The Lattice of Verbal Fuzzy Numbers in IP-Traffic Prediction

Elisabeth Rakus-Andersson

The School of Engineering, Blekinge Institute of Technology
S-37179 Karlskrona, Sweden
Elisabeth.Andersson@bth.se

Abstract. Fuzzy numbers constitute a helpful tool for some users who deal with imprecise data. To facilitate a procedure of performing operations on fuzzy numbers for these users who do not possess deep mathematical qualifications, we introduce a special space of fuzzy numbers verbally defined. This consists of families of fuzzy numbers in the *L-R* form provided that every family has a common mean value. We not only discuss some properties of the space, but also propose new conceptions of the arithmetical operations on the verbally expressed fuzzy numbers. The classification of the IP-traffic can act as a practical aspect, which clarifies the theoretical assumptions formulated in the paper.

1 Introduction

An interdisciplinary cooperation between scientists representing different domains of knowledge is sometimes complicated because of different ways of expressing the obtained results. For example, when a mathematician defines a fuzzy number, he or she will remember the theoretical information about it as definitions, theorems and the like. A fuzzy number in the well-known *L-R* representation appears in the mathematician's imagination as a triple containing a mean and two spreads that can be included in a sequence of operations.

Contrary to this impression a practician formulates the meaning of a fuzzy number by using verbal expressions as, e. g., "almost exactly 4" or "near 4" without taking care of the mathematical appearance of such descriptions. Moreover, after performing several operations he also expects an answer in a verbal form because the values of borders do not often give him the clear knowledge about the result. Since the classical operations on fuzzy numbers [2, 3, 4, 5, 7, 8, 9, 10, 12, 16] usually extend the spreads of the results in the inconvenient way, we suggest accepting the concept of a number space with members, which can be determined verbally (Section 2). In accordance with a list that is stated for more and more imprecise numbers with a common mean value we get in advance a design of spreads associated with these numbers. Every number is now identified by its name, which gives us the reason to christen the set of such numbers as "the space of fuzzy numbers verbally defined."

The space is countable and has its own total order that differs from some methods of ranking fuzzy numbers proposed by [1, 6, 11, 13, 14, 15, 20, 21]. To keep the result of executed operations on verbal fuzzy numbers within some borders, we suggest introducing of new arithmetic over the space in Section 3.

Section 4 reveals a trial of a practical application of verbally defined fuzzy numbers. By interpolating a set of points with vague coordinates describing the package sizes and associated densities in the DNS traffic, we derive the formula of a fuzzy polynomial, which identifies this sort of Internet protocols.

The space of verbal fuzzy numbers has a regular structure, which makes easy to find the numerical representation of an event verbally expressed. This encourages us in making the first attempt of proposing fuzzy sets as Yager's probabilities of events in which a certain variable takes a value being a verbal fuzzy number. This approach is discussed in Section 5.

The construction of the space of verbal fuzzy numbers should help the possible user of applicable fuzzy theory to interpret his data and results in the form he understands best, i.e., as some expressions belonging to the natural language. We assume that the user does not need a great accuracy in his investigations, but he wants to obtain an easily accessible orientation in his problem.

2 The Verbal Fuzzy Number Space

Some experiments carried out on a sample of objects lead to imprecise values of the observed parameters. The scientists often seek a way of expressing the data when it is not known exactly. The idea of a fuzzy number seems to be useful in the considered case. Let us recall the conception of a fuzzy number M in L-R representation as the following definition [2, 3, 21].

Definition 1
A fuzzy number M is of L-R type if there exist reference functions L (for left) and R (for right), and scalars $\alpha>0$, $\beta>0$ with

$$y = \mu_M(x) = \begin{cases} L(\frac{m-x}{\alpha}) \text{ for } x \le m \\ R(\frac{x-m}{\beta}) \text{ for } x \ge m, \end{cases} \quad (1)$$

where m, called the mean value of M, is a real number, and α and β are called the left and right spreads, respectively. The functions L and R map $R^+ \to [0,1]$. The function L should satisfy $L(0)=1$, $L(x)<1$ for every $x>0$; $L(x)>0$ for every $x<1$; $L(1)=0$. The same conditions refer to the function R. Symbolically, M is denoted by (m, α, β).

2.1 Fuzzy Numbers Verbally Defined and Their Mathematical Adaptation

Let us first suppose that $M=(m, \alpha, \beta)$, $m\in Z$, represents the fuzzy number named as, e.g., "close to *m*" or "very close to *m*". We try to designate the difference between "close" and "very close" by establishing the appropriate values of the spreads α and β for both fuzzy numbers. Let us suggest a list containing the most popular, verbal expressions, which are likely used for the imprecise numbers by some scientists. We denote the list by $LIST_m$ and state its contents as [17, 18, 19]

$$LIST_m = \{M_0 = \text{exactly}\, m,\ M_1 = \text{almost exactly}\, m,\ M_2 = \text{very close to}\, m,\ M_3 = \text{close to}\, m,\ M_4 = \text{rather close to}\, m,\ M_5 = \text{appreciatively}\, m,\ M_6 = \text{perhaps}\ m\}. \quad (2)$$

Users who need more expressions can extend the contents of the list.

Each notion of the $LIST_m$ is proposed to have an *L*-*R* representation suggested as

$$M_{d,i} = (m, d\cdot i, d\cdot i),\ m\in Z,\ i = 0, 1, 2, 3, 4, 5, 6,\ d\in R\,, \quad (3)$$

where d is a fix multiplier accommodating the breadth of the spreads $d\cdot i$ to a concrete problem. The assumption $m\in Z$ seems to fit best in the case of imprecise numbers instead of $m\in R$. If m is a real number it can be given as the number with a very high decimal precision, and we will probably not need any fuzzy representation of such m. If m belongs to Z it will be reasonable to decide the magnitude of d as 0.1.

If a user sometimes needs a higher accuracy of the mean m, e.g., as a real number with one digit in the decimal expansion, it will be possible to arrange the fuzzy numbers as $M_{d,i} = (m.k, d\cdot i, d\cdot i),\ m\in Z,\ k = 0,...,9,\ i = 0, 1, 2, 3, 4, 5, 6,\ d = 0.01$.

By accepting, for $x\in R$, that

$$L(x) = R(x) = -x + 1 \quad (4)$$

we derive new formulas of the membership functions for $M_{d,i},\ i = 1, ..., 6$, expanded as

$$y = \mu_{M_{d,i}}(x) = \begin{cases} L(\frac{m-x}{d\cdot i}) = -\frac{m-x}{d\cdot i} + 1 = \frac{x + d\cdot i - m}{d\cdot i} & \text{for } x\in[m - d\cdot i, m] \\ R(\frac{x-m}{d\cdot i}) = -\frac{x-m}{d\cdot i} + 1 = \frac{d\cdot i + m - x}{d\cdot i} & \text{for } x\in[m, m + d\cdot i]. \end{cases} \quad (5)$$

The formula (5), found on the basis of (4) is comparable to Fullér's presentation of the membership function of a triangular fuzzy number. Let us analyse his approach to the *L*-*R* fuzzy number as [7]

$$y = \mu_{M_{d,i}}(x) = \begin{cases} 1 - \dfrac{|m-x|}{d \cdot i} = 1 - \dfrac{+(m-x)}{d \cdot i} = \dfrac{x + d \cdot i - m}{d \cdot i} & \text{for } x \in [m - d \cdot i, m] \\ 1 - \dfrac{|m-x|}{d \cdot i} = 1 - \dfrac{-(m-x)}{d \cdot i} = \dfrac{d \cdot i + m - x}{d \cdot i} & \text{for } x \in [m, m + d \cdot i] \end{cases} \tag{6}$$

to conclude that the membership function of $M_{d,i}$ in (5) and (6) is the same in spite of the different origins.

The following picture shows two adjacent groups of fuzzy numbers for $m = 3$ respectively $m = 4$, which are generated by (3).

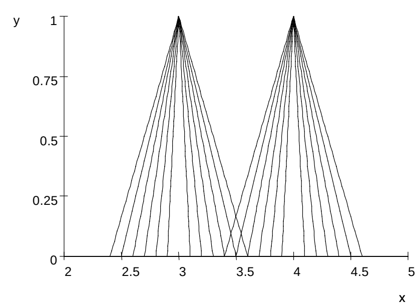

Fig 1. The fuzzy numbers belonging to $LIST_3$ and $LIST_4$

2.2 The Order in the Space of Verbal Fuzzy Numbers

Let us consider the space of fuzzy numbers of the type $M_{d,i} = (m, d \cdot i, d \cdot i), \quad m \in Z$, $i = 0, 1, 2, 3, 4, 5, 6,\ d = 0.1$ denoted by *VBFN*(*LIST*). The space contains the fuzzy numbers verbally described at the first stage and mathematically formalized afterwards.

Many scientists defined the order among elements in the newly created spaces [6, 11, 13, 14, 20, 21]. The space *VBFN*(*LIST*) also has its own order, which can be suggested as the effect of the operation $\leq$.

The operation $\leq$ is treated as the relation $\Re$ with the property
$(m, d \cdot i, d \cdot i) \Re (n, d \cdot j, d \cdot j)$ if and only if $(m, d \cdot i, d \cdot i) \leq (n, d \cdot j, d \cdot j)$, for
$(m, d \cdot i, d \cdot i), (n, d \cdot j, d \cdot j) \in VBFN(LIST)$.

Definition 2

We say that

a) $(m, d \cdot i, d \cdot i) < (n, d \cdot j, d \cdot j)$ iff $m < n$,
$(m, d \cdot i, d \cdot i) \in LIST_m$, $(n, d \cdot j, d \cdot j) \in LIST_n$ $i, j = 0, 1, 2, 3, 4, 5, 6,\ d = 0.1$. The values of *i* and *j* have no effect if $m \neq n$;

b) $(m, d \cdot i, d \cdot i) < (m, d \cdot j, d \cdot j)$ iff $i > j$, $(m, d \cdot i, d \cdot i), (m, d \cdot j, d \cdot j) \in LIST_m$;

c) $(m, d \cdot i, d \cdot i) = (n, d \cdot j, d \cdot j)$ iff $m = n$ and $i = j$.

Using the above definition we can ascertain that, e.g., (3, 0.2, 0.2)<(5, 0.3, 0.3) while (4, 0.3, 0.3)<(4, 0.1, 0.1).

If $(m, d \cdot i, d \cdot i)\Re(m, d \cdot i, d \cdot i)$ then $(m, d \cdot i, d \cdot i) \leq (m, d \cdot i, d \cdot i)$ and $\Re$ is reflexive.

We assume that $(m, d \cdot i, d \cdot i)\Re(n, d \cdot j, d \cdot j)$, which results in the inequality $(m, d \cdot i, d \cdot i) \leq (n, d \cdot j, d \cdot j)$. If we also accept the premise $(n, d \cdot j, d \cdot j)\Re$ $(m, d \cdot i, d \cdot i)$, which implies $(m, d \cdot i, d \cdot i) \geq (n, d \cdot j, d \cdot j)$, we shall conclude that $(m, d \cdot i, d \cdot i) = (n, d \cdot j, d \cdot j)$. The last statement confirms that $\Re$ is the anti-symmetric relation.

Further, the assumption $(m, d \cdot i, d \cdot i)\Re(n, d \cdot j, d \cdot j)$ builds the statement $(m, d \cdot i, d \cdot i) \leq (n, d \cdot j, d \cdot j)$. If even $(n, d \cdot j, d \cdot j)\Re(p, d \cdot k, d \cdot k)$, which brings $(n, d \cdot j, d \cdot j) \leq (p, d \cdot k, d \cdot k)$, then $(m, d \cdot i, d \cdot i) \leq (p, d \cdot k, d \cdot k)$. The last inequality is equivalent to $(m, d \cdot i, d \cdot i)\Re$ $(p, d \cdot k, d \cdot k)$, and this proves that $\Re$ is the transitive relation. The numbers $(m, d \cdot i, d \cdot i)$, $(n, d \cdot j, d \cdot j)$, $(p, d \cdot k, d \cdot k)$ belong to *VBFN*(*LIST*), $i, j, k = 0, \ldots, 6$.

Possessing the features of reflexivity, anti-symmetry and transitivity, the relation $\Re$ introduces the partial order into *VBFN*(*LIST*), and we can declare that the space of verbal fuzzy number is a poset.

Let us further investigate if the poset (*VBFN*(*LIST*), $\leq$) has the total order.

Definition 3

If (*SPACE*, $\Re$) is a poset, we say that *SPACE* is totally ordered if for all $x, y \in SPACE$ either $x\Re y$ or $y\Re x$.

According to Definition 2 it can be checked that two verbal fuzzy numbers coming from *SPACE* = *VBFN*(*LIST*) satisfy either $(m, d \cdot i, d \cdot i) \leq (n, d \cdot j, d \cdot j)$ or $(n, d \cdot j, d \cdot j) \leq (m, d \cdot i, d \cdot i)$ that inserts the total order into *VBFN*(*LIST*).

The drawing below illustrates the order among the verbal fuzzy numbers from *VBFN*(*LIST*).

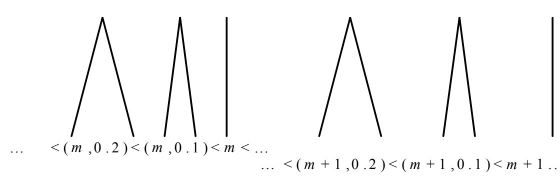

Fig.2 The order among fuzzy numbers from *VBFN*(*LIST*)

Since we always are able to determine both the least and the greatest element for every pair of members of the space *VBFN*(*LIST*), we then will assert that the verbal fuzzy numbers constitute the lattice (*VBFN*(*LIST*), $\leq$).

3 The Operations on Verbal Fuzzy Numbers

The next problem to solve is to implement four arithmetic operations performed on verbal fuzzy numbers. Let us first discuss one disadvantage concerning the classical definition of the product of two fuzzy numbers.

We thus quote the definition as follows regarding [2, 3]: "The product of two arbitrary fuzzy numbers $A=(m_A,\alpha_A,\beta_A)_{L-R}$ and $B=(m_B,\alpha_B,\beta_B)_{L-R}$ is equal to $(m_A,\alpha_A,\beta_A)_{L-R}\cdot(m_B,\alpha_B,\beta_B)_{L-R} \cong (m_A\cdot m_B, m_A\cdot\alpha_B+m_B\cdot\alpha_A, m_A\cdot\beta_B+m_B\cdot\beta_A)$ if $m_A>0$, $m_B>0$." Suppose that, in the particular case, $A = (m_A, \alpha, \alpha)$ and $B = (m_B, \alpha, \alpha)$. Then the product $A\cdot B \cong (m_A\cdot m_B, \alpha(m_A+m_B), \alpha(m_A+m_B))$ has spreads, which mostly depend on the magnitudes of m_A, m_B. It seems to be logical that the spreads of the product should be influenced by the spreads of the multiplied numbers in the substantial grade instead. By multiplying (2, 0.3, 0.3) and (3, 0.3, 0.3) we get (6, 1.5, 1.5) while the multiplication of (15, 0.3, 0.3) and (30, 0.3, 0.3) leads to (600, 13.5, 13.5). In both cases the multiplied numbers have the same small spreads, which means that they show the same grade of imprecision. It is difficult to explain the large difference between the spreads in the obtained results since the larger mean of a fuzzy number can be associated with "little fuzziness" in it.

Further, if we use the fuzzy numbers $M_{d,i}$, $i = 0, 1, 2, ..., 6,$ defined by (2), (3) and (5) as the data in the operations of addition, subtraction, multiplication and division we would like to cause the appearance of a result, e.g., *r* that yields the value resembling an arbitrary value (m, $d\cdot i$, $d\cdot i$) belonging to $LIST_m$. We expect that *r* should be included in the class of numbers specified by (2). By performing the operations on fuzzy numbers due to [2, 3, 13, 16, 21] we recognize that we never get the result *r* with the spreads satisfying the condition (3).

To demonstrate this allegation let us add "close to 5" to "appreciatively 3", $d = 0.1$, when applying the classical definition of addition of two fuzzy numbers [2, 3]. Hence, $(5, 0.3, 0.3)_{L-R}+(3, 0.5, 0.5)_{L-R} = (8, 0.8, 0.8)_{L-R}$. The result of the addition is the number with the larger spreads than the members of $LIST_{m=8}$. The multiplication of these numbers will bring the effect $(5, 0.3, 0.3)_{L-R}\cdot(3, 0.5, 0.5)_{L-R} \approx (5\cdot3, 5\cdot0.5+3\cdot0.3, 5\cdot0.5+3\cdot0.3) = (15, 3.4, 3.4)$, which is only an attempt of appreciating the multiplication effect as an *L-R* number [22]. Moreover, we are not sure how to name verbally the number with so huge spreads. Even testing the operations, recommended by [4, 5], we get the number $(5, 0.3, 0.3)\cdot(3, 0.5, 0.5) = (5\cdot3, \max(0.3\cdot3, 0.5\cdot5), \max(0.3\cdot3, 0.5\cdot5)) = (15, 2.5, 2.5)$, which still preserves the large values of the spreads, and we note that both α and β also will exceed the borders of the numbers placed in (3).

3.1 The Max and Min as the Operations on Verbal Fuzzy Numbers

By performing the operations of addition, subtraction, multiplication and division we wish to obtain other numbers looking like $M_{d,i}$, in which the magnitudes of spreads are determined by spreads of input data. We thus propose the following operations on the fuzzy numbers from *FN*(*LIST*) by defining them as [17, 18]

$$M_{d,i} + N_{d,j} = (m, d \cdot i, d \cdot i) + (n, d \cdot j, d \cdot j) = (m + n, \max(d \cdot i, d \cdot j), \max(d \cdot i, d \cdot j))$$
$$M_{d,i} - N_{d,j} = (m, d \cdot i, d \cdot i) - (n, d \cdot j, d \cdot j) = (m - n, \max(d \cdot i, d \cdot j), \max(d \cdot i, d \cdot j))$$
$$M_{d,i} \cdot N_{d,j} = (m, d \cdot i, d \cdot i) \cdot (n, d \cdot j, d \cdot j) = (m \cdot n, \min(d \cdot i, d \cdot j), \min(d \cdot i, d \cdot j)) \qquad (7)$$
$$M_{d,i} \div N_{d,j} = (m, d \cdot i, d \cdot i) \div (n, d \cdot j, d \cdot j) = (m \div n, \min(d \cdot i, d \cdot j), \min(d \cdot i, d \cdot j))$$
for $m, n \in Z,\ i, j = 0,1,2,3,4,5,6,\ d = 0.1$.

If we want to preserve the result of division as a verbal fuzzy number with the mean belonging to *Z*, then we should use either the floor function $\lfloor m \div n \rfloor =$ the greatest integer less than or equal to $m \div n$, or the ceiling function $\lceil m \div n \rceil =$ the least integer greater than or equal to $m \div n$. The choice depends on the magnitude of a decimal expansion in the quotient $m \div n$.

These operations, which are intuitively interpreted as similar to the operations occurring in the case of the largest *t*-norm for multiplication (min) and the smallest *s*-norm for addition (max) [22], warrant that their results always will be the fuzzy numbers from (2). Let us note that the list of $M_{d,i}$ has been created according to the scientists' advice (especially physicians). Hence, the users who define the entries of data in the way that is comfortable for them – in other words by choosing any description belonging to $LIST_m$, $m \in Z$, – expect that they obtain a result *r* as another fuzzy number interpreted as one of the expressions from, e.g., $LIST_n$, $n \in Z$.

The sets of numbers generated by (7) are close. We further deduce that there exists an identity element equal to the fuzzy number (0, 0, 0) in *VBFN*(*LIST*) over the addition given by (7) because (*m*, *i*, *i*)+(0, 0, 0) = (*m*+0, max(*i*, 0), max(*i*, 0)) = (*m*, *i*, *i*). Since (*m*, *i*, *i*)·(1, 0.6, 0.6) = (*m*·1, min(*i*, 0.6), min(*i*, 0.6)) = (*m*, *i*, *i*) we can even accept the fuzzy value of (1, 0.6, 0.6) as the identity element for the multiplication introduced by (7) over the space *VBFN*(*LIST*). No inverses can be found with respect to either the addition or multiplication operators. The associative law is satisfied for the max and min operations; therefore it is also valid in the case of the addition and the multiplication of the verbal fuzzy numbers. Hence, the space *VBFN*(*LIST*), having such features as closure, associativity and the existence of the identity element for addition respectively multiplication, is a semi-group with these operations brought into it by the connections (7).

If both *d* and *i* are fix numbers then the set of fuzzy numbers of the type (*m*, *p*, *p*), $m \in Z$, $p = d \cdot i$ is a constant, will keep all the properties, which are characteristic of a ring. The set is called $FN(z, p)$ and it is a subset of $VBFN(LIST)$.

Theorem 1

The set of fuzzy numbers $FN(z, p) = \{(z, p, p) : z \in Z, p = i \cdot d \text{ is fix}\} \subset VBFN(LIST)$ with the operations of addition and multiplication defined by (7) is a ring.

Proof

$FN(z, p)$ is a commutative group with respect to addition of fuzzy numbers in accordance to (7) since the following conditions are satisfied:

1) For any $(m,p,p)\in FN(z,p)$ and $(n,p,p)\in FN(z,p)$,
$(m,p,p)+(n,p,p)=(m+n,\max(p,p),\max(p,p))=(m+n,p,p)\in FN(z,p)$ – the closure of $FN(z,p)$.
2) For any $(m,p,p)\in FN(z,p)$, $(n,p,p)\in FN(z,p)$ and $(s,p,p)\in FN(z,p)$, $((m,p,p)+(n,p,p))+(s,p,p)=(m,p,p)+((n,p,p)+(s,p,p))$ – associativity.
3) There is an element $(0,p,p)\in FN(z,p)$ ($0\in Z$) such that $(m,p,p)+(0,p,p)=(0,p,p)+(m,p,p)=(m,p,p)$ for all $(m,p,p)\in FN(z,p)$ – identity.
4) For all $(m,p,p)\in FN(z,p)$ there is $(-m,p,p)\in FN(z,p)$, ($-m\in Z$) satisfying $(m,p,p)+(-m,p,p)=(0,p,p)$ – inverse.
5) For any $(m,p,p)\in FN(z,p)$ and $(n,p,p)\in FN(z,p)$, $(m,p,p)+(n,p,p)=(n,p,p)+(m,p,p)$ – commutativity.

The set $FN(z,p)$ with multiplication fulfils conditions 1)–3) according to the following statements:

1) For any $(m,p,p)\in FN(z,p)$ and $(n,p,p)\in FN(z,p)$, $(m,p,p)\cdot(n,p,p)=(mn,\min(p,p),\min(p,p))=(mn,p,p)\in FN(z,p)$ – the closure of $FN(z,p)$.
2) For any $(m,p,p)\in FN(z,p)$, $(n,p,p)\in FN(z,p)$ and $(s,p,p)\in FN(z,p)$, $((m,p,p)\cdot(n,p,p))\cdot(s,p,p)=(m,p,p)\cdot((n,p,p)\cdot(s,p,p))$ – associativity.
3) There is an element $(1,p,p)\in FN(z,p)$ ($1\in Z$) such that $(m,p,p)\cdot(1,p,p)=(1,p,p)\cdot(m,p,p)=(m,p,p)$ for all $(m,p,p)\in FN(z,p)$ – identity.

Finally we check the reliability of the distributive law:
$(m,p,p)\cdot((n,p,p)+(s,p,p))=((m,p,p)\cdot(n,p,p))+((m,p,p)\cdot(s,p,p))$.

Since the set $FN(z,p)$ satisfies all the axioms above, it will be interpreted as the ring. It is easy to learn that the sets $FN(z,0), FN(z,0.1), FN(z,0.2),..., FN(z,0.6)$, $z\in$Z, d=0.1 are disjoint in pairs. They are parts in the equation *VBFN*(*LIST*) = $FN(z,0)\cup FN(z,0.1)\cup\ldots\cup FN(z,0.6)$, which constitutes the partition of the space *VBFN*(*LIST*). We notice that $FN(z,p)$ becomes the set Z (also the ring) if i=0. It would be desirable to find in the nearest future such operations, applied to the verbal fuzzy numbers from the entire set *VBFN*(*LIST*), which allow it to be the ring.

3.2 Other Norms as the Operations on Verbal Fuzzy Numbers

Since the comparisons of two values via max and min operations always give one of the values as the result, it seems to be suitable to seek other operations on verbal fuzzy numbers. They should yield smoother (not max or min) effects as the concatenation operations.

We recall the formula of the Dubois-Prade parametric t-norm given as [22]

$$DP_\lambda(x, y) = \frac{xy}{\max(x, y, \lambda)}, \; 0 \le \lambda \le 1. \tag{8}$$

If $\lambda = 0$ the formula (8) becomes

$$DP(x, y) = \frac{xy}{\max(x, y)}. \tag{9}$$

and provides us with the same output values as minimum.

The parametric Dubois-Prade *s*-norm is formulated as [22]

$$DP_\lambda^{co}(x, y) = 1 - \frac{(1-x)(1-y)}{\max(1-x, 1-y, \lambda)}, \; 0 \le \lambda \le 1. \tag{10}$$

It is stated in the shorter form

$$DP^{co}(x, y) = 1 - \frac{(1-x)(1-y)}{\max(1-x, 1-y)} \tag{11}$$

if $\lambda = 0$. The effects of using the Dubois-Prade *s*-norm, stated according to (11), resemble the results obtained for maximum.

By making the appropriate choice of the λ parameter to get the close set of results, we wish to propose another group of basic operations on the fuzzy numbers from *VBFN*(*LIST*). Let us suggest them as [19]

$$\begin{aligned}
&(m, 0.1\cdot i, 0.1\cdot i) + (n, 0.1\cdot j, 0.1\cdot j) = (m+n, DP_\lambda^{co}(0.1\cdot i, 0.1\cdot j), DP_\lambda^{co}(0.1\cdot i, 0.1\cdot j)),\\
&(m, 0.1\cdot i, 0.1\cdot i) - (n, 0.1\cdot j, 0.1\cdot j) = (m-n, DP_\lambda^{co}(0.1\cdot i, 0.1\cdot j), DP_\lambda^{co}(0.1\cdot i, 0.1\cdot j)),\\
&(m, 0.1\cdot i, 0.1\cdot i) \times (n, 0.1\cdot j, 0.1\cdot j) = (m\cdot n, DP_\lambda(0.1\cdot i, 0.1\cdot j), DP_\lambda(0.1\cdot i, 0.1\cdot j)),\\
&(m, 0.1\cdot i, 0.1\cdot i) \div (n, 0.1\cdot j, 0.1\cdot j) = (m\div n, DP_\lambda(0.1\cdot i, 0.1\cdot j), DP_\lambda(0.1\cdot i, 0.1\cdot j))\\
&\text{for } m, n \in Z,\; i, j = 0,1,2,3,4,5,6.
\end{aligned} \tag{12}$$

The floor or ceiling functions should be applied in the case of division to guarantee that the mean of a result is a member of *Z*.

The thorough analysis of the topological properties of the space *VBFN*(*LIST*) with the Dubois-Prade norms – made as in the case of min and max – ought to reveal if these norms act better in the aspiration to accept *VBFN*(*LIST*) as the ring.

As the example producing the effect of operations (7), we demonstrate the calculus concerning the extraction of the DNS-histogram class in the Internet traffic.

4 The Interpolating Function in Identification of IP Traffic

To overcome privacy issues and the unreliability of using port information, an approach to the classification of Internet traffic can be made by using information obtained from IP itself. Basing our observations on datagram size histograms, the objective is to extract characteristic data for various application protocols such as HTTP, NNTP, DNS, FTP, etc. The histograms, plotted in Fig. 3, are examples of DNS traffic distributions, showing a characteristic mix of package sizes in the range 0-12 bins.

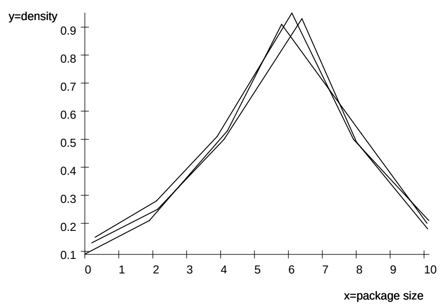

Fig. 3 The examples of the IP datagram sizes carrying DNS traffic

If we, by analysing the larger sample of histograms, evaluate the coordinates of the most characteristic points in the polygons as (x_0, y_0) = (rather close to 3, rather close to 40), (x_1, y_1) = (close to 6, close to 90), (x_2, y_2) = (very close to 8, very close to 60) (when scaling the second coordinate by multiplying the values by 10^2), then we will able to test the operations suggested by (7) in the Lagrange polynomial [19] to obtain the general equation of a function that interpolates the points belonging to the same class of Internet traffic.

We express the polynomial $p_n(x)$ as

$$p_n(x) = \sum_{k=0}^{n} y_k L_{n,k}(x) , \tag{13}$$

in which $L_{n,k}$ is the Lagrange coefficient polynomial expanded as

$$L_{n,k} = \prod_{\substack{j=0 \\ j \neq k}}^{n} (x - x_j) \Bigg/ \prod_{\substack{j=0 \\ j \neq k}}^{n} (x_k - x_j) . \tag{14}$$

The polynomial $p_2(x)$ expressing a relationship between the DNS package size and the density has as the entries the fuzzy numbers x_0, x_1, x_2, y_0, y_1, y_2 given above in the verbal form. After their mathematical interpretation in accordance to (3) we create the function for any (x, p, p) belonging to the interval [0, 12], $d = 0.1$, $p = 0.1 \cdot i$, $i = 0, \ldots, 6$, as

$$y = p_2(x) = (40,0.4,0.4) \cdot \frac{((x,p,p)-(6,0.3,0.3))((x,p,p)-(8,0.2,0.2))}{((3,0.4,0.4)-(6,0.3,0.3))((3,0.4,0.4)-(8,0.2,0.2))}$$
$$+ (90,0.3,0.3) \cdot \frac{((x,p,p)-(3,0.4,0.4))((x,p,p)-(8,0.2,0.2))}{((6,0.3,0.3)-(3,0.4,0.4))((6,0.3,0.3)-(8,0.2,0.2))}$$
$$+ (60,0.2,0.2) \cdot \frac{((x,p,p)-(3,0.4,0.4))((x,p,p)-(6,0.3,0.3))}{((8,0.2,0.2)-(3,0.4,0.4))((8,0.2,0.2)-(6,0.3,0.3))}.$$

We perform the recommended operations to obtain the function

$$y = p_2(x) = (-6x^2 + 69x - 108, l, l)$$

where

$$l = \max(\min(0.4, \min(\max(p,0.3), \max(p.0.2))),$$
$$\min(0.3, \min(\max(p,0.4), \max(p.0.2))),$$
$$\min(0.2, \min(\max(p,0.4), \max(p.0.3)))).$$

The function y should improve the accuracy in the process of interpolation when more points are used in (13).

If, for example, the package size is assumed to be x = "almost exactly 4" = (4, 0.1, 0.1) then the corresponding density of appearing of this package can be prognosticated as $y = (0.72 \cdot 10^2, 0.2, 0.2) \approx$ "very close to 0.7". Moreover, we can discuss if the function $y = p_2(x)$ expanded above is an indicator of the DNS traffic in the interval under consideration.

5 The Yager Fuzzy Probability of Verbal Fuzzy Numbers

Let us now assume that some values of the package sizes represent the variable X, which is measured in units among b and v, $X = \{b, b+1, \ldots, v\}$, $b, v \in N$, ranked in the ascending order. By investigating a number of samples coming from the Internet traffic (e.g., DNS-class histograms) we appreciate the aggregated values of the variable X within these histograms as the verbal fuzzy numbers $(m, d \cdot i, d \cdot i)$, $m \in \{b, b+1, \ldots, v\}$, $i = 1, 2, \ldots, 6$.. For instance, if the variable X = "the package size in the DNS traffic" is measured in the scale from 3 to 8 bins, then one of the generalized values of X will be appreciated as x = "close to 6"=(6, 0.3, 0.3).

Let us appreciate Yager's probability of the event that the variable X takes as a value the number "close to 6" = (6, 0.3, 0.3), d=0.1. We thus refer to Yager's probability definition formulated as follows [22].

Definition 4

Let $M_{d,i}^{\alpha}$ be the α-level set of a fuzzy number $M_{d,i}$. Then the probability of a fuzzy event $M_{d,i}$ represented by the fuzzy number $M_{d,i}$ can be defined as

$$\widetilde{P}(M_{d,i}) = \sum_{\alpha\in[o,1]} \alpha \Big/ P(M_{d,i}^{\alpha}) . \tag{15}$$

The α-level of $M_{d,i}$ is a non-fuzzy set $M_{d,i}^{\alpha}$ determined as

$$M_{d,i}^{\alpha} = \{x : \mu_{M_{d,i}}(x) \geq \alpha\}. \tag{16}$$

The variable $X = \{x_1, x_2, ..., x_v\} = \{b, b+1, ..., v\}$ has its classical discrete probability distribution $P(x_1) = p_1$, $P(x_2) = p_2$, …, $P(x_v) = p_v$. We calculate the mean value $\mu = \sum_{k=1}^{v} x_k p_k$ and the standard deviation $\sigma = \sqrt{\sum_{k=1}^{v} (x_k - \mu)^2 p_k}$ for the variable X.

Let us suppose that the probability distribution of X resembles the normal continuous probability distribution, which is even observed in the example from Section 4. The example confirms that the packages, which come via the Internet traffic, are mostly concentrated in the neighbourhood of the mean in the considered interval.

Further, we assume that the α-levels are decided for $\alpha = \alpha_1, \alpha_2, ..., \alpha_t$. In accordance with (5), for d=0.1, we obtain $M_{0.1\cdot i}^{\alpha_s} = [x_{0.1\cdot i,left}^{\alpha_s}, x_{0.1\cdot i,right}^{\alpha_s}]$, $s = 1,2,...,t$, where

$$\begin{aligned} x_{0.1\cdot i,left}^{\alpha_s} &= 0.1\cdot i\cdot \alpha_s - 0.1\cdot i + m \\ x_{0.1\cdot i,right}^{\alpha_s} &= 0.1\cdot i + m - 0.1\cdot i\cdot \alpha_s . \end{aligned} \tag{17}$$

The probability value of $M_{d,i}^{\alpha_s}$ is developed as

$$P(M_{d,i}^{\alpha_s}) = P(x_{0.1\cdot i,left}^{\alpha_s} \leq X \leq x_{0.1\cdot i,right}^{\alpha_s}) = \Phi(\frac{x_{0.1\cdot i,right}^{\alpha_s} - \mu}{\sigma}) - \Phi(\frac{x_{0.1\cdot i,left}^{\alpha_s} - \mu}{\sigma}) \tag{18}$$

provided that Φ is the probability distribution function for the normal distribution.

Finally, we propose the Yager probability of the verbal fuzzy number $M_{d,i}$ in the form of a fuzzy set

$$\widetilde{P}(M_{d,i}) = \sum_{s=1}^{t} \alpha_s \Big/ P(M_{d,i}^{\alpha_s}) . \tag{19}$$

If the probabilities of the variable values do not fit to the normal distribution we should find another probability model in order to evaluate the probabilities of the α-levels.

We prove the method discussed above to find the fuzzy set that approximates the probability of the event "the package size in the DNS traffic"="close to 6"=(6, 0.3, 0.3) when accepting the scale of the package sizes as X={3, 4, 5, 6, 7, 8}. Apart from it, we proceed from the assumption that most packages in DNS traffic in the investigated interval come with the largest densities about the mean value of the interval.

Assuming that $P(3)=0.05$, $P(4)=0.10$, $P(5)=0.35$, $P(6)=0.35$, $P(7)=0.10$, $P(8)=0.05$ we compute $\mu=5.5$ and $\sigma=1.12$. We adapt the discrete distribution to the continuous normal distribution $N(\mu,\sigma)$ to calculate the probabilities of the following α-levels:

$$P(6,0.3,0.3)^{0.2}) = P(5.76 \le X \le 6.24) \approx 0.1622\ ,$$
$$P(6,0.3,0.3)^{0.4}) = P(5.82 \le X \le 6.18) \approx 0.1188\ ,$$
$$P(6,0.3,0.3)^{0.6}) = P(5.88 \le X \le 6.12) \approx 0.0757\ ,$$
$$P(6,0.3,0.3)^{0.8}) = P(5.94 \le X \le 6.06) \approx 0.0398\ ,$$
$$P(6,0.3,0.3)^{1}) = P(5.99 \le X \le 6.01) \approx 0.0036\ .$$

Hence, $P(6,0.3,0.3) = {}^{0.2}\!/_{0.1622} + {}^{0.4}\!/_{0.1188} + {}^{0.6}\!/_{0.0757} + {}^{0.8}\!/_{0.0398} + {}^{1}\!/_{0.0036}$.

The probabilities, which are the elements of the support, seem to have too small values. It is a consequence of using the continuous distribution instead of the discrete one. Since it is the first study on this topic, a closer analysis may be needed in the future. For instance, we may assume including a particular scaling factor into the probability formula as a remedy of improving the obtained results. On the other hand we can define seven verbal fuzzy numbers for each positive integer assigned to X, which leads to 42 verbal fuzzy numbers in the case of six basic sizes. The probability of meeting an arbitrary number of these 42 without making any assumptions about their fuzziness is about 0.02. The last value is only the very rough evaluation of the "close to 6" probability, but it converges to $p=0.0398$, which possesses a high membership degree equal to 0.8.

6 Conclusions

The mathematicians who work on some applications of mathematics to other subjects should respect other scientists' opinions; especially if they have the different abilities in formulating their problems in the way they understand best. It means that researchers, non-professional mathematically, often use the verbal expressions to describe a value they need. Being aware of it we have made a trial of introducing a new fuzzy space with verbal fuzzy numbers in the particular *L-R* form.

It makes possible to choose required values by using simple linguistic formulations and replacing them by appropriate fuzzy numbers from the space. The operations, defined in the space, keep the results within the same space without enlarging the borders of the numbers. It is necessary to emphasize that the operations tested in the paper differ from the classical rules based on the extension principle. There exist some subspaces in the space of fuzzy numbers that are rings. The fact of discovering such subspaces reinforces the efforts to look for a total space of fuzzy numbers, which fulfils the conditions of the ring as a counterpart of the set Z. The assumptions, concerning the verbal fuzzy numbers, lead to the conclusion that the order in the space, which is a countable set, can be introduced in the unique way.

By testing the non-fuzzy Lagrange interpolating polynomial with verbal fuzzy numbers assisting the algorithm, we have found a function that can help in recognizing some properties of the DNS packages in IP traffic extraction. By means of the function we are able to predict the density corresponding to an arbitrary package size belonging to the interval of package sizes under consideration.

It may be interesting for researchers dealing with the area of information retrieval to prove the last model for the prediction of the number of keywords used in a query attempt. If we assume that the interpolated points are the pairs of the type (number of keywords, number of query attempts) then we, by applying the formula of an interpolating function, can calculate the estimate of the number of keywords that are going to be used at a specific attempt.

The first trial of approximating the probability of verbal fuzzy numbers was also made as a tool of creating the fuzzy distribution with fuzzy sets.

7 Prognoses

We wish to improve the suggested operations in such a way that they do not lead to dominant values of the spreads in the numbers obtained as results. Unfortunately, both maximum and minimum operations have the tendency to the creation of sharp effects. We will try to invent other operations in the proposed fuzzy space, which should be fit better for the theoretical and practical assumptions.

Acknowledgements

The author thanks The Swedish Royal Academy of Sciences for the grant received in 2003 and allocated for development of Fuzzy Mathematics in applied branches. The author is grateful to the reviewers for their valuable remarks that have contributed in improving this paper.

References

1. Detyniecki, M., Yager, R.: Ranking Fuzzy Numbers Using α-weighted Valuations. Internat. J. Uncertainty, Fuzziness Knowledge-Based Systems 8 5 (2000) 573-591

2. Dubois, D., Prade, H.: Fuzzy Real Algebra. Some Results. Fuzzy Sets and Systems 2 (1978a) 327-348
3. Dubois, D., Prade, H.: Operations on Fuzzy Numbers. Int. J. Systems Sci., Vol. 9 (6) (1978b) 613-626
4. Dug Hun Hong, Sungho Lee, Hae Young Do: Fuzzy Linear Regression Analysis for Fuzzy Input-Output Data Using Shape-Preserving Operations. Fuzzy Sets and Systems 122(3) (2001) 513-526
5. Dug Hun Hong: Shape Preserving Multiplications of Fuzzy Numbers. Fuzzy Sets and Systems 123 (2001) 81-84
6. Facchinetti, G., Ghiselli Ricci, R., Muzziol, S.: Note of Ranking Fuzzy Triangular Numbers. Internat. J. Intell. Syst. 13 (1998) 613-622
7. Fullér, R.: On Product-Sum of Triangular Fuzzy Numbers. Fuzzy Sets and Systems 41 (1991) 83-87
8. Fullér, R., Majlender, P.: On Weighted Possibilistic Mean and Variance of Fuzzy Numbers. Fuzzy Sets and Systems 136 (2003) 363-374
9. Goetschel, R., Woxman, W.: Topological Properties of Fuzzy Numbers. Fuzzy Sets and Systems 10 (1983) 87-99
10. Goetschel, R., Woxman, W.: Elementary Fuzzy Calculus. Fuzzy Sets and Systems 18 (1986) 31-34
11. Iskander, M. G.: Comparison of Fuzzy Numbers Using Possibility Programming. Comments and New Concepts. Computers & Mathematics with Applications 43 (2002) 833-840
12. Kaufmann, A., Gupta, M. M.: Introduction to Fuzzy Arithmetic Theory and Application. Van Nostrand Reinhold, New York (1991)
13. Kun-lun Zhang, Hirota K.: On Fuzzy Number Lattice. Fuzzy Sets and Systems 92 (1997) 113-122
14. Liang-Hsuan Chen, Hai-Wen Lu: An Approximate Approach for Ranking Fuzzy Numbers Based on Left and Right Dominance. Computers & Mathematics with Applications 41 (2001) 1589-1602
15. Matarazzo, B., Munda, G.: New Approaches for the Comparison of *L-R* Fuzzy Numbers: a Theoretical and Operational Analysis. Fuzzy Sets and Systems 118 (2001) 407-418
16. Mizumoto, M., Tanaka, K.: Some properties of Fuzzy Numbers. In: Gupta, M. M., Ragade, R. K., Yager, R. (eds.): Advances in Fuzzy Set Theory and Applications, North-Holland Publishing Company (1979)
17. Rakus-Andersson, E.: Factor Analysis with Qualitative Factors as Fuzzy Numbers. In: Abraham, A., Ruiz-del-Solar, J., Köppen, M. (eds.): Soft Computing Systems, Design, Management and Applications. IOS Press Ohmsha, Amsterdam Berlin Oxford Tokyo Washington, DC (2002) 106-114
18. Rakus-Andersson, E.: The Newton Interpolation Method with Fuzzy Numbers as the Entries. In: Rutkowski, L., Kacprzyk, J. (eds.): Neural Networks and Soft Computing Proceedings of the Sixth International Conference on Neural Network and Soft Computing. Springer-verlag – Advances in Soft Computing (2003a) 310-315
19. Rakus-Andersson, E.: The Dubois-Prade Norms as the Operations on Fuzzy Numbers in Lagrange Interpolation Polynomial. In: Kaynak, O. (ed.): Proceedings of the 10th IFSA World Congress. Bogaziąi University Library Cataloging-in Publication Data (2003b) 449-452
20. Tran, L., Duckstein, L.: Comparison of Fuzzy Numbers Using a Fuzzy Distance Measure. Fuzzy Sets and Systems 130 (2002) 331-341
21. Wang, X., Kerre, E.: Reasonable Properties for the Ordering of Fuzzy Quantities (I), (II). Fuzzy Sets and Systems 118 (2001) 375-385, 387-405
22. Zimmermann, H. J.: Fuzzy Set Theory and Its Applications, 3rd edn. Kluwer Academic Publish, Boston (1996)