International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 139-156 June 2018-Dec 2018 www.csjournals.com

Detecting Malicious Behavior Of Network Through Automated Discovery Of Fuzzy Decision Rules

[1]Janki Arora, [2]Mukesh Yadav
[1]Student, Assistant Professor[2]
Department of Computer Science & Engineering, Gurgaon Institute of Technology and Management, Gurgaon(Hr.), India
jankiarora1@rediffmail.com , mukesh.cs@ieee.org

Abstract: The advances in data collection have generated an urgent need for techniques that can intelligently and automatically analyze and mine knowledge from huge amounts of data. The Knowledge Discovery in Databases (KDD) is the process of extracting the knowledge from huge data collection. Data mining is a step of KDD in which patterns or models are extracted from data by using some automated techniques. Discovering knowledge in the form of classification rules is one of the most important tasks of data mining. Recently there have been several applications of genetic algorithms for successful discovery of concise and comprehensible rules with high predictive accuracy.

Further Fuzzy Logic is a well proven tool to handle the vagueness and uncertainty inherent to decision making systems. To be able to make effective decisions with uncertain and vague information, fuzzy logic has to be incorporated in the discovery of classification rules.

This paper addresses the discovery of knowledge in the form of Fuzzy Classification Rules (FCRs) using an evolutionary approach to detect intrusion in a network. In this work, application of GA in the fuzzy environment is preferred over other searching techniques as it is a promising optimization technique to find better or optimal solutions due to global searching. The proposed GA design for discovery of FCRs includes the following:

1. A triangular fuzzy membership function is applied for fuzzifying the attributes.
2. An appropriate encoding scheme to represent FCRs is suggested.
3. An effective fitness function for measuring the goodness of fuzzy rules is devised.
4. Genetic operators with syntactic constraints have been designed for exchanging and creating new genetic material for the rules.

The proposed approach is applied to KDD Cup 1999 datasets from UCI Machine Learning Repository and the experimental results are presented. The discovered rules in the form of FCRs establish the effectiveness of the proposed system.

Keywords: Intrusion Detection System (IDS), Fuzzy Classification Rules (FCRs), Genetic Algorithm (GA), Knowledge Discovery in Databases (KDD).

1. Introduction

1.1 Need of Automation

In the last few decades have seen a staggering growth in the amount of data being generated and gathered by humans. Databases are precious treasures. A database not only stores and provides data but also contains hidden precious knowledge, which can be very important (Wong and Leung, 2000; Dunham, 2003). Contributing factors include the computerization of many business, scientific and government transactions, and advances in data collection tools ranging from scanned text and image

platforms to satellite remote sensing systems. In addition, popular use of the World Wide Web as a global information system has flooded us with a tremendous amount of data and information.

1.2 Knowledge Discovery in Databases and Data Mining

Knowledge discovery in databases (KDD) and data mining are swiftly evolving areas of research that are at the intersection of many disciplines. KDD is the process of finding useful information and patterns in data. It starts with the understanding of the problem and concludes with the analysis and assessment of the results. The input to this process is the data, and the output is the useful information required by the users (Fayyad et al., 1996a; Mitra et al., 2002).

1.2.1 The KDD Process: The KDD process includes two steps:

a) Pre-processing [or Data Preparation] step: The goal of data preparation methods is to transform the data to facilitate the application of given data mining algorithms.

b) Post processing [or Knowledge Refinement] step: The goal of knowledge refinement methods is to validate and refine discovered knowledge. The KDD process is both interactive and iterative, involving numerous steps with many decisions being made by the user.

KDD is iterative because the output of each step is often feedback to previous steps as shown in Fig.1.1 and typically many iterations of this process are necessary to extract high-quality knowledge from data.

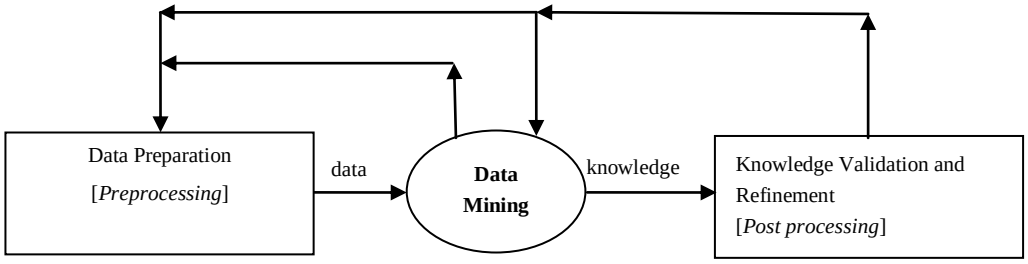


Fig.1.1: The iterative nature of the knowledge discovery process.

KDD as a process is depicted in Fig.1.2 and consists of an interactive sequence of the following steps (Han and Kamber, 2001):

- **Data cleaning**: Used to remove noise and inconsistent data. Erroneous data may also be corrected or removed.
- **Data integration**: Used to combine data from multiple data sources.
- **Data selection**: During this process data relevant to the analysis task are retrieved from the database.
- **Data transformation**: In this process data are transformed or consolidated into forms appropriate for mining by performing operations like summary or aggregation.
- **Data mining**: It is an essential process where intelligent methods are applied in order to extract data patterns.
- **Pattern evaluation**: Used to identify the truly interesting patterns representing knowledge based on some interestingness measures.

- **Knowledge presentation**: In this visualization and knowledge representation techniques are used to present the mined knowledge to the user.

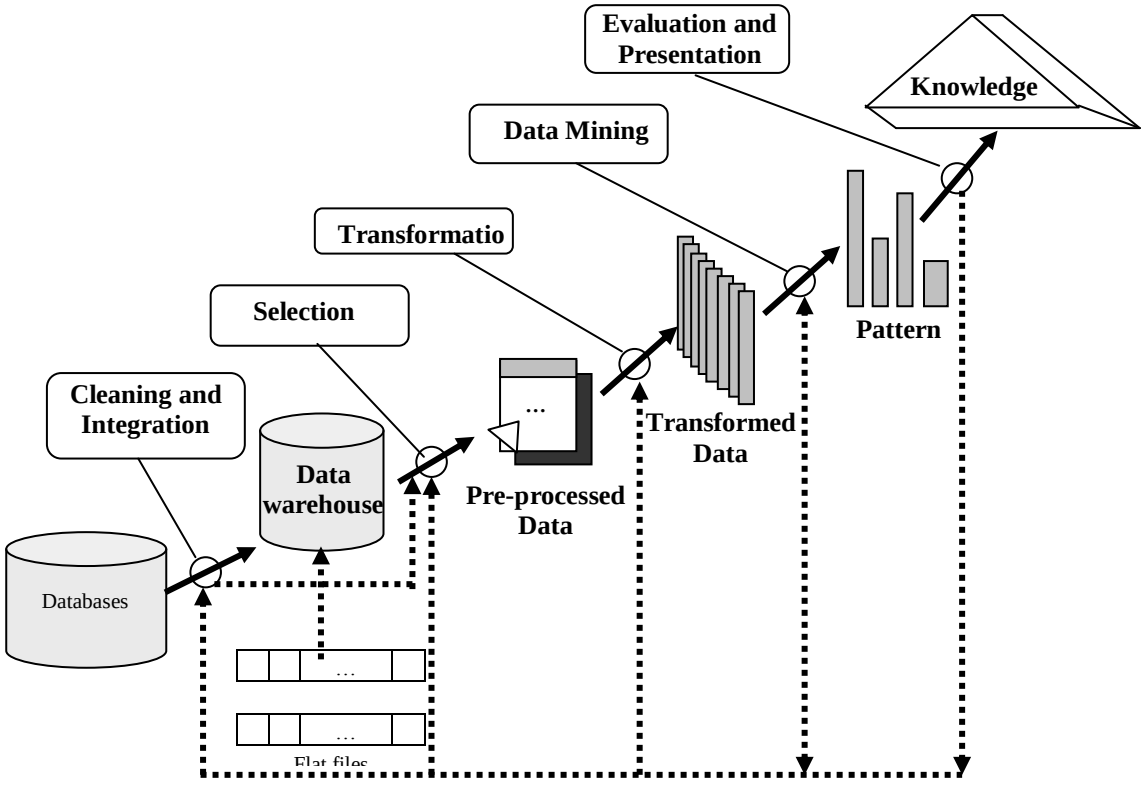


Fig.1.2: Data mining as a step in the process of knowledge discovery(Han and Kamber, 2001).

1.2.2 Data Mining Tasks

In literature, data mining tasks are also referred to as data mining outcomes or types and can be classified into two categories: descriptive and predictive. Descriptive mining tasks characterize general properties of the data in the database. Examples include association rule discovery and clustering. On the other hand, predictive mining tasks perform inference on the current data in order to make predictions. Examples of predictive mining tasks include classification and regression.

In the Fig.1.3(E) groups (A) and (C) behave similarly.

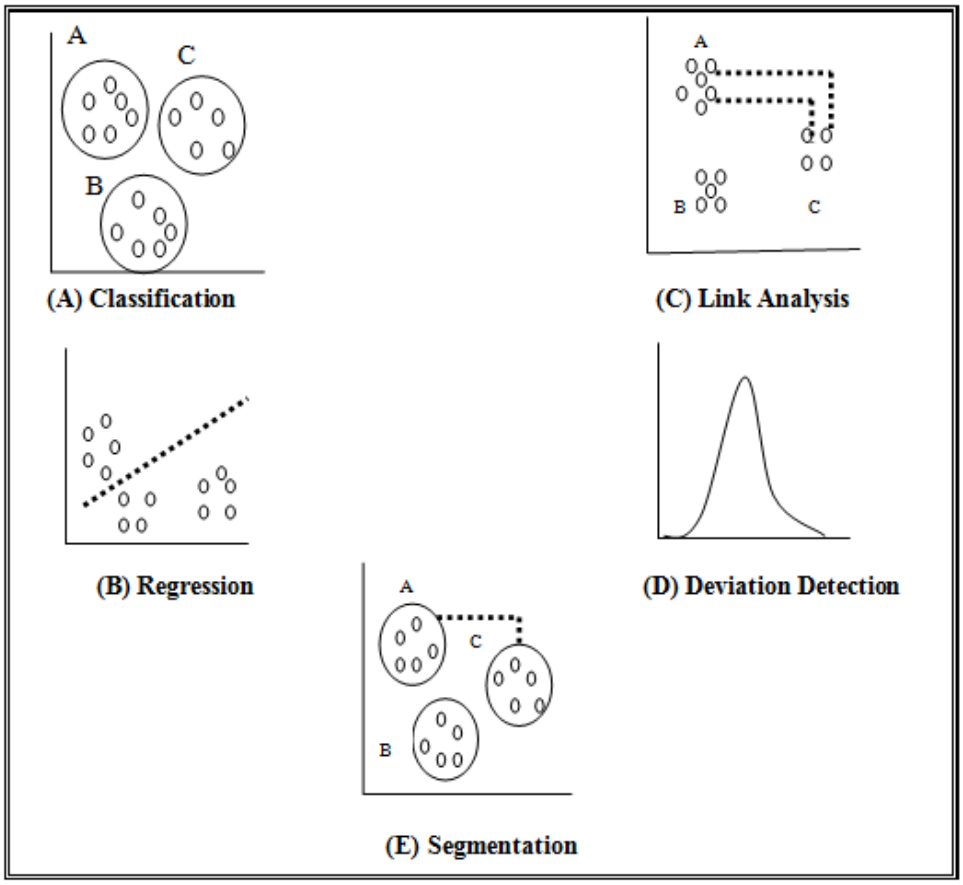


Fig.1.3: Data mining tasks.

These data mining tasks are typically used in combination with each other, either in parallel or as a part of sequential operation.

1.2.3 Data Mining Methodologies

These are the steps which one would follow to do mining. Once the outcomes are determined then data mining can be done. There are two approaches for data mining: top down and bottom up. One could combine the two and have a hybrid approach. The top down approach starts with some idea or a pattern or hypothesis. In the bottom up approach of data mining there is no hypothesis to be tested. This is much harder as the tool has to examine the data and then come up with pattern. The bottom up approach could be directed or undirected. The hybrid approach is a combination of both top down and bottom up mining. In this the tool can switch between top down and bottom up mining and again between directed and undirected mining.

1.2.4 Data Mining Techniques

There are numerous Data mining techniques which lay down the basis for an important part of data mining and include algorithms and techniques applied for mining data. Various Data mining techniques illustrated in Fig.1.4.

Data Mining Techniques	Data Mining Methods				
	Classification	Regression	Segmentation	Link Analysis	Deviation Detection
Inductive Logic Programming	X	X			
Genetic Algorithms	X	X	X		
Neural Networks	X	X	X		
Statistical Methods	X	X	X	X	X
Decision Trees	X		X		
Hidden Markov Models	X				

Fig.1.4: Data mining techniques

These techniques have been taken from statistics, data management and machine learning and there is some overlap between these techniques.

1.3 Classification: Classification is the process of learning a mapping between elements from feature space to one of several predefined classes. Once a classification model is built from training data, it can be used to classify new data. Once the model is built, then it can be used to classify new data. Example, in a banking application, customers who apply for a credit card may be classify as a "good risk", a "fair risk" or a "poor risk". Hence, this type of activity is also called supervised learning.

1.4 Objective

The simple structure of production rules fails to capture various aspects of the real world knowledge. The objective of this thesis is to discover knowledge (classification rules) in the form that can handle uncertainty, vagueness and ambiguity. Knowledge discovery is a computationally intensive task. In this thesis, a genetic approach is proposed for mining rules in the form of Fuzzy Classification Rules (FCRs) which are capable of dealing with large databases.

1.5 Evolutionary Algorithms

The Evolutionary Algorithms (EAs) is an umbrella term that describes computer-based problem-solving systems that use computational models of evolutionary processes as key elements in their design and implementation. Various EAs that have been put forwarded and widely reported are shown in Fig.1.5.

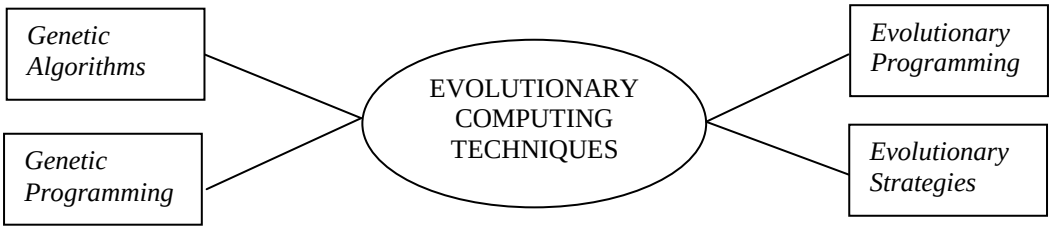


Fig 1.5: Evolutionary Computing Approaches

2. BACKGROUND AND LITERATURE REVIEW

2.1 Methods for Discovering FCR: Data classification represents an important theme and is perhaps the most commonly applied data mining technique (Fayyad et al., 1996). The classification problem becomes very hard when the numbers of possible different combinations of parameters are so high that techniques based on exhaustive search of the parameter space rapidly become computationally infeasible. Thus, it is natural to devote attention to a heuristic approach to the classification problem (Falco et al., 2005). Traditionally, classifier algorithms focused either on accuracy or interpretability. But fuzzy classifier takes into account both performance and interpretability so as to keep rule bases small and comprehensible (Roubos et al., 2001).

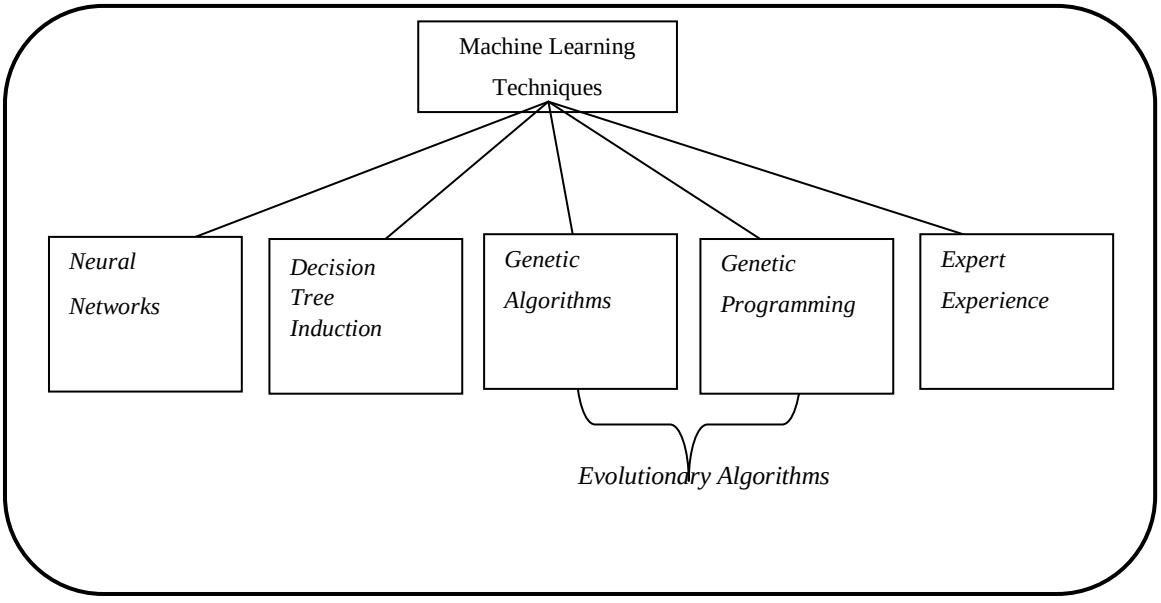


Fig.2.1: Taxonomy of Machine Learning Techniques

Fig.2.1 represents various Machine Learning techniques that have been applied for automated discovery of FCRs.

2.1.1 Neural Networks

Neural Networks (Lippmann, 1987) can also be used to generate fuzzy decision rules as both neural networks and fuzzy systems are functionally equivalent despite of different structures. This functional equivalence has been explained and proved by Buckley et al. (1993) in the sense that any continuous, layered feedforward neural net can be approximated to any degree of accuracy by a fuzzy logic system and any continuous, discrete fuzzy logic system can be approximated to any degree of accuracy by a three-layered, feedforward neural net. Hayashi and Imura (1990) suggested a two-step procedure to extract fuzzy rules. In the first step, a neural net is trained from sample data and in the subsequent step an algorithm is used to automatically extract fuzzy rules from the trained neural net. Many other methods of generating fuzzy rules through neural networks have been suggested. Kosko (1992) proposed a system called Fuzzy Cognitive Maps, which integrates neural network and fuzzy logic. A fuzzy logic system can also be constructed directly with neurons representing fuzzy logic and linguistic terms. Lin and Lee (1991) proposed a neural-network-based fuzzy logic system which consists of five layers. The first layer is the input layer. Each node in this layer represents a linguistic variable. The second and fourth layers contain term nodes which act as membership functions to

represent the terms of the respective linguistic variable. The third layer is the rule node layer. Each node with its connection represents a fuzzy rule. The fifth layer is the output layer. On this layer two nodes are used for each output variable. One is for the desired output and the other is for the actual output. The system can be constructed from training examples with a hybrid learning scheme where a self-organized clustering learning method is used to locate fuzzy membership functions for the input and output terms. In this a competitive learning method is used to determine the connection of rule nodes and a supervised learning method is used to tune membership functions in order to improve the system's performance.

Though neural network approach is suitable in building a fuzzy logic system with a relatively small number of numerical variables but lack of analytical guidance in determining the network configuration and trap of local optimal in the learning process limits its applicability in fuzzy environment.

2.1.2 Decision Tree Induction

Another widely used machine learning method is the induction of decision trees (Quinlan, 1986; Safavian and Landgrebe, 1991). The search method decision tree induction employs fuzzy entropy to find the most efficient decision nodes (Weber, 1992; Yuan and Shaw, 1995). Although in most of the cases this method works well but still is rendered inefficient for generating fuzzy decision tree, as it may not be able to generate best tree due to one-step ahead node splitting without backtracking. Moreover, the best tree may not be able to yield the best set of rules.

2.1.3 Evolutionary Algorithms

The abovementioned search techniques, neural networks and decision tree induction have problems of being trapped into local optimal. Thus, application of EAs for discovering comprehensible IF-THEN classification rules in the fuzzy environment is preferred over other searching techniques. There has been very extensive research on EAs for discovering FCRs (Cordon et al., 2004; Walter and Mohan, 1999). Simple taxonomy of Fig.2.2.represents three different evolutionary approaches used for discovering FCRs.

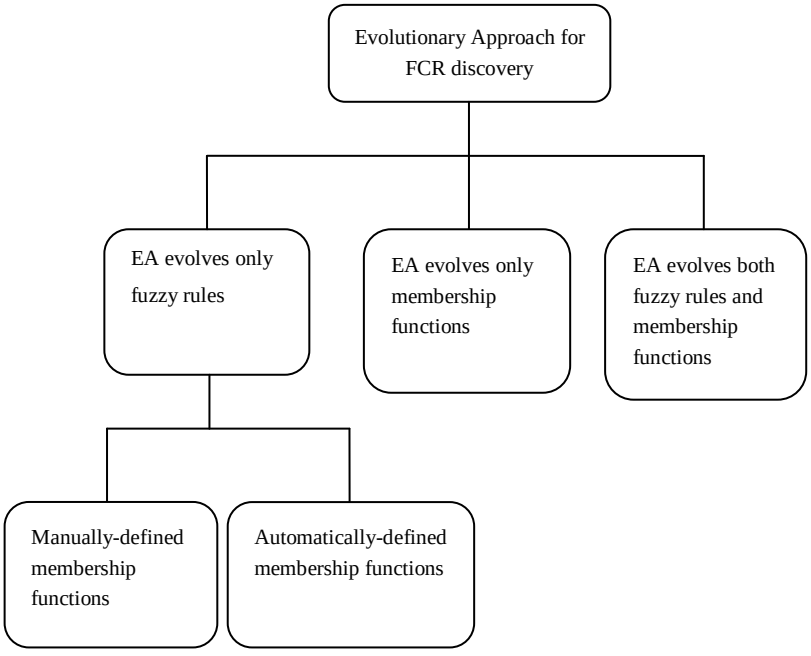


Fig.2.2: Taxonomy of using an EA for discovering FCRs

3. Research Directions

Several directions for the future research which aim at discovering truly interesting FCRs are:

1) Reduction of number of Misclassifications
   Though researchers have been able to discover compact and accurate classifiers, further experiments are needed to reduce the number of misclassifications as accuracy of the model depends on the misclassification rate (Roubous et al., 2001).

2) Improvement of Selection mechanism

   Mansoori et al. (2008) has stated that there is needed to improve the selection mechanism as the non-random selection mechanism cannot utilize rules co-operation while producing offspring. Moreover, Mendes et al. (2001) have suggested that further experimentation must be carried out for designing "Intelligent" operator which selects tree nodes to be pruned on the basis of predictive power unlike "Blind" operator used for tree pruning which selects nodes randomly. Also, performance needs to be improved as current algorithm may not cope with data sets having large number of attributes

3) Fuzzification of other relational operators
   As evolutionary system called RFCRD proposed by Akbarzadeh (2008) fuzzified only "greater than" and "less than" operator so fuzzification of other relational operators can be an interesting search. Another important research extension lies in the replacement of the GP algorithm with the Gene Expression Programming (GEP) algorithm. This substitution must be carried out because of difference in evolutionary strategies of both techniques. GEP allows evolution of multi-genic chromosomes i.e. each gene codes for a different sub-expression or a specific classification rule set. For instance, if a problem includes n classes i.e. parse tree must yield n classification rule sets than in case of GP, system must run n times. On the other hand, by using multi-genic GEP system, any given run would simultaneously evolve n classification rule sets.
4) Comparative study of results
   Another significant future search would be comparative study of accuracy of FCRs evolved by various machine learning techniques with the results explored by other traditional rule induction and decision-tree-induction methods. Like in case of genetic system proposed by Romao et al. (2002) future direction lies in comparative study of degree of interestingness of rules discovered by GA and by another Data Mining algorithm.
5) Reducing time complexity of Evolutionary Algorithms
   Evolutionary algorithms have been widely applied in data mining. However, use of genetic algorithms is somewhat restricted due to their typically large running time.
6) Use of filtering techniques
   One of the very important research directions is to improve the efficiency of GAs by employing filtering techniques to reduce search space and augmenting GAs with cache like memory to reduce the number of fitness evaluations.

4. DISCOVERY OF FUZZY DECISION RULES

This section presents a classification algorithm based on evolutionary approach that discovers comprehensible decision rules in the form of fuzzy classification rules (FCRs) to detect the malicious behavior of the network. The main intent behind integrating fuzzy logic (FL) with genetic algorithm (GA) is to cope with real world cognitive uncertainties such as vagueness and ambiguity involved in classification problems. Moreover, instead of using numeric values or ranges in rules, three fuzzy linguistic variables as small, medium and large are used for making the discovered rules more comprehensible. The proposed approach has flexible chromosome encoding, where each chromosome corresponds to an FCR. Appropriate genetic operators are suggested and a suitable weighted fitness

function is proposed that incorporates the basic constraints on FCRs to measure the goodness of result sets.

4.1Fuzzy Classification Rules

The standard PRs in the form **If** P **Then** D**,** are the most predominant method of representing the discovered knowledge**.** The PRs, however, are unable to handle the inherent uncertainty and vagueness prevalent in the real world knowledge.

Traditional concept hierarchies usually represent knowledge using crisp description. A crisp structure assumes that the particular element belongs to its crisp set or discourse either with degree 1 or degree 0 i.e. with only one out of two possible degrees of membership: 1 or 0. However, a concept description for human knowledge is vague generally and a crisp description of concept cannot represent human knowledge completely. For a concept, fuzzy description is more appropriate than crisp description in real world with clearly defined boundaries. Thus, for a concept fuzzy description reflecting partial belonging of one item to another is more appropriate than crisp description in real world (Chian et al., 2004). In a fuzzy structure, a particular element belongs to a fuzzy set to a certain degree called the degree of membership, typically represented by a real-valued number in the interval [0…1]. For instance, consider a rule for customer credit application approval:

$$\text{If}(\text{years\_employed} \geq 2)\text{AND}(\text{income} \geq 50\text{K})$$

$$\text{Thencredit} = \text{approved}$$

This rule says that customer applications with a job experience for atleast two years and has high income of atleast $50,000 are approved but not if his income is $49,000.

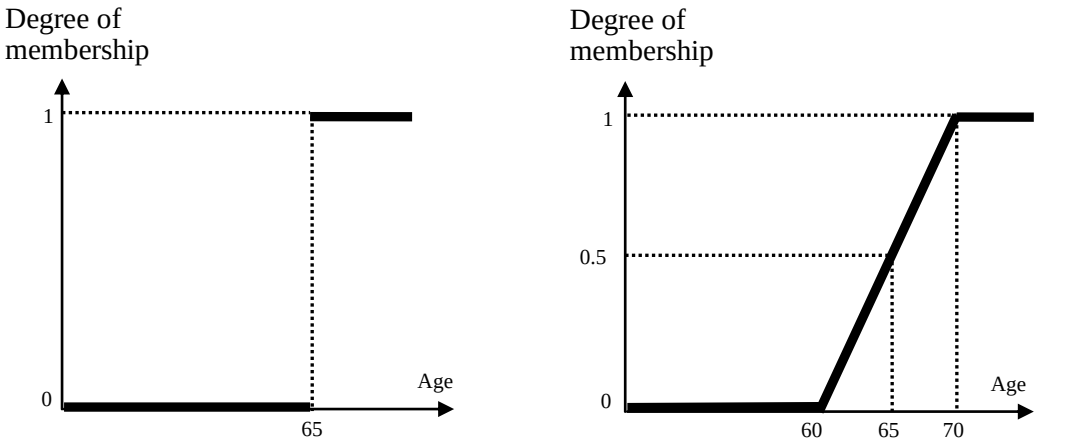
Such harsh thresholding seems unfair so an alternate solution is to overcome the shortcomings of traditional approach is to integrate fuzzy logic (FL) with a genetic approach for classification task to discover FCRs. The concept of fuzzy logic was proposed by Prof. Lotfi A. Zadeh. FL is a methodology for computing with words. Computing means manipulation of numbers and symbols. The FL was developed to overcome the inefficiency of classical sets and computer logic which arises due to its inability to manipulate data from vague [not clear] human ideas. FL allows computers to determine the difference among the data with shades of grays.

The process of FL similar to human reasoning which leads to an obvious benefit of incorporating domain knowledge of user thereby, resulting in a more comprehensible membership function with fuzzy thresholds despite of gradual boundaries. Thus, rather than yielding "precise cut-offs" between categories, FL uses truth values between 0.0 and 1.0 to represent degree of membership that a certain value has in a given category. For instance, income $49,000 is more or less high although not as high as an income of $50,000.

Thus, concept of fuzzy sets can be considered as a generalization of the concept of crisp sets, as in case of fuzzy sets the degree of membership can take on any value in the continuous interval [0...1], whereas in case of crisp sets the degree of memberships can take on only the value 0 or 1. For example, consider two alternative definitions of the set of old people.

A crisp definition could be: a person belongs to the set of old people if and only if$\text{age} \geq 65\text{years}$. Fig.3.1(a). illustrates this definition. The problem with this definition is that there is a very abrupt change in the status of a person, from not oldto old. In the day of person's 65-year birthday the person status abruptly changes from not old to old, this clearly is not a reasonable description of reality.

In contrast, a fuzzy definition would recognize that people gradually get older. A possible definition of fuzzy set old is illustrated in Fig.3.1(b). In this example a person whose age is less than 60 years is considered old to a degree of 0. From that age onwards the degree of oldness of a person gradually increases, until the person reaches the age of 70 years. From that age onwards the person is considered old to a degree of 1. This is a more reasonable description of reality than the crisp description of Fig.3.1(a).



(a) crisp definition of old (b) fuzzy definition of old

Fig.4.1: Crisp versus fuzzy definitions of the set of old people

Fig 4.1(b) illustrated how one can fuzzify an attribute value such as old. In practice one fuzzifies all the values of an attribute. For instance the attribute age can be fuzzified into three linguistic values namely, young, middle-aged and old.

Proposed System

The proposed system comprised of two parts as shown in Fig 4.2:

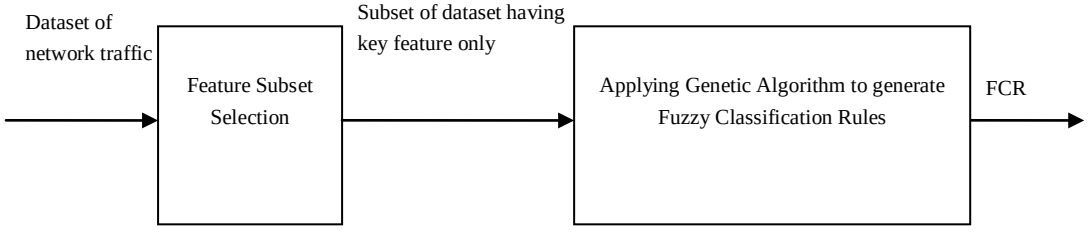


Fig.4.2. The proposed system

1) **Feature Subset Selection:** A reduction in the number of key features is necessary to improve performance in terms of learning time, classification accuracy, and comprehensibility of the learned rules.
2) **Applying GA to generate FCR:** In this section GA approach is presented for the automated discovery of FCRs as the underlying knowledge representation. The evolutionary system is able to acquire information from datasets and extract comprehensible classification rules for each available class, given the values of some attributes, called predicting attributes.
3) The major steps of this evolutionary system can be formalized as follows (Falco et al., 2005):

Discovery of FCRs (D)

Input: Dataset D

Output: FCRs

i. Generate at random an initial population of rules representing potential solutions to the classification problem by implementing following steps over the dataset D.
    1. Normalize the population.
    2. Fuzzify the population.
    3. Encode the population.

ii. Evaluate each rule on the basis of an appropriate fitness function.

iii. Select the rules to undergo the mechanism of reproduction.

iv. Apply the genetic operators, such as crossover and mutation, to generate new rules.

v. Reinsert these offspring to create the new current population.

vi. Repeat steps (ii) to (v) until no further improvements occur or a fixed maximum number of generations have been reached.

The overall proposed system for automated mining of rules is represented in Fig.3.3. as:

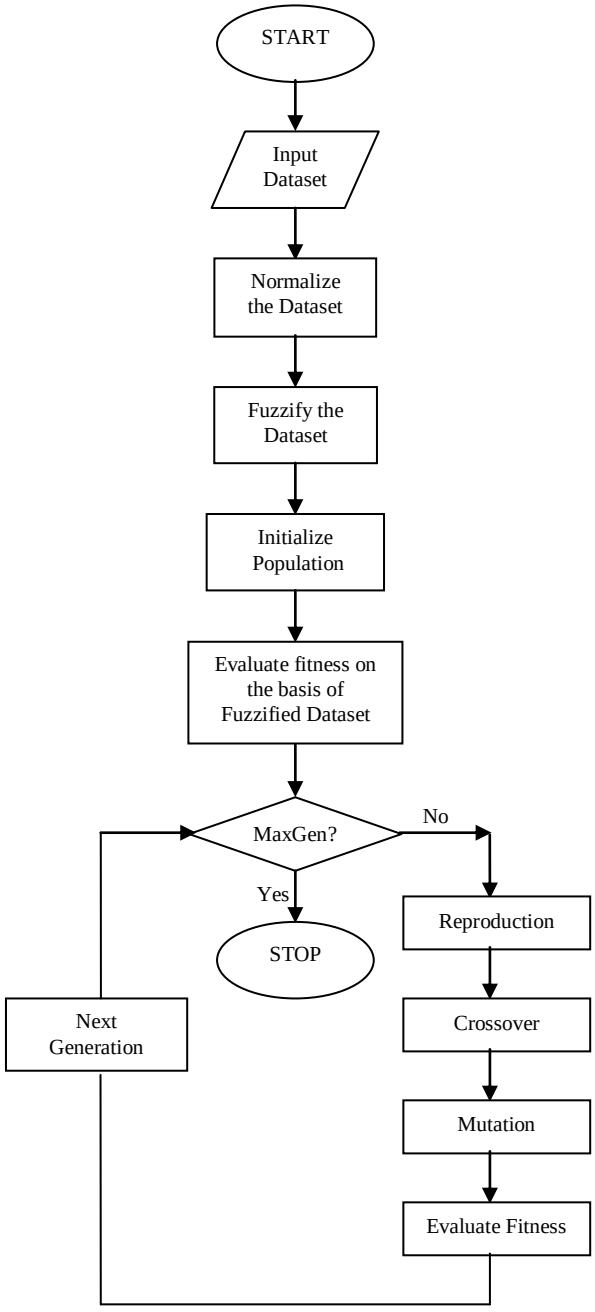


Fig.3.3. The proposed system for rule mining

5. EXPERIMENTAL SETUP AND RESULTS:

This section focuses on the evaluation of the proposed system over real-valued datasets taking into account metrics like predictive accuracy and comprehensibility used for evaluating the rule. Predictive accuracy measures the accuracy of the rules extracted from the dataset whereas, comprehensibility is measured by the number of attributes involved in the rule and tries to quantify the understandability of

the rule. Experiments were carried out taking into consideration these measures as the objective of underlying dissertation and the experimental results obtained demonstrate the effectiveness of the proposed system for automated discovery of fuzzy decision rules to detect network intrusion.

5.1 Experimental Setup: The proposed GA approach is implemented using GALIB247 on a Pentium core 2 duo processor with Ubuntu release 10.10 as operating system. The performance of the suggested approach is validated on KDD Cup 99 real-valued datasets publically available at UCI [University of California at Irvine] machine learning repository and its corresponding site is ftp://ftp.ics.uci.edu/pub/machine-learning-databases/. GALIB247 is freely available at http://lancet.mit.edu/ga/dist/.The implementation code is compiled using g++ compiler.

5.1.1 Description of the Datasets: Since 1999, KDD Cup 99 has been the most wildly used data set for the evaluation of anomaly detection methods. This data set is prepared by Stolfo et al. and is built based on the data captured in DARPA'98 IDS evaluation program. DARPA'98 is about 4 gigabytes of compressed raw (binary) tcpdump data of 7 weeks of network traffic, which can be processed into about 5 million connection records, each with about 100 bytes. The two weeks of test data have around 2 million connection records. KDD training dataset consists of approximately 4,900,000 single connection vectors each of which contains 41 features and is labeled as either normal or an attack. KDD training dataset is very large so I am using 10% KDD training dataset. The feature subset selection is performed on 10% KDD training dataset so that 41 attributes reduces to 8 attributes. Finally the redundant records are eliminated from the dataset using Microsoft Office Excel 2007.The removal of redundant records reduces the number of instances from 494021 to 33517.

5.1.2 Simulation Parameters: In order to evaluate the presented methodology the simulation is carried using the simulation parameters listed in Table 5.1.

Table 5.1 Simulation Parameters

Sr. No.	Parameters	Values
1	Population Size (N_{pop})	100
2.	Crossover Probability (P_c)	0.66
3.	Mutation Probability (P_m)	0.1
4.	Maximum Generations(max_gen)	100

The parameters are tuned in the few initial runs of the GAs and the proposed algorithm was terminated when the best Fitness did not change continually throughout 10 generations.

5.1.3 Experimental Results: The desired rule set was discovered by carrying out the simulation process using the simulation parameters given in Table 4.1. Table 5.2 shows the discovered rule set.

Table 5.2. Experiment Results

KDD CUP 99 DATASET			
Sr No.	Rules: PRs	Coverage	Fitness
1.	if (((Src_Bytes =Small) || (Src_Bytes =Medium)) && ((Num_Failed_Logins =Small) || (Num_Failed_Logins = Medium)) && (Root_Shell =On) && (Num_Access_Files =Small) && ((Srv_Count =Small) || (Srv_Count =Medium)) && (Serror_Rate =Small)) then Attack.	0.946643	0.667197
2.	if (((Src_Bytes =Small) || (Src_Bytes =Large)) && ((Root_Shell =On) && ((Srv_Count =Small) || (Srv_Count =Medium)) && ((Serror_Rate =Small) || (Serror_Rate =Medium)) then Attack.	0.00144207	0.667197
3.	If (((Duration=Small) || (Duration=Medium)) && ((Src_Bytes =Small) || (Src_Bytes =Large)) && ((Num_Failed_Logins =Small) || (Num_Failed_Logins =Medium) && (Root_Shell =On) && ((Num_Access_Files =Small) || (Num_Access_Files =Large)) && ((Srv_Count=Small) || (Srv_Count =Medium)) && (Serror_Rate=Small) && (Same_Srv_Rate =Large)) then Attack.	0.00546992	0.667197
4.	If (((Duration=Small) || (Duration=Medium)) && (Num_Failed_Logins =Small) && (Root_Shell =On) &&(Srv_Count =Small)) then Attack.	0.0362506	0.66718

4.1.4 Predictive Accuracy

Table 4.3 lists the predictive accuracy of the discovered rule sets for different initial population size and maximum generation of the experimental datasets.

Table 5.3. Predictive Accuracies of Experimental Dataset

Intial Population Size	Maximum Number of Generations	Predictive Accuracy	Final number of rules generated	Time Taken (in milliseconds)
50	50	0	1	18979
75	50	95	3	28690
75	75	90	1	41971
75	100	95	4	59997
100	100	98	4	79161
200	200	98	4	452906
500	500	98	5	2450272

6. CONCLUSION AND FUTURE CONSIDERATIONS

In recent years there has been increasing interest in applying Evolutionary Algorithms (EAs) to Knowledge Data Discovery. The work presented in this thesis has demonstrated successful application of GAs for automated discovery of Fuzzy Decision Rules to detect malicious behavior of a network. The underlying knowledge representation is capable of handling uncertainty and vagueness inherent to decision making support systems.

A genetic approach is proposed for the discovery of decision rules in the form of Fuzzy Classification Rules (FCRs) that can efficiently cope with vague data which do not have crisp boundaries between them. The proposed scheme has flexible chromosome encoding and appropriate crossover and mutation operators have been described. Keeping in view the basic constraints on FCRs, an appropriate fitness function is formulated so as to make the task of rule mining easier. This work has integrated fuzzy logic with genetic algorithm based approach for the automated discovery of Fuzzy Decision Rules to identify an attack in a network. The performance of this proposed algorithm is tested across KDD Cup 99 datasets and the results are quite encouraging and have established the effectiveness of the proposal. The scheme provides a mechanism to discover concise and comprehensible classification rules in the form of FCRs.

It generally takes time for techniques to mature and become robust and effective for use in real world problems. There is always some scope for the improvement. The proposed work can also be further explored in the light of the following suggestions:

a) Comparative study of various types of fuzzy rules

A comparative analysis of the two kinds of fuzzy rule-types Mamdani-type fuzzy rule and Sugeno-type fuzzy rule can be carried out; as in our proposed work we have explored Mamdani-type fuzzy rule only.

b) Discovery of Production Rules with Exceptions

One of the most important extensions of present work would be development of EAs for the automated discovery of Fuzzy Censored Production Rules (FCPRs) (Saroj and Bharadwaj, 2009) from large datasets.

c) Reducing Time and Space Complexity

An efficient feature selection method can be used to produce attributes that are more appropriate to produce desired output in a comparative less amount of time or memory.

REFERENCES

[1] A. A. Freitas,2002, "Data Mining and Knowledge Discovery with Evolutionary Algorithm", Natural Computing Series, Springer-Verlag, New York, USA.

[2] A. A. Freitas, 2003, "A survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery", Advances in Evolutionary Computation Theory and Applications, Springer-Verlag, New York, USA, pp. 819-845.

[3] B. Kosko, 1992, "Neural Networks and Fuzzy Systems", Prentice-Hall, Englewood Cliffs, NJ.

[4] B. Liu, W. Hsu and S. Chen, 1997, "Using General Impressions to Analyze Discovered Classification Rules", In Proceedings of the 3rdInternationalConference on Knowledge Discovery and Data Mining(KDD-97), Newport Beach, CA, USA, AAAI Press, PortlandOregon, USA, pp. 31-36.

[5] C. Mota, H. Ferreira and A. Rosa, 1999, "Independent and simultaneous evolution of fuzzy sleep classifiers by genetic algorithms", Proc. Genetic and Evolutionary Computation Conf. (GECCO-99), Morgan Kaufmann, pp. 1622-1629.

[6] C. T. Lin and C. S. G. Lee, 1991, "Neural-Network-based fuzzy logic control and decision system", IEEE Trans. Comput., pp. 1320-1336.

[7] D. E. Goldberg, 1989, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley Publishing Company, Inc. MA, New York.

[8] D. Walter and C. K. Mohan, 2000, "ClaDia: A fuzzy classifier system for disease diagnosis", In Proc. Congress on Evolutionary Computation (CEC-2000), La Jolla, CA, USA, vol. 2.

[9] E. Noda, Alex A. Freitas and H.S. Lopes, July 1999, "Discovering Interesting Prediction Rules with a Genetic Algorithm", In Proc. Congress on Evolutionary Computation (CEC-99), Washington D.C., USA, pp. 1322-1329.

[10] Eghbal G. Mansoori, Mansoor J. Zolghadri and Seraj D. Katebi, Aug.2008, "SGERD: A Steady-State Genetic Algorithm For Extracting Fuzzy Classification Rules From Data," IEEE Trans. Fuzzy Syst., vol.16, no. 4, pp. 1061-1071.

[11] Fries, T.P. , July 2010,"Evolutionary optimization of a fuzzy rule-based network intrusion detection system", IEEE, Fuzzy Information Processing Society (NAFIPS), 2010 Annual Meeting of the North American.

[12] F. Rothlauf, 2002, Representations for Genetic and Evolutionary Algorithms, Physica-Verlag, Heidelberg, Germany.

[13] G. Helmer, J. S. K. Wong, V., Honavar, V., and L. Miller, February 2002, "AutomatedDiscovery of Concise Predictive Rules for Intrusion Detection," Journalof System Software, vol. 60, pp. 165-175.
[14] G. J. Klir and B. Yuan, 1995, "Fuzzy Sets and Fuzzy Logic: Theory and Applications", Prentice-Hall.
[15] H. Ishibuchi, T. Nakashima and T. Murata, 1999, "Performance evaluation of fuzzy classifier systems for multi-dimensional pattern classification problems", IEEE Trans. Syst.,Man, Cybern., Part B, vol. 29, pp. 601-618.
[16] H. Ishibuchi, T. Nakashima and T. Kuroda, 2000, "A Hybrid Fuzzy GBML Algorithm for Designing Compact Fuzzy Rule-Based Classification Systems", In Proceedings of the 9th IEEE Int. Conf. Fuzzy Systems (FUZZ IEEE-2000), San Antonio, TX, USA, pp. 706-711.
[17] H. M. Chen and S. Y. Ho, 2001, "Designing an Optimal Evolutionary Fuzzy Decision Tree for Data Mining", In Proc. of the Genetic and Evolutionary Computation Conference (GECCO-2001), San Francisco, California, USA, Morgan Kaufmann, San Francisco, California, USA, pp. 943-950.
[18] I. D. Falco, A. D. Cioppa, A. Iazzetta and E. Tarantion, 2005, "An Evolutionary Approach for Automatically Extracting IntelligibleClassification Rules", Knowledge and Information Systems, Springer-Verlag, New York, USA, vol. 7, no. 2, pp. 179-201.
[19] J. Hanand M. Kamber, 2001, "Data Mining: Concepts and Techniques",Morgan Kaufmann, San Francisco, California, USA.
[20] J. A. Roubos, M. Setnes and J. Abonyi, May 2001, "Learning fuzzy classification rules from labeled data," IEEE Trans. Fuzzy Syst., vol. 8, no. 54, pp.509-522.
[21] J. Gopalan, R. Alhajj and K. Barker, June 2006, "Discovering accurate and interesting classification rules using genetic algorithm", Proc. Int. Conf. on Data Mining, pp. 389-395.
[22] J. J. Buckley, Y. Hayashi and E. Czogala, 1993, "On the equivalence of neural nets and fuzzy expert systems", Fuzzy Sets and Systems, pp. 129-134.
[23] J. R. Quinlan, 1986, "Introduction of Decision Trees", Machine Learning, pp. 81-106.
[24] JiaLimin, Zhang Ruyan, Zhang Yong, Xing Zongyi and CaiGuoqiang, Sept. 2008, "Approach of Fuzzy Classification Based on Hybrid Co-Evolution Algorithm," In Proc. 4th IEEE Int. Conf. on Ntwrk Comp. & Adv. Info. Mgt., Gyeongju, vol. 2, pp. 266-271.
[25] K. A. Crockett, Z. Bandar and A. Al-Attar, May 2000, "Soft decision trees: A new approach using non-linear fuzzification", Proc. 9th IEEE Int. Conf. Fuzzy Systems (FUZZ IEEE-2000), San Antonio, TX, USA.
[26] K. A.De Jong, W. M. Spearsand D. F. Gordon, 1993, "Using Genetic Algorithms for Concept Learning",Machine Learning, Kluwer Academic Publishers, Hingham, MA, USA, vol. 13, issue 2-3, pp. 161-188.
[27] K. K. Bharadwaj and Basheer M. Al-Maqaleh, 2005, "Evolutionary approach for automated discovery of censored production rules", Enformatika, vol. 10, pp. 147-152.
[28] K. K. Bharadwaj and N. K. Jain, 1992, "Hierarchical Censored Production Rules (HCPRs) System", Data and Knowledge Engineering, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, vol. 8, no. 1, pp. 19-34.
[29] Li W, 2004, "A Genetic Algorithm Approach to Network Intrusion Detection", SANS Institute, USA.
[30] Li-Xin Wang and Jerry M. Mendel, Dec. 1992, "Generating Fuzzy Rules by Learning from Examples," IEEE Trans. on Syst. Man, Cybernetics, vol. 22, no. 6, pp. 1414-1427.
[31] Lotfi A. Zadeh, May 1996, "Fuzzy Logic = Computing with Words", IEEE Trans. Fuzzy Syst., vol. 4, no. 2, pp. 103-111.
[32] M. Delgado, F. V. Zuben and F. Gomide, 1999, "Modular and hierarchical evolutionary design of fuzzy systems", Proc. Genetic and Evolutionary Computation Conf. (GECCO-99), Morgan Kaufmann, pp. 180-187.
[33] M. H. Dunham, 2003, "Data Mining Introductory and Advanced Topics", Pearson Education.

[34] M. H. Marghny and I.E. El-Semman, Dec. 2005, “Extracting fuzzy classification rules with gene expression programming”, In Proc. Int. Conf. on AI, Machine Learning, AIML 2005, Cairo, Egypt.

[35] M. J. Pazzani, 2000, “Knowledge Discovery from data”, IEEE Intelligent Systems, vol. 15, no. 2, pp. 10-13.

[36] M. L. Wongand K. S. Leung, 2000, “Data Mining Using Grammar Based Genetic Programming and Applications”,Kluwer Academic Publishers,Norwell, MA, USA.

[37] N. Xiong and L. Litz, 1999, “Generating linguistic fuzzy rules for pattern classification with genetic algorithms”, Proc. 3rd European Conf. Principles of Data Mining and Knowledge Discovery (PKDD-99), Lecture Notes in AI 1704, Springer-Verlag, pp. 574-579.