n-1) \) is the expected reward over the remaining \( n-1 \) transitions when the system begins in State \( i \) and first moves to State \( j \). The general form of the total expected reward after \( n \) transitions, starting in State \( i \) can be expressed as:

\[
v(n) = \sum_{j=1}^{N} p_{ij} r_{ij} + \sum_{j=1}^{N} p_{ij} v_j (n - 1) \tag{6.4.13}
\]

where \( p_{ij} \) are elements of \( P \). If \( q_i = \sum_{j=1}^{N} p_{ij} r_{ij} \) represents the expected reward from the next transition when \( i \) is the current state, and the remaining expected reward over \( n-1 \) transitions \( v(n-1) = \sum_{j=1}^{N} v_j (n - 1) \) then:

\[
v(n) = q + pv(n - 1), \quad \text{where} \quad q' = [q_1 \quad q_2 \quad \cdots \quad q_N] \tag{6.4.14}
\]

Suppose that the annual PNE fair started in June, and the initial state of May was prior to the event and could not exert any influence on the break-in ratio in June. Thus, let \( v(0) = 0 \), then obtain \( v(n) = q_i = \sum_{j=1}^{N} p_{ij} r_{ij} \), allowing to evaluate the effectiveness of the decisions in crime prevention actions. The matrices of transition probabilities are generated from questionnaires, which reflect human knowledge and tend to be subjective. The matrices of rewards are emphasized on empirical equations and contain the crispy values (i.e., the estimated number of new break-in cases), which are based on statistics and tend to be objective. Therefore, the Markov Chain model combines both the statistics and the knowledge, shares both objectiveness and subjectiveness and equation 6.4.13 can be interpreted as the following pseudo code:
In the next section, the reward matrix method is examined to show the effectiveness of the crime preventive precautions in an empirical model.

6.4.3.2 Empirical Break-in Analysis

Suppose the volume of the visitors to PNE fair has a certain influence on the number of the break-in cases because of its attraction. Statistically, every 10,000 visitors per month results in 3 new break-in events; a 5 hour increase per day in police patrol reduces 2 break-in events; a higher alert level of block-watching reduces the break-in events by 10 percent. The block-watching affects both influences of the visitor volume and the police patrol but in a contrarily way: any higher alert reduces the offence increase by the visitor volume, but degrades the efforts of the police patrol. The tendency toward the precautions is pessimistic and to eliminate overrating of the preventive actions. Thus, the model is defined as follows:

\[
\Delta R = (3 \times V_{10k} - 2 \times P_{5h}) \times 10 \times K_{bw} \%
\]

(6.4.11)

\(\Delta R\): the reduced break-in events per month by the preventive actions,

\(V_{10k}\): volume of visitors per 10,000 persons,

\(P_{5h}\): action 1, 5 hours of police patrol per day,

\(K_{bw}\): action 2, alert level of block-watching.
Factors | Domain | Low | Mid | High
--- | --- | --- | --- | ---
Visiting Volume (10,000 persons) | 0–100 | 1 | 6 | 10
Police Patrol (hours/day) | 0–24 | 5 | 10 | 16
Block-watching (alert level) | 0–10 | 3 | 5 | 7

Table 6.10 Fuzzy domain and set: visiting volume, police patrol, and block-watching

<table>
<thead>
<tr>
<th>Period (calendar month)</th>
<th>Fuzzy Operator</th>
<th>Visiting Volume (10,000 persons)</th>
<th>Police Patrol (hours/day)</th>
<th>Block-watching (alert level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>May</td>
<td>about</td>
<td>0.8</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>June</td>
<td>about</td>
<td>8</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>July</td>
<td>about</td>
<td>12</td>
<td>15</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 6.11 Comparison: visiting volume, police patrol, and block-watching

Then we approximate the fuzzy numbers $V_{lk}$, $P_{sh}$, $K_{bw}$, and $\Delta R$ by equation (6.4.11) as follows:

- **May:**
  
  If $V_{lk} = (0.8, 1, 1.2)$, $P_{sh} = (3, 4, 5)$, $K_{bw} = (0, 1, 2)$, then $\Delta R = (0, 0.14, 0.32)$

- **June:**
  
  If $V'_{lk} = (7.2, 8, 8.8)$, $P_{sh} = (8.1, 9, 9.1)$, $K_{bw} = (2, 3, 4)$, and if we continuously keep the preventive actions at level 3, the reduced Break-In-Frequency in August can be assumed to be about the same, that is $\Delta R' = (3.67, 6.12, 9.1)$

- **July:**
  
  If $V'_{lk} = (10.8, 12, 13.2)$, $P_{sh} = (13.5, 15, 16.5)$, $K_{bw} = (5, 6, 7)$, then if we continuously keep the preventive actions at level 6, the reduced Break-In-Frequency in August can be assumed to be about the same, that is $\Delta R^2 = (13.5, 18, 23.1)$

**6.4.3.3 Transition Matrix**

The transition matrices in the Markov chains are used to predict the break-in ratio when the less/more effective crime preventive actions have been taken. Let
where:

\[ P^1 = \begin{bmatrix}
P_{11}^1 & P_{12}^1 & P_{13}^1 \\
P_{21}^1 & P_{22}^1 & P_{23}^1 \\
P_{31}^1 & P_{32}^1 & P_{33}^1
\end{bmatrix} \quad P^2 = \begin{bmatrix}
P_{11}^2 & P_{12}^2 & P_{13}^2 \\
P_{21}^2 & P_{22}^2 & P_{23}^2 \\
P_{31}^2 & P_{32}^2 & P_{33}^2
\end{bmatrix} \] (6.4.9)

\[ P^1: \] the transition probabilities less effective preventive actions,
\[ P^2: \] the transition probabilities more effective preventive actions;
\[ P_{ij}^1: \] when less effective actions, the probability that if the break-in ratio is \( i \)th state on the first stage, then it moves to \( j \)th state on the second stage,
\[ P_{ij}^2: \] when more effective actions, the probability that if the break-in ratio is \( i \)th state on the first stage, then it moves to \( j \)th state on the second stage;

\( i=j=1 \): the break-in ratio keeps low from July to August, so the break-in trend is Good,
\( i=j=2 \): the break-in ratio keeps mid from July to August, so the break-in trend is Fair,
\( i=j=3 \): the break-in ratio keeps high from July to August, so the break-in trend is Poor.

The matrices define the probabilities of the transition from state to state. To determine each value in the matrix, a widely used method is surveying the investigators or criminologists for their estimation. For example, what is the transition probability of August in "low/mid/high" state, respectively, if break-in preventive actions have been taken in a less/more effectively way. Thus, the matrices can be obtained below:

\[ P^1 = \begin{bmatrix}
0.1 & 0.3 & 0.6 \\
0.1 & 0.4 & 0.5 \\
0.25 & 0.35 & 0.4
\end{bmatrix} \quad P^2 = \begin{bmatrix}
0.15 & 0.35 & 0.55 \\
0.1 & 0.25 & 0.65 \\
0.05 & 0.2 & 0.75
\end{bmatrix} \] (6.4.10)

\( P^1 \) contains the transition probabilities from the current period of June to the next period of August with less effective preventive actions. \( P^2 \) contains the transition probabilities
from the current period of July to the next period of August with more effective preventive actions. The attitude of the estimations is pessimistic due to the significant increase of the burglary rate offences during the PNE fair. So that the matrices exhibit the concern that the possibilities of burglary rates shift higher from the Low state to the High state. The investigators or criminologists predicted a worsen situation in August even if the preventive actions had been taken.

6.4.3.4 Reward Matrix

In the previous section, the investigators or criminologists subjectively showed their pessimistic estimations. In this section, the rewards or payoffs are introduced in the Markov chains modal to measure the effectiveness of the preventive actions that had been taken in July. If the objective efforts prevail over the pessimistic estimations, then a decline of the burglary rate become feasible in the period of August.

Let \( r_{ij} = (r_{ij}^k - \Delta R^k) \) be the reward associated with a transition from State \( i \) to State \( j \), which may be interpreted as the reward from the transition itself or the reward for being in State \( i \) (or State \( j \)) during one time period. Suppose there are \( N \) states in the system. We define the matrix of rewards, \( R \), to be:

\[
R^k = \sum_{i=1}^{N} \sum_{j=1}^{N} (r_{ij}^k - \Delta R^k) = \begin{pmatrix}
    r_{11}^k - \Delta R^k & r_{12}^k - \Delta R^k & \cdots & r_{1N}^k - \Delta R^k \\
    r_{21}^k - \Delta R^k & \cdots & \cdots & \cdots \\
    \vdots & \ddots & \ddots & \ddots \\
    r_{N1}^k - \Delta R^k & \cdots & \cdots & r_{NN}^k - \Delta R^k
\end{pmatrix}
\]

The symbol \( k \) is used to identify decisions that have taken place in July. Then the reward matrices can be simplified for the stages before or after the actions as follows [13]:

\[
R^1 = \|r_{ij}^1 - \Delta R^1\|, R^2 = \|r_{ij}^2 - \Delta R^2\|, \text{ or}
\]

\[
R^k = \|r_{ij}^k - \Delta R^k\|, \text{ where } i=j=1, 2, 3. k=1, 2
\]

(6.4.12)
i: the state of the system in the current month (i.e., in June or July), including: Low, Mid, High

j: the state of the system in future (i.e., in August), including: Low, Mid, High

k: stands for the decision of preventive actions available to the investigator in a period, specifically,

k=1: before the decision;
k=2: after the decision.

\( r_{ij}^k \): is the predictive number of new break-in cases, specifically,

\( r_{ij}^1 \): before the decision, the predictive number of new break-in cases in August;
\( r_{ij}^2 \): after the decision, the predictive number of new break-in cases in August.

\( \Delta R^k \): is the reduced number by the preventive actions, specifically,

\( \Delta R^1 \): before the decision, the reduced number by the preventive actions in August,
\( \Delta R^2 \): after the decision, the reduced number by the preventive actions in August,

\( r_{ij}^k-\Delta R^k \): is the corrected predictive number of new break-in cases from reaching state \( j \) in next period from \( i \) at the current time, specifically,

\( r_{ij}^1-\Delta R^1 \): before the decision, the difference between the new cases and the reduced cases in August;
\( r_{ij}^2-\Delta R^2 \): after the decision, the difference between the new cases and the reduced cases in August.

\( \Delta R \) is a matrix of the reduced break-in cases per month by the break-in preventive actions based on the empirical model in equation 6.4.11. Previously were obtained \( \Delta R^1=(3.67, 6.12, 9.1) \) before the decision, and \( \Delta R^2=(13.5, 18, 23.1) \) after the decision.

To obtain the values of \( r_{ij}^k \), the investigators or criminologists are surveyed through questionnaires or votes, for instance, how many break-in cases will be expected in August in Low/Mid/High state, respectively, and if current number of break-in cases is Low, then the values of \( r_{11}^1, r_{12}^1, r_{13}^1 \) can be initially estimated as follows:

\[
\begin{bmatrix}
    r_{11}^1 & r_{12}^1 & r_{13}^1 \\
    r_{21}^1 & r_{22}^1 & r_{23}^1 \\
    r_{31}^1 & r_{32}^1 & r_{33}^1
\end{bmatrix} = \begin{bmatrix}
    (5,7,9) & (11,12,14) & (17,19,20) \\
    (6,7,8) & (12,14,16) & (14,15,17) \\
    (8,9,10) & (9,11,12) & (16,18,20)
\end{bmatrix}
\]

\[
\begin{bmatrix}
    r_{11}^2 & r_{12}^2 & r_{13}^2 \\
    r_{21}^2 & r_{22}^2 & r_{23}^2 \\
    r_{31}^2 & r_{32}^2 & r_{33}^2
\end{bmatrix} = \begin{bmatrix}
    (14,15,16) & (16,18,19) & (18,19,21) \\
    (14,15,16) & (17,18,19) & (20,23,25) \\
    (15,16,17) & (14,18,15) & (22,25,27)
\end{bmatrix}
\]
After applying the fuzzy arithmetic equation 5.6 to the fuzzy number, the reward matrices can be obtained:

\[
R^1 = \left| r^1_j - \Delta r^1 \right| = \begin{bmatrix}
    r^1_{11} - \Delta r & r^1_{12} - \Delta r & r^1_{13} - \Delta r \\
    r^1_{21} - \Delta r & r^1_{22} - \Delta r & r^1_{23} - \Delta r \\
    r^1_{31} - \Delta r & r^1_{32} - \Delta r & r^1_{33} - \Delta r
\end{bmatrix} = \begin{bmatrix}
    (4.1,0.885,3.33) & (1.9,5.881,0.33) & (7.9,12.88,16.33) \\
    (3.1,0.884,3.33) & (2.9,7.881,23) & (4.9,8.88,133) \\
    (1.1,2.886,33) & (0.1,4.888,33) & (6.9,11.88,1633)
\end{bmatrix}
\]

\[
R^2 = \left| r^2_j - \Delta r^2 \right| = \begin{bmatrix}
    r^2_{11} - \Delta r & r^2_{12} - \Delta r & r^2_{13} - \Delta r \\
    r^2_{21} - \Delta r & r^2_{22} - \Delta r & r^2_{23} - \Delta r \\
    r^2_{31} - \Delta r & r^2_{32} - \Delta r & r^2_{33} - \Delta r
\end{bmatrix} = \begin{bmatrix}
    (-9.1,,-3,,-2.5) & (-7.1,0,5.5) & (-5.1,,-1.7,5) \\
    (-9.1,,-3,,-2.5) & (-6.1,0,5.5) & (-3.1,5,11.5) \\
    (-8.1,,-2,3.5) & (-9.1,0,1.5) & (-1.1,7,13.5)
\end{bmatrix}
\]

6.4.3.5 Total Expected Reward Vector

The investigators and criminologists predicted a rise of the burglary offences from May to June, and the empirical equation indicated a burglary increase if there is a lack of crime preventive actions. Therefore, if no effective preventive actions had been taken from June to July, the rising numbers of burglary offences can be predicted in August as follows

- If the state of break-in ratio was Low in August:
  \[\nu_1 = q_1 = \sum_{j=1}^{3} p_{j} r_{j} = [0.1 \times (-4.1) + 0.3 \times (1.9) + 0.6 \times (7.9),
  0.1 \times 0.88 + 0.3 \times 5.88 + 0.6 \times 12.88, 0.1 \times 5.33 + 0.3 \times 10.33 + 0.6 \times 16.33] = [0.49, 9.58, 13.43]\]

- If the state of break-in ratio was Mid in August:
  \[\nu_2 = q_2 = \sum_{j=1}^{3} p_{j} r_{j} = [0.1 \times (-3.1) + 0.4 \times (2.9) + 0.5 \times (4.9),
  0.1 \times 0.88 + 0.4 \times 7.88 + 0.5 \times 8.88, 0.1 \times 4.33 + 0.4 \times 12.33 + 0.5 \times 13.33] = [3.3, 7.68, 12.03]\]

- If the state of break-in ratio was High in August:
The number of the burglary offences rose from a normal level in June due to the PNE fair. The results show that it is more likely the burglary rate would rise at the same rate. However, if effective preventive actions had been taken in the period of July, then the rising trend could possibly be averted and the declining numbers of burglary offences can be predicted in August as follows:

- **If the state of break-in ratio was Low:**

  \[
  \nu(1) = \sum_{j=1}^{3} p_j r_j = [0.15 \times (-9.1) + 0.35 \times (-7.1) + 0.55 \times (-5.1), 0.15 \times (-3) + 0.35 \times 0 + 0.55 \times 1, 0.15 \times 2.5 + 0.35 \times 5.5 + 0.55 \times 7.5] = [-6.665, 0.1, 6.425]
  \]

- **If the state of break-in ratio was Mid:**

  \[
  \nu(2) = \sum_{j=1}^{2} p_j r_j = [0.1 \times (-9.1) + 0.25 \times (-6.1) + 0.65 \times (-3.1), 0.1 \times (-3) + 0.25 \times 0 + 0.65 \times 5, 0.1 \times 2.5 + 0.25 \times 5.5 + 0.65 \times 11.5] = [-4.45, 2.95, 9.1]
  \]

- **If the state of break-in ratio was High:**

  \[
  \nu(3) = \sum_{j=1}^{3} p_j r_j = [0.05 \times (-8.1) + 0.2 \times (-9.1) + 0.75 \times (-1.1),
  0.05 \times 5 + 0.2 \times 0 + 0.75 \times 7, 0.05 \times 3.5 + 0.2 \times 1.5 + 0.75 \times 13.5] = [-3.05, 10.025, 0.475]
  \]

The predictions show that the efforts of the preventive actions prevail over the pessimistic estimations by the investigators or criminologists, and the empirical model can measure the effectiveness of the preventive actions. The results assured that the preventive actions would be very effective in reducing the burglary rate during the PNE summer fair event.

Two sets of fuzzy numbers for each state are listed in the Table 6.12 as follows:
### State of Break-ins

<table>
<thead>
<tr>
<th>State of Decision</th>
<th>No Decision</th>
<th>Has Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>[4.9, 9.58, 13.43]</td>
<td>[-6.655, 0.1, 6.425]</td>
</tr>
<tr>
<td>Mid</td>
<td>[3.3, 7.68, 12.03]</td>
<td>[-4.45, 2.95, 9.1]</td>
</tr>
<tr>
<td>High</td>
<td>[2.45, 8.96, 4.498]</td>
<td>[-3.05, 10.025, 0.475]</td>
</tr>
</tbody>
</table>

Table 6.12 Comparison: state vector of break-ins in with/without decision

The fuzzy numbers reflect the quality of the decision at different stages. In the next section, the methods of ranking fuzzy numbers are described. The first method is a simple criterion-based ranking method; the second is an approximate approach that reflects the optimistic or pessimistic perspectives of a decision maker.

### 6.4.4 Decision Quality Analysis

In this section, the quality of the decisions is evaluated. Kaufmann and Gupta (1988)'s approximation methods are widely adopted to compare fuzzy numbers. These methods utilize either the basic arithmetic or fuzzy arithmetic methods. Both utilities are illustrated.

#### 6.4.4.1 Illustration: Linear Ordering of Fuzzy Numbers

Kaufmann and Gupta (1988) developed the three criteria for ordering the T.F.N.s. Here we apply the same methods to evaluate the predictions of the break-in trend for PNE fair. First we have the following six T.F.N.s:

\[ v(1) = [4.9, 9.58, 13.43], v(2) = [3.3, 7.68, 12.03], v(3) = [2.45, 8.96, 4.498] \]
\[ v'(1) = [-6.655, 0.1, 6.425], v'(2) = [-4.45, 2.95, 9.1], v'(3) = [-3.05, 10.025, 0.475] \]

By calculating the removal, the ordinary representatives can be obtained by equation 5.9:

\[ v(1) = [4.9, 9.58, 13.43] \rightarrow \hat{v}(l) = \frac{49 + 2 \times 9.58 + 13.43}{4} = 937 \]
\[
\begin{align*}
\mathbf{v}(2) &= [3.3, 7.68, 12.03] \rightarrow \mathbf{\hat{v}}(2) = \frac{3.3 + 2 \times 7.68 + 12.03}{4} = 7.67, \\
\mathbf{v}(3) &= [2.45, 8.96, 4.498] \rightarrow \mathbf{\hat{v}}(3) = \frac{2.45 + 2 \times 8.96 + 4.498}{4} = 6.22, \\
\mathbf{v}'(1) &= [-6.655, 0.1, 6.425] \rightarrow \mathbf{\hat{v}}'(1) = \frac{-6.655 + 2 \times 0.1 + 6.425}{4} = -0.01, \\
\mathbf{v}'(2) &= [-4.45, 2.95, 9.1] \rightarrow \mathbf{\hat{v}}'(2) = \frac{-4.45 + 2 \times 2.95 + 9.1}{4} = 2.64, \\
\mathbf{v}'(3) &= [-3.05, 10.025, 0.475] \rightarrow \mathbf{\hat{v}}'(3) = \frac{-3.05 + 2 \times 10.025 + 0.475}{4} = 4.37,
\end{align*}
\]

Thus, the following ordered classes can be obtained under first criterion by the removal:

class 1: \{ \mathbf{v}'(1) \}, the ordinary representative -0.01,
class 2: \{ \mathbf{v}'(2) \}, the ordinary representative 2.64,
class 3: \{ \mathbf{v}'(3) \}, the ordinary representative 4.37,
class 4: \{ \mathbf{v}(3) \}, the ordinary representative 6.22,
class 5: \{ \mathbf{v}(2) \}, the ordinary representative 7.67,
class 6: \{ \mathbf{v}(1) \}, the ordinary representative 9.37.

Thus, we obtain the linear ordering:

\[\mathbf{v}'(1) < \mathbf{v}'(2) < \mathbf{v}'(3) < \mathbf{v}(3) < \mathbf{v}(2) < \mathbf{v}(1)\]

<table>
<thead>
<tr>
<th>State of Rate</th>
<th>Low</th>
<th>Mid</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State of Decision</strong></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Number of new break-in cases</strong></td>
<td>4.9, 9.58, 13.43</td>
<td>-6.655, 0.1, 6.425</td>
<td>3.3, 7.68, 12.03</td>
</tr>
<tr>
<td><strong>Vector</strong></td>
<td>(\mathbf{v}(1))</td>
<td>(\mathbf{v}'(1))</td>
<td>(\mathbf{v}(2))</td>
</tr>
<tr>
<td><strong>Removal Criterion</strong></td>
<td>9.37</td>
<td>-0.01</td>
<td>7.67</td>
</tr>
<tr>
<td><strong>Class</strong></td>
<td>6</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><strong>Rank</strong></td>
<td>(\mathbf{v}'(1) &lt; \mathbf{v}'(2) &lt; \mathbf{v}'(3) &lt; \mathbf{v}(3) &lt; \mathbf{v}(2) &lt; \mathbf{v}(1))</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>State (with better action precautions)</strong></td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6.13 Kaufmann and Gupta’s Criteria: Linear ordering of fuzzy number

Results and Explanations:

- First, under all three initial circumstances (Low, Mid, High), the trend of burglary offences declined in July. The divergence between Yes and No is distinctive, which demonstrates that the preventive actions are very effective in preventing burglary offences. If the police maintain current break-in prevention efforts, the break-in ratio would probably continue to decline in August as well.

- Second, the actions work better in an area with the record of a lower break-in ratio. That is, if the break-in ratio was very low in May and the preventive actions would be taken in August, then the influence of PNE fair would be possibly minimized and the break-in risk would remain very low as well.

- Third, if no actions were be taken, then the influence of PNE fair would become more significant in an area with a lower break-in ratio in May than in an area with higher rate.

- The explanation for the finding is that the break-in ratio increases more significantly from a lower point, and the preventive actions such as block-watching and police patrol might have been taken previously. Therefore, no more precautions can be put into the community.

- Fourth, it is less effective in high break-in ratio areas, but these actions still exert a positive influence by preventing burglary offences because the removal value is distinctive from the one without the decision.

6.4.4.2 Illustration: Approximate Ranking of fuzzy numbers

According to Equations (5.20), (5.21) and (5.22), the rankings are obtained as follows:

The maximizing sets (right-side utilities) can be obtained as follows:

$$U_M(1) = \frac{[c_1 - \inf(x)]}{[\sup(x) - \inf(x)] - (b_1 - c_1)} = \frac{13.43 - (-6.655)}{[13.43 - (-6.655)] - (4.9 - 13.43)} = 0.84$$

$$U_M^-(1) = \frac{[c_1' - \inf(x)]}{[\sup(x) - \inf(x)] - (b_1' - c_1')} = \frac{6.425 - (-6.655)}{[13.43 - (-6.655)] - (0.1 - 6.425)} = 0.50$$

$$U_M(2) = \frac{[c_2 - \inf(x)]}{[\sup(x) - \inf(x)] - (b_2 - c_2)} = \frac{12.03 - (-4.45)}{[12.03 - (-4.45)] - (7.68 - 12.03)} = 0.79$$
The minimizing sets (left-side utilities) are obtained as follows:

\[
U_M(2) = \frac{c_2 - \text{inf}(x)}{[\text{sup}(x) - \text{inf}(x)] - (b_2' - c_2')} = \frac{9.1 - (-4.45)}{[12.03 - (-4.45)] - (2.95 - 9.1)} = 0.60
\]

\[
U_M(3) = \frac{c_3 - \text{inf}(x)}{[\text{sup}(x) - \text{inf}(x)] - (b_3' - c_3')} = \frac{4.498 - (-3.05)}{[10.025 - (-3.05)] - (8.96 - 4.498)} = 0.88
\]

\[
U_M'(3) = \frac{c_3' - \text{inf}(x)}{[\text{sup}(x) - \text{inf}(x)] - (b_3' - c_3')} = \frac{0.475 - (-3.05)}{[10.025 - (-3.05)] - (10.025 - 0.475)} = 1.00
\]

When \(\lambda = 0.5\), then the total utilities can be obtained as follows:

\[
U_T(1) = \lambda U_M(1) + (1 - \lambda)[1 - U_G(1)] = 0.5 \times 0.84 + 0.5 \times (1 - 0.34) = 0.75
\]

\[
U_T'(1) = \lambda U_M'(1) + (1 - \lambda)[1 - U_G'(1)] = 0.5 \times 0.50 + 0.5 \times (1 - 0.75) = 0.37
\]

\[
U_T(2) = \lambda U_M(2) + (1 - \lambda)[1 - U_G(2)] = 0.5 \times 0.79 + 0.5 \times (1 - 0.42) = 0.69
\]

\[
U_T'(2) = \lambda U_M'(2) + (1 - \lambda)[1 - U_G'(2)] = 0.5 \times 0.60 + 0.5 \times (1 - 0.69) = 0.45
\]

\[
U_T(3) = \lambda U_M(3) + (1 - \lambda)[1 - U_G(3)] = 0.5 \times 0.88 + 0.5 \times (1 - 0.39) = 0.74
\]

\[
U_T'(3) = \lambda U_M'(3) + (1 - \lambda)[1 - U_G'(3)] = 0.5 \times 1.00 + 0.5 \times (1 - 0.50) = 0.75
\]
Here obtain $U' T(1) < U T(1), U' T(2) < U T(2), U' T(3) > U T(3)$, thus $v'(1) < v(1), v'(2) < v(2), but v'(3) > v(3)$. Compared to the predictions in August without the preventive decisions, the predicted risks with the decision declined in all categories except the high category. Therefore, the preventive actions can be considered to be effective. The results are listed in the follow table 6.14:

<table>
<thead>
<tr>
<th>State of Rate</th>
<th>Low</th>
<th>Mid</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action Period*</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Number of new break-in cases</td>
<td>4.9, 9.58, 13.43</td>
<td>-6.655, 0.1, 6.425</td>
<td>3.3, 7.68, 12.03</td>
</tr>
<tr>
<td>Extrema</td>
<td>Sup(x) 13.43</td>
<td>Inf(1-x) -6.655</td>
<td>12.03</td>
</tr>
<tr>
<td>Ranking Utilities (λ=0.5)</td>
<td>* Note: The condition in August is considered as same as the one either in June if with no the decision or in July if with the decision.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_M = \text{sup}(\mu(x), \mu_M(x))$</td>
<td>0.84 0.50</td>
<td>0.79 0.60</td>
<td>0.88 1.00</td>
</tr>
<tr>
<td>$U_G = \text{sup}(\mu(x)-\mu_G(x))$</td>
<td>0.34 0.75</td>
<td>0.42 0.69</td>
<td>0.39 0.50</td>
</tr>
<tr>
<td>$U_T = \lambda U_M + (1-\lambda)(1- U_G)$</td>
<td>0.75 0.37</td>
<td>0.69 0.45</td>
<td>0.74 0.75</td>
</tr>
<tr>
<td>Quality Index</td>
<td>50.03%</td>
<td>33.80%</td>
<td>-0.70%</td>
</tr>
<tr>
<td>Quality of the actions (with better action precautions)</td>
<td>Good</td>
<td>Fair</td>
<td>Poor</td>
</tr>
</tbody>
</table>

Table 6.14 Linear Ordering of Fuzzy Number: left and right dominance

Results and Explanations:

- Under the preventative actions, the break-in risk is predicted to decline in a low or mid risk area. This means that the preventive actions are more effective in low break-in ratio areas.

- The explanation for this is that the preventive actions, such as block-watching and police patrol, might have not been taken previously in the region, the precautious efforts are not exhausted and effectively deters burglars from seeking targets within this community.

- According to this approximate method, the preventive actions are not effective in high risk areas, and this averts the conclusion by Kaufmann and Gupta's three
This approximate ranking is more accurate and supposed to be more reliable than those criteria.

- The explanation for this results are that the preventive actions, such as block-watching and police patrol, might have been implemented and these efforts have been exhausted previously. There are not many preventative actions that can be put into the community; or the targets are saturated in the regions, and that stalls the increase burglaries.

- The **Quality Indexes** indicate that the quality of the decision is **good, fair, and good** for the region with **low, mid, and high break-in ratios** respectively.

The study findings demonstrate that the Markov chains model has a strong capacity to deal with imprecise or uncertain data when it is incorporated with fuzzy set theory, fuzzy numbers and arithmetic, and linear approximate ranking. The model is implemented to support typical statistical methods in the social sciences such as questionnaires and surveys.

### 6.5 Summary

In Chapter 6, the case studies demonstrate the strengths devoted by fuzzy logic models and Markov chain model to *Crime Pattern Theory*, and illustrate several methods that analyze *break-in risks*. The risks can be estimated, predicted, and evaluated either in geographic locations or in time series. The study confirms that the fuzzy logic is an adaptive method that can resolve a variety of decision-support problems ranging from simple to complex. Additionally, the fuzzy logic decision support system can manipulate both objective attributes and subjective human knowledge factors.

The effectiveness of a method is determined by the nature of the problem to resolve. But for the complex problems, the best approach is to use *trial-and-error* method and find out the best combination of different methods. A *top-down* method is recommended as follows:

1. Identify the objective for all target problems,
2. Dissemble the objectives to each individual task (i.e., each environmental factors),
3) Apply trial-and-error method to find out an appropriate decision-support method (i.e., scalable monotonic chain, COA, Markov chain, etc.),
4) Apply trial-and-error method to find out an appropriate integration method such as sum of total, arithmetic mean, fuzzy logic defuzzification methods etc.

The top-down method is possibly time-consuming but will save time in developing the simple method instead of a complicated method, for example, use scalable monotonic chain instead of Markov chain, use fuzzy logic instead of neural networks.

The applicability of fuzzy logic and Markov Chain model is not particularly limited to a residence or a small area, and can be extended in larger scales (i.e., community, city, and county; bus lines, sky train stations, airport, and parks). The complexity of the techniques ranges from easy calculating the pattern of a single attribute to sophisticated analyzing the patterns of mixed factors of physical and sociological cues/signals. However, applicability of fuzzy logic models needs validation from criminological theories. The correlations between environmental factors and burglary risk are strong in high crime blocks and weak in low crime blocks, though all blocks reside in Northern Hastings district. These High crime blocks are usually close to major roads, commercial shops, schools, parks, and are where offenders’ routine activities take place. The offenders are familiar with the cues and signals in high crime blocks and unlikely to break and enter in low crime blocks. Therefore, the findings in high crime blocks might not be applicable in low crime blocks, even though the cues and signals are similar in both blocks.
7 Conclusions

7.1 Problems in Newlands’ Study and Improvements

In Table 4.1, Newlands designed Fence Height as \{under 3', 3' – 4', 4' – 6', 6'\}, the ambiguity of fence heights at \{4'\} or \{6'\} is a design error in methodology unless he did not actually classify raw data of fence height in such a way. This is a common human error in data classification. It reflects ambiguity and vagueness of criminological data: difficulty in classifying data near categorical boundary and difficulty in finding a boundary for categories. Although this is not a significant problem in fuzzy set design, this categorical error needs to be eliminated in traditional classification (crisp set).

![Figure 7.1 Clustering in high crime blocks and low crime blocks](image)

Newlands counted burglary events by street blocks, then based on break-in count he classified the street blocks in high crime blocks and low crime blocks. As illustrated in Figure 7.1, the street blocks on the left-hand side of the dark line are low crime blocks, those on the right-hand side are high crime blocks. He found that high crime blocks were primarily located along Hastings Street and near a Pub near P.N.E. The explanation is that offenders were familiar and attracted to those areas, where their activity nodes resided. Hot spots such as a Pub near P.N.E. and shops along Hastings Street might be
their activity nodes, where burglary offenses clustered. The findings of this approach do comply with Crime Pattern Theory.

The pitfalls for this approach are the difficulties in clustering burglary events and identifying hot spots or activity nodes. First, if two events are adjacent but locate on two different blocks, they may not be clustered together. Second, there are ambiguous possibilities to cluster events by distance or street block, and the choices in classification are error-prone. Third, computing resources would be wasted in generating meaningless clusters, and removing those clusters from final results is another burdensome task. The solutions are:

- First, identify possible hot spots based on a preemptive concept of Crime Pattern Theory, for example, Pubs, fast-food franchises, shops, bus stops, schools, parks etc. It has been proven that activity nodes are usually surrounding those hot spots.
- Second, cluster burglary events surrounding individual hot spots,
- Third, apply fuzzy set theory to those hot spots and resolve the ambiguity problem in clustering,
- Fourth, apply fuzzy logic theory to unveil burglary patterns in the clusters.

This approach will accelerate the clustering process and reduce ambiguity among adjacent clusters by preemptively eliminating meaningless clusters. Figure 7.2 illustrates the clustering process based on known hot spots e.g. Pandora Park, shops along Hastings Street, Callister Park and P.N.E.

Figure 7.2 Clustering based on preempted concept of hot spot
Most pattern recognition algorithms are time-resource consuming. Future study on \textit{Neural Networks}, \textit{Generic Algorithms}, or other \textit{proximity methods} can benefit from the preempted approach. By applying preempted concept of \textit{hot spot}, the computation would be sped up and resource would be invested in more meaningful clustering.

\textbf{7.2 Strengths and Limitations}

One of the main advantages of using fuzzy logic to deal with imprecise data, is that the \textit{rule} format knowledge bases are easy to examine and understand. The linguistic syntaxes such as \textit{must, below, possibly, if-then, “AND”}, and \textit{“OR”} exhibit natural simplicity, in which the knowledge base is easy to update and maintain, because the \textit{rule} format of knowledge bases is similar to the way that humans cognize and describe. It also demonstrates an inherent flexibility, where all \textit{rules} contribute to the final result if the rule conditions match or partially match. That is, the final result is derived by aggregating the hypotheses of all the rules in the risk assessment, while only one rule is activated in response to its conditions being true in conventional rule-based systems. The flexibility of the fuzzy inference is displayed when dealing with incomplete and inconsistent data. The parallel computing can be implemented based on the flexibility of the reasoning process. The rules of knowledge bases can be computed in multiple threads or modules in parallel, which may physically occur in multiple CPUs or PCs (Figure 7.3). The investigators or insurers have the ability to generate, store and process the rules separately and individually. This means the gaps of organizational structures, physical location, job shifts and schedules can be surpassed.

Through the aid of the Markov Chain model, fuzzy logic further deals with rules under multiple possibilities, multiple states, and multiple concerns of confidence, which fuses the objective knowledge (empirical model) and the subjective knowledge (users’ inputs of prediction from surveys or votes). Many existing empirical models or equations can be converted into an intelligent decision-making system.
The main shortcoming of this approach is that the membership functions, rules, transition matrices, and reward matrices have to be specified manually, which is time-consuming due to the trial-and-error process. Further, the elicitation of rules from human experts can be an expensive and error-prone procedure. Moreover, the transplantation of the rules is difficult. Therefore, they cannot adapt automatically to changes in the operations. New rules have to be manually altered if the jurisdiction conditions changes. However, the use of intelligent hybrid systems (e.g., neural networks, genetic algorithms) may ease this difficulty.

The second problem is common for the complex knowledge/decision support systems and is known as the combinatorial rule explosion problem. The fuzzy rules are usually stated as the intersection of the input subsets. The number of the rules changes exponentially with the number of the inputs. For example, in a two-input and one-output system, the Building Age input has four subsets (i.e., Under Const., New, Average, Old), and the Break-In-Frequency input has three subsets (i.e., One+, Two+, Three+). If a rule matrix contained all possible rules, then the total number of rules is 12. In a three-input and one-output system, an additional input Fence Type has six subsets (i.e., None, Open Wooden, Closed Wooden, Chain Link, Metal, Cement-Rock), then the total number of rules is 72. In a six-input and one-output system, where any input has five subsets, then the total number of rules is 15625. The combinatorial rule explosion problem significantly downgrades system performance. Researchers have found some effective
methods to ease the problem, but a careful design is strongly recommended to reduce the system complexity and only select those rules that make a significant contribution to the inference output. [6]

7.3 Fuzzy Logic Model Versus Probabilistic Model

Fuzzy logic was developed to handle real-world situations where either numeric data were incomplete or inconsistent, or were supplemented with linguistic characterization (e.g. expert opinion). Recently, some researchers have launched their efforts to purely numeric data. But Burr pointed out that those efforts actually benefit from the maturity of statistically motivated methods for numeric data [33, pp122]. Further he compared fuzzy logic methods and statistical methods in time-series forecasting and ironically concluded that fuzzy logic was developed for deterministic models [33, pp126]. As a result of that fuzzy logic was not intended to compete with statistical methods for numeric data, Burr’s conclusion is partially correct, and this is why probabilistic models outperform fuzzy logic models for numeric data. But his conclusion is insufficient, because his fuzzy logic models did not include any stochastic component and his fuzzy forecaster does not use all of training data. However, fuzzy logic models have natural simplicity to linguistic characterization of phenomena and have advantages on criminological data. On the other hand, hybrid synergism such neuro-fuzzy offers a new and complementary way to model uncertainty and deserves more attentions from researchers.

7.4 Contributions

The methodology and modeling of fuzzy logic and Markov chains has been discussed and applied in a variety of cases related to break-in offence patterns. The major contribution of this study is that the approach can be used in realistic industrial applications and academic research. The prototypes have been implemented in MATLAB and Visual C++, and are ready to be embedded into a particular application.
7.5 Future Research

The first area of future research is synergism of AI technologies. This study shows that the fuzzy logic decision-making system has a strong capacity to process imprecise and vague data, linguistic rules, and ranking data or data sets. But it has less of a capability to search for patterns in data sets or to be trained to learn a pattern and predict a trend, which is the domain of neural networks, or genetic algorithms, or neuro-fuzzy methods.

A second area of future research is to search for patterns in a data warehouse, a process called data mining. Algorithms such as inferring rudimentary rules, statistical modeling, instance-based learning must be studied. Data warehousing is being increasingly used as the basis for a decision support system.

A third area of future research involves the study of geographic information systems (GIS). A Geographic Information System (GIS) is a computer system used to assemble, store, manipulate, and display data which contains physical locations (geographic coordinates) of features and information about those features (attribute data). The spatial analysis of the crime pattern theory requires GIS technology. Fuzzy logic theory has a strong capability to solve clustering and proximity problems.

Fourth, future research is needed to study the decision-making process of offenders. Our research analyzed the physical environmental restraints of residences. Bennett and Wright discussed offender decision-making mechanism on the decision to offend, offence planning, and displacement [1]. Through the aid of AI technologies, the decision-making mechanism will be understood in-depth based on knowledge, past experience, degree of rationality, perception of opportunity, motivation of gains, etiological effects (i.e., drug, alcohol, illness). This approach will benefit from the technique of knowledge-base intelligent systems. New research directions can start with different decision models such as the rational decision model, the intuitive decision model, and the cognitive decision model.
Bibliography


[9] Patricia L. Brantingham, “*Environmental Design – Added Dimension to Policing?*” *Liaison*, Communication Division, Programs Branch, Ministry of the Solicitor General, 6:5 (2-8), Ottawa, ON, 1980


