Algebra of Morphological Filter on Intuitionistic Fuzzy Hypergraphs
With application in Summarization of Documents

Dhanya P.M1*, Sreekumar A2, Jathavedan M3
1, 2,3 Department of Computer Applications,
Cochin University of Science and Technology,
Kochi, India.

Ramkumar P.B4
4Department of Basic Sciences and Humanities,
Rajagiri School of Engineering and Technology,
Rajagiri Valley,
Kakkanad, Kochi, India.

DCS, RSET, Kochi, India

Abstract—Graph morphology has been an area of interest over many years and has its applications in text and image processing. Later came hypergraphs with more number of nodes in a hyperedge. Intuitionistic Fuzzy hypergraphs (IFHG) are further developments in graph theory where a membership degree and non membership degree is given for every node and hyperedge. Further many morphological operations can be applied on the sub graphs of this IFHG. Morphological operations include dilation, erosion, adjunction, opening, closing and filtering. The purpose of this paper is to define union, intersection, complement operations on morphological filter designed on IFHG. An application of this type of modeling in the field of document summarization is also shown in this paper.

Keywords—Intuitionistic Fuzzy; hypergraph; filter; summarization;

I. INTRODUCTION

Let $[H^e, (\mu^e, \gamma^e), (\mu^h, \gamma^h), H^h, H^e]$ be an intuitionistic fuzzy hypergraph with membership degree $\mu^h$ and non membership degree $\gamma^h$ defined on the set of nodes $H^h$ and membership degree $\mu^e$ and non membership degree $\gamma^e$ defined on a set of hyperedges $H^e$. Depending on the membership degree $\mu^h$, the node can be treated as high priority, medium priority and low priority. The non membership degree $\gamma^h$ <= 1 - $\mu^h$. Similarly depending on the membership degree $\mu^e$, the hyperedge can be treated as high priority, medium priority and low priority. The sum of the membership degree and non membership degree of the node is less than or equal to 1 [1], i.e., $\mu^h + \gamma^h <= 1$. So also, the sum of the membership degree and non membership degree of the hyperedge is less than or equal to 1 [1], i.e., $\mu^e + \gamma^e <= 1$. If all the nodes in a hyperedge have $\mu^h > 0.5$, then $\mu^e$ is the supremum of all $\mu^h$ in that edge. In such a case $\gamma^e <= 1 - \mu^e$. If there is at least one node with $\gamma^h > 0.5$, then the $\gamma^e$ of that edge is the supremum of all $\gamma^h$ in that edge. In such a case $\mu^e <= 1 - \gamma^e$. Morphological operations like opening, closing, half opening, half closing of hypergraphs [2] are defined which can be extended to weighted hypergraphs also. Morphological operations like dilation, erosion, adjunction are already defined on hypergraphs [3] and morphological dilation is defined on Intuitionistic Fuzzy Hypergraphs [4]. Morphological Erosion of IFHG [5] is also done, where erosion with respect to nodes and hyperedges are considered. This paper introduces algebraic operations on morphological filtering of IFHG. In order to apply morphological filter, we define sub IFHGs $X_{IF}$ and $Y_{IF}$ which are obtained from $\alpha - \beta$ cut on $H_{IF}$. Here $\delta$ is the dilation operator and $\varepsilon$ is the erosion operator.

II. RELATED WORKS

Structure of Intuitionistic Fuzzy graphs (IFG) [6], strong arc, weak arc, strong path, strongest path are studied earlier, where the authors introduced the concept of arc being called as a bridge. Some properties of complete IFG are also discussed. Complement of an IFG, operations like union, join, cartesian product, composition are analyzed, where the authors also studied about the properties of self complementary IFG [7]. Cardinality of IFG, the concepts like bipartite, complete bipartite, strong arc, strength of connectedness, dominating set, number, independent set, number are studied. The authors studied some properties of minimal independent, minimal dominating IFG [8], where these concepts can be applied to network applications. A method to find the shortest hyperpath [9] in an Intuitionistic fuzzy weighted hypergraph is proposed where ranking is done using scores and accuracy. Now coming to the field of Intuitionistic Fuzzy hypergraphs, concepts of IFHG [1], strength of an edge, ($\alpha$, $\beta$) cut, incidence matrix, adjacent level between two vertices, stronger edge, dual IFHG, its incidence matrix are already defined. These ($\alpha$, $\beta$) cuts can be used for hypergraph partitioning. Operations like complement, join, union, intersection, ringsum, cartesian product, composition are defined for intuitionistic fuzzy graphs [10], where the authors proposed to apply these operations in clustering techniques. Isomorphism between two intuitionistic fuzzy...
directed hypergraphs (IFDHG) is also discussed [11], where the authors have introduced IFDHG, its order, in-degree, out-degree, homomorphism, weak isomorphism and co-weak isomorphism between two IFDHGs. Application of IFHG in radio coverage network is also suggested by [12], in which their model can be used to determine station programming and develop marketing strategies. The authors also suggested the use of IFHG for clustering computer networks. An application with Intuitionistic Fuzzy sets for career choice[13] which is a decision making system was developed, where the system represented the performance of students using membership μ, non membership υ and hesitation margin π. They applied normalized Euclidean distance to determine the apt career choice.

Now let us see some fuzzy based works in Text processing. A lot of methods have been developed for text summarization using fuzzy based systems. A number of parameters like sentence position in paragraph, sentence length, similarity to title, similarity of keyword, similarity to text concept, proper noun, and sentence cohesion are used in fuzzy systems. The authors [14] have compared the score of vector based method and fuzzy method given by five judges and the fuzzy based summary gave a summary which rejects 77% of the concepts as opposed to 66% performance by the vector based method. In another method, the vector features [15] created for each sentence in the document include title feature, sentence length, term weight, sentence position, sentence to similarity sentence, numerical data etc. The results compared with word summarizer, copernic summarizer has shown a better result. Almost the same set of features are used by a triangular membership function [16] which fuzzifies each score to three values: low, medium and high. A parallel summary using latent semantic analysis is also taken and both are merged to get the final summary. Experimental results have shown an average precision of 89%. A comparison of fuzzy system is done with neural network with features like cue phrases, legal vocabulary, paragraph structure, citation, term weight, named entity, similarity to neighboring sentences, absolute location etc and a better result is shown. Almost the same set of features are used for fuzzy system is done with neural network with features like cue phrases, legal vocabulary, paragraph structure, citation, term weight, named entity, similarity to neighboring sentences, absolute location etc and a better result is shown.

Fig.1 (a) H (b) X (c) H-with Node numbers (d) Y

IV. ALGEBRA OF FILTER

A morphological filter is nothing but an opening filter (γ) which is obtained as (δ o ε) or a closing filter (Φ) which is obtained as (ε o δ). Opening filter can again be classified as i) opening filter w.r.t nodes δ(ε(Y)) for which the filtrate are nodes ii) opening filter w.r.t hyperedges δ(ε(Y)) for which the filtrate are hyperedges. Closing filter can be classified as i) closing filter w.r.t nodes ε(δ(Y)) for which the results are nodes ii) closing filter w.r.t hyperedges ε(δ(Y)) for which the results are hyperedges. Now let us apply algebraic operations like union, intersection, complement operations on these filters.

A. Proposition A: Let H be a parent IFHG, X and Y be sub IFHGs, δ be the dilation operator and ε be the erosion operator, then

\[ \delta^{\circ}(e^{\circ}(X \cup Y)) = \delta^{\circ}(e^{\circ}(X^{\circ})) \cup \delta^{\circ}(e^{\circ}(Y^{\circ})) \]  

(1)

where δ(ε(ε(X \cup Y)) is an opening filter w.r.t hyperedges.
Proof: Consider L.H.S of eq(1). Let \( \epsilon \) be an arbitrary node in \( \varepsilon^n(X \cup Y) \), i.e.,
\[
v \notin (X \cup Y) \quad; \quad \epsilon \in (X \cup Y)
\]
(2)
Let \( \epsilon \) be the edge which contains this \( \epsilon \), i.e,
\[
e \subseteq \delta^e(X \cup Y)
\]
(3)
Since \( \epsilon \in (X \cup Y) \), this \( \epsilon \) may be present either in X or Y. Consider R.H.S of eq(1). Let \( \epsilon \) be a node in \( \varepsilon^n(X) \), i.e,
\[
v \notin X^\epsilon \quad; \quad \epsilon \in X
\]
(4)
Now \( \delta^e(X^\epsilon) = \{ e_1, e_2, \ldots, e_k \} \), which contain \( \epsilon \). Let \( \epsilon \) be a node in \( \varepsilon^n(Y) \), i.e,
\[
u \notin Y^\epsilon \quad; \quad \epsilon \in Y
\]
(5)
Now \( \delta^e(X^\epsilon) = \{ e_1, \ldots, e_k \} \), let \( \epsilon \) be an arbitrary edge in \( \{ e_1, \ldots, e_k \} \).
This \( \epsilon \) can be either a member of \( \delta^e(X^\epsilon) \) or \( \delta^e(Y^\epsilon) \), i.e,
\[
e \subseteq \delta^e(X^\epsilon) \cup \delta^e(Y^\epsilon)
\]
(6)
Hence Eq(1) is implied from eq(3) and eq(6).

Example 4.1 Consider the IFHGs given in Fig 1. Let us find the R.H.S of eq(1), where we get \( \varepsilon^n(X) = \{ e_31, e_32, e_33, e_35 \} \). Thus we get \( \delta^e(X^\epsilon) = \{ e_31, e_32, e_33, e_35 \} \). Again in R.H.S of eq(7) we can find \( \varepsilon^n(Y) = \{ e_31, e_36, e_37 \} \). Thus we get \( \delta^e(Y^\epsilon) = \{ e_36, e_37 \} \). Now \( \delta^e(X^\epsilon) \cup \delta^e(Y^\epsilon) = \{ e_31, e_32, e_33, e_35, e_36, e_37 \} \). Let us consider the L.H.S of eq(7), where \( \delta^e(X \cup Y) = \{ e_31, e_32, e_33, e_35, e_36, e_37 \} \). The results are shown in Fig.2(a).

\[
\begin{align*}
\text{Fig 2 : Results of (a) Proposition A (b) Proposition B (c) Proposition C}
\end{align*}
\]

\[\text{B. Proposition B: Let } H \text{ be a parent IFHG, } X \text{ and } Y \text{ be sub IFHGs, } \delta \text{ be the dilation operator and } \epsilon \text{ be the erosion operator, then}
\]
\[
\delta^e(X \cup Y) \subseteq \delta^e(X^\epsilon) \cup \delta^e(Y^\epsilon)
\]
(7)
where \( \delta^e(X \cup Y) \) is an opening filter w.r.t to hyperedges.

Proof: Consider L.H.S of eq(7). Let \( \epsilon \) be an arbitrary node in \( \varepsilon^n(X \cup Y) \), i.e,
\[
v \notin (X \cup Y) \quad; \quad \epsilon \in (X \cup Y)
\]
(8)
Let \( \epsilon \) be the edge which contains this \( \epsilon \). Now we can say that
\[
e \subseteq \delta^e(X \cup Y)
\]
(9)
Consider R.H.S of eq(7). Let \( \delta^e(X^\epsilon) = \{ e_1, \ldots, e_p \} \). Let \( \delta^e(Y^\epsilon) = \{ e_p, \ldots, e_k \} \). Let \( \epsilon \) be an arbitrary edge in \( \{ e_1, \ldots, e_k \} \).
This \( \epsilon \) will be a member of \( \delta^e(X^\epsilon) \) and \( \delta^e(Y^\epsilon) \), i.e,
\[
e \subseteq \delta^e(X^\epsilon) \cup \delta^e(Y^\epsilon)
\]
(10)
Eq(7) is implied from eq(9) and eq(10).

Example 4.2 Consider the IFHGs given in Fig 1. Let us find the R.H.S of eq(7), where we get \( \varepsilon^n(X) = \{ n_31, n_33, n_35 \} \). Thus we get \( \delta^e(X^\epsilon) = \{ n_31, n_33, n_35 \} \). Again in R.H.S of eq(7) we can find \( \varepsilon^n(Y) = \{ n_31, n_36, n_37 \} \). Thus we get \( \delta^e(Y^\epsilon) = \{ n_31, n_36, n_37 \} \). Now \( \delta^e(X^\epsilon) \cup \delta^e(Y^\epsilon) = \{ n_31, n_33, n_35, n_36, n_37 \} \). Let us consider the L.H.S of eq(7), where \( \delta^e(X \cup Y) = \{ n_31, n_33, n_35, n_36, n_37 \} \). The results are shown in Fig.2(b).

C. Proposition C: Let \( H \) be a parent IFHG, \( X \) and \( Y \) be sub IFHGs, \( \delta \) be the dilation operator and \( \epsilon \) be the erosion operator, then
\[
\delta^e(X \cup Y) \subseteq \delta^e(X^\epsilon) \cup \delta^e(Y^\epsilon)
\]
(11)
where \( \delta^e(X \cup Y) \) is an opening filter w.r.t to nodes.

Proof: Consider L.H.S of eq(11).Let \( \epsilon \) be an arbitrary edge in \( \varepsilon^n(X \cup Y) \). \( \epsilon \) be an arbitrary node in \( \varepsilon^n(X \cup Y) \). By definition of \( \delta^e(X \cup Y) \), this \( \epsilon \) can be a node either of \( X \) or \( Y \) . i.e, \( \epsilon \in X \) or \( \epsilon \in Y \). Consider R.H.S of eq(11). Let \( \{ \epsilon_1, \epsilon_2, \ldots, \epsilon_p \} \) be the nodes in \( \delta^e(X^\epsilon) \). By definition of \( \delta^e(X^\epsilon) \), a node in the above set belongs to \( \epsilon \). ie \( \epsilon \in X \) or \( \epsilon \in Y \). Let \( \{ \epsilon_{p+1}, \epsilon_{p+2}, \ldots, \epsilon_q \} \) be the nodes in \( \delta^e(Y^\epsilon) \). By definition of \( \delta^e(Y^\epsilon) \), a node in the above set belongs to \( Y \). ie \( \epsilon \in Y \). A node \( \epsilon \) in \( \{ \epsilon_1, \epsilon_2, \ldots, \epsilon_p \} \) is either an element of \( X \) or \( Y \). Thus \( \epsilon \in X \) or \( \epsilon \in Y \). This implies eq(11).

Example 4.3 Consider IFHGs in Fig.1. Let us find the R.H.S of eq(11),where we get \( \varepsilon^n(X) = \{ c_31, c_32, c_33 \} \). Thus \( \delta^e(X^\epsilon) = \{ c_31, c_32, c_33 \} \). Again \( \varepsilon^n(Y) = \{ c_36, c_37 \} \). Thus \( \delta^e(Y^\epsilon) = \{ c_36, c_37 \} \). Let us consider the L.H.S of eq(11), where \( \delta^e(X \cup Y) = \{ c_31, c_32, c_33, c_36, c_37 \} \). The results are shown in Fig.2(c).
Considering L.H.S, we get $e^n(X \cup Y)^n = \{ e_0, e_1, e_{10}, e_{11}, e_{12} \}$. Thus $\delta^b(e^n(X \cup Y)^b) = \{ n_1, n_8, n_9, n_{12}, n_{13}, n_{14}, n_{15}, n_{17}, n_{16}, n_{19}, n_{20}, n_{32}, n_{33}, n_{35}, n_{36}, n_{37} \}$. The results are shown in Fig.2(c).

**D. Proposition D:** Let H be a parent IFHG, X and Y be sub IFHGs, $\delta$ be the dilation operator and $\varepsilon$ be the erosion operator, then

$$\delta^b(e^n(X \cup Y)^b) \subset \delta^b(e^n(Y^n)) \cap \delta^b(e^n(X^n))$$  \hspace{1cm} (12)

where $\delta^b(e^n(X \cup Y)^b)$ is an opening filter w.r.t to nodes.

**Example 4.4:** Consider the IFHGs given in Fig.1. Let us find the L.H.S of eq(12). Here $e^n(X \cup Y)^b$ is the set of edges which contains $(X \cup Y)^b$ only. ie, $e^n(X \cup Y)^b = \{ e_{11} \}$. Now $\delta^b(e^n(X \cup Y)^b)$ is the set of nodes within $\{ e_{11} \}$. Thus $\delta^b(e^n(X \cup Y)^b) = \{ n_{13}, n_{14}, n_{18}, n_{19} \}$. Now consider R.H.S of eq(12). Here $\delta^b(e^n(Y^n)) = \{ n_1, n_{8}, n_{13}, n_{14}, n_{17}, n_{18}, n_{19}, n_{26}, n_{27}, n_{32} \}$. Also $\delta^b(e^n(Y^n)) = \{ n_8, n_9, n_{13}, n_{14}, n_{15}, n_{18}, n_{19}, n_{20}, n_{28}, n_{29}, n_{30} \}$. Thus $\delta^b(e^n(X^n)) \cap \delta^b(e^n(Y^n)) = \{ n_{13}, n_{14}, n_{18}, n_{19}, n_{32} \}$. Thus we can see that L.H.S $\subset$ R.H.S. The results are given in Fig.3(a) and 3(b).

Fig 3 : (a) L.H.S of Proposition D  , 3(b) R.H.S. of Proposition D

So far we have seen opening filters. Now let us see closing filters, which are of four types.

**E. Proposition E:**  Let H be a parent IFHG, X and Y be sub IFHGs, $\delta$ be the dilation operator and $\varepsilon$ be the erosion operator, then

$$\varepsilon^b(\delta^b(X \cup Y)^b) = \varepsilon^b(\delta^b(X^n)) \cup \varepsilon^b(\delta^b(Y^n))$$  \hspace{1cm} (13)

where $\varepsilon^b(\delta^b(X \cup Y)^b)$ is a closing filter w.r.t to hyperedges.

**Proof:** Consider the L.H.S of eq(13). Let $v$ be an arbitrary node in $\delta^b(X \cup Y)^b$. According to the definition of $\delta^b$, $v \in X$ or $v \in Y$. Let $h$ be an edge which contains this $v$. By definition of $\varepsilon^b$, this is an edge in both $X$ and $Y$. ie, $\varepsilon^b(h) \in X \cup Y$. Consider R.H.S of eq(13). Let $\{ h_1, h_2, \ldots, h_p \}$ be the edges in $\varepsilon^b(\delta^b(X^n))$. By definition of $\varepsilon^b$, an edge $h$ in $\{ h_1, h_2, \ldots, h_p \}$ is an element of X. ie, $h \in X$. Let $\{ h_{p+1}, h_{p+2}, \ldots, h_q \}$ be edges in $\varepsilon^b(\delta^b(Y^n))$. By definition of $\varepsilon^b$, an edge $h$ in $\{ h_1, h_2, \ldots, h_q \}$ will be edge of $X$ or $Y$. ie $h \in X$ or $h \in Y$. Eq(13) is implied from this.

**Example 4.5:** Consider the IFHGs given in Fig.1. In L.H.S of eq(13), $\delta^b(X \cup Y)^b$ are the nodes within $(X \cup Y)^b$, ie, $\delta^b(X \cup Y)^b = \{ n_5, n_6, n_9, n_{13}, n_{14}, n_{15}, n_{17}, n_{18}, n_{19}, n_{20}, n_{31}, n_{32}, n_{33}, n_{35}, n_{36}, n_{37} \}$. Now $\varepsilon^b(\delta^b(X \cup Y)^b) = \{ e_6, e_7, e_{10}, e_{11}, e_{12} \}$. Consider R.H.S of eq(13). Here $\delta^b(Y^n)$ are the nodes within $X^n$. ie, $\delta^b(Y^n) = \{ n_7, n_8, n_{12}, n_{13}, n_{14}, n_{15}, n_{17}, n_{18}, n_{19}, n_{31}, n_{33}, n_{35}, n_{36} \}$. Thus $\varepsilon^b(\delta^b(Y^n))$ is the set of edges which consists of these nodes only. ie $\varepsilon^b(\delta^b(Y^n)) = \{ e_6, e_7, e_{10}, e_{11}, e_{12} \}$. Now $\delta^b(Y^n)$ are the nodes in $Y^n$. ie, $\delta^b(Y^n) = \{ n_7, n_8, n_{12}, n_{13}, n_{14}, n_{15}, n_{17}, n_{18}, n_{19}, n_{31}, n_{33}, n_{35}, n_{36}, n_{37} \}$. Thus $\varepsilon^b(\delta^b(Y^n))$ is the set of edges which contain the above nodes only. ie, $\varepsilon^b(\delta^b(Y^n)) = \{ e_6, e_7, e_{10}, e_{11}, e_{12} \}$. Therefore $\varepsilon^b(\delta^b(X^n)) \cup \varepsilon^b(\delta^b(Y^n)) = \{ e_6, e_7, e_{10}, e_{11}, e_{12} \}$. The results are shown in Fig 4(a).

**F. Proposition F:** Let H be a parent IFHG, X and Y be sub IFHGs, $\delta$ be the dilation operator and $\varepsilon$ be the erosion operator, then

$$\varepsilon^b(\delta^b(X \cup Y)^b) = \varepsilon^b(\delta^b(X^n)) \cup \varepsilon^b(\delta^b(Y^n))$$  \hspace{1cm} (14)

where $\varepsilon^b(\delta^b(X \cup Y)^b)$ is a closing filter w.r.t to hyperedges.

**Proof:** Consider L.H.S of eq(14). Let $v$ be a node in $\delta^b(X \cup Y)^b$. By definition of $\delta^b$, it is a node in both $X$ and $Y$. Let $e$ be an edge which contains this $v$. By definition of $\varepsilon^b$, this is an edge in both $X$ and $Y$. ie, $e \in X \cup Y$. Consider R.H.S of eq(14). Let $\{ u_1, u_2, \ldots, u_p \}$ be nodes in $\delta^b(X^n)$. By definition of $\delta^b$, these nodes are in $X$. Let $\{ e_1, e_2, \ldots, e_p \}$ be the edges which consist of these nodes. By definition of $\varepsilon^b$, these edges are in $X$. Let $\{ u_{p+1}, u_{p+2}, \ldots, u_q \}$ be nodes in $\delta^b(Y^n)$. By definition of $\delta^b$, these nodes are in $Y$. Let $\{ e_{p+1}, e_{p+2}, \ldots, e_q \}$ be the edges which consist of these nodes. By definition of $\varepsilon^b$, these edge is in $Y$. Now an edge $e$ in $\{ e_1, e_2, \ldots, e_q \}$ is present both in $X$ and $Y$. ie $e \in X, e \in Y$. 

8
Example 4.6 Consider L.H.S of eq(14). There $\delta^0(X \cap Y)^p = \{n_1, n_2, n_18, n_9, n_3\}$. Now $\varepsilon^0(\delta^0(X \cap Y)^p)$ is the set of edges which contain the above nodes only. i.e. $\varepsilon^0(\delta^0(X \cap Y)^p) = \{e_1\}$. Consider R.H.S of eq(14). $\delta^0(X)$ is the set of nodes in $X^p$. i.e., $\delta^0(X) = \{n_1, n_2, n_12, n_13, n_14, n_17, n_18, n_19, n_31, n_35, n_36, n_38\}$. Now $\varepsilon^0(\delta^0(X))$ is the edges which contain the above nodes only. i.e., $\varepsilon^0(\delta^0(X)) = \{e_9, e_{10}, e_{11}\}$. Likewise $\delta^0(Y)$ is the set of nodes in $Y^p$ only = $\{n_2, n_3, n_15, n_18, n_19, n_9, n_32, n_36, n_37\}$. Now $\varepsilon^0(\delta^0(Y))$ is the set of edges which contains these nodes only. i.e., $\varepsilon^0(\delta^0(Y)) = \{e_7, e_{11}, e_{12}\}$. Now $\varepsilon^0(\delta^0(X)) \cap \varepsilon^0(\delta^0(Y)) = \{e_{11}\}$. The results are shown in Fig 4(b).

G. Proposition G: Let H be a parent IFHG, X and Y be sub IFHGs, $\delta$ be the dilation operator and $\varepsilon$ be the erosion operator, then

$$\varepsilon^0(\delta^0(X \cup Y)^p) = \varepsilon^0(\delta^0(X)^p) \cup \varepsilon^0(\delta^0(Y)^p)$$  \hspace{1cm} (15)

where $\varepsilon^0(\delta^0(X \cup Y)^p)$ is a closing filter w.r.t. to nodes.

Proof: Consider L.H.S of eq(15). Let $\{e_1, e_2, ..e_n\}$ be the edges in $\delta^0(X \cup Y)^p$. Let $v$ be an arbitrary node in any edge in $\{e_1, e_2, ..e_n\}$. According to the definition of $\varepsilon^0(\delta^0(X \cup Y)^p)$, this $v$ can be contained in $\{e_1, e_2, ..e_n\}$ but not in $\{e_1, e_2, ..e_n\}'$ Consider R.H.S of eq(15). Let $\{e_1, e_2, ..e_n\}$ be the edges in $\delta^0(X)$. Let $\{n_1, n_2, ..n_p\}$ be the nodes in $\{e_1, e_2, ..e_n\}$. According to the definition of $\varepsilon^0(\delta^0(X))$ these nodes are present in $\{e_1, e_2, ..e_p\}$ but not in $\{e_1, e_2, ..e_p\}'$. Let $\{e_{p+1}, e_{p+2}, ..e_k\}$ be the edges in $\delta^0(Y)$. Let $\{n_{p+1}, n_{p+2}, ..n_k\}$ be the nodes in $\{e_{p+1}, e_{p+2}, ..e_k\}$. According to the definition of $\varepsilon^0(\delta^0(Y))$ these nodes are present in $\{e_{p+1}, e_{p+2}, ..e_k\}$ but not in $\{e_{p+1}, e_{p+2}, ..e_k\}'$. Any edge $v$ in $\{e_1, e_2, ..e_n\}$, is present in $\{e_1, e_2, ..e_p\}$ but not in $\{e_1, e_2, ..e_p\}'$.

Example 4.7 Consider L.H.S of eq(15). Here $\delta^0(X \cup Y)^p$ is the set of edges which contains nodes in $(X \cup Y)^p = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_{10}, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}\}$. Thus $\delta^0(X \cup Y)^p$ is the set of edges in $X^p$ = All edges in $X$. Now $\varepsilon^0(\delta^0(X \cup Y)^p)$ is the set of nodes in $X^p$ but not in their complement = All nodes in $X$. Consider R.H.S of eq(15). Here $\delta^0(Y)^p$ is the set of edges which contains nodes in $Y^p = \{e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_{10}, e_{11}, e_{12}\}$. Thus $\delta^0(Y)^p$ is the set of edges in $Y^p$ but not in their complement = All nodes in $Y$. Now $\varepsilon^0(\delta^0(Y)^p)$ is the set of edges which contains nodes in $Y^p$ = All edges in $Y$. Now $\varepsilon^0(\delta^0(Y)^p)$ is the set of nodes in $Y^p$ but not in their complement = All nodes in $Y$. Hence $\varepsilon^0(\delta^0(Y)^p)$ are the nodes in the above edges, but not in their complement = $\{n_2, n_3, n_15, n_18, n_19, n_20, n_21, n_22, n_23, n_24, n_25, n_26, n_27, n_28, n_29, n_31, n_32, n_35, n_36, n_37\}$. Thus L.H.S $\subset$ R.H.S. The results are shown in Fig 5.

H. Proposition H: Let H be a parent IFHG, X and Y be sub IFHGs, $\delta$ be the dilation operator and $\varepsilon$ be the erosion operator, then

$$\varepsilon^0(\delta^0(X \cap Y)^p) \subset \varepsilon^0(\delta^0(X)^p) \cap \varepsilon^0(\delta^0(Y)^p)$$  \hspace{1cm} (16)

where $\varepsilon^0(\delta^0(X \cap Y)^p)$ is a closing filter w.r.t. to nodes.

Example 4.8 Consider L.H.S of eq(16). Here $\delta^0(X \cap Y)^p$ is the set of edges which consists of any element in $(X \cap Y)^p = \{e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_{10}, e_{11}, e_{12}\}$. Now $\varepsilon^0(\delta^0(X \cap Y)^p)$ is the set of nodes in the above edges but not in their complement . i.e $\varepsilon^0(\delta^0(X \cap Y)^p) = \{n_2, n_3, n_15, n_18, n_19, n_20, n_21, n_22, n_23, n_24, n_25, n_26, n_27, n_28, n_29, n_31, n_32, n_35, n_36, n_37\}$). Now consider R.H.S of eq(16). $\delta^0(X)^p$ is the set of edges which contain any node in $X^p$ . i.e. $\delta^0(X)^p = $ All edges in $H$. Now $\varepsilon^0(\delta^0(X)^p) = \varepsilon^0(\delta^0(Y)^p) = \varepsilon^0(\delta^0(Y)^p)$ are the nodes in the above edges, but not in their complement = $\{n_2, n_3, n_15, n_18, n_19, n_20, n_21, n_22, n_23, n_24, n_25, n_26, n_27, n_28, n_29, n_31, n_32, n_35, n_36, n_37\}$. Thus L.H.S $\subset$ R.H.S. The results are shown in Fig 5.

V. APPLICATION OF FILTER IN MULTI DOCUMENT SUMMARY

Multi document summarization is an important area in the field of Natural Language processing. There are many summarization methods developed so far. In this section let us see how multiple documents can be represented using an IFHG assigning membership and non membership values to the nodes and hyperedges.

A. Modeling Multiple documents using IFHG

Let $[H_F, (\mu_a, \gamma_a), (\mu_c, \gamma_c), H^p, H^p']$ be an Intuitionistic Fuzzy hypergraph as mentioned in Section 3. Let $H^p$ be the
hyperedges, where a hyperedge $e_i$ represents a document $D_i$. Let $H^n$ be the nodes where a node $v_j$ represents a keyword $w_j$ in the document. The same is shown in Fig 6 (a). Therefore the number of hyperedges in the $H^n$ will be same as the number of documents considered for summarization and the number of nodes in a hyperedge will be same as the number of keywords in that document. The membership value for a node $v_j$ depends on the normalized TF-IDF value and the priority of that word. Irrespective of the domain there are words which are having high priority. E.g.: Words like important, famous, beautiful, attractive, relevant etc are given high priority. So these words will have $\mu_n > 0.5$. For such words $\gamma_n = 1 - \mu_n$. There are words which are having very low priority. Some sample words include notorious, expensive, least, badly etc. These words are having $\gamma_n > 0.5$. So for such words $\mu_n < 1 - \gamma_n$. Rest of the words are given $\mu_n = \gamma_n = 0.5$ as they are medium priority words. Once all these nodes (words) are assigned with both membership degree ($\mu_n$) and non-membership degree ($\gamma_n$), the hyperedges are also assigned membership degree ($\mu_e$) and non-membership degree ($\gamma_e$) as per the rules in Section I.

Fig 6 : (a) IFHG of Documents (b) Partitions

B. System Architecture

For the formation of IFHG, all the documents are subjected to preprocessing steps like stop word removal and stemming. Keywords in a document are found out by considering the tf-idf of the words. Only words with $tf \cdot idf > \theta$ are considered for the construction of the IFHG. Once an IFHG is formed as mentioned in Section V.A, it can be partitioned in to sub IFHGs based on the absence of overlapping nodes between the hyperedges as shown in Fig 6 (b). Such partitions are now document clusters. For each cluster, sub IFHG $X_i$ is created by applying an ($\alpha, \beta$) cut. This $X_i$ is subjected to opening and closing filters as mentioned in Section IV. The results of filter w.r.t hyperedges will yield good documents and results of filter w.r.t nodes will yield priority keywords. These priority words combined with priority documents yield good summaries. If needed, the number of parameters considered for assigning membership degree can be increased. The architecture of the system is shown in Fig 7. The algorithm is shown in Fig 8.

Algorithm 1 Multidocument summary using IFHG

1: Input : a documents, $D_i, \theta$
2: Output : Summary
3: for each document $D_i = 1$ to $n$ do
4: Remove stop words
5: Perform Stemming of words
6: Find $tf \cdot idf$ for all words
7: Find all words $v_j$ in $D_i$ with $tf \cdot idf > \theta$
8: end for
9: $P =$ number of hypernodes = $\sum$ Number of words in all the $D_i$
10: $J =$ number of hyperedges = $n$
11: for each $k = 1$ to $P$ do
12: Assign $\mu_n$, $\gamma_n$ for all nodes as per rules in section V.A
13: end for
14: for each $m = 1$ to $J$ do
15: Assign $\mu_e$, $\gamma_e$ for all edges as per rules in section V.A
16: end for
17: Clusters $\rightarrow$ Find sub IFHGs with non overlapping hypernodes
18: for each cluster $C_i = 1$ to $c$ do
19: Find sub IFHG $X_i$, with ($\alpha, \beta$) cut where $\alpha > \gamma_0$
20: Apply opening filter w.r.t. nodes $\delta(\theta(X_i)^p)/\delta(\theta(X_i)^n)$ to find relevant documents.
21: Apply closing filter w.r.t. hyperedges $\theta(X_i)^e$ to find keywords in summary.
22: Find Opening_Summary_Filter($C_i$) $= \text{Combine}(\delta^p(\theta(X_i)^p)/\delta^p(\theta(X_i)^n), \delta^e(\theta(X_i)^e))$
23: Apply closing filter w.r.t. nodes $\delta(\theta(X_i)^n)$ to find relevant documents.
24: Find Closing_Summary_Filter($C_i$) $= \text{Combine}(\delta^p(\theta(X_i)^p), \delta^e(\theta(X_i)^e))$
25: end for
26: for two clusters $C_i, C_j$ do
27: Find Opening_Summary_Filter($C_i$) $\cup$ Opening_Summary_Filter($C_j$) to create union of two summaries
28: Find $\delta(\theta(X_i)^p)$ $\cup$ $\delta(\theta(X_j)^p)$ to create union of keywords in summaries
29: Find Closing_Summary_Filter($C_i$) $\cup$ Closing_Summary_Filter($C_j$) to create union of two summaries
30: Find $\delta(\theta(X_i)^p)$ $\cup$ $\delta(\theta(X_j)^p)$ to find all the keywords of two summaries
31: end for

In the above algorithm we can see that summary can be generated either by combining an opening filter with respect to hyperedges along with opening filter with respect to nodes or
by using a closing filter with respect to the hyperedges along with closing filter with respect to the nodes. Similarly keywords can be found by using an opening filter or by using a closing filter.

VI. RESULT ANALYSIS

The proposed system with multiple documents modeled as IFHG was tested with different number of input documents. The system developed in python, and the hypergraph created using python pygraph works for both Malayalam documents and English documents. The difference between two systems lies in the stemming phase, where in Malayalam, a stemmer developed using Tree based method [20] is used. Porter stemmer is being used for English documents. Rests of the developed using Tree based method [20] is used. Porter stemmer is being used for English documents. Rests of the

Developing Intuitionistic Fuzzy Hypergraphs (IFHG) for document summarization. The results obtained with various test cases are given in Table I. The results of the IFHG system is compared with results of 50 human summarizers and Rouge-L, Rouge-2, Rouge-1 scores are calculated and tabulated in Table II.

<table>
<thead>
<tr>
<th>No of Docs</th>
<th>Total No of Lines</th>
<th>Total No of Words</th>
<th>No of nodes created</th>
<th>No of Summarized Lines</th>
<th>Percent age of Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>43</td>
<td>1975</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>125</td>
<td>4453</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>135</td>
<td>5749</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>257</td>
<td>9497</td>
<td>27</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>393</td>
<td>13,372</td>
<td>48</td>
<td>110</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>444</td>
<td>25,852</td>
<td>50</td>
<td>125</td>
</tr>
</tbody>
</table>

Advantages of IFHG based Summary system:

- Even though the number of words in the document may be large, the IFHG created will be having only less number of nodes as we are considering only keywords for constructing the nodes. This reduces the space required for storing the data structure.
- Priority levels are not limited to high, medium and low. Different levels of priorities like very high, high, medium high, low, very low can be set with the help of the membership and non membership degrees of the words as well as the documents.
- Varying the threshold of ($\alpha$, $\beta$) cut results in different summaries with different priority levels. Increasing the threshold of ($\alpha$, $\beta$) cut results in a very short summary, where by reducing it results in a lengthy summary.
- Range oriented ($\alpha$, $\beta$) cuts like $\text{value}_1 < \alpha < \text{value}_2$ are also possible which results in different sets of summaries.
- Since filters w.r.t to nodes result in keyword list of documents, union/intersection of filters result in union/intersection of keywords. Since filters w.r.t hyperedges result in selection of good documents, they combined with node filters result in summary. Union/intersection of summaries are also possible.

TABLE II. ROUGE SCORES

<table>
<thead>
<tr>
<th>No of docs</th>
<th>Rouge-L</th>
<th>Rouge-2</th>
<th>Rouge-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F</td>
</tr>
<tr>
<td>3</td>
<td>0.87</td>
<td>0.93</td>
<td>0.90</td>
</tr>
<tr>
<td>4</td>
<td>0.84</td>
<td>0.96</td>
<td>0.89</td>
</tr>
<tr>
<td>5</td>
<td>0.90</td>
<td>0.96</td>
<td>0.93</td>
</tr>
<tr>
<td>6</td>
<td>0.91</td>
<td>0.96</td>
<td>0.94</td>
</tr>
</tbody>
</table>

VII. CONCLUSION

The proposed work is a novel method which models multiple documents using Intuitionistic Fuzzy hypergraphs. The system is tested with documents of varying size and has shown better results when compared to human summaries. Documents in various domains are considered for summarization. The priority levels, ($\alpha$, $\beta$) cuts, range oriented ($\alpha$, $\beta$) cuts and union/intersection operations applied on the filters give different types of summaries suitable for different applications. Here IFHG takes only less space since care is taken to reduce the number of nodes. We have also developed a system which models sentences as hyperedges and words as nodes. Such IFHG modeling and filtering can be done in other areas like mobile networking, social networking, image processing etc. Modeling medical reports as IFHG and creating medical report summary is a future enhancement of this work.

References


