ORIGINAL ARTICLE

A study of cluster hypergraphs and its properties

Ananta Maity[1] · Kousik Das[2] · Sovan Samanta[3] · Sukumar Mondal[1] · Vivek Dubey[4]

Abstract
In this study, cluster hypergraphs are introduced to generalize the concept of hypergraphs, where cluster nodes are allowed. Few related terms and properties on cluster hypergraphs are discussed. Some operations, including the Cartesian product, union, intersection, etc., are studied. Different types of matrix representations and isomorphism are also proposed on cluster hypergraphs. The notion of an effective degree for nodes is introduced to capture the group/ cluster effects. At last, the area of applications and future directions with conclusions is deployed.

Keywords Hypergraphs · Cluster hypergraphs · Degree and effective degree · Basic operations · Matrix representation · Isomorphisms

1 Introduction

The graphs are perfect representations of relations among data. The great mathematician L. Euler (1707–1783), known as the father of Graph Theory, first provided the concepts of graphs in 1736 by solving the famous Konigsberg bridge problem. A graph (Harary 1972; Sabidussi 1959) is a set of vertices with a relation among vertex. There are huge developments in graph theory and its applications in competitions, colorings, social networks (Das et al. 2018; Yan and Ding 2009), etc. But a hypergraph is a generalization of the graph in which any subset of a vertex set is an edge rather than two vertices. Specially, Berge (1961,1973, 1989) introduced hypergraphs as a generalization of graphs.

Buroscha and Ceccherini (1996) characterized cube-hypergraphs where each hyperedge contains three vertices. Tyshkevich and Zverovich (1996) introduced a new multivalued function, the line hypergraph. The function generalizes two classical concepts at once, namely, of the line graph and the dual hypergraph. In terms of this function, proofs of some known theorems online graphs can be unified, and their more general versions can be obtained. Sonntag and Teichert (2000) defined hypertrees, and they extended the notion to competition hypergraphs (Sonntag and Teichert 2004). Bollabos and Scott (2000) discussed three-uniform hypergraphs. There are various research works on hypergraphs (Djokovic 1973; Frankl and Katona 2000; Katona and Kierstead 1999). Bergen studied about the infection in hypergraphs (Bergen et al. 2018). But all these previous work's concepts of a group of nodes as a single node ignored. Samanta et al. (2020a) have been introduced cluster hypergraphs to generalize the concept of hypergraphs and to capture the notion of cluster nodes. The author considered the simple node as well as the cluster node in hypergraph where a simple node is as a usual node, and the cluster node is containing more than one simple node. The author also proposed the various types of the coloring of cluster hypergraphs (Samanta et al. 2020b).

✉ Vivek Dubey
vivek112000@gmail.com

Ananta Maity
amaity4499@gmail.com

Kousik Das
kousikmath@gmail.com

Sovan Samanta
ssamantavu@gmail.com

Sukumar Mondal
smnlkhan@gmail.com

[1] Department of Mathematics, Raja N. L. Khan Women's College (Autonomous), Midnapore, West Bengal 721102, India

[2] Department of Mathematics, D. J. H. School, Dantan, West Bengal 721451, India

[3] Department of Mathematics, Tamralipta Mahavidyalaya, East Medinipur, West Bengal 721636, India

[4] Amrita Vishwa Vidyapeetham, Bangalore, India
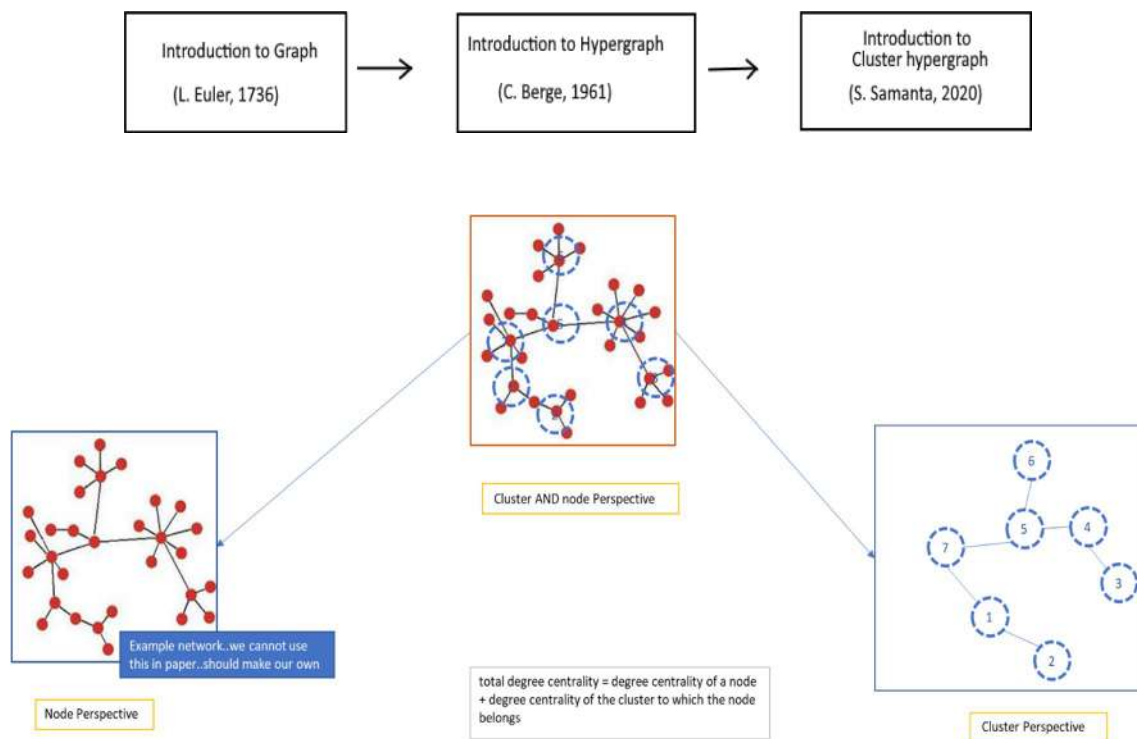
Current concepts of graphs use simple nodes and clusters separately to find the influence of/on a node. But this approach is not enough to capture more complicated situation, when a node is part of different clusters. For example, node 4 is part of two additional clusters: 8, 9. In this paper, we develop a general approach to address the issue for hyper cluster graphs. In this study, we introduced various terms and properties of cluster hypergraphs with the application. Thus, the major contributions of this study are as follows.

- The degree and effective degree of cluster hypergraphs has been introduced.
- Operations (Cartesian product, union, intersection) on two cluster hypergraphs have been proposed.
- Various matrices on cluster hypergraphs have been presented.
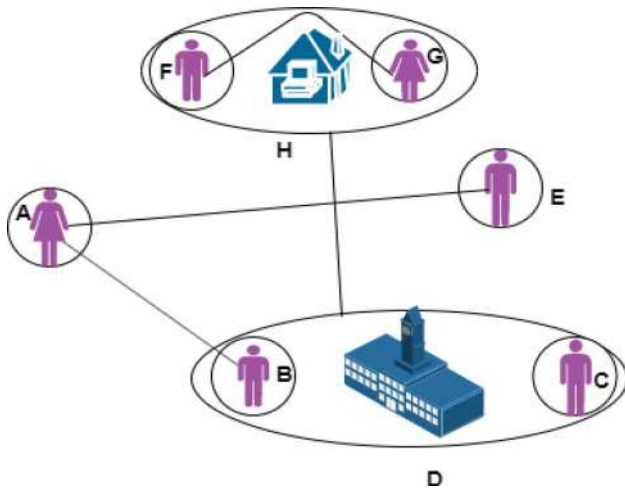- Type-I and Type-II isomorphism have been introduced.
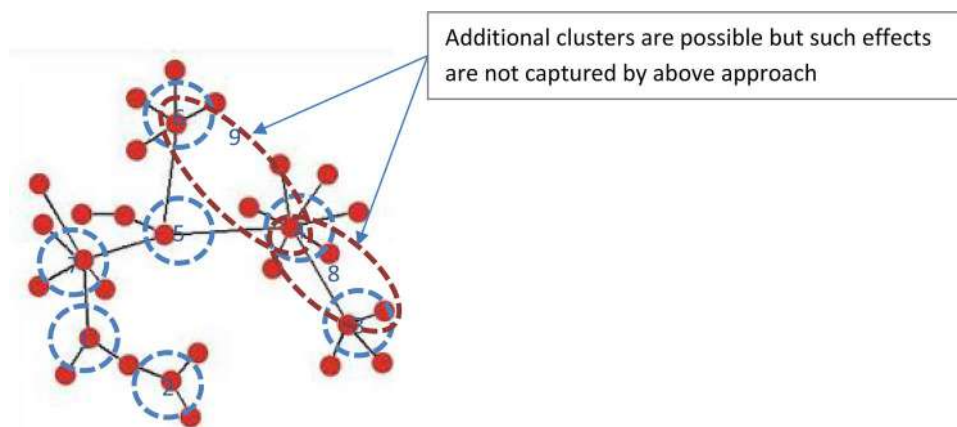
Fig. 1 Cluster hypergraph

The rest of the sections are organized as follows. In Sect. 2, preliminaries on cluster hypergraphs are described. In Sect. 3, the degree and effective degree of cluster hypergraphs have basic operations, such as a Cartesian product, union and intersection, are defined. In Sect. 4, different types of matrix representation with examples are displayed. In Sect. 5, isomorphisms on cluster hypergraphs are studied. In Sect. 6, the conclusion and future direction are described.

2 Preliminaries

Definition 2.1 Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set and let $E = \{e_1, e_2, \ldots, e_m\}$ be a family of subsets of X such that

$$e_i \neq \phi (i = 1, 2, \ldots, m)$$

$$\bigcup_{\{i=1\}}^{m} e_i = X$$

The pair (X, E) is called a hypergraph with vertex set X and hyperedge set E. The elements $\{x_1, x_2, \ldots, x_n\}$ of X are vertices of hypergraph H, and the sets $\{e_1, e_2, \ldots, e_m\}$ are hyperedges of hypergraph H.

Definition 2.2 Let X be a non-empty set and V_X be a subset of $P(X)$ such that $\phi \notin V_X$ and $X \subset V_X$. Now, E be a multi-set whose elements belong to $P(P(X))$ such that

- $\phi \notin E$
- For each element $e \in E$, there exists at least one element $v \in V_X$ such that $v \in e$.

Then, $G = (V_X, E)$ is said to be cluster hypergraph where V_X is said to be vertex set, and E is said to be multi-hyperedge set.

A **k-cluster** hypergraph is defined as follows:

Definition 2.3 Let X be a non-empty set and V_X be a subset of $P^k(X)$, $k = 1, 2, 3 \ldots$ such that $\phi \notin V_X$ and $X \subset V_X$. Now, E be a multi-set whose elements belong to $P(V_X)$ such that

- $\phi \notin E$
- For each element $e \in E$, there exists at least one element such that $v \in e$.

Then, $G = (V_X, E)$ is said to be k-cluster hypergraph where V_X is said to be vertex set, and E is said to be multi-hyperedge set. Generally, for $k = 1, -1$ cluster hypergraphs are assumed as cluster hypergraphs.

Example 2.4 In Fig. 1, $X = \{A, B, C, E, F, G\}$, $V_X = \{\{A\}, \{B\}, \{C\}, D = \{B, C\}, \{E\}, \{F\}, \{G\}, H = \{F, G\}\}$ and $E = \{\{\{A\}, \{B\}\}, \{\{A\}, \{E\}\}, \{\{B, C\}, \{B\}\}, \{\{B, C\}, \{C\}\}, \{\{F\}, \{G\}\}, \{\{F, G\}, \{F\}\}, \{\{F, G\}, \{G\}\}, \{\{F, G\}, \{B, C\}\}\}$. It can be easily verified that for each element e in E, there exists an element $v \in V_X$ such that $v \in e$. For example, $\{\{F, G\}, \{F\}\}$ in E,, there exists an element $\{F, G\} \in V_X$ such that $\{F, G\} \in \{\{F, G\}, \{F\}\}$.

Remark 2.5

- It is assumed that each node inside a cluster node is automatically connected to the cluster node, but these inside nodes may not be connected to each other.
- In a virtual representation (Fig. 2) of any cluster hypergraph, the cluster nodes are assumed as separate nodes,
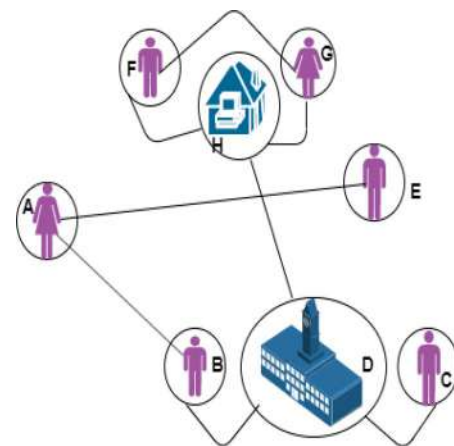

Fig. 2 Virtual representation of Fig. 1

and the connections to the inside nodes are shown in the representation.

- The maximal nodes are those nodes which are not contained in any other cluster nodes. The elements of X are termed as simple nodes. A simple node may be termed as a maximal node if it does not belong to any other nodes.
- Depending on the cluster node sizes and their edges, cluster hypergraphs are classified into some different categories.

3 Degree and effective degree

Definition 2.5 *Let $X = \{x_1, x_2, \ldots, x_m\}$ be a non-empty set and $G = (V_X, E)$ be a cluster hypergraph where $V_X = \{v_1, v_2, \ldots, v_n\}$ be set of nodes, that is, $v_i \in P(X)$, $i = 1, 2, \ldots, n$ and $E = \{e_1, e_2, \ldots, e_k\}$ be the set of edges, that is, $e_i \in P(P(X))$, $i = 1, 2, \ldots, k$. w, the degree of node v_i which is contained in the edges e_{ij}, $j = 1, 2, \ldots, p$ which is denoted as $d(v_i)$ and defined as $d(v_i) = \left| \sum_{j=1}^{p} e_{ij} \right|$ in the virtual representation of G.*

Note that an edge containing single vertex will contribute once to the degree of containing vertex. Similarly, the edges containing more than one vertex, contribute once to each of its vertices.

Example 2.6 In the cluster hypergraph (Fig. 1), the degree of node $\{D\}$ is 3.

A number of connections to a node are termed as a degree of a node. If a node is contained in another cluster node (group), then the node may have separate effect of connections. In this study, a separate term is defined for a node, an effective degree of a node.

Definition 2.7 *Let $X = \{x_1, x_2, \ldots, x_m\}$ be a non-empty set and $G = (V_X, E)$ be a cluster hypergraph where $V_X = \{v_1, v_2, \ldots, v_n\}$ be set of nodes, that is, $v_i \in P(X)$, $i = 1, 2, \ldots, n$ and $E = \{e_1, e_2, \ldots, e_k\}$ be the set of edges, that is, $e_i \in P(P(X))$, $i = 1, 2, \ldots, k$. Also, let, cv_i is cluster node containing simple node v_i. Now, an effective degree of a simple node v_i is denoted as $ed(v_i)$ d defined as*

$$ed(v_i) = d(v_i) + \frac{\sum_{i=1}^{l} d(cv_i)}{l},$$

where l is the number of cluster nodes containing v_i.

Example 2.8 In the cluster hypergraph (Fig. 1), degree of node $\{C\}$ is 1, but an effective degree of node $\{C\}$ is $1 + \frac{2}{1} = 3$.

Theorem 2.9 *Let $G = (V_X, E)$ be a cluster hypergraph where $|X| = k$. Then, the total degree of the cluster hypergraph is less than $(2^k - 1)2^{2^k - 2}$.*

Proof Since $G = (V_X, E)$ is a cluster hypergraph where $|X| = k$. Thus, maximum number of nodes of G is $2^k - 1$. Now, in cluster hypergraphs, there may be edges containing a single node, containing double nodes and so on. Thus, the contribution of edges of containing n nodes is $n \times$ (number of edges). The maximum contribution of all edges is
$$\binom{2^k - 1}{1} + 2 \times \binom{2^k - 1}{2} + 3 \times \binom{2^k - 1}{3} + \ldots + (2^k - 1)$$
$$\times \binom{2^k - 1}{2^k - 1} = (2^k - 1)2^{2^k - 2}.$$ Hence the result.

Corollary 2.10 *Let $G = (V_X, E)$ be a cluster hypergraph where $|V_X| = k$. Then, the total degree of the cluster hypergraph is less than $k \times 2^{k-1}$.*

Theorem 2.11 *Let $G = (V_X, E)$ be a cluster hypergraph where $|X| = k$. Then, the total effective degree of the cluster hypergraph is less than $\delta + (2^k - 1)(2^k - 2)$, where δ is the sum of maximum degrees of nodes, and l is the number of cluster nodes in the hypergraph.*

Proof Let $G = (V_X, E)$ be a cluster hypergraph where $X = \{x_1, x_2, \ldots, x_k\}$. Thus, the maximum number of nodes of G is 2^{k-1}. Now, an effective degree is $ed(x_i) = d(x_i) + \frac{\sum_{i=1}^{l} d(cx_i)}{l}$ ere l is the number of cluster nodes containing x_i. Here, sum of the effective degree is
$$\sum_{i=1}^{k} ed(x_i) = \sum_{i=1}^{k} \left[d(v_i) + \frac{\sum_{i=1}^{l} d(cv_i)}{l} \right] =$$
$$\sum_{i=1}^{k} d(v_i) + \sum_{i=1}^{k} \frac{\sum_{i=1}^{l} d(cv_i)}{l}.$$ Let us suppose the total number of cluster nodes in the hypergraph is l. Now, the degree of a cluster node is always less than $2^k - 2$. Thus, sum of the degrees of all cluster nodes is less than or equal $l \times (2^k - 2)$. One simple node may belong to all cluster nodes. Thus, sum of all effective degrees is less than or equal to $\delta + (2^k - 1)(2^k - 2)$, where $\delta = (2^k - 1)2^{2^k - 2}$ and l is the number of cluster nodes in the hypergraph. Hence, the result is true.
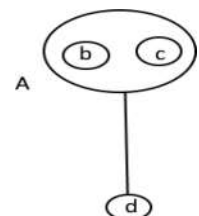
Fig. 3 Cluster hypergraph G

Fig. 4 Cluster hypergraph H

Fig. 5 Cluster hypergraph G × H

Fig. 6 Cluster hypergraph G

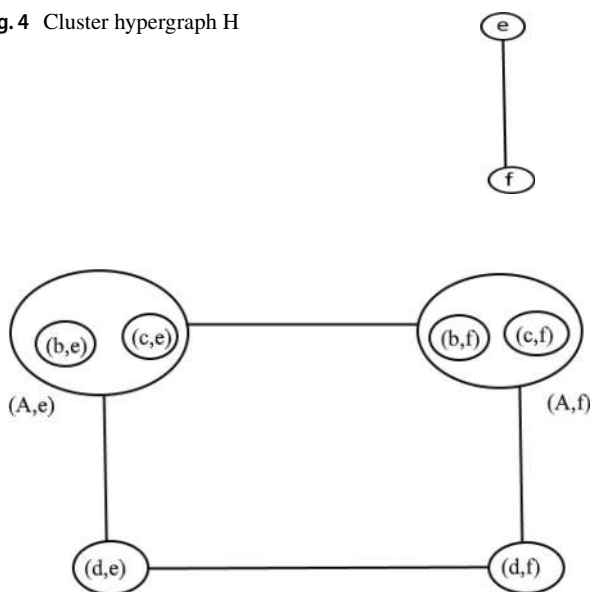**Fig. 7** Cluster hypergraph H

4 Operations on cluster hypergraphs

The basic operations, that is, cartesian product, union, the intersection of cluster hypergraphs are defined and studied some properties.

Definition 3.1 *Suppose* $G = (V_1, E_1)$ *and* $H = (V_2, E_2)$. *are two cluster hypergraphs. Then, the Cartesian product of G and H is denoted by* $G \times H = (V, E)$ *and defined as vertex set* $E = \{ (u_i, v_j), u_i \in V_1 \text{ and } v_j \in V_2 \}$ *and e is an edge* $\in E$ *of* $G \times H$ *if* $e = (u_i, v_j)(u_k, v_l)$ *where either* $i = k$ *and* $v_j v_l \in E_2$ *or* $j = l$ *and* $u_i u_j \in E_1$.

Example 3.2 Consider two cluster hypergraphs (Figs. 3, 4) $G = (V_1, E_1)$ and $H = (V_2, E_2)$, where $V_1 = \{\{b\}, \{c\}, A = \{b, c\}, \{d\}\}$,, $E_1 = \{\{b, A\}, \{c, A\}, \{A, d\}\}$ $V_2 = \{\{f\}, \{e\}\}$, $E_2 = \{\{f, e\}\}$. Then, the cross-product (Fig. 5) of G and H is $G \times H$, whose vertex set is $\{\{(b, e)\}, \{(b, f)\}, \{(c, e)\}, \{(c, f)\}, \{(A, e)\}, \{(A, f)\}, \{(d, e)\}, \{(d, f)\}\}$, and edge set is $\{\{\{(b, e)\}, \{(A, e)\}\}, \{\{(A, e)\}, \{(c, e)\}\}, \{\{(A, e)\}, \{(d, e)\}\}, \{\{(A, e)\}, \{(A, f)\}\}, \{\{(A, f)\}, \{(b, f)\}\}, \{\{(A, f)\}, \{(c, e)\}\}, \{\{(A, f)\}, \{(d, f)\}\}, \{\{(d, f)\}, \{(d, e)\}\}$

Proposition 3.3 *If* $G = (V_1, E_1)$ *and* $H = (V_2, E_2)$ *be two cluster hypergraph, then e is product* $G \times H$ *also a cluster hypergraph.*

Proof Let $G = (V_1, E_1)$ and $H = (V_2, E_2)$ are two cluster hypergraphs with $G \times H = (V, E)$. Then, $\emptyset \notin E_1$, $\emptyset \notin E_2$ and for each element e_1 in E_1, there exists at least one element v_1
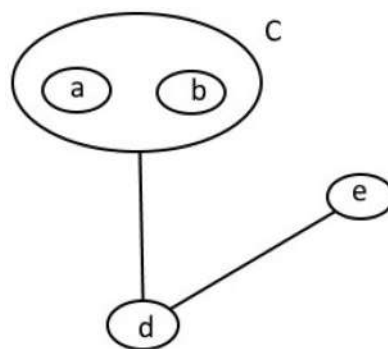
in V_1 such that $v_1 \in e_1$ and for each element e_2 in E_2, there exists at least one element v_2 in V_2 such that $v_2 \in e_2$. Therefore, $\phi \notin V$, $\emptyset \notin E$ and for each element e $\in E$, there exists at least one element v $\in V$ such that v $\in$ e. Hence, all the conditions of the definition of cluster hypergraph satisfy in $G \times H$,; therefore, $G \times H$ is a cluster hypergraph. Hence the result.

Definition 3.4 *The union of two cluster hypergraphs* $G = (V_1, E_1)$ *and* $H = (V_2, E_2)$ *is denoted by* $G \cup H = (V, E)$, *where* $V = \{v : v \in V_1 \cup V_2 \text{ such that } v \notin u \text{ where } u \in V_1 \cup V_2\}$ *and* $E = \{(u, v) : u, v \in V \text{ and } (u, v) \in E_1 \cup E_2\}$.

Example 3.5 Consider two cluster hypergraphs (Figs. 6, 7) $G = (V_1, E_1)$ and $H = (V_2, E_2)$,

Fig. 8 Cluster hypergraph $G \cup H$

w h e r e $V_1 = \{\{a\}, \{b\}, C = \{a,b\}, \{d\}, \{e\}\}$, $E_1 = \{\{\{a\}, C\}, \{\{b\}, C\}, \{\{d\}, C\}, \{\{d\}, \{e\}\}\}$ a n d $V_2 = \{\{a\}, \{g\}, \{h\}, I = \{g,h\}\}, E_2 = \{\{\{a\}, I\}, \{\{g\}, I\}, \{\{h\}, I\}\}$.

The vertex set of $G \cup H$ (Fig. 8) is $\{\{a\}, \{b\}, C = \{a,b\}, \{d\}, \{e\}, \{g\}, \{h\}, I = \{g,h\}\}$, and edge set is $\{\{\{a\}, C\}, \{\{b\}, C\}, \{\{d\}, C\}, \{\{d\}, \{e\}\}, \{\{a\}, I\}, \{\{g\}, I\}, \{\{h\}, I\}\}$.

Proposition 3.6 *If $G = (V_1, E_1)$ and $H = (V_2, E_2)$ be two cluster hypergraphs, then $G \cup H$ is a cluster hypergraph.*

Proof Since $G = (V_1, E_1)$ and $H = (V_2, E_2)$ are two cluster hypergraphs with $G \cup H = (V, E)$. Then, $\emptyset \notin E_1, E_2$ and for each element e_1 in E_1, there exists at least one element v_1 in V_1 such that $v_1 \in e_1$ and for each element e_2 in E_2, there exists at least one element v_2 in V_2 such that $v_2 \in e_2$. Then, $\emptyset \notin E$ ($\because \emptyset \notin E_1, E_2$) and for each element e in E, there exists at least one element v in V such that $v \in e$. Therefore, $G \cup H$ is a cluster hypergraph. Hence the result.

Definition 3.7 *The intersection of two cluster hypergraphs $G = (V_1, E_1)$ and $H = (V_2, E_2)$ is $G \cap H = (V, E)$, where $V = V_1 \cap V_2$ and $E = \{(u, v) : u, v \in V$ and $(u, v) \in E_1 \cap E_2\}$.*

Example 3.8 Consider two cluster hypergraphs (Figs. 9, 10) $G = (V_1, E_1)$ and $H = (V_2, E_2)$, where $V_1 = \{\{a\}, \{b\}, C = \{a, b\}, \{d\}, \{e\}\}$, $E_1 = \{\{\{a\}, C\}, \{\{b\}, C\}, \{\{d\}, C\}\}$, $\{\{d\}, \{e\}\}\}$ and $V_2 = \{\{a\}, \{b\}, C = \{a, b\}, \{d\}, \{f\}, \{g\}, H = \{f,g\}\}$, $E_2 = \{\{a\}, C\}, \{\{b\}, C\}, \{\{d\}, C\}, \{\{f\}, H\}\}$, $\{\{g\}, H\}, \{C,H\}\}$.

The vertex set of $G \cap H$ (Fig. 11) is $V = \{\{a\}, \{b\}, C = \{a, b\}, d\}$, and edge set is $E = \{\{\{a\}, C\}, \{\{b\}, C\}, \{\{d\}, C\}\}\}$.
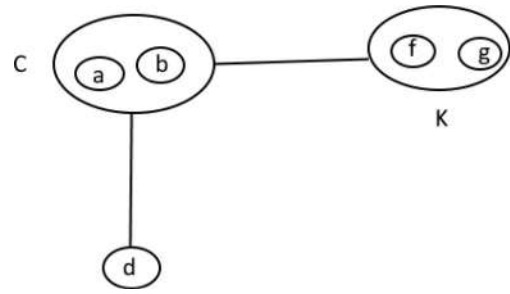

Fig. 10 Cluster hypergraph H

Proposition 3.9 *If $G = (V_1, E_1)$ and $H = (V_2, E_2)$ be two cluster hypergraphs, then $G \cap H$ is also a cluster hypergraph if $V_1 \cap V_2$, $E_1 \cap E_2 \neq \emptyset$.*

Proof Since $G = (V_1, E_1)$ and $H = (V_2, E_2)$ are two cluster hypergraphs with $G \cap H = (V, E)$. Then, $\emptyset \notin E_1, E_2$ and for each element e_1 in E_1, there exists at least one element v_1 in V_1 such that $v_1 \in e_1$ and for each element e_2 in E_2, there exists at least one element v_2 in V_2 such that $v_2 \in e_2$. Then, $\emptyset \notin V, E$ and for each element e in E, there exists at least o element v in V such that $v \in e$. Therefore, $G \cap H$ is a cluster hypergraph. Hence the result.

5 Matrix representation of cluster hypergraphs

Matrix representation is the easiest and convenient way to represent graphs. We represent cluster hypergraph as four types of matrices.

- Adjacency matrix
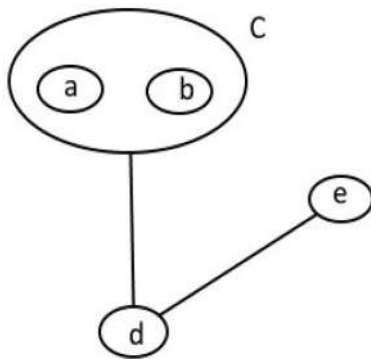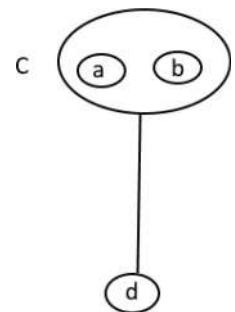- Incidence matrix
- Location matrix
- Circuit matrix

Fig. 9 Cluster hypergraph G

Fig. 11 Cluster hypergraph $G \cap H$

Definition 4.1 *The adjacency matrix of a cluster hypergraph with n nodes (total number of simple and cluster nodes) is an $n \times n$ symmetric matrix $X = \left[x_{ij}\right]$, where*

$$x_{ij} = (l, m) \text{ in which } l = \begin{cases} 1, & \text{if two nodes are adjacent} \\ 0, & \text{if two nodes are not adjacent} \end{cases}$$

and

If one cluster node is adjacent to another maximal node, m is related to the indirect path from simple nodes of one cluster node to the simple nodes of another cluster node or to the maximal node itself).

$$m = \begin{cases} & \text{if one simple node inside a cluster is indirectly} \\ 1, & \text{connected to another simple node inside other clusters} \\ & \text{or to the maximal node itself} \\ 0, & \text{if two nodes are not indirectly connected} \end{cases}$$

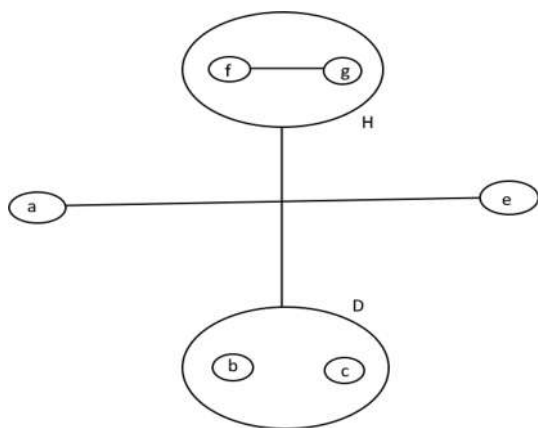If two clusters are connected, then the nodes among the cluster are indirectly connected.

Fig. 12 Cluster hypergraph

Fig. 13 Cluster hypergraph with virtual representation (right)
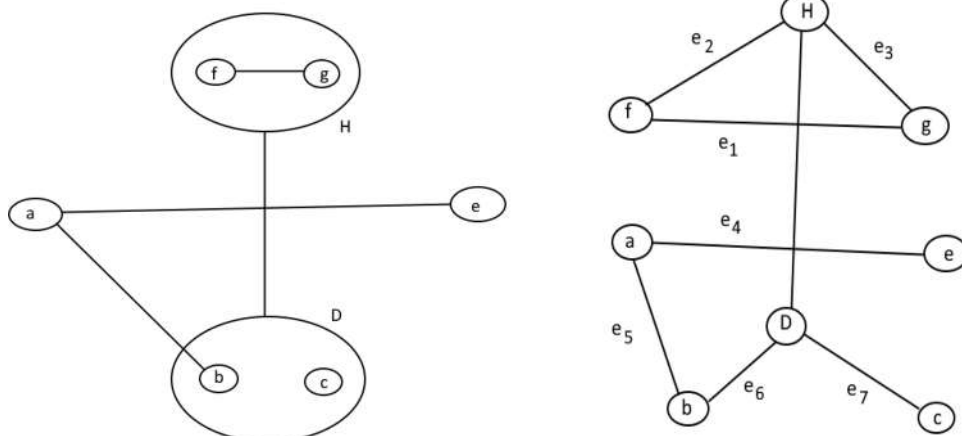
Example 4.2 Consider a cluster hypergraph (Fig. 12) $G = (V_X, E)$, where $V_X = \{\{a\}, \{b\}, \{c\}, D = \{b, c\}, \{e\}, \{f\}, \{g\}, H = \{f,g\}\}$ and $E = \{\{\{a\}, \{e\}\}, \{\{b\}, D\}, \{\{c\}, D\}, \{\{f\}, \{g\}\}, \{\{f\}, H\}, \{\{g\}, H\}, \{D, H\}\}$.

The corresponding adjacency matrix of Example 4.2 is given below:

	a	b	c	D	e	f	g	H
a	(0,0)	(0,0)	(0,0)	(0,0)	(1,0)	(0,0)	(0,0)	(0,0)
b	(0,0)	(0,0)	(0,0)	(1,0)	(0,0)	(0,1)	(0,1)	(0,1)
c	(0,0)	(0,0)	(0,0)	(1,0)	(0,0)	(0,1)	(0,1)	(0,1)
D	(0,0)	(1,0)	(1,0)	(0,0)	(0,0)	(0,1)	(0,1)	(1,0)
e	(1,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
f	(0,0)	(0,1)	(0,1)	(0,1)	(0,0)	(0,0)	(1,0)	(1,0)
g	(0,0)	(0,1)	(0,1)	(0,1)	(0,0)	(1,0)	(0,0)	(1,0)
H	(0,0)	(0,1)	(0,1)	(1,0)	(0,0)	(1,0)	(1,0)	(0,0)

Observations 4.3

- The entries along the principal diagonal are all (0,0) if the graph has no self-loop.
- If the entry value is (1,0), then there is a direct edge between two nodes.
- If the entry values are (0,1), then the nodes are indirectly connected through some cluster to maximal node adjacency.

Note 4.4 For complete CCCH, the entries for the adjacency matrix will be either (1,0) or (0,1) for all non-diagonal entries.

Definition 4.5 *Let us consider a cluster hypergraph with m nodes and n edges. The incidence matrix $A = \left[a_{ij}\right]$ is an $m \times n$ matrix define as,*

$$a_{ij} = \begin{cases} 1, & \text{if } j\text{-th edge is incident on } i\text{-th node} \\ 0, & \text{otherwise} \end{cases}$$

Example 4.6 Consider a cluster hypergraph (Fig. 13) $G = (V_X, E)$, where $V_X = \{\{a\}, \{b\}, \{c\}, D = \{b, c\}, \{e\}, \{f\}, \{g\}, H = \{f, g\}\}$. and $E = \{\{\{a\}, \{e\}\}, \{\{a\}, \{b\}\{\{b\}, D\}, \{\{c\}, D\}, \{\{f\}, \{g\}\}, \{\{f\}, H\}, \{\{g\}, H\}, \{D, H\}\}$.

The corresponding incidence matrix of Example 4.6 is given below:

	e_1	e_2	e_3	e_4	e_5	e_6	e_7
a	0	0	0	1	1	0	0
b	0	0	0	0	1	1	0
c	0	0	0	0	0	0	1
D	0	0	0	0	0	1	1
e	0	0	0	1	0	0	0
f	1	1	0	0	0	0	0
g	1	0	1	0	0	0	0
H	0	1	1	0	0	0	0

Observations 4.7

- Since every edge is incidence on precisely two vertices, each column of A has exactly two 1's.
- The number of 1's is in a row equal to the degree of the corresponding vertex.

Note 4.8 For (m, n) uniform cluster hypergraph, the n entries of a column of the incidence matrix for will be 1.

Definition 4.9 *Let us consider a cluster hypergraph with m simple nodes and n cluster nodes. The location matrix defined as an $m \times n$ matrix $A = [a_{ij}]$ such that*

$$a_{ij} = \begin{cases} 1, & \text{if } i\text{-th simple node is located in } j\text{-th cluster node} \\ 0, & \text{otherwise} \end{cases}$$

Example 4.10 In Fig. 12, the corresponding location matrix is given below:

	D	H
a	0	0
b	1	0
c	1	0
e	0	0
f	0	1
g	0	1

Observations 4.11

- Number of rows is the total number of simple nodes.
- Number of columns is the total number of cluster nodes.
- The number of 1's in each column is the number of simple nodes, contained in the corresponding cluster nodes.

Note 4.12 For (m, n) uniform cluster hypergraph, m-entries for each column in the location matrix will be 1.

Definition 4.13 *A chain is defined as the finite alternating sequence of nodes and edges, beginning and ending with nodes such that each edge is incident with the nodes preceding and following it. In a virtual representation of cluster hypergraph G, a chain is defined to be a sequence $(V_1, E_1, V_2, E_2, \ldots, E_q, V_{q+1})$ such that*

- (i) $V_1, V_2, \ldots, V_q$ *are vertices of G, which may be repeated.*
- (ii) $E_1, E_2, \ldots, E_q$ *are all distinct edges.*
- (iii) *each edge is incident with the vertices preceding and following it.*
- (iv) $V_1 = V_{q+1}$.

Then, this chain is called a circuit.

Definition 4.14 *Let us consider a cluster hypergraph G with the number of different circuits in a virtual representation of G be m, and the number of edges in G is n. Then, the circuit matrix $B = [b_{ij}]$ of G is a $m \times n$ matrix define as follows:*

$$b_{ij} = \begin{cases} 1, & \text{if } i\text{-th circuit includes } j\text{-th edge} \\ 0, & \text{otherwise} \end{cases}$$

Example 4.15 For this above cluster hypergraph (Fig. 14), different circuits are $\{e_1, e_2, e_3\}, \{e_4, e_5, e_6\}, \{e_4, e_7, e_8\}, \{e_1, e_2, e_4, e_5, e_6, e_3\}, \{e_1, e_2, e_4, e_7, e_8, e_3\}$.

These circuits are denoted by A, B, C, D and E, respectively.

The corresponding matrix of Fig. 10 is given below:

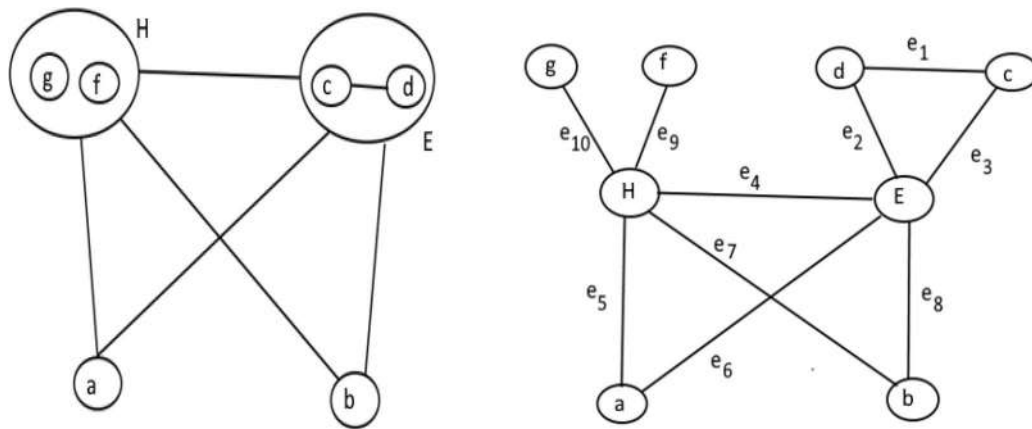|	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
A	1	1	1	0	0	0	0	0	0	0
B	0	0	0	1	1	1	0	0	0	0
C	0	0	0	1	0	0	1	1	0	0
D	1	1	1	1	1	1	0	0	0	0
E	1	1	1	1	0	0	1	1	0	0

Fig. 14 Cluster hypergraph with virtual representation (right)

Observations 4.16

- The number of 1's in a row is equal to the number of edges in the corresponding circuit.
- A column of all zeros corresponds to a non-circuit edge.

6 Isomorphism on cluster hypergraph

Isomorphism is an important property of cluster hypergraphs. We define isomorphism on cluster hypergraphs.

Definition 5.1 *The two cluster hypergraphs $G = (V_X, E)$ and $G' = (V'_X, E')$ are said to be Type-I isomorphic if there exists a bijective mapping $f : V_X \to V'_X$ such that there exists an edge $(a, b) \in E$ if and only if $(f(a), f(b)) \in E'$ for all $a, b \in V_X$.*

Observations 5.2

- There are one–one correspondences between nodes and edges among two cluster hypergraphs.
- There should be the same number of nodes between two cluster hypergraphs.
- The degree of a node under the isomorphic image should be the same.
- The effective degree of a node under the isomorphic image should be the same.
- The matrices, that is, adjacency matrices, incidence matrixes, location matrices and circuit matrices under isomorphic image should be the same.
- There should be the same number of maximal nodes under isomorphism.

Definition 5.3 *Two cluster hypergraphs $G = (V_X, E)$ and $G' = (V'_X, E')$ are said to be Type-II isomorphic if there exists a bijective mapping $f : V_X \to V'_X$ such that for all*

maximal nodes $a, b \in V_X$ there exists edge $(a, b) \in E$ if and only if there exist maximal nodes $f(a), f(b) \in V'_X$ such that $(f(a), f(b)) \in E'$.

Observation 5.4

- Type-I isomorphism implies Type-II isomorphism between the two cluster hypergraphs but not conversely.
- Adjacency matrices of Type-I and Type-II isomorphism between the cluster hypergraphs may not be the same.

7 Area of applications

A. This study of cluster hypergraphs introduces a new area of graph theory and networks. The properties, matrix representations and isomorphism of cluster hypergraphs will be main backgrounds of the future studies in the area.

There are a number of applications of cluster hypergraphs in searching, centrality measures, etc. in social networks. There is a big problem to search for a particular type of small amount of data or a small network from a big amount of data or a big network. It is possible to search such small networks by using cluster hypergraphs. First represent the virtual presentation of cluster hypergraphs to the desired particular types of network and fit all possible type-2 isomorphism mappings to all such small networks of a big network. This can be done by adjacency matrices or degree sequence properties of cluster hypergraphs. Thus, isomorphism properties of cluster hypergraphs can be applied for filtering the required data as advanced searching. This study also analyzes the node's importance in the network, especially the importance of a cluster node in a social network where a group of nodes is taken as a cluster node. The

Fig. 15 Research network
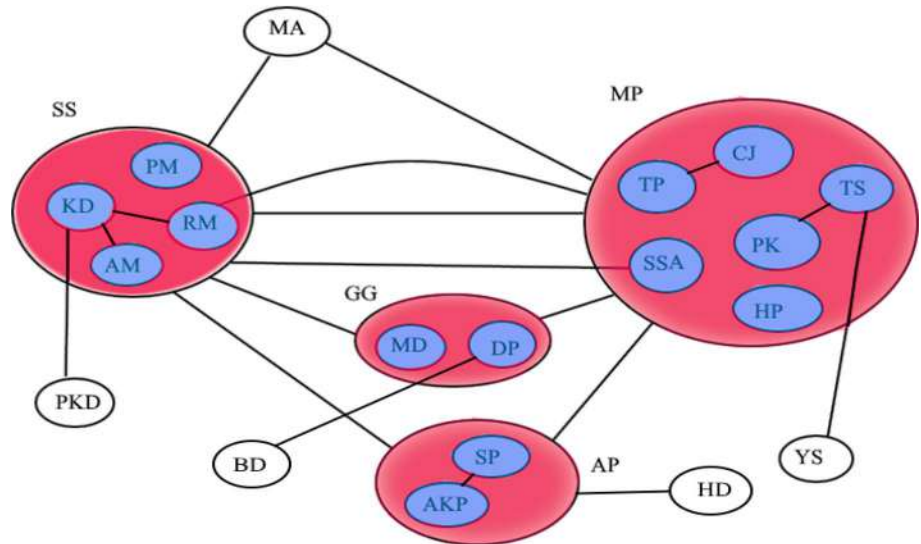
degree of cluster indicates the direct impact of the cluster in the network.

We considered a small research network (Fig. 15) among researchers and supervisors in an area of a specific subject. These persons are taken as cluster nodes or simple nodes. Here, a supervisor is considered as a cluster node, and a researcher is taken as a simple node. The edges are considered if there is a paper between them. Also, a supervisor always has an edge over his/her researchers, which belongs to the same cluster. But the problem is to find "who is influential in this network?". Hence, we have to find some measures like degree centrality, effective degree centrality of nodes in the network. Here, degree centrality indicates

the direct influence of a node in the network, and effective degree centrality indicates the effective influence of node considering the belongingness of node in a cluster in a network (Fig. 16).

All the necessary calculations are done (Table 1) and found the direct influence and effective direct influence of all nodes in the network using the definition of degree and effective degree, respectively. In this study of the applications, we compared (Fig. 17) the results, and it is observed that

- The direct influence and effective direct influence of maximal clusters are the same.

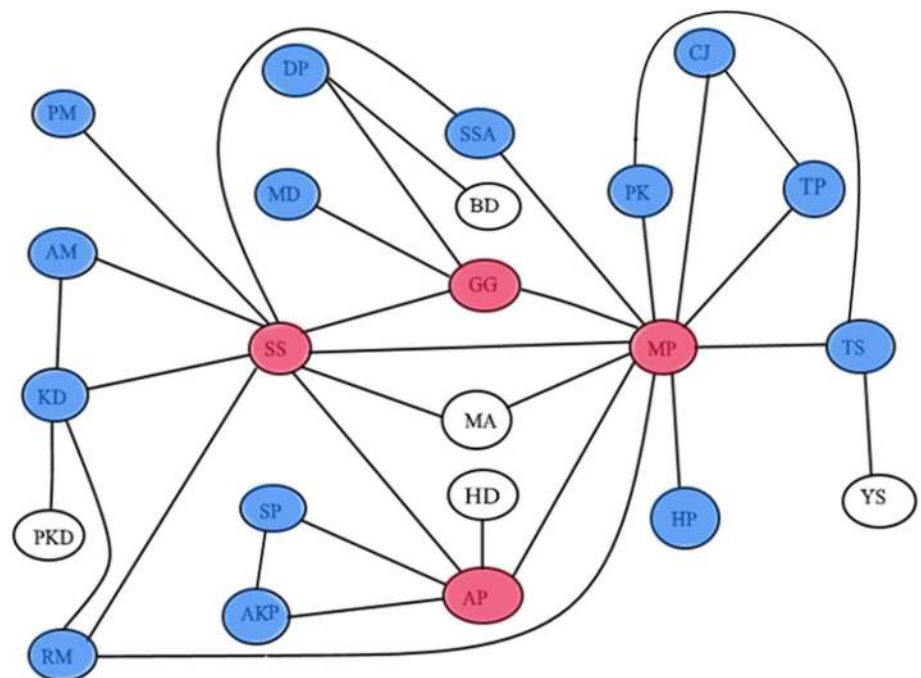**Fig. 16** Virtual representation of Fig. 19

Table 1 Result of direct and effective direct influence of researchers

Node	Direct influence	Effective direct influence
PM	1	10
AM	2	11
KD	4	13
RM	3	12
SS	**9**	–
DP	2	6
MD	1	5
GG	**4**	–
SP	2	7
AKP	2	7
AP	**5**	–
PK	2	13
HP	1	12
CJ	2	13
TP	2	13
TS	3	14
SSA	2	13
MP	**11**	–
PKD	1	1
MA	2	2
YS	1	1
HD	1	1
BD	1	1

The bold values indicate that the nodes are cluster nodes

- The effective direct influence of non-maximal nodes may be greater than direct influence.
- The non-maximal nodes within a cluster are also influential when these nodes belong to a bigger cluster with more nodes.
- The cluster nodes and nodes in the cluster are much more influential than the simple maximal node that are not in a cluster.
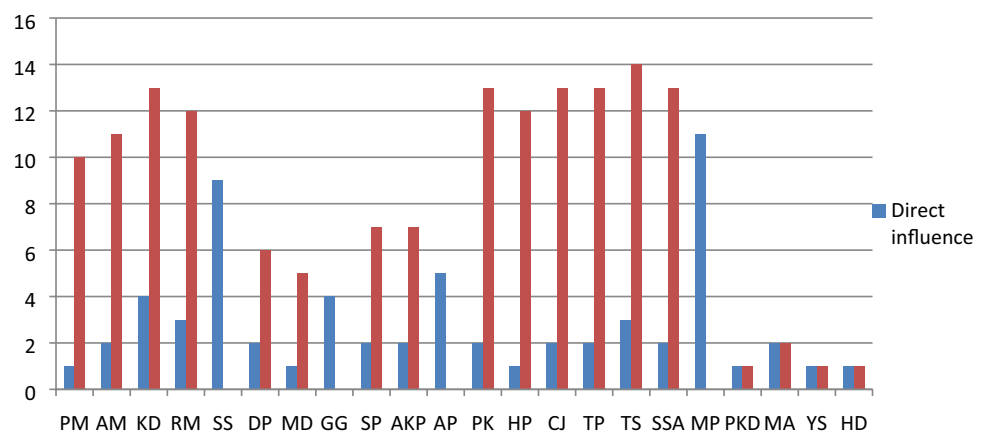
Figure 17 shows that the effective degree brings additional information that is in clusters, and hence effective degree is higher in each case. But, the relative importance of node (ranking) and its normalized inform us that effective degree shows a different perspective. The relative importance of the node changes in the network due to the influence of clusters. This is shown in Fig. 18a and b for the above example problem.

B. Another application

We illustrate the application of the above-described approach to a real-life problem. We study the publications for COVID 19 and how authors, departments (consisting of different authors), and papers (authors collaborated on papers) are interrelated. For this purpose, a small data have been collected and have been shown in Table 2. Further, we take the perspective that an author is a node, a paper is a cluster of nodes, and a department is a hyperedge. Thus, an influential department is one that works on different influential papers. Further, it is often observed that nodes associated with a persuasive paper to gain from the association. We provide a means to compute the impact of such an association with a paper/s on individual nodes.

In Tables 2 and 3, paper-author and department-author details of ten collected paper are displayed.

Here, each individual author is taken as simple node. Each paper may be written by more than one author. The collection of authors per paper is considered as cluster. Now, same department authors are connected by hyperedges. The traditional graph where two authors are connected if they share a common paper, is shown in Fig. 19. The cluster hypergraph of assumed data is shown in Fig. 20. Here, small black circles are simple nodes. Big circles are clusters and blue lines indicate the hyperedges.

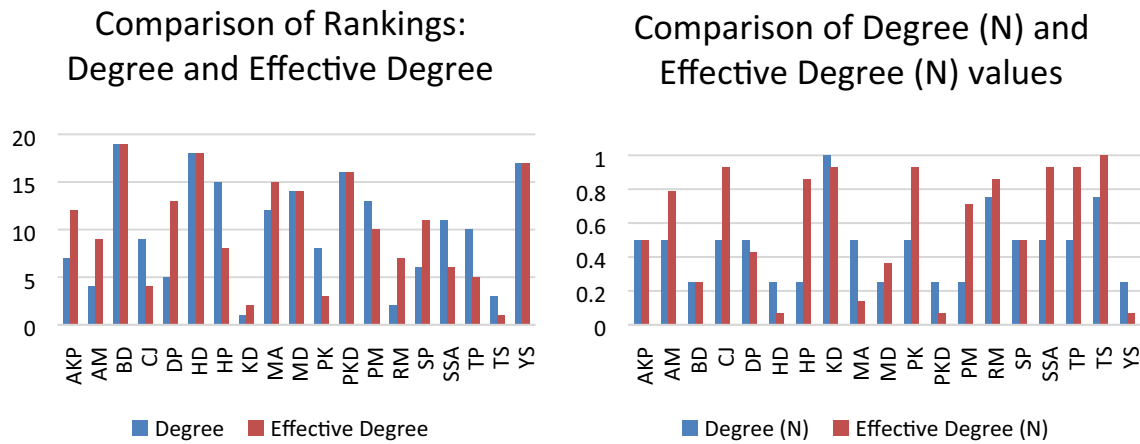Now, effective degrees of simple nodes of the cluster hypergraphs are calculated in Table 4.

Fig. 17 Comparison of the result of Table 1

Comparison of Rankings: Degree and Effective Degree

Comparison of Degree (N) and Effective Degree (N) values

Fig. 18 **a** Left and **b** right

Table 2 Paper-author representation

	S.S	M.A	M.P	R.A.B	H.R	Q.X	Y.X	W.A.D	G.G	T.P	S.M	S.P	T.K.P
Paper-1	1	1	1	0	0	0	0	0	0	0	0	0	0
Paper-2	1	0	1	1	1	0	0	0	0	0	0	0	0
Paper-3	0	0	0	0	0	1	1	0	0	0	0	0	0
Paper-4	0	0	1	0	1	0	0	0	0	0	0	0	0
Paper-5	0	1	0	0	0	0	0	1	0	0	0	0	0
Paper-6	0	0	1	0	0	0	0	0	1	0	0	0	0
Paper-7	0	0	0	1	1	0	0	0	0	0	0	0	0
Paper-8	0	0	0	0	0	1	0	0	1	0	0	1	0
Paper-9	0	0	1	0	0	0	0	0	0	0	1	0	1
Paper-10	1	0	1	0	0	0	0	0	0	1	1	0	0

Table 3 Department and author representation

	S.S	M.A	M.P	R.A.B	H.R	Q.X	Y.X	W.A.D	G.G	T.P	S.M	S.P	T.K.P
Dept-1	1	0	0	1	1	0	0	0	0	0	1	0	0
Dept-2	0	1	1	0	0	0	0	0	1	1	0	1	1
Dept-3	0	0	0	0	0	1	0	0	0	0	0	0	0
Dept-4	0	0	0	0	0	0	1	0	0	0	0	0	0
Dept-5	0	0	0	0	0	0	0	1	0	0	0	0	0

Figure 21a and b show the relative ranking using the currently accepted degree centrality measure and compare it to the proposed effective degree.

We observe that the normalized effective degree values are bigger than those computed using the normalized degree. This, we believe, is due to the fact that since all values (effective degree) increase by some amount (from the base—the degree values), and the largest value has also increased similarly. This has the effect of increasing every

value. If the highest effective degree value were significantly higher, we would have seen a reduction in the normalized value (or vice versa). We also see a change in the ranking of influential nodes. This is a significant difference. We find that when additional information is considered, the influence of the node increases. For example, an author that is associated with an influential paper would appear more influential that other authors who are not connected to such clusters.
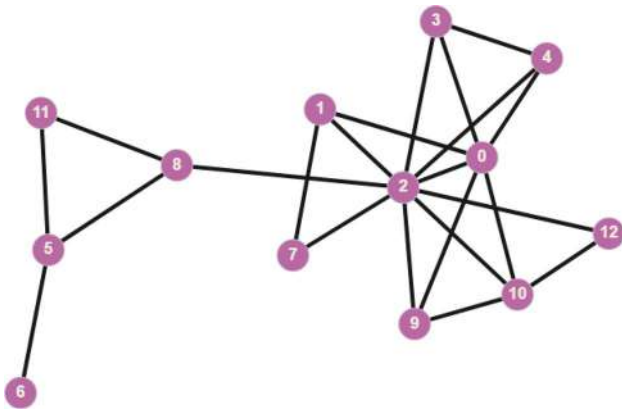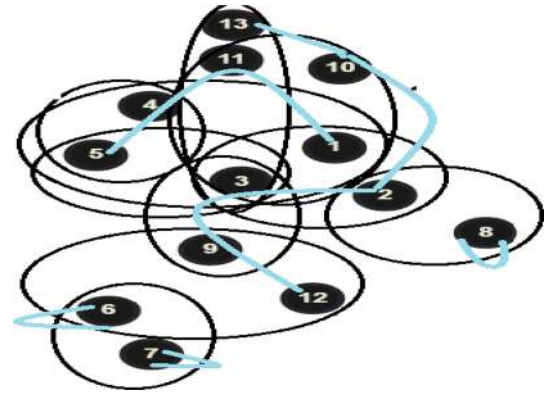
Fig. 19 Network on author-author representation

Fig. 20 Cluster hypergraph of considered data

Table 4 Comparison of table of degree and effective degree

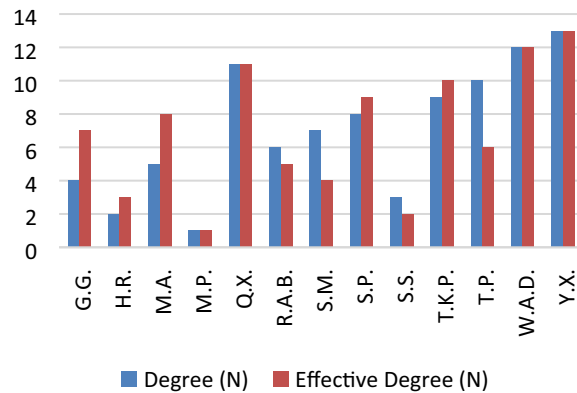| Node	Names	Degree	Effective degree
1	S.S	7	$= d(S.S) + \left(d(P_1) + d(P_2) + d(P_{10})\right)/3 = 7 + (3 + 4 + 4)/3 \approx 10.67$
2	M.A	10	$= d(M.P.) + \left(d(P_2) + d(P_{10}) + d(P_1) + d(P_9) + d(P_4) + d(P_6)\right)/6 = 10 + (4 + 4 + 3 + 3 + 2 + 2)/6 = 13$
3	M.P	6	$= d(M.A.) + \left(d(P_1) + d(P_5)\right)/2 = 6 + (3 + 2)/2 = 8.5$
4	R.A.B	7	$= d(H.R.) + \left(d(P_2) + d(P_4) + d(P_7)\right)/3 = 7 + (4 + 2 + 2)/3 \approx 9.67$
5	H.R	6	$= d(R.A.B.) + \left(d(P_2) + d(P_7)\right)/2 = 6 + (4 + 2)/2 = 9$
6	Q.X	6	$= d(G.G.) + \left(d(P_6) + d(P_8)\right)/ = 6 + (2 + 3)/2 = 8.5$
7	Y.X	4	$= d(Q.X.) + \left(d(P_8) + d(P_3)\right)/2 = 4 + (3 + 2)/2 = 6.5$
8	W.A.D	3	$= d(Y.X.) + d(P_3) = 3 + 2 = 5$
9	G.G	5	$= d(S.P.) + d(P_8) = 5 + 3 = 8$
10	T.P	5	$= d(T.K.P) + d(P_9) = 5 + 3 = 8$
11	S.M	5	$= d(T.P.) + d(P_{10}) = 5 + 4 = 9$
12	S.P	3	$= d(W.A.D) + d(P_5) = 3 + 2 = 5$
13	T.K.P	6	$= d(S.M.) + \left(d(P_{10}) + d(P_9)\right)/ = 6 + (4 + 3)/2 = 9.5$
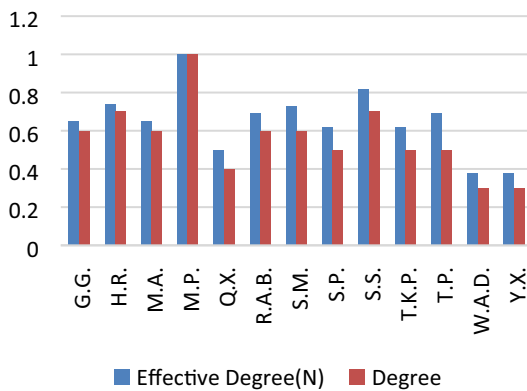
Fig. 21 **a** Left and **b** right

8 Conclusion

In this paper, some properties of cluster hypergraphs have been developed. The effective degree is defined and is taken as a replacement of degree as the effective degree is more convenient to represent as the effects of other persons who form a cluster/groups are considered separately along with direct impact. A comparison with the currently available approach has been mentioned for describing the limitations of these approaches, and this also serves as the motivation for developing the proposed method. Two realistic applications have been shown. For the second application, we collect data on COVID-19-related publications. In both applications, we find that our approach not only captures more information but also changes the influence ranking of the nodes. Computationally, our method has similar complexity as that of existing methods. Thus, the approach in the paper improves upon currently available approaches to addressing the problem of determining the influence of a node in cluster hypergraphs.

References

Berge C (1961) Fárbung von Graphen deren sämtliche bzw. ungerade Krense starr sind Math. Nat. Reihe, Wiss. Z. Martin Luther Univ, Halle Wittenberg, pp 114–1157

Berge C (1973) Graphs and hypergraphs. Elsevier, New York

Berge C (1989) Hypergraphs. North-Holland, Amsterdam

Bergen R, Fallat S, Gorr A, Ihringer F, Meagher K, Purdy A, Yang B, Yuad G (2018) Infection in hypergraphs. Discrete Appl Math 237:43–56

Bollabosand B, Scott AD (2000) Judicious partitions of 3-uniform hypergraphs. Eur J Combin 21:289–300

Buroscha G, Ceccherini PV (1996) characterization of cube-hypergraphs. Discrete Math 152:55–68

Das K, Samanta S, Pal M (2018) Study on centrality measures in social networks: a survey. Social Netw Anal Min 8(13):1–11

Djokovic DZ (1973) Distance preserving subgraphs of hypercubes. J Combin Theory Ser B 14:263–267

Frankl P, Katona GY (2000) Extremal k-edge-hamiltonian hypergraphs. Electron Notes Discrete Math 5:130–133

Harary F (1972) Graph theory, 3rd edn. Addison-Wesley, Reading, MA

Katona GY, Kierstead HA (1999) Hamiltonian chains in hypergraphs. J Graph Theory 30(3):205–212

Sabidussi G (1959) The composition of graphs. Duke Math J 26:693–696

Samanta S, Muhiudin G, Alanazi AM, Das K (2020a) A mathematical approach on representations of competitions: Competition cluster hypergraphs. Math Probl Eng. https://doi.org/10.1155/2020/2517415

Samanta S, Lee JG, Naseem U, Khan SK, Das K (2020b) Concepts on coloring of cluster hypergraphs. Math Probl Eng. https://doi.org/10.1155/2020/3705156

Sonntag M, Teichert HM (2000) Sum numbers of hypertrees. Discrete Math 214:285–290

Sonntag M, Teichert HM (2004) Competition hypergraphs. Discrete Appl Math 143:324–329

Tyshkevich RI, Zverovich VE (1996) Line hypergraphs. Discrete Math 161:265–283

Yan E, Ding Y (2009) Applying centrality measures to impact analysis: a co-authorship network analysis. J Am Soc Inform Sci Technol 60(10):2107–2118