Notes on Intuitionistic Fuzzy Sets
Print ISSN 1310–4926, Online ISSN 2367–8283
2023, Volume 29, Number 4, 411–417
DOI: 10.7546/nifs.2023.29.4.411-417

n-Dimensional intuitionistic fuzzy index matrix representation of multidimensional data partitioning methods

Veselina Bureva[1] **and Krassimir T. Atanassov**[2,3]

[1] Laboratory of Intelligent Systems, "Prof. Dr. Assen Zlatarov" University
1 "Prof. Yakimov" Blvd., 8010 Burgas, Bulgaria
e-mail: `vbureva@btu.bg`

[2] Department of Bioinformatics and Mathematical Modelling,
Institute of Biophysics and Biomedical Engineering,
Bulgarian Academy of Sciences
105 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
e-mail: `krat@bas.bg`

[3] Laboratory of Intelligent Systems, "Prof. Dr. Assen Zlatarov" University
1 "Prof. Yakimov" Blvd., 8010 Burgas, Bulgaria

Received: 12 November 2023 **Revised:** 2 December 2023
Accepted: 11 December 2023 **Online first:** 14 December 2023

Abstract: In the present research paper, the data partitioning operations are investigated. Their functionalities are discussed. n-dimensional intuitionistic fuzzy index matrix is used to represent the multidimensional data partitioning. Attributes selection and values selection by predicates is presented.
Keywords: Big data, Intuitionistic fuzzy index matrix, OLAP.
2020 Mathematics Subject Classification: 03E72.

1 Introduction

Nowadays the Big Data paradigm is imposed as useful framework for storing huge amounts of data. It has the property to provide better parallel processing. Different algorithms, leading its beginning from statistics and data mining, are transformed into helpful Big data procedures. OLAP (Online Analytical Processing) analyses is a technique that is adapted in the field of big data systems. Obviously, OLAP cubes perform five basic operations: *Roll-Up*, *Drill-Down*, *Slice*, *Dice* and *Pivot*. *Roll-up* and *Drill-Down* are typical operations navigating through the hierarchy of the OLAP cube dimensions and executing aggregation functions over the data. *Pivot* rotates the OLAP cube in different sides. *Slice* and *Dice* operations leads its beginning from the relational operations *Selection* and *Projection*. *Slice* operation selects appropriate rows according to the predefined data condition. *Dice* operation selects appropriate columns according to the predefined data condition. In series of papers these and other operations are represented by index matrices (see, e.g., [1, 3, 4, 6–11]).

The partitioning provides fragmentation of the data for parallel executing in distributed systems. The strategies for data duplication (replication) are already well known. The processes of data fragmentation, replication, and allocation are very important for distributed databases. In the relational databases there are several types of data fragmentation: vertical fragmentation, horizontal fragmentation and mixed or hybrid fragmentation. The OLAP cubes use partitioning for better performance. Typically, parallel execution of queries to different partitions is used. Another reason for cube partitioning is to provide separate data for appropriate users. The partitions are constructed using filtering of a fact table for multiple partitions or using tables, views, or named queries. The separate partitions can be performed using slice and dice operations for slicing the OLAP cube dimensions.

2 Operator $C(A, P)$

In the current research the n-dimensional structure of the OLAP cube will be presented in the form of n-dimensional index matrix and respectively n-dimensional intuitionistic fuzzy index matrix.

Let everywhere below by $|X|$ we denote the cardinality of set X.

Let us have the n-dimensional IM (see [3])

$$A = [K_1, \ldots, K_n, \{a_{k_{1,i_1},\ldots,k_{n,i_n}}\}],$$

where for each j $(1 \leq j \leq n)$ and for each s $(1 \leq s \leq |K_j|)$: $k_{j,i_s} \in K_j$, and $a_{k_{j,i_s}} \in S$, where S is a fixed set of objects. Let over $S \cup \{*\}$ be defined the predicate P, where "$*$" denotes the empty symbol, and for each $a \in S$: $P(*)$ is false, $P(a)$ hsas truth valud "false" or "true".

Let us define operator C by:

$$C(A, P) = [K_1, \ldots, K_n, \{C(a_{k_{1,i_1},\ldots,k_{n,i_n}}, P)\}],$$

where

$$C(a_{k_{1,i_1},\ldots,k_{n,i_n}}, P) = \begin{cases} a_{k_{1,i_1},\ldots,k_{n,i_n}}, & \text{if } P(a_{k_{1,i_1},\ldots,k_{n,i_n}}) \text{ is true} \\ *, & \text{otherwise} \end{cases}.$$

Therefore, the IM

$$B = C(A, P)$$

contains only these elements of A that satisfy predicate P, while the remaining elements of B are "$*$".

Let (for operation @ see [1]):

$$D = @B = [L_1, \ldots, L_m, \{b_{l_{1,q_1},\ldots,l_{m,q_m}}\}],$$

where $m \leq n$, for each p $(1 \leq p \leq m)$ there is j $(1 \leq j \leq n)$ such that $L_p \subseteq K_j$. Therefore, D is an IM that contains only these elements of A that satisfy predicate P, and for each p there is q so that $b_{l_{p,q_p}} \neq *$.

Let the evaluation of each element $a_{k_{1,i_1},\ldots,k_{n,i_n}}$ be an Intuitionistic Fuzzy Pair (IFP, see [2,5]), i.e., there is an evaluation function $V : S \to \{\langle a, b\rangle | a, b, a + b \in [0, 1]\}$ so that

$$V(a_{k_{1,i_1},\ldots,k_{n,i_n}}) = \langle \mu_{k_{1,i_1},\ldots,k_{n,i_n}}, \nu_{k_{1,i_1},\ldots,k_{n,i_n}} \rangle.$$

We must mention that if $\mu_{k_{1,i_1},\ldots,k_{n,i_n}} = 1, \nu_{k_{1,i_1},\ldots,k_{n,i_n}} = 0$, then the IFP $a_{k_{1,i_1},\ldots,k_{n,i_n}}$ is called a *tautological IFP (TIFP)* and when $\mu_{k_{1,i_1},\ldots,k_{n,i_n}} \geq \nu_{k_{1,i_1},\ldots,k_{n,i_n}}$, then the IFP $a_{k_{1,i_1},\ldots,k_{n,i_n}}$ is called an *intuitionistic fuzzy tautological IFP (IFTIFP).*

Now, the operator C for predicate P can be defined, e.g., as follows

$$\begin{aligned} C(a_{k_{1,i_1},\ldots,k_{n,i_n}}, P) &= \begin{cases} a_{k_{1,i_1},\ldots,k_{n,i_n}}, & \text{if } a_{k_{1,i_1},\ldots,k_{n,i_n}} \text{ is an IFTIFP} \\ *, & \text{otherwise} \end{cases} \\ &= \begin{cases} a_{k_{1,i_1},\ldots,k_{n,i_n}}, & \text{if } \mu_{k_{1,i_1},\ldots,k_{n,i_n}} \geq \nu_{k_{1,i_1},\ldots,k_{n,i_n}} \\ *, & \text{otherwise} \end{cases}. \end{aligned}$$

3 Partitioning operations example

The Bookshop example is used to represent the different types of partitioning operations. Dimensions hierarchical concepts are: "*Books*", "*Bookshops*" and "*Locations*". The structure of the Bookshop-cube is presented on the Figure 1. The "*Location*" dimension has the hierarchy *Town – Country*, the dimension "*Bookshops*" has the hierarchy *Bookshop Name – Regional Manager – Owner* and the hierarchy of the dimension "*Books*" is *Title – Publisher – Genre*. The loaded data into the OLAP cube is visualized in Figure 2.

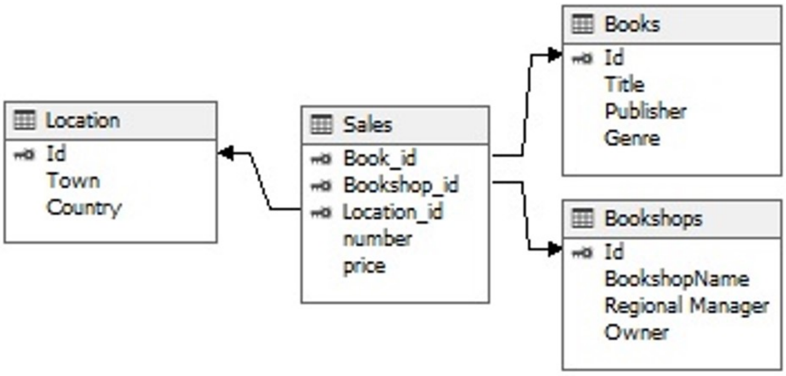

Figure 1. The dimension hierarchies

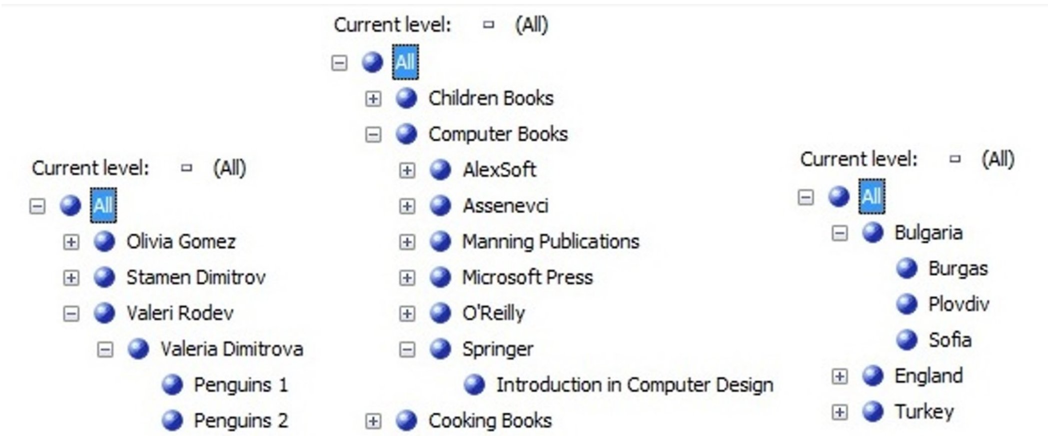

Figure 2. Browsing dimensions *Bookshops*, *Books* and *Locations*

Two variants of the OLAP cube containing the dimensions *Books*, *Bookshops* and *Locations* are constructed. The first contains the measure degree of membership for books that are sold out. The second contains the measure degree of non-membership for books that are sold out. Thereafter the operation *Drill-Across* is performed using Power BI Excel Add-in (Power Pivot, OLAP Add-in). The result has the following form (Figure 3):

	Column Labels					
	Olivia Gomez		Stamen Dimitrov		Valeri Rodev	
Row Labels	S_Degree_Membership	S_Degree_NoNMembership	S_Degree_Membership	S_Degree_NoNMembership	S_Degree_Membership	S_Degree_NoNMembership
Children Books	**0.27**	**0.13**	**0.17**	**0.07**	**0.13**	**0.14**
Bulgaria	0.20	0.10	0.50	0.20		
England	0.60	0.30				
Turkey					0.40	0.43
Computer Books	**0.20**	**0.13**	**0.25**	**0.08**	**0.27**	**0.07**
Bulgaria	0.35	0.15	0.35	0.05		
England			0.25	0.05	0.25	0.05
Turkey	0.25	0.25	0.15	0.15	0.55	0.15
Cooking Books	**0.14**	**0.10**	**0.32**	**0.02**	**0.10**	**0.02**
Bulgaria	0.23	0.20	0.70	0.05	0.20	0.05
England	0.20	0.10	0.25	0.00	0.10	0.00
Turkey						

Figure 3. Result of Drill-Across operation

In this case the aggregations are calculated using $average()$ function (in the standard case the $sum()$ function is used). Standard partitioning operations are made using the OLAP cube with one measure. Here, we present the partitioning options after performing *Drill-Across* operation over two OLAP cubes or OLAP cubes having two-valued measures. An example of possible partitioning is to filter data by condition: $(\mu > 0.30) \wedge (\nu < 0.20)$. The information is filtered and the resulting dataset is:

	Column Labels					
	Olivia Gomez		Stamen Dimitrov		Valeri Rodev	
Row Labels	**S_Degree_Membership**	**S_Degree_NoNMembership**	**S_Degree_Membership**	**S_Degree_NoNMembership**	**S_Degree_Membership**	**S_Degree_NoNMembership**
Children Books			**0.17**	**0.07**		
Bulgaria			0.50	0.20		
England						
Turkey						
Computer Books	**0.12**	**0.05**	**0.12**	**0.02**	**0.18**	**0.05**
Bulgaria	0.35	0.15	0.35	0.05		
England						
Turkey					0.55	0.15
Cooking Books			**0.23**	**0.02**		
Bulgaria			0.70	0.05		
England						
Turkey						

Figure 4. Filtered data

The partitioned OLAP cube, presented in the Figure 4, contains all values that satisfies the predicate. The second OLAP cube partition will continue to store all values that not satisfy the assigned predicate. We can remove all rows containing the value "$Null$" (Figure 5). Thereafter 6 rows are deleted.

	Column Labels					
	Olivia Gomez		Stamen Dimitrov		Valeri Rodev	
Row Labels	**S_Degree_Membership**	**S_Degree_NoNMembership**	**S_Degree_Membership**	**S_Degree_NoNMembership**	**S_Degree_Membership**	**S_Degree_NoNMembership**
Children Books			**0.17**	**0.07**		
Bulgaria			0.50	0.20		
Computer Books	**0.12**	**0.05**	**0.12**	**0.02**	**0.18**	**0.05**
Bulgaria	0.35	0.15	0.35	0.05	0.55	0.15
Cooking Books			**0.70**	**0.05**		
Bulgaria			0.70	0.05		

Figure 5. Partition that satisfies the predicate with removed the empty rows

If we have the condition to store the sales by the three different owners separately then the partitions will have the following form (after removing empty rows for each of them) (Figure 6).

	Column Labels	
	Olivia Gomez	
Row Labels	**S_Degree_**	**S_Degree_NoNMembership**
Computer Books	**0.12**	**0.05**
Bulgaria	0.35	0.15

	Stamen Dimitrov	
Row Labels	**S_Degree_Membership**	**S_Degree_NoNMembership**
Children Books	**0.17**	**0.07**
Bulgaria	0.50	0.20
Computer Books	**0.12**	**0.02**
Bulgaria	0.35	0.05
Cooking Books	**0.70**	**0.05**
Bulgaria	0.70	0.05

	Valeri Rodev	
Row Labels	**S_Degree_**	**S_Degree_NoNMembership**
Computer Books	**0.18**	**0.05**
Bulgaria	0.55	0.15

Figure 6. Divide the partition that satisfies the predicate by owners

In the same way, we can filter the input data for appropriate bookshops owner and to reduce the empty rows (Figure 7). For example: $\neg$(*Olivia Gomez* $\wedge$ *Stamen Dimitrov*). In this case the predicates are applied to the values of the dimensional attributes.

	Valeri Rodev	
Row Labels	S_Degree_Membership	S_Degree_NoNMembership
Children Books	**0.13**	**0.14**
Bulgaria		
England		
Turkey	0.40	0.43
Computer Books	**0.27**	**0.07**
Bulgaria		
England	0.25	0.05
Turkey	0.55	0.15
Cooking Books	**0.10**	**0.02**
Bulgaria	0.20	0.05
England	0.10	0.00
Turkey		

	Valeri Rodev	
Row Labels	S_Degree_Membership	S_Degree_NoNMembership
Children Books	**0.13**	**0.14**
Turkey	0.40	0.43
Computer Books	**0.27**	**0.07**
England	0.25	0.05
Turkey	0.55	0.15
Cooking Books	**0.10**	**0.02**
Bulgaria	0.20	0.05
England	0.10	0.00

Figure 7. Filter the input data for appropriate bookshops owner and to reduce the empty rows

Frequently used example is to filter the *Time* dimension. Let select the intuitionistic fuzzy evaluations for sales of the books for years that are: $(year > 2020) \wedge (year < 2022)$ (Figure 8).

Figure 8. Filter the input data for books sales IFE for $year > 2020 \wedge year < 2022$

4 Conclusion

In the current research work an example for predicates-based partitioning of the n-dimensional index matrix is presented. Operator C(A,P) applies the predicates over the IM. An example of OLAP cube partitioning after Drill-Across operation execution is presented. Appropriate examples over the OLAP cube are visualized.

Acknowledgement

This research was funded by Asen Zlatarov University—Burgas, Project: Center of Excellence UNITE BG05M2OP001-1.001-0004 /28.02.2018 (2018–2023). The authors declare no conflict of interest.

References

[1] Atanassov, K. (2014). *Index Matrices: Towards an Augmented Matrix Calculus*. Studies in Computational Intelligence Series, Vol. 573, Springer, Cham.

[2] Atanassov, K. (2017). *On Intuitionistic Fuzzy Logics*. Springer, Cham.

[3] Atanassov, K. (2018). n-Dimensional extended index matrices. *Advanced Studies in Contemporary Mathematics*, 28(2), 245–259.

[4] Atanassov, K., & Bureva, V. (2020). Operations over 3-Dimensional Extended Intuitionistic Fuzzy Index Matrices. *Advanced Studies in Contemporary Mathematics*, 30(3), 335–356.

[5] Atanassov, K., Szmidt, E., & Kacprzyk, J. (2013). On intuitionistic fuzzy pairs. *Notes on Intuitionistic Fuzzy Sets*, 19(3), 1–13.

[6] Bureva, V., Sotirov, S., Sotirova, E., & Atanassov, K. (2021). Index Matrix Representation of MapReduce Algorithm. *Proceedings of 2021 Big Data, Knowledge and Control Systems Engineering (BdKCSE)*, Sofia, Bulgaria, 1–4, doi: 10.1109/BdKCSE53180.2021.9627306.

[7] Bureva, V., Traneva, V., Sotirova, E., & Atanassov, K. (2017). Index Matrices And OLAP-Cube Part 2: An presentation of the OLAP-analysis by Index Matrices. *Advanced Studies in Contemporary Mathematics*, 27(4), 647–672.

[8] Bureva, V., Traneva, V., Sotirova, E., & Atanassov, K. (2019). Index Matrices and OLAP cube. Part 4: A Presentation of the OLAP Drill Across Operation by Index Matrices. *Advanced Studies in Contemporary Mathematics*, 29(1), 109–123.

[9] Bureva, V., Traneva, V., Sotirova, E., & Atanassov, K. (2019). Index Matrices and OLAP cube. Part 5: Index Matrix Operations over OLAP cube. *Advanced Studies in Contemporary Mathematics*, 30(1), 69–88.

[10] Traneva, V., Bureva, V., Sotirova, E., & Atanassov, K. (2017). Index Matrices And OLAP-Cube Part 1: Application of the Index Matrices to presentation of operations in OLAP-Cube. *Advanced Studies in Contemporary Mathematics*, 27(2), 253–278.

[11] Traneva, V., Bureva, V., Sotirova, E., & Atanassov, K. (2018). Index matrices and OLAP-cube. Part 3: A presentation of the OLAP “InterCube Set” and “Data cube” operations by index matrices. *Advanced Studies in Contemporary Mathematics (Kyungshang)*, 28(3), 423–448.