Fuzzy qualitative link analysis for academic performance evaluation

Boongoen, Tossapon; Shen, Qiang; Price, Christopher John

Published in:
International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems

DOI:
10.1142/S0218488511007131

Publication date:
2011

Citation for published version (APA):
Boongoen, T., Shen, Q., & Price, C. J. (2011). Fuzzy qualitative link analysis for academic performance evaluation. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *19*(3), 559-585. https://doi.org/10.1142/S0218488511007131

Fuzzy Qualitative Link Analysis for Academic Performance Evaluation

Tossapon Boongoen, Qiang Shen and Chris Price

Abstract

Many approaches have been developed for academic performance evaluation using various fuzzy techniques. Initial methods rely greatly on experts' specification of analytical parameters, without making use of valuable information embedded in collected data. Given this insight, fuzzy rule induction has recently been studied as a data-driven alternative. Despite its efficiency and reported performance, the fuzzy subsethood metric representing the strength of relations between system variables is only used at a coarse level, with the underlying semantics being unfortunately distorted via a local re-scaling scheme. To overcome this problem, a novel fuzzy classification system is introduced in this paper, in which the existing measure is used to its full potential via the methodology of qualitative link analysis. With a network representation where variables and their relations are encoded as graph nodes and edges, the classification of a new instance conceptually becomes a problem of link-based similarity estimation that can be effectively resolved using the proposed fuzzy qualitative model. This new approach has been evaluated against the existing rule-based method, revealing significant advantages of the present work.

Index Terms

Academic performance evaluation, link analysis, order-of-magnitude reasoning, fuzzy sets.

I. Introduction

A student's learning achievement is typically determined by the performance levels in accordance with pre-specified educational objectives. It is crucial that an evaluation system provides a fair academic grading and useful information for individual improvement. An ideal system should also be transparent and efficient such that its results are readily communicatable with staff and students, and can be regularly reviewed and enhanced. Many existing approaches obtain these desirable characteristics using methods developed on the theory of fuzzy set [66]. Indeed, different fuzzy-based models have been employed for evaluation based on numerical scores obtained in an assessment [6], [9], [21], [37], [48], [60], [62] and for assessing prior educational achievement based on evidence such as academic certificates [18], [24]. This research focuses on the former which can be further categorized into expert and example-assisted approaches.

For evaluation of students' answer script, initial methods such as those introduced in [9], [21], [37] have been designed around a standard set of parameters formally specified in accordance with learning and teaching experts' opinion. Principally, these techniques identify for each question the performance level that best matches an awarded score. Following that, the final grade is justified by comparing the aggregated outcome of question-specific levels and a pre-defined grade scale. To obtain a refined evaluation mechanism, several recently developed models have included the use of newly proposed analytical parameters in addition to awarded scores. For instance, the method proposed in [60] attaches to each question-specific score a confidence degree of the evaluator, which is interpreted as a fuzzy number. Also, the techniques of [62] and [6] focus on fuzzy membership functions that are used to represent three quality criteria of the questions: difficulty, importance and complexity.

Despite their usefulness, these expert-directed models are based solely on domain experts' opinion and judgment. They are usually restrictive to or biased towards a definite human's mindset, and not easily adaptable to a changing environment. Taking this into account, acquiring such knowledge from collected data is a practical and well-known alternative. For instance, a fuzzy rule-based framework, named Weighted Subsethood-Based Algorithm (WSBA), has recently been put forward for academic performance evaluation [48]. It exploits the measure of 'fuzzy subsethood' to form a weighted rule set, which has proven more effective than a benchmark fuzzy rule-based system such as

T. Boongoen, Q. Shen and C. Price are with the Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK e-mail:{tsb,qqs,cjp}@aber.ac.uk

Neuro-Fuzzy Classification (NEFCLASS) [45]. However, its core information metric is used only at a coarse level in classifying a new instance. Furthermore, the semantics of such measure is explicitly altered by the introduction of a subjective re-scaling scheme.

Having identified such shortcomings, this paper presents a new fuzzy classification system for academic performance evaluation, in which the fuzzy subsethood measure is exploited to its full potential whilst the corresponding semantics is well preserved for an accurate outcome. Unlike the conventional fuzzy classification approach [32], [44] that is implemented as a collection of 'If-Then' rules, the proposed framework uniquely represents a classification problem as a link or social network [26], [61]. In particular, fuzzy variables (both attributes and classes) and their associations are qualitatively implemented as graph nodes and edges, respectively. The strength or weight of each edge is effectively determined by the fuzzy subsethood estimated between two involving nodes. With this novel terminology, the classification of a new case is accomplished using a link-based similarity algorithm that gauges the similarity between nodes representing the examined instance and classes.

Many advanced link-based similarity techniques such as SimRank [33] and PageSim [41] have proven more accurate, but greatly less efficient than simple SNA (Social Network Analysis) methods [40]. The present research follows the previous study in [12] to obtain an computationally inexpensive and effective model via the aggregation of different simple link-pattern measures. Note that, to overcome the problem with inaccurate quantitative descriptions, a fuzzy extension of the crisp-interval order-of-magnitude model is employed to qualitatively represent link measures, their relevance (or importance) degrees and aggregations. This modification aims to reduce the native difficulties with pure symbolic calculus and crisp distinctions amongst qualitative descriptors [3], [53]. Specifically, fuzzy sets [66] provide flexible, but standard numerical semantics of order-of-magnitude labels, with mathematically-sound operations.

The rest of this paper is organized as follows. Section II presents an overview of existing fuzzy systems that have been developed for academic performance evaluation. Following that, Section III introduces the proposed link analysis approach to fuzzy classification. This includes the fundamental link network formalism and the corresponding representation of fuzzy variables and relations. Then, Section IV describes the qualitative link-based similarity measures and the associated fuzzy order-of-magnitude model. The assessment of the application of this new model to academic performance evaluation is detailed in Section V. The paper is concluded in Section VI, with the perspective of further work.

II. Related Work

This section presents the ideas of various fuzzy systems for academic performance evaluation, from which the current research is developed. In addition, the methodology of link analysis is exhibited as the basis of the proposed approach.

A. *Fuzzy Systems for Academic Performance Evaluation*

A number of methods found in the literature can be categorized into expert-directed and example-assisted approaches, each of which is reviewed in the following subsections.

1) Expert Directed Approach: A large number of performance evaluation models are based entirely on parameters specified by experts or experienced evaluators. These input variables are typically designed as fuzzy sets that represent uncertain/overlapping levels of performance, with regard to the importance and difficulty of the examining or practical questions. This expert-system alike approach is inspired by the initial model of [9], which is proposed to evaluate students' performance from answer scripts. It relies upon the use of the concept of similarity between fuzzy sets. Formally, the similarity $S(Q,Q') \in [0,1]$ between fuzzy sets Q and Q' is defined as

$$S(Q,Q') = \frac{\sum\limits_{\forall\, x_i \in X} \mu_Q(x_i)\mu_{Q'}(x_i)}{\max\left(\sum\limits_{\forall\, x_i \in X} \mu_Q(x_i), \sum\limits_{\forall\, x_i \in X} \mu_{Q'}(x_i)\right)} \tag{1}$$

where X denotes the set of domain elements.

At the outset of this evaluation model, an evaluator is required to give fuzzy marks Q_j for the j-th question into a 'fuzzy grade sheet', which is a table containing rows for question numbers and columns for awarding marks in term of fuzzy values. These question-specific marks are then matched with 'Standard Fuzzy Sets (SFS)' that are predefined with membership values corresponding to different levels of student performance. Principally, SFS are devised by experts with respect to the standard fixed by an educational authority such as an academic institution. The example of SFS employed in [9] is shown in Table I, in which the following levels of student performance are defined: Excellent (A), Very Good (B), Good (C), Satisfactory (D) and Unsatisfactory (F).

TABLE I
STANDARD FUZZY SETS (SFS) OF STUDENT'S PERFORMANCE.

Linguistic terms	**Fuzzy sets**
Excellent	{0/0, 0/20, 0.8/40, 0.9/60, 1/80, 1/100}
Very Good	{0/0, 0/20, 0.8/40, 0.9/60, 0.9/80, 0.8/100}
Good	{0/0, 0/20, 0.8/40, 0.9/60, 0.9/80, 0.8/100}
Satisfactory	{0.4/0, 0.4/20, 0.9/40, 0.6/60, 0.2/80, 0/100}
Unsatisfactory	{1/0, 1/20, 0.4/40, 0.2/60, 0/80, 0/100}

Specifically to the j-th question, the awarded mark Q_j is compared against each of the performance levels specified in SFS. This matching is accomplished using the similarity measure previously defined by Eq. 1. The grade G_j corresponding to the fuzzy mark Q_j is estimated as

$$G_j = \underset{\forall\theta\in\Theta}{\operatorname{argmax}} S(Q_j, \theta) \tag{2}$$

where Θ is the set of performance levels specified in SFS, i.e. $\Theta = \{A, B, C, D, F\}$.

Having obtained grades of all questions, the total score $TS \in [0, 100]$ is calculated using the so-called 'mid-grade' point of the grade awarded. Table II shows an example set of mid-grade points that is used in the study of [9]. In essence, the total score TS is estimated as follows, where $T(Q_j)$ is the mark assigned to the j-th question, $\sum_{\forall j} T(Q_j) = 100$ and $P(G_j)$ is the mid-point of grade G_j:

$$TS = \frac{1}{100}\left[\sum_{\forall j} T(Q_j) \times P(G_j)\right] \tag{3}$$

TABLE II
GRADES AND THEIR CORRESPONDING MID-GRADE POINTS.

Linguistic terms	**Grades/Scores**	**Mid-grade points**
Excellent	$90 \leq A \leq 100$	95
Very Good	$80 \leq B < 90$	85
Good	$50 \leq C < 70$	60
Satisfactory	$30 \leq D < 50$	40
Unsatisfactory	$0 \leq F < 30$	15

Based on the resulting TS, a new final grade will be determined based on crisp interval values referring to the level of performance. Despite its general applicability, this model has been criticized for the lack of cohesive associations between fuzzy marks in a fuzzy grade sheet and those standard levels of performance. The use of mid-grade points to derive the total score may create an unexpected result, hence reducing its reliability [48]. As suggested by [21], this method may be inefficient with the pair-wise matching between fuzzy marks and performance levels of SFS.

To overcome such drawbacks, a new model to evaluating students' answer script is introduced by [21]. In particular, the 'degrees of satisfaction' are defined in advance by experts with respect to levels of performance, from which the 'maximum degree of satisfaction' per level can be obtained. According to [27], examples of degrees of satisfaction and the maximum degree of satisfaction are listed in Table III. This includes eleven levels of student performance exploited in [21].

Initially, fuzzy marks Q_j are given for the j-th question in the 'extended fuzzy grade sheet'. Unlike the work of [9], the values of Q_j are awarded with respect to each of the performance levels. Then, the degree of satisfaction D_j can be estimated from Q_j as

TABLE III
PERFORMANCE LEVELS AND CORRESPONDING DEGREES OF SATISFACTION.

Performance levels	**Degrees of satisfaction**	**Maximum satisfaction degrees**
Extremely Good (EG)	100%	1.00
Very Very Good (VVG)	91-99%	0.99
Very Good (VG)	81-90%	0.90
Good (G)	71-80%	0.80
More or less Good (MG)	61-70%	0.70
Fair (F)	51-60%	0.60
More or less Bad (MB)	41-50%	0.50
Bad (B)	25-40%	0.40
Very Bad (VB)	10-24%	0.24
Very Very Bad (VVB)	1-9%	0.09
Extremely Bad (EB)	0%	0.00

$$D_j = \frac{\sum\limits_{\forall\pi\in\Pi} \mu_\pi(Q_j)F(\pi)}{\sum\limits_{\forall\pi\in\Pi} \mu_\pi(Q_j)} \tag{4}$$

where $\Pi = \{EG, VVG, VG, G, MG, F, MB, B, VB, VVB, EB\}$ is a set of performance levels, $F(\pi)$ is the maximum satisfaction degree of the performance level $\pi \in \Pi$, and $\mu_\pi(Q_j)$ is the membership value awarded of Q_j to the level π.

Following that, the total score TS can be summarized from all questions as follows:

$$TS = \sum_{\forall j} T(Q_j) \times D_j \tag{5}$$

where $T(Q_j)$ is the mark assigned to the j-th question. Based on the resulting TS, a final grade is determined using the standard satisfaction levels in Table III. It is clear that this method is more efficient than the initial model of [9]. However, the use of an extended fuzzy grade sheet to specify fuzzy marks may not scale up to the case where many fuzzy values needed to be considered for each of the questions [48]. This problem is intensified when the number of questions or modes of assessment increases.

Other expert-directed methods found in the literature also adopt a rather similar idea, perhaps with more analytical parameters and certain innovative evaluation criteria. An alternative to assess students' performance is introduced in [37], based on the notion of 'fuzzy expected values'. This work does not require an evaluator to fill out a fuzzy grade sheet as original student scores are represented by crisp values. Such scores are transformed into fuzzy values using the expert-specified fuzzy partitions, which represent an expected percentage of students who will be given a certain level of performance, i.e. grade A, B, C, D or F.

Also, motivated by the concept of certainty factor that is commonly used in traditional expert systems, a method which evaluates students' answer scripts using fuzzy numbers is reported in [60]. In this research, each of the fuzzy numbers used is associated with a confidence degree of the evaluator. Likewise, several recently developed models have also included additional parameters that allow the refined performance assessment to be achieved. For instance, the techniques of [6] and [62] focus on modelling fuzzy membership functions that are used to represent three factors of the questions given to students: difficulty, importance and complexity. The method to generate such fuzzy sets automatically is introduced in [5]. Despite their potential, different evaluation models discussed thus far rely heavily on the subjective knowledge of domain experts without making use of available data regarding past assessments. This makes such a system inflexible to changing educational standards and evaluation criteria.

2) Example Assisted Approach: The problem mentioned above can be overcome to a great extent using the knowledge captured from examples of students' scores and the corresponding levels of performance. This is a useful alternative to, or as a complementary of, the expert-directed approach when historical data is readily available. An example assisted system is adaptable to the shifted decision trends over time. It provides a platform for aggregating multiple experts' opinion as well as a pool of past cases collected from different sources. While the results obtained from a conventional expert-driven model is restricted to a single viewpoint, those obtained from the data-driven counterpart produces a more robust conclusion across different principles and styles of human judgement.

In practice, the task of performance evaluation can be implemented by a 'fuzzy rule-based system (FRBS)', which is transparent and effective with multiple attributes of imprecise data. Beyond its original application to the control domain [39], [55], FRBS has recently received a great deal of attention for its contributions to general classification and pattern recognition problems [44]. It can be perceived as an approximator of nonlinear mappings from non-fuzzy input vectors to non-fuzzy output values [32]. Principally, designing FRBS is to find a compact set of fuzzy 'If-Then' classification rules that capture the input-output behavior of the domain problem presented by training examples. Many methods have been devised to formulate and learn such fuzzy rules from quantitative data: heuristic procedures [1], [31], neuro-fuzzy approach [45], data clustering technique [2], association rules [22], fuzzy nearest neighbor method [36], genetic algorithm [30] and a rough-fuzzy approach [52].

Particularly to the problem of academic performance evaluation, the concept of inducing fuzzy rules from historical data has been realized by the Weighted Subsethood-Based Algorithm (WSBA) [48]. Intuitively, the use of linguistic rule models such as 'If assignment score is *very poor* and exam score is *average*, Then the final result is *poor*' helps representing the natural way by which humans make judgements and decisions. Also, examples can be used to create a fuzzy model which integrates information from historical data with expert opinions. WSBA is built upon the measure of 'fuzzy subsethood' that represents the degree to which a fuzzy set is a subset of another fuzzy set [47], [65]. Formally, the subsethood value $\vartheta(C, A) \in [0, 1]$ of fuzzy set A to fuzzy set C is defined as

$$\vartheta(C, A) = \frac{\sum\limits_{x \in U} \nabla(\mu_C(x), \mu_A(x))}{\sum\limits_{x \in U} \mu_C(x)} \tag{6}$$

where ∇ denotes a t-norm operator.

For a classification problem with K classes and M linguistic variables (or attributes), the weight $w(C_k, A_{ij}) \in [0, 1]$ between the j-th linguistic term A_{ij} of linguistic variable $A_i, i = 1 \ldots M$ and class $C_k, k = 1 \ldots K$ can be estimated by

$$w(C_k, A_{ij}) = \frac{\vartheta(C_k, A_{ij})}{\max\limits_{t=1 \ldots l} \vartheta(C_k, A_{it})} \tag{7}$$

given that the linguistic variable A_i has l possible linguistic terms of $A_{i1}, \ldots, A_{il}$.

Unlike the conventional formation of If-Then rules, the aforementioned weight metric is particularly used to reflect differences between the relative contributions made by individual linguistic terms of each variable towards the conclusion (or class). Examples of such fuzzy rules, each of which corresponds to a specific class, are given below:

Rule 1: **If** (A_1 is $w(C_1, A_{11})A_{11}$ OR $w(C_1, A_{12})A_{12}$) AND ... AND (A_M is $w(C_1, A_{M1})A_{M1}$ OR $w(C_1, A_{M2})A_{M2}$ OR $w(C_1, A_{M3})A_{M3}$) **Then** class is C_1

...

Rule K: **If** (A_1 is $w(C_K, A_{11})A_{11}$ OR $w(C_K, A_{12})A_{12}$) AND...AND (A_M is $w(C_K, A_{M1})A_{M1}$ OR $w(C_K, A_{M2})A_{M2}$ OR $w(C_K, A_{M3})A_{M3}$) **Then** class is C_K

According to [48], the resulting ruleset can be simply defined by

$$\xi(C_k) = \mathop{\nabla}_{i} \left(\mathop{\triangle}_{j} \; w(C_k, A_{ij}) \times \mu_{A_{ij}}(x) \right), \; k = 1 \ldots K \tag{8}$$

where $\xi(C_k) \in [0, 1]$ is the score awarded to class C_k. Also, $\triangle$ and ∇ denotes logical disjunction and conjunction, which are implemented by maximum and minimum, respectively.

WSBA has been applied to the problem of performance evaluation, where targeted classes $C_1 \ldots C_K$ correspond to K different levels of performance with variables $A_1 \ldots A_M$ representing M assessment criteria. After obtaining all level-specific scores $\xi(C_k), k = 1 \ldots K$, the level (or grade) with the maximum score is finally awarded. Note that WSBA makes use of locally (variable-specific) normalized subsethood measures as defined by Eq. 7, instead of their native values (see Eq. 6). In so doing, with respect to each variable, the maximum subsethood measure of a given term and a certain class becomes 1 regardless of the underlying magnitude of the subsethood value. This may degrade the effectiveness of WSBA as the core metric used to capture relations between variable terms and targeted classes is semantically distorted. Driven by this observation, the work here is developed in order to improve the example-assisted performance evaluation, using the social network formalism which is effective for representing observed information objects (variable terms and classes) and preserving the semantics of their interactions.

B. Link Analysis

The methodology of link analysis [26] has attracted a great deal of attention in the past decade with successful applications in a variety of domains; for instance, in database management [7], intelligence data analysis [14] and consensus clustering [29]. To further enhance system maintainability and service quality, finding similar objects has become a significant issue that is extensively studied in the communities of database, information retrieval and recommender for many years [8], [16], [23], [38]. Initial attempts to justify similarity made use of text-based methods, such as the cosine similarity and the TFIDF (Term Frequency-Inverse Document Frequency) model [50]. These approaches require large storage and long computing time due to the need of full-text comparison [41]. Seeking alternatives is necessary, especially for the World Wide Web where text-based methods may be inapplicable for pages with little texts and a large amount of multimedia objects [41], and for intelligence data analysis where content-based approaches can be misleading due to fraud descriptions of terrorists' name, address, appearance and contact details [14], [58].

However, in environments where information objects are linked in accordance to their relations, the similarity can be evaluated upon the structure of such links. Empirical results have shown that link-based similarity measures can enhance the performance of the classic content-based counterparts [17], [40]. Also, in the bibliometrics field, link-based similarity measures such as bibliographic coupling [34], Amsler [4], co-citation [54] and SimRank [33] have been developed to disclose the proximity of scientific paper publications from their cross-citation patterns. With fairly similar purposes and terminologies, Connected-Triple [35], PageSim [41] and Connected-Path [14] algorithms were recently established to reveal duplicated author names in publication databases, similar web pages and aliases in intelligence data collections, respectively. Similar link-based statistical information can be applied to improve supervised learning models that aim to resolve identity problems in terrorism-related and publication datasets [28], [59]. In addition, a random walk and other social network analysis techniques have been widely explored for a variety of domains, including: publications [56], [57], email messages [43], collaborative recommendations [25] and films [42].

Through the use of link analysis, detection of similar or duplicated objects is done by examining relation patterns amongst references of objects, which can be formally specified as an undirected graph $G(V, E)$. It is composed of two sets, the set of vertices V and that of edges E. Let X and R be the set of all references and that of their relations in the dataset, respectively. Then, vertex $v_i \in V$ denotes reference $x_i \in X$ and each edge $e_{ij} \in E$ linking vertices $v_i \in V$ and $v_j \in V$ corresponds to a relation $r_{ij} \in R$ between references $x_i \in X$ and $x_j \in X$. Let O be the set of real-world objects each being referred to by at least one member of set X, a pair of references (x_i, x_j) are aliases when both references correspond to the same underlying real-world object: $(x_i \equiv o_k) \wedge (x_j \equiv o_k), o_k \in O$. In practice, disclosing an alias/duplicated pair in the graph G is to find a pair of vertices (v_i, v_j), whose similarity $s(v_i, v_j)$ is significantly high. Intuitively, the higher $s(v_i, v_j)$ the greater the possibility that vertices v_i and v_j, and hence corresponding references x_i and x_j, constitute the actual alias pair. Each edge $e_{ij} \in E$ possesses statistical information $f_{ij} \in \{1, \ldots, \infty\}$, representing the frequency of any relation occurring between references x_i and x_j within the given dataset. By representing the multiplicity of each link (i.e. relation) as the frequency count, the graph terminology used in this paper is of the simple type (i.e. no parallel edges), without losing any potential link information.

Particularly to the publication databases as described in [35], as an example of using aforementioned terminology, a set of author references (i.e. names) and their relations can be presented as the graph shown in Fig. 1, where $X = \{A, B, C, D, E\}$, $R = \{r_{AB}, r_{AC}, r_{AD}, r_{BE}, r_{CD}\}$, and r_{ij} denotes the fact that references x_i and x_j are authors of the same paper (i.e. co-authors). Accordingly, edge e_{AD} is presented with $f_{AD} = 2$ since references A and D are co-authors of two different papers (i.e. $paper_2$ and $paper_3$). In the contrary, the frequency statistics f_{AC} of edge e_{AC} is 1 as references A and C have only one joint publication, $paper_2$. Effectively, given O being the set of real-world author entities, a pair of references, such as (A, B), may then be hypothesized, based on their similarity, as the alias pair (i.e. $(A \equiv o_k) \wedge (B \equiv o_k), o_k \in O$).

(a) relations in database table

publications	author names
$paper_1$	A and B
$paper_2$	A, C and D
$paper_3$	A and D
$paper_4$	B and E
$paper_5$	C and D

(b) relations in a graph format

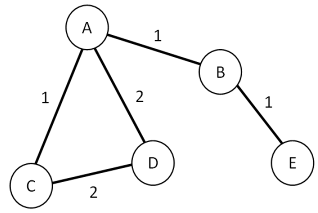

Fig. 1. Relations between author references and publications, presented in: (a) database table format and (b) graphical format.

III. A Link Analysis Approach to Fuzzy Classification

Despite the success of link analysis, very few studies in the literature have attempted to marry this concept with fuzzy sets [64]. The present research introduces an innovative, hybrid reasoning framework in which the network representation scheme and link-based analysis are integrated for fuzzy classification. This follows the future directions set out in [46] for qualitative reasoning – integration of models from different domains and hybrid modeling. Principally, involving variable terms and classes are represented as graph nodes, while their observed associations are encoded as corresponding edges. Then, the likelihood that a new data instance belonging to a specific class is determined via the notion of link-based similarity measure.

Suppose that a fuzzy classification system (FS) includes M attribute variables of $A_i, i = 1 \dots M$, where each $A_i = \{A_{i1}, \dots, A_{i\beta_i}\}$ and β_i is the number of terms defined for A_i. These terms are defined as fuzzy sets by experts or the results of historical data analysis. In particular, the term A_{ij} of a variable A_i is represented mathematically by a definite membership function on the corresponding universe of discourse U^{A_i}, i.e. $\mu_{A_{ij}}(x), x \in U^{A_i}$. An FS aims to discover the relation patterns between these variable terms and K classes of $C_k, k = 1 \dots K$, each of which is also specified as a fuzzy set with a membership function $\mu_{C_k}(y), y \in U^C$. These patterns are obtained from inducing the information provided with training instances $D = \{d_1, \dots, d_N\}$. Each $d_t \in D$ is represented by a vector $d_t = (x_1^t, \dots, x_M^t, y^t)$, where $x_i^t \in U^{A_i}, i = 1 \dots M$, is the real value of attribute A_i of instance d_t, and $y^t \in U^C$ is the conclusion value of d_t.

Conceptually, a fuzzy classification system FS resembles a social network [61] and can be represented as a graph $G = (V, W)$. While vertices in a social network are actual people each with their own definition of the world, each vertex $v_p \in V$ of an FS denotes a fuzzy variable that possesses a domain-specific mathematical definition, i.e. membership function. Edges amongst individuals in a social network can be of different types, ranging from their mutual interests to shared community/transportation services. They are normally weighed in accordance with the frequencies that two involving parties attend a joint event. Likewise, an edge $w_{pq} \in W$ in FS stands for the relation between fuzzy variables represented by vertices $v_p, v_q \in V$. In essence, the corresponding weight $|w_{pq}|$ can be extracted from a collection of training instances D.

By following the graph-based model of [29] which has been successfully applied to the cluster ensemble problem, the graph G representing an FS is 'bipartite' with $V = V^A \cup V^C$, where V^A is a set of vertices corresponding to linguistic variable terms and the V^C set contains vertices representing different classes. For simplicity, it is assumed that the weight of any edge w_{pq} between vertices of the same type (i.e. $v_p, v_q \in V^A$ or $v_p, v_q \in V^C$) is irrelevant, i.e. $w_{pq} \notin W$. This is in line with existing FRBSs, especially WSBA [48] where only term-class relations are considered. Note that the formalism proposed herein can be effectively extended to the case where term-term and class-class relations are also examined.

Let vertices $v_p \in V^A$ and $v_q \in V^C$ correspond to the variable term A_{ij} and class C_k, respectively. The weight of edge w_{pq} based on a training instance $d_t \in D$ is denoted as $|w_{pq}(d_t)| \in [0, 1]$ and estimated as follows:

$$|w_{pq}(d_t)| = \mu_{A_{ij}}(x_i^t) \nabla \mu_{C_k}(y^t) \tag{9}$$

given that ∇ denotes a t-norm operator which is interpreted here as minimum. Following that, the final weight of edge w_{pq}, which is equivalent to the fuzzy subsethood measure, can be specified as

$$|w_{pq}| = \frac{\sum_{\forall\, d_t \in D} |w_{pq}(d_t)|}{\sum_{\forall\, d_t \in D} \mu_{C_k}(y^t)} \tag{10}$$

where $|w_{pq}| \in (0, 1]$ is obtained by

$$|w_{pq}| = \frac{|w_{pq}|}{\max_{\forall\, w' \in W} |w'|} \tag{11}$$

Having generated such a bipartite graph G, the likelihood that a test instance $d_s = (x_1^s, \ldots, x_M^s)$ belonging to any specific class $C_k, k = 1 \ldots K$ is determined using the link-based similarity method. To accomplish this, edges between a new class-like vertex representing d_s and those corresponding to variable terms are first amended to G. For the edge w_{pq} between vertices $v_p, v_q \in V$ that represent d_s and a specific linguistic term A_{ij}, its weight is calculated by

$$|w_{pq}| = \frac{\mu_{A_{ij}}(x_i^s)}{\sum_{\forall\, i',j'} \mu_{A_{i'j'}}(x_{i'}^s)} \tag{12}$$

Based on this updated graph, the likelihood that a test instance d_s is of class C_k is justified by the link-based similarity measure $sim(v_a, v_b)$, where $v_a, v_b \in V$ representing d_s and C_k. Intuitively, the higher the similarity is, the greater the likelihood becomes. At the end of this link-based classification process, the test instance d_s is assigned to a class C_* such that

$$sim(d_s, C_*) = \max_{k=1 \ldots K} sim(d_s, C_k) \tag{13}$$

Different from any conventional FRBS, the proposed framework designs a fuzzy classification problem as link analysis, where a social network representation is employed to model variable terms, classes and their relations. Similar to WSBA, system parameters which are relations (or weighted edges) are also acquired from training instances. However, while WSBA simply aggregates relations without considering their attached semantics, the link-based model preserves such information that is better summarized via a link-based similarity method. The underlying approach is general for performing classification and suits the task of academic performance evaluation well. It resembles the concept of data-driven templates used for a recognition of human motion [19]. In particular, numerical inputs are similarly mapped onto qualitative spaces that enable a process of classification.

IV. Qualitative Link-Based Similarity Measures

In spite of their notable performance, a number of advanced link-based similarity techniques (e.g. SimRank [33], Random-Walk [25], [43] and PageSim [41]) are inefficient, especially for large problems. In contrast, several simple SNA (Social Network Analysis) methods, such as Connected-Triple [49] and Jaccard coefficient [51], are less computationally expensive, but with comparatively lower performance [40]. Their inaccuracies are mainly due to the limited amount of information employed in the underlying similarity evaluation. According to the recent study of [12], the effectiveness of an SNA-like technique can be substantially boosted, without degrading the advantage of efficiency, via the aggregation of different simple link-pattern measures. In particular, to overcome the problem with inaccurate descriptions, an order-of-magnitude model is employed to qualitatively represent link measures, their relevance (or importance) degrees and aggregations.

The current research exploits a fuzzy extension of the crisp-interval qualitative model [12]. In order to overcome the inherent difficulties with inadequate symbolic calculus and crisp distinction amongst qualitative descriptors (as pointed out in [3], [53]), fuzzy sets [66] are incorporated to provide flexible, but standard numerical semantics of order-of-magnitude labels, with mathematically-sound operations. This granular technology, similarly exploited for information retrieval [15], is effective to deliver an analysis outcome that is both unambiguous and interpretable.

A. Link Pattern Measures and Qualitative Labels

Link analysis is based on examining relation patterns amongst objects in a given link network, which can be formally specified as an undirected graph $G = (V, W)$. Many existing techniques, such as SimRank [33] and Connected-Triple [49], have concentrated on the numerical count of shared neighboring objects (or the normalized count for SimRank). Let a vertex $v_i \in V$ represents an object of interest and $N_{v_i} \subset V$ be a set of vertices directly linked to v_i, called neighbors of v_i. The similarity $sim(v_i, v_j)$ between $v_i, v_j \in V$ is determined by the cardinality of $N_{v_i} \cap N_{v_j}$, the set of neighbors shared by both v_i and v_j. Intuitively, the higher the cardinality is, the greater the similarity of these entities becomes.

Despite its simplicity, such an approach does not take into account edges' weights that may help to refine the underlying similarity measure. Hence, the weighted cardinality (WCT) metric of [29], originally developed for the problem of consensus clustering, is adopted here. Given a graph with n vertices, the measure $WCT(v_i, v_j) \in [0, n-2]$ between v_i, v_j is defined as

$$WCT(v_i, v_j) = \sum_{\forall v_h \in N_{v_i} \cap N_{v_j}} \nabla(|w_{ih}|, |w_{jh}|) \tag{14}$$

where ∇ is a t-norm operator, conveniently interpreted as minimum.

The uniqueness of shared neighbors (UQ) is another useful alternative to the conventional, unweighted cardinality. According to [11], [14], the uniqueness measure $UQ(v_i, v_j)^{v_h}$ between vertices $v_i, v_j \in V$ can be approximated from each joint vertex $v_h \in N_{v_i} \cap N_{v_j}$ as follows:

$$UQ(v_i, v_j)^{v_h} = \frac{|w_{ih}| + |w_{jh}|}{\sum\limits_{\forall v_s \in V} |w_{hs}|} \tag{15}$$

Following that, the overall $UQ(v_i, v_j) \in [0, n-2]$ is estimated by

$$UQ(v_i, v_j) = \sum_{\forall v_h \in N_{v_i} \cap N_{v_j}} UQ(v_i, v_j)^{v_h} \tag{16}$$

A common drawback of the numerical measures previously presented is the inability to achieve coherent and natural interpretation through existing seemingly fine-grained scales. Exploring a link network with crisp numerically valued criteria is typically considered inflexible comparing to the use of interval and linguistic descriptors. Specifically, a wrong interpretation of a link measure may occur if there exists a unduly high property value within a link network. A more accurate and naturally expressive measure is to exploit qualitative labels like 'low' and 'high'.

To tackle such an important shortcoming, a link measure (e.g. WCT and UQ) is gauged in accordance with its specific fuzzy order-of-magnitude (FOM) spaces, an extension of the crisp-interval OM introduced in [12]. Let $FOM(\pi) = (L(\pi), F(\pi))$ be an FOM space of the link measure π, with $L(\pi)$ and $F(\pi)$ respectively denoting the set of qualitative labels and the set of fuzzy sets which specify the labels. For example, in particular to $FOM(WCT)$, Fig. 2 presents the fuzzy sets $F(WCT)$ that are defined for the labels $L(WCT) = \{Low\ (L)$, $Medium\ (M)$, $High\ (H)\}$. Similarly, for $FOM(UQ)$, Fig. 3 shows the fuzzy sets $F(UQ)$ which have been specified for qualitative labels of the UQ measure, i.e. $L(UQ) = \{Very\ Low\ (VL)$, $Low\ (L)$, $Moderate\ (M)$, $High\ (H)$, $Very\ High\ (VH)\}$.

B. Similarity Evaluation via Relevance-Based Aggregation

Many data analysis systems achieve an accurate conclusion by aggregating values of different domain attributes. In general, each examined variable (or link measure) π is assigned a different degree of relevance (or weight) W_π. This may be given by domain experts in according with their past experiences or estimated from past data if such knowledge is not readily available. With the original OM framework [12], a weight can be naturally expressed using the order-of-magnitude label set such as $OM(W) = \{None, +, ++, +++\}$ or $OM(W) = \{0, 1, 2, 3\}$. However, these crisp-interval descriptors are generally ineffective. The specification and subsequent manipulation of weights can be more efficiently handled using the fuzzy-set based approach. For the current study, $FOM(W)$ is defined

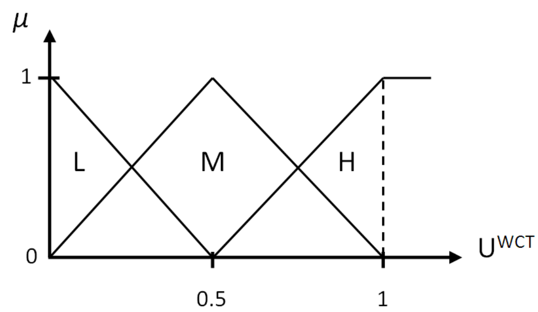

Fig. 2. Definition of fuzzy sets that correspond to qualitative labels $L(WCT)$ defined on $U^{WCT} = [0, n-2]$.

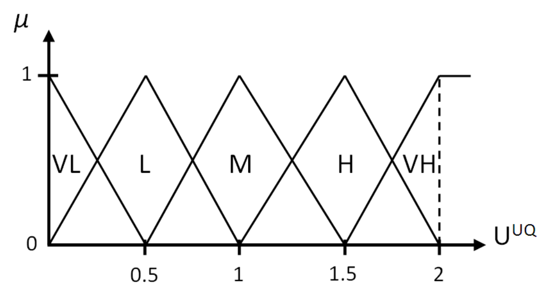

Fig. 3. Definition of fuzzy sets that correspond to qualitative labels $L(UQ)$ defined on $U^{UQ} = [0, n-2]$.

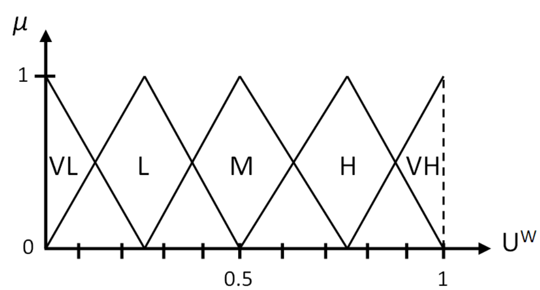

Fig. 4. Definition of fuzzy sets that correspond to qualitative labels $L(W)$ defined on $U^W = [0, 1]$.

by the label set $L(W) = \{Very\ Low\ (VL),\ Low\ (L),\ Moderate\ (M),\ High\ (H),\ Very\ High\ (VH)\}$ and the corresponding fuzzy sets $F(W)$ shown in Fig. 4.

The development of a qualitative reasoner usually involves a number of variables that are represented with qualitative labels of different granularity, defined on dissimilar universes of discourse. Therefore, prior to the aggregation process, the usual homogenization process conducted in the conventional OM model is similarly required to map fuzzy-set based variables onto the unified scale $U^* = [0, 1]$. This homogenization can be summarized as follows:

- *Step 1*: For a variable π whose universe of discourse $U^\pi = [p, p']$, $p, p' \in \mathcal{R}$, $p' > 1$, truncate $[p, p']$ to $[p, \delta]$ such that

$$\mu_f(\delta) = \mu_f(\gamma),\ \forall\ f \in F(\pi), \delta < \gamma \leq p' \tag{17}$$

  In particular to $FOM(WCT)$ and $FOM(UQ)$, the corresponding universes U^{WCT} and U^{UQ} are truncated from $[0, n-2]$ and $[0, n-2]$ to $[0, 1]$ and $[0, 2]$, respectively. In contrast, the universe of discourse U^W, of the qualitative space $FOM(W)$ is left unchanged as it is already defined on the unified scale $U^* = [0, 1]$.
- *Step 2*: The resulting universe $U^\pi = [p, \delta]$ is aligned with the unified scale $U^* = [0, 1]$. Principally, each value $x \in U^\pi$ is mapped to its corresponding value $x^* \in U^*$ by

$$x^* = \frac{x - p}{\delta - p} \tag{18}$$

- *Step 3*: Each fuzzy set $f \in F(\pi)$ that is defined for a variable π on the universe of discourse U^{π}, is linearly mapped onto the unified universe of discourse U^* such that

$$\mu_{f^*}(x^*) = \mu_f(x) \tag{19}$$

where f^* specified on the unified universe of discourse U^* is the equivalent fuzzy set of f defined on the original U^{π}.

Following the above pre-processing procedure, the aggregated outcome Ω that is also a fuzzy set in $U^* = [0,1]$ can be estimated using the weighted average function φ:

$$\Omega = \varphi(WCT, UQ, W_{WCT}, W_{UQ}) = \frac{(WCT \times W_{WCT}) + (UQ \times W_{UQ})}{W_{WCT} + W_{UQ}} \tag{20}$$

The membership function of Ω is denoted by $\mu_{\Omega}(t), \forall t \in U^*$, where t is an ordinary weighted average that is calculated as

$$t = \varphi(x, x', w, w') = \frac{xw + x'w'}{w + w'} \tag{21}$$

where $x \in WCT$, $x' \in UQ$, $w \in W_{WCT}$ and $w' \in W_{UQ}$. By the extension principle, the membership function of Ω can be defined by

$$\mu_{\Omega}(t) = \sup\Big(\min(\mu_{WCT}(x), \mu_{W_{WCT}}(w)), \min(\mu_{UQ}(x'), \mu_{W_{UQ}}(w'))\Big) \tag{22}$$

Although mathematically rigorous, finding the exact membership function $\mu_{\Omega}(t)$ can be complex and computationally expensive. Recognizing this, a discrete approximate method that makes use of the α-cut fuzzy arithmetic, is exploited to aggregate fuzzy sets (see [20], [67] for more details). In particular, the α-cut of a variable WCT and that os its weight W_{WCT} are

$$(WCT)_{\alpha} = \{(x, \mu_{WCT}(x)) | x \in WCT, \mu_{WCT}(x) \geq \alpha\} \tag{23}$$

$$(W_{WCT})_{\alpha} = \{(w, \mu_{W_{WCT}}(w)) | w \in W_{WCT}, \mu_{W_{WCT}}(w) \geq \alpha\} \tag{24}$$

These α-cuts are crisp intervals and can be expressed in continuous closed form as follows:

$$\begin{aligned}(WCT)_{\alpha} = [(a)_{\alpha}, (b)_{\alpha}] = \Big[&\min\{x \in WCT | \mu_{WCT}(x) \geq \alpha\}, \\ &\max\{x \in WCT | \mu_{WCT}(x) \geq \alpha\}\Big]\end{aligned} \tag{25}$$

$$\begin{aligned}(W_{WCT})_{\alpha} = [(c)_{\alpha}, (d)_{\alpha}] = \Big[&\min\{w \in W_{WCT} | \mu_{W_{WCT}}(w) \geq \alpha\}, \\ &\max\{w \in W_{WCT} | \mu_{W_{WCT}}(w) \geq \alpha\}\Big]\end{aligned} \tag{26}$$

where $(a)_{\alpha}$ and $(c)_{\alpha}$ are the *left endpoints* of $(WCT)_{\alpha}$ and $(W_{WCT})_{\alpha}$. Similarly, $(b)_{\alpha}$ and $(d)_{\alpha}$ are the *right endpoints* of $(WCT)_{\alpha}$ and $(W_{WCT})_{\alpha}$, respectively.

Following the above, $(\Omega)_{\alpha}$ can be obtained such that

$$(\Omega)_{\alpha} = \Big[\min \varphi(x, x', w, w'), \max \varphi(x, x', w, w')\Big] \tag{27}$$

Due to the monotonicity of the function φ, this equation is simplified as

$$\begin{aligned}(\Omega)_{\alpha} = \Big[&\min_{w \in \{(c)_{\alpha}, (d)_{\alpha}\}, w' \in \{(c')_{\alpha}, (d')_{\alpha}\}} \varphi_L(w, w'), \\ &\max_{w \in \{(c)_{\alpha}, (d)_{\alpha}\}, w' \in \{(c')_{\alpha}, (d')_{\alpha}\}} \varphi_R(w, w')\Big]\end{aligned} \tag{28}$$

where $(a)_\alpha \leq x \leq (b)_\alpha$, $(c)_\alpha \leq w \leq (d)_\alpha$, $a, x, b \in WCT$ and $c, w, d \in W_{WCT}$. Similarly, $(a')_\alpha \leq x' \leq (b')_\alpha$, $(c')_\alpha \leq w' \leq (d')_\alpha$, $a', x', b' \in UQ$ and $c', w', d' \in W_{UQ}$. Also, $\varphi_L(w, w')$ and $\varphi_R(w, w')$ are defined by

$$\varphi_L(w, w') = \frac{(a)_\alpha w + (a')_\alpha w'}{w + w'} \tag{29}$$

$$\varphi_R(w, w') = \frac{(b)_\alpha w + (b')_\alpha w'}{w + w'} \tag{30}$$

Having obtained the aggregated fuzzy set Ω, it is necessary to interpret its value using pre-defined labels, such that consistent comparison can be achieved. Given a standard space of $FOM(Y)$, Ω is mapped onto the ordered label set $L(Y) = \{VL, L, M, H, VH\}$ that is specified on the universe of discourse $U^* = [0, 1]$. The corresponding fuzzy sets $F(Y) = \{f_{VL}, f_L, f_M, f_H, f_{VH}\}$ are identical to those of $FOM(W)$ (see Fig. 4). From this, Ω is matched against each label of $L(Y)$ as follows:

$$\Omega = (l, \beta_l(\Omega)), \forall\, l \in L(Y) \tag{31}$$

where $\beta_l(\Omega)$ is defined by

$$\beta_l(\Omega) = \max_{\forall t \in U^*} \Big(\min(\mu_{f_l}(t), \mu_\Omega(t)) \Big) \tag{32}$$

As an example, Fig. 5 exhibits two aggregated measures Ω_1 and Ω_2 that can be mapped onto $FOM(Y)$ such that: $\Omega_1 = (VL, 0), (L, 0), (M, 0.7), (H, 0.6), (VH, 0)$ and $\Omega_2 = (VL, 0), (L, 0.6), (M, 0.7)$, $(H, 0)$, $(VH, 0)$.

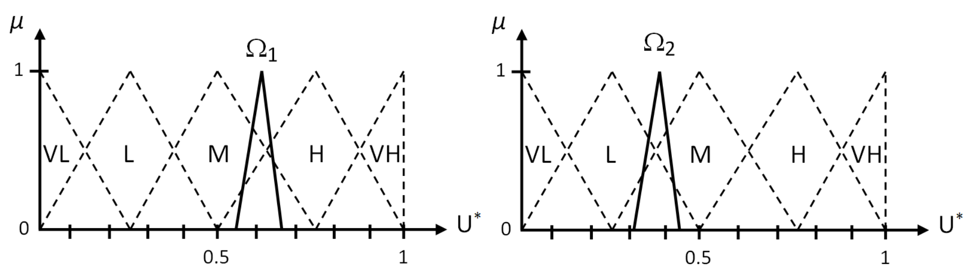

Fig. 5. Descriptions of aggregated measures Ω_1 and Ω_2 based on the qualitative space $FOM(Y)$ that is defined on $U^* = [0, 1]$.

V. Performance Evaluation

This experimental study is set to investigate the effectiveness of fuzzy qualitative link analysis approach, against a comparable method such as WSBA [48], for the task of academic performance evaluation. The following presents the problem definition which sets the scene for this evaluation, including details of the investigated datasets, and the discussion about the evaluation results.

A. Problem Formulation and Investigated Datasets

As with any data-driven learning problem, selecting a 'norm' group to be used as the basis for comparison is very important. Learned rules (or link network) can only be as good as the given set of data; they are simply a generalized version of the data. Hence, the example-assisted approach possesses an inherent limitation, regardless of the method employed (statistical, fuzzy rules or fuzzy qualitative link analysis). Nevertheless, it is reasonable to assume that there exists considerable amount of historical data which is representative to use (as is the case for any established educational organization), even though for the matter of illustrative convenience a relatively small dataset is adopted for training.

This evaluation is to first report the performance of the proposed model on the conventional dataset (i.e. 'SAP50A') as used in [48]. It contains 50 training instances and 15 testing cases. According to the problem formulation defined therein, the level of student's academic performance is expressed using five linguistic terms that can be regarded

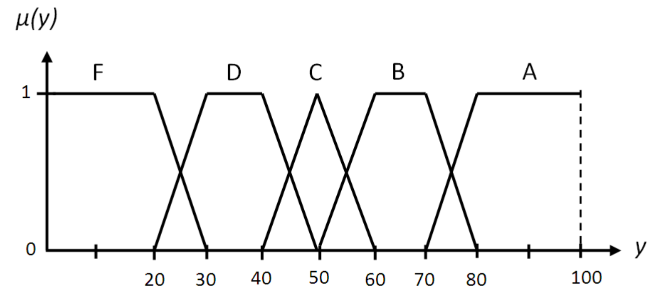

Fig. 6. Membership functions corresponding to five levels/grades of academic performance: F, D, C, B and A.

as classes: C_1 = Unsatisfactory (F), C_2 = Satisfactory (D), C_3 = Average (C), C_4= Good (B), C_5 = Excellent (A). Fig. 6 shows fuzzy sets defined for these performance classes, i.e. $\mu_{C_k}(y), y \in U^C, k = 1 \dots 5$ with $U^C = [0, 100]$.

The resulting performance level is determined by three collective scores of assignment, test and final exam. These scores are treated as attribute variables of a classification system – assignment, test and final exam which are denoted as A_1, A_2 and A_3, respectively. A set of five scoring terms, A_{11} = F, A_{12} = D, A_{13} = C, A_{14} = B and A_{15} = A, is used to represent the domain of variable A_1. They are defined as fuzzy sets shown in Fig. 7(a), i.e. membership functions $\mu_{A_{1j}}(x_1), x_1 \in U^{A_1}, j = 1 \dots 5$ with $U^{A_1} = [0, 100]$. Variables A_2 and A_3 are similarly defined each by five terms, with the corresponding fuzzy sets being depicted in Fig. 7(b) and (c).

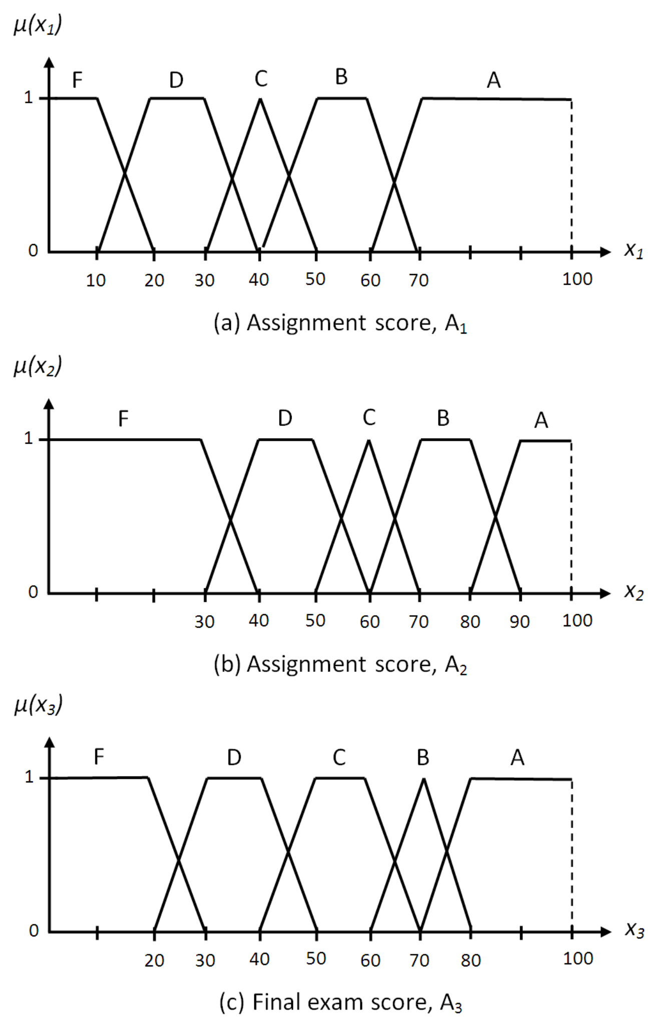

Fig. 7. Membership functions corresponding to five different terms (A, B, C, D and F) of (a) assignment score A_1, (b) test score A_2 and (c) final exam score A_3.

In addition to the experiment using a conventional set of data, this empirical study is also conducted on a randomly-generated dataset, which is referred to as 'RANDOM' hereafter. While the SAP50A dataset concentrates on typical patterns of students' performance that are relatively consistent across different scores, the RANDOM dataset presents patterns with scores being mostly fluctuated and inconsistent. This is relevant as different scores can be assessed at different occasions when illness and other unforeseen conditions may apply. Any performance assessment model should be evaluated with regard to both normal and atypical cases, in order to correctly appreciate its robustness.

The RANDOM dataset consists of 30 data instances, 20 of which are used as training samples and the rest as test instances. As shown in Table IV, each data instance consists of three basic scores, A_1, A_2, A_3, that are selected from the range of $[0, 100]$. Note that the probability of different scores to be chosen is uniform in this particular evaluation. The total performance score, $TS \in U^C$, is estimated as the statistical means of those collective values [48]. Given TS, the final performance grade GR is awarded in accordance with the standard classification shown below.

- $GR = F$ (Unsatisfactory), for the total score between [0-25]
- $GR = D$ (Satisfactory), for the total score between (25-45]
- $GR = C$ (Average), for the total score between (45-55]
- $GR = B$ (Good), for the total score between (55-75]
- $GR = A$ (Excellent), for the total score between (75-100]

TABLE IV
RANDOM DATASET: TRAINING AND TEST DATA INSTANCES.

No	Assignment	Test	Final exam	Total Score (TS)	Grade (GR)
Training data					
1	75.00	50.00	80.00	68.33	B
2	26.00	81.00	57.00	54.67	C
3	68.00	78.00	1.00	49.00	C
4	90.00	74.00	2.00	55.33	C
5	47.00	57.00	31.00	45.00	D
6	47.00	54.00	50.00	50.33	C
7	99.00	56.00	76.00	77.00	A
8	100.00	92.00	49.00	80.33	A
9	81.00	42.00	26.00	49.67	C
10	36.00	23.00	60.00	39.67	D
11	69.00	77.00	17.00	54.33	C
12	48.00	16.00	90.00	51.33	C
13	37.00	100.00	9.00	48.67	C
14	76.00	0.00	57.00	44.33	D
15	8.00	87.00	90.00	61.67	B
16	94.00	67.00	24.00	61.67	B
17	22.00	40.00	56.00	39.33	D
18	98.00	60.00	5.00	54.33	C
19	57.00	96.00	27.00	60.00	B
20	56.00	7.00	39.00	34.00	D
Test Data					
1	63.00	80.00	61.00	68.00	B
2	39.00	68.00	76.00	61.00	B
3	88.00	92.00	55.00	78.33	A
4	54.00	70.00	33.00	52.33	C
5	73.00	20.00	5.00	32.67	D
6	16.00	54.00	48.00	39.33	D
7	48.00	100.00	25.00	57.67	B
8	28.00	48.00	23.00	33.00	D
9	82.00	17.00	59.00	52.67	C
10	59.00	98.00	70.00	75.67	A

B. Evaluation Results and Discussion

The objective of this evaluation is to provide evidence that the proposed approach can produce results similar to the expected grades that are provided as ground truth. According to the work of [48], this is derived from the

TABLE V
THE RESULTING LEVELS OF PERFORMANCE GENERATED BY DIFFERENT EXAMPLE-DIRECTED MODELS ON THE SAP50A DATASET.

Test No	**Expected Grade**	**Predicted Grade by Examined Methods**			
		WSBA	Ω_1	Ω_2	Ω_3
1	F	F	F	F	F
2	F	F	F	F	F
3	F	F	F	F	F
4	D	D	D	D	D
5	D	D	D	D	D
6	D	C*	D	C*	D
7	C	C	C	C	C
8	C	C	C	C	C
9	C	C	C	C	C
10	B	B	C*	C*	C*
11	B	B	B	B	B
12	B	B	B	B	B
13	A	A	A	A	A
14	A	A	A	A	A
15	A	A	A	A	A

* Indicates that a predicted grade is different from the expected value.

TABLE VI
THE RESULTING LEVELS OF PERFORMANCE GENERATED BY DIFFERENT EXAMPLE-DIRECTED MODELS ON THE RANDOM DATASET.

Test No	**Expected Grade**	**Predicted Grade by Examined Methods**			
		WSBA	Ω_1	Ω_2	Ω_3
1	B	C*	C*	B	C*
2	B	C*	C*	C*	B
3	A	A	A	A	A
4	C	B*	C	C	C
5	D	C*	C*	C*	C*
6	D	D	D	D	D
7	B	C*	B	C*	B
8	D	C*	D	D	D
9	C	C	D*	D*	D*
10	A	B*	A	A	A

* Indicates that a predicted grade is different from the expected value.

average of collected scores. To obtain a robust conclusion, three different link analysis models, namely Ω_1, Ω_2 and Ω_3, are assessed using three different weight settings (see Section IV):

- For Ω_1, W$_{WCT} = VH$ and W$_{UQ} = VH$
- For Ω_2, W$_{WCT} = M$ and W$_{UQ} = VH$
- For Ω_3, W$_{WCT} = VH$ and W$_{UQ} = M$

Table V presents the experimental results obtained by WSBA and the three link analysis models on the SAP50A dataset. These results suggest that the proposed methods are as effective as WSBA for the typical patterns of students' performance. On the other hand, based on Table VI that reports the findings on the RANDOM dataset, WSBA often (7 out of 10, to be exact) generates the final grades that are different from the expected values and the Ω_3 method is the most accurate amongst examined techniques. Specifically, Ω_3 outputs a correct grade for seven test instances, whilst Ω_1 and Ω_2 are similarly accurate for six cases. This finding suggests that the proposed link analysis approach is more effective than the rule-based counterpart, and robust to the perturbation of its weighting parameters.

Tables VII and VIII show the grade-specific likelihood measures on the RANDOM dataset, generated by WSBA and the fuzzy qualitative link analysis methods respectively. These results provide additional information with respect to the strength of student's performance belonging to a specific grade. This can be very useful in differentiating smoothly student performances over boundary cases, giving a second opinion in deciding on borderline performances. Aiding for this qualitative revision, a linguistic approach employed by the link analysis-based models is naturally interpretable and more effective as a communication means than the numerical metric used by WSBA. Since the quality of a learning model is data dependant, the proposed model may be utilized in conjunction with other example-assisted and expert-directed alternatives. This practice can further improve the

TABLE VII
LIKELIHOOD OF AWARDED GRADES GENERATED BY WSBA.

Test No	Likelihood of Awarded Grade				
	F	D	C	B	A
1	0.00	0.04	0.42	0.19	0.00
2	0.00	0.00	0.13	0.04	0.00
3	0.00	0.05	0.39	0.21	1.00
4	0.00	0.04	0.28	0.34	0.00
5	0.00	0.10	0.53	0.05	0.00
6	0.00	0.31	0.13	0.11	0.00
7	0.00	0.05	0.39	0.27	0.00
8	0.00	0.18	0.22	0.13	0.00
9	0.00	0.44	0.50	0.05	0.00
10	0.00	0.00	0.00	0.12	0.00

TABLE VIII
LIKELIHOOD OF AWARDED GRADES GENERATED BY LINK-BASED METHODS: AGG_1, AGG_2 AND AGG_3. NOTE THAT THE MATCHING DEGREES TO STANDARD LABELS (VL, L, M, H, VH) ARE GIVEN IN BRACKETS.

Method	Test No	Likelihood of Awarded Grade				
		F	D	C	B	A
AGG_1	1	VL(1.0)	VH(0.4), H(0.9)	VH(0.74), H(0.76)	VH(0.66), H(0.84)	VH(0.34), H(0.84)
	2	VL(1.0)	VH(0.3), H(0.8)	VH(0.68), H(0.82)	VH(0.66), H(0.84)	VH(0.34), H(0.84)
	3	VL(1.0)	H(0.5), M(1.0)	VH(0.2), H(0.7)	VH(0.2), H(0.7)	VH(0.84), H(0.66)
	4	VL(1.0)	VH(0.46), H(0.96)	VH(0.56), H(0.94)	VH(0.48), H(0.98)	M(0.24), L(0.74)
	5	VL(1.0)	VH(0.22), H(0.72)	VH(0.56), H(0.94)	VH(0.3), H(0.8)	H(0.1), M(0.6)
	6	VL(1.0)	VH(0.86), H(0.64)	VH(0.6), H(0.9)	VH(0.68), H(0.82)	VH(0.4), H(0.9)
	7	VL(1.0)	VH(0.6), H(0.9)	VH(0.78), H(0.72)	VH(0.8), H(0.7)	H(0.36), M(0.86)
	8	VL(1.0)	VH(0.8), H(0.7)	VH(0.62), H(0.88)	VH(0.46), H(0.96)	H(0.08), M(0.58)
	9	VL(1.0)	VH(0.6), H(0.9)	VH(0.3), H(0.8)	H(0.38), M(0.88)	VH(0.16), H(0.66)
	10	VL(1.0)	H(0.28), M(0.78)	H(0.36), M(0.86)	VH(0.28), H(0.78)	VH(0.38), H(0.88)
AGG_2	1	VL(1.0)	VH(0.44), H(0.94)	VH(0.72), H(0.78)	VH(0.76), H(0.74)	VH(0.5), H(1.0)
	2	VL(1.0)	VH(0.5), H(1.0)	VH(0.78), H(0.72)	VH(0.76), H(0.74)	VH(0.52), H(0.98)
	3	VL(1.0)	H(0.5), M(1.0)	VH(0.14), H(0.64)	VH(0.14), H(0.64)	VH(0.78), H(0.72)
	4	VL(1.0)	VH(0.5), H(1.0)	VH(0.56), H(0.94)	VH(0.5), H(1.0)	M(0.28), L(0.78)
	5	VL(1.0)	VH(0.26), H(0.76)	VH(0.52), H(0.98)	VH(0.26), H(0.76)	H(0.06), M(0.56)
	6	VL(1.0)	VH(0.9), H(0.6)	VH(0.62), H(0.88)	VH(0.78), H(0.72)	VH(0.44), H(0.94)
	7	VL(1.0)	VH(0.74), H(0.76)	VH(0.86), H(0.64)	VH(0.84), H(0.66)	H(0.38), M(0.88)
	8	VL(1.0)	VH(0.86), H(0.64)	VH(0.7), H(0.8)	VH(0.54), H(0.96)	H(0.06), M(0.56)
	9	VL(1.0)	VH(0.58), H(0.92)	VH(0.24), H(0.74)	H(0.38), M(0.88)	VH(0.12), H(0.62)
	10	VL(1.0)	H(0.3), M(0.8)	VH(0.36), M(0.86)	VH(0.36), H(0.86)	VH(0.46), H(0.96)
AGG_3	1	VL(1.0)	VH(0.36), H(0.86)	VH(0.76), H(0.74)	VH(0.6), H(0.9)	VH(0.18), H(0.68)
	2	VL(1.0)	VH(0.1), H(0.6)	VH(0.58), H(0.92)	VH(0.6), H(0.9)	VH(0.16), H(0.66)
	3	VL(1.0)	VH(0.02), H(0.52)	VH(0.28), H(0.78)	VH(0.26), H(0.76)	VH(0.88), H(0.62)
	4	VL(1.0)	VH(0.42), H(0.92)	VH(0.54), H(0.96)	VH(0.44), H(0.94)	M(0.2), L(0.7)
	5	VL(1.0)	VH(0.18), H(0.68)	VH(0.6), H(0.9)	VH(0.32), H(0.82)	H(0.12), M(0.62)
	6	VL(1.0)	VH(0.82), H(0.68)	VH(0.58), H(0.92)	VH(0.58), H(0.92)	VH(0.38), H(0.88)
	7	VL(1.0)	VH(0.46), H(0.96)	VH(0.7), H(0.8)	VH(0.78), H(0.72)	H(0.34), M(0.84)
	8	VL(1.0)	VH(0.74), H(0.76)	VH(0.54), H(0.96)	VH(0.38), H(0.88)	H(0.08), M(0.58)
	9	VL(1.0)	VH(0.6), H(0.9)	VH(0.36), H(0.86)	H(0.38), M(0.88)	VH(0.22), H(0.72)
	10	VL(1.0)	H(0.26), M(0.76)	H(0.36), M(0.86)	VH(0.22), H(0.72)	VH(0.3), H(0.8)

accuracy of the final solution.

VI. CONCLUSIONS

This paper has presented a novel fuzzy qualitative classification system for academic performance evaluation using the link analysis methodology. Unlike the conventional approach where fuzzy rules are used to encode information provided by training data, the proposed model considers involving variables, classes and their relations as elements of a social network that can be modelled as a weighted graph. This unique scheme allows the core information metric (i.e. subsethood) to be interpreted via multiple perspectives, each of which is realized by a specific link measure.

A fuzzy qualitative model of similarity estimation is introduced to generate similarity measures in an accurate and efficient manner. Note that the theory of fuzzy sets is employed to extend the conventional order-of-magnitude based qualitative reasoner, such that qualitative descriptors are well defined quantitatively and the results of their manipulation are less ambiguous. Empirical studies show that the fuzzy qualitative link analysis is more effective than the state-of-the-art fuzzy rule-based technique for academic performance evaluation. It is also robust to the perturbation of its parameters, i.e. weights assigned to different link measures. In addition, the resulting linguistic descriptions of grade-specific likelihood are useful for further revisions regarding the formal partition of performance levels and students' improvement.

Despite the promising findings, the potential of link analysis approach for a general classification problem is to be further investigated. For this challenging task, fuzzy terms that are pre-defined in the current research, may be obtained using data-driven mechanisms extensively developed in the literature. As for the present work with a single-type link, its performance may be enhanced via a refinement of link representation. In particular, each link type may correspond to a specific order-of-magnitude of the underlying link strength that the link association is to capture. Also, it is interesting to examine the behavior of the proposed approach with respect to other link measures. In particular to the estimation of fuzzy subsethood metric where the minimum is currently employed, the utilization of other aggregation operators such as OWA (Ordered Weighted Averaging) [63] and its data-dependent variants [10], [13] may form another piece of important future research.

References

[1] S. Abe and M. S. Lan. A method for fuzzy rules extraction directly from numerical data and its application to pattern classification. *IEEE Transactions on Fuzzy Systems*, 3(1):18–28, 1995.

[2] S. Abe and R. Thawonmas. A fuzzy classifier with ellipsoidal regions. *IEEE Transactions on Fuzzy Systems*, 5(3):358–368, 1997.

[3] A. H. Ali, D. Dubois, and H. Prade. Qualitative reasoning based on fuzzy relative orders of magnitude. *IEEE Transactions on Fuzzy Systems*, 11(1):9–23, 2003.

[4] R. Amsler. *Application of citation-based automatic classification*. Unpublished manuscript, The University of Texas at Austin, Linguistics Research Center, Austin, TX, 1972.

[5] S. M. Bai and S. M. Chen. Automatically constructing grade membership functions of fuzzy rules for students' evaluation. *Expert Systems with Applications*, 35(3):1408–1414, 2008.

[6] S. M. Bai and S. M. Chen. Evaluating students' learning achievement using fuzzy membership functions and fuzzy rules. *Expert Systems with Applications*, 34:399–410, 2008.

[7] I. Bhattacharya and L. Getoor. Collective entity resolution in relational data. *ACM Transactions on KDD*, 1(1), 2007.

[8] M. Bilenko and R. J. Mooney. Adaptive duplicate detection using learnable string similarity measures. In *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pages 39–48, 2003.

[9] R. Biswas. An application of fuzzy sets in students' evaluation. *Fuzzy Sets and Systems*, 74:187–194, 1995.

[10] T. Boongoen and Q. Shen. Clus-DOWA: A new dependent OWA operator. In *Proceedings of IEEE International Conference on Fuzzy Sets and Systems*, pages 1057–1063, 2008.

[11] T. Boongoen and Q. Shen. Detecting false identity through behavioural patterns. In *Proceedings of International Crime Science Conference*, 2008.

[12] T. Boongoen and Q. Shen. Order-of-magnitude based link analysis for false identity detection. In *Proceedings of the 23rd International Workshop on Qualitative Reasoning*, pages 7–15, 2009.

[13] T. Boongoen and Q. Shen. Nearest-neighbor guided evaluation of data reliability and its applications. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, DOI:10.1109/TSMCB.2010.2043357.

[14] T. Boongoen, Q. Shen, and C. Price. Disclosing false identity through hybrid link analysis. *AI and Law*, 18(1):77–102, 2010.

[15] G. Bordogna, G. Pasi, and R. R. Yager. Soft approaches to distributed information retrieval. *International Journal of Intelligent Systems*, 34:105–120, 2003.

[16] M. Buechi, A. Borthwick, A. Winkel, and A. Goldberg. ClueMaker: A language for approximate record matching. In *Proceedings of International Conference on Information Quality*, pages 207–223, 2003.

[17] P. Calado, M. Cristo, M. A. Gonçalves, E. S. de Moura, B. A. Ribeiro-Neto, and N. Ziviani. Link-based similarity measures for the classification of web documents. *Journal of American Society for Information Science and Technology*, 57(2):208–221, 2006.

[18] C. Carlsson, R. Fullér, and S. Fullér. OWA operators for doctoral student selection problem. In *The ordered weighted averaging operators: theory and applications*, pages 167–177, Norwell, MA, USA, 1997. Kluwer Academic Publishers.

[19] C. S. Chan and H. Liu. Fuzzy qualitative human motion analysis. *IEEE Transactions on Fuzzy Systems*, 17(4):851–862, 2009.

[20] P. T. Chang, K. C. Hung, K. P. Lin, and C. H. Chang. A comparison of discrete algorithms for fuzzy weighted average. *IEEE Transactions on Fuzzy Systems*, 14(5):663 – 675, 2006.

[21] S. M. Chen and C. H. Lee. New methods for students' evaluation using fuzzy sets. *Fuzzy Sets and Systems*, 104:209–218, 1999.

[22] M. De Cock, C. Cornelis, and E. E. Kerre. Elicitation of fuzzy association rules from positive and negative examples. *Fuzzy Sets and Systems*, 149(1):73–85, 2005.

[23] W. Cohen and J. Richman. Learning to match and cluster large high-dimensional data sets for data integration. In *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pages 475–480, 2002.

[24] C. Fourali. Fuzzy logic and the quality of assessment of portfolios. *Fuzzy Sets and Systems*, 68:123–139, 1994.

[25] F. Fouss, A. Pirotte, J. M. Renders, and M. Saerens. Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 19(3):355–369, 2007.

[26] L. Getoor and C. P. Diehl. Link mining: a survey. *ACM SIGKDD Explorations Newsletter*, 7(2):3–12, 2005.

[27] C. Gipps and G. Stobart. *Assessment: A teachers' guide to the issue*. Hodder and Stoughton, London, 1993.
[28] P. Hsiung, A. Moore, D. Neill, and J. Schneider. Alias detection in link data sets. In *Proceedings of International Conference on Intelligence Analysis*, 2005.
[29] N. Iam-on, T. Boongoen, and S. Garrett. LCE: A link-based cluster ensemble method for improved gene expression data analysis. *Bioinformatics*, DOI:10.1093/bioinformatics/btq226.
[30] H. Ishibuchi, T. Nakashima, and T. Murata. Three-objective genetics-based machine learning for linguistic rule extraction. *Information Sciences*, 136(1-4):109–133, 2001.
[31] H. Ishibuchi, K. Nozaki, and H. Tanaka. Distributed representation of fuzzy rules and its application to pattern classification. *Fuzzy Sets and Systems*, 52(1):21–32, 1992.
[32] H. Ishibuchi and T. Yamamoto. Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 13(4):428–435, 2005.
[33] G. Jeh and J. Widom. Simrank: A measure of structural-context similarity. In *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 538–543, 2002.
[34] M.M. Kessler. Bibliographic coupling between scientific papers. *American Documentation*, 14(1):10–25, 1963.
[35] S. Klink, P. Reuther, A. Weber, B. Walter, and M. Ley. Analysing social networks within bibliographical data. In *Proceedings of International Conference on Database and Expert Systems Applications*, pages 234–243, 2006.
[36] L. I. Kuncheva and J. C. Bezdek. A fuzzy generalized nearest prototype classifier. In *Proceedings of the 7th International Fuzzy Systems Association World Congress*, volume III, pages 217–222, 1997.
[37] C. K. Law. Using fuzzy numbers in educational grading system. *Fuzzy Sets and Systems*, 83:311–323, 1996.
[38] S. Lawrence, C. L. Giles, and K. D. Bollacker. Autonomous citation matching. In *Proceedings of International Conference on Autonomous Agents*, pages 392–393, 1999.
[39] C. C. Lee. Fuzzy logic in control systems: fuzzy logic controller part i and part ii. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2):404–435, 1990.
[40] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 58(7):1019–1031, 2007.
[41] Z. Lin, I. King, and M. R. Lyu. Pagesim: A novel link-based similarity measure for the world wide web. In *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence*, pages 687–693, 2006.
[42] B. Malin. Unsupervised name disambiguation via social network similarity. In *Proceedings of SIAM Workshop on Link Analysis, Counterterrorism and Security*, pages 93–102, 2005.
[43] E. Minkov, W. W. Cohen, and A. Y. Ng. Contextual search and name disambiguation in email using graphs. In *Proceedings of International Conference on Research and Development in Information Retrieval*, pages 27–34, 2006.
[44] T. Nakashima, G. Schaefer, Y. Yokota, and H. Ishibuchi. A weighted fuzzy classifier and its application to image processing tasks. *Fuzzy Sets and Systems*, 158:284–294, 2007.
[45] D. Nauck and R. Kruse. A neuro-fuzzy method to learn fuzzy classification rules from data. *Fuzzy Sets and Systems*, 89(3):277–288, 1997.
[46] C. J. Price, L. Travé-massuyès, R. Milne, L. Ironi, K. Forbus, B. Bredeweg, M. H. Lee, P. Struss, N. Snooke, P. Lucas, M. Cavazza, and G. M. Coghill. Qualitative futures. *Knowledge Engineering Review*, 21(4):317–334, 2006.
[47] K. A. Rasmani and Q. Shen. Modifying fuzzy subsethood-based rule models with fuzzy quantifiers. In *Proceedings of IEEE International Conference on Fuzzy Systems*, pages 1679–1684, 2004.
[48] K. A. Rasmani and Q. Shen. Data-driven fuzzy rule generation and its application for student academic performance evaluation. *Applied Intelligence*, 25(3):305–319, 2006.
[49] P. Reuther and B. Walter. Survey on test collections and techniques for personal name matching. *International Journal on Metadata, Semantics and Ontologies*, 1(2):89–99, 2006.
[50] G. Salton. *Automatic Text Processing*. Addison-Wesley, 1989.
[51] G. Salton and M. J. McGill. *Introduction to modern information retrieval*. McGraw-Hill, New York, 1983.
[52] Q. Shen and A. Chouchoulas. A rough-fuzzy approach for generating classification rules. *Pattern Recognition*, 35:2425–2438, 2002.
[53] Q. Shen and R. Leitch. On extending the quantity space in qualitative reasoning. *AI in Engineering*, 7:167–173, 1992.
[54] H. Small. Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24:265–269, 1973.
[55] M. Sugeno. An introductory survey of fuzzy control. *Information Sciences*, 36(1/2):59–83, 1985.
[56] J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos. Relevance search and anomaly detection in bipartite graphs. *ACM SIGKDD Explorations Newsletter*, 7(2):48–55, 2005.
[57] H. Tong, C. Faloutsos, and J. Y. Pan. Fast random walk with restart and its applications. In *Proceedings of IEEE International Conference on Data Mining*, pages 613–622, 2006.
[58] A. G. Wang, H. Atabakhsh, T. Petersen, and H. Chen. Discovering identity problems: A case study. In *Proceedings of IEEE International Conference on Intelligence and Security Informatics*, pages 368–373, 2005.
[59] C. Wang, V. Satuluri, and S. Parthasarathy. Local probabilistic models for link prediction. In *Proceedings of IEEE International Conference on Data Mining*, pages 322–331, 2007.
[60] H. Y. Wang and S. M. Chen. Evaluating students' answerscripts using fuzzy numbers associated with degrees of confidence. *IEEE Transactions on Fuzzy Systems*, 16(2):403–415, 2008.
[61] S. Wasserman and K. Faust. *Social network analysis: methods and applications*. Cambridge University Press, 1994.
[62] S. Weon and J. Kim. Learning achievement evaluation strategy using fuzzy membership function. In *Proceedings of ASEE/IEEE Frontiers in Education Conference*, pages 10–13, 2001.
[63] R. R. Yager. On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man, and Cybernetics*, 18:183–190, 1988.

[64] R. R. Yager. Intelligent social network analysis using granular computing. *International Journal of Intelligent Systems*, 23:1197–1220, 2008.
[65] Y. Yuan and M. J. Shaw. Induction of fuzzy decision trees. *Fuzzy Sets and Systems*, 69:125–139, 1995.
[66] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
[67] H. J. Zimmermann. *Fuzzy Set Theory - And Its Applications*. Kluwer, MA, 2001.