Analysis on Secure Big Data Capturing and Transmission in Cloud Storage through Novel Deg-Greedy Algorithm

K. Thiagarajan¹*, R. Satheesh Babu² and K. Saranya³

¹Department of Mathematics, PSNA College of Engineering and Technology, Dindigul - 624622, Tamil Nadu, India; vidhyamannan@yahoo.com
²Department of Mathematics, Bharathiyar University, Coimbatore - 641046, Tamil Nadu, India; r.satheeshbabu@yahoo.co.in
³Department of IT, PSNA College of Engineering and Technology, Dindigul - 624622, Tamil Nadu, India; sharancse16@gmail.com

Abstract

This paper presents Novel Deg-Greedy algorithm based on Dominating Nodes (DN) and cooperative jamming scheme for secure transmission of big data in wireless cloud networks. Each node has a transmission radius in a multi-hop wireless cloud network that can transmit a message to all of its neighbours positioned in the radius. A Dominating Node (DN) based virtual backbone plays an important role in wireless networks for efficient routing and broadcasting. In the proposed algorithm, the same message is sent by a source node sends to all the leading nodes in the network. A dominating node is selected based on degree based evaluation of the given cloud network. The dominating nodes broadcast the requesting big data message to its path, if it finds the requested information along the path it sends the content to the destination node. During this path of flow there may presence of eavesdroppers and it is prevented by emitting jamming signal by adopting cooperative jamming scheme. The novel algorithm is analyzed by splitting the network as Strongly Perfect Network (SPN), Strongly Weakly Perfect Network (SWPN) and Weekly Strongly Perfect Network (WSPN).

Keywords: Big Data, Cooperative Jamming and Broadcasting, Dominating Node (DN), Wireless Cloud Network

1. Introduction

Wireless networks consist of nodes that without any static infrastructure are able to communicate with each other through the wireless link.[1] It has the capability to communicate directly with another node inside the surroundings of all nodes. The nodes destined for other nodes are able to forward the packets. Among the networks that are used in disaster rescues, wireless conferences in the hall, battlefields, monitoring objects in a possibly remote or dangerous environment, wireless internet, etc. there are examples such as ad hoc, packet ratio, local area and sensor networks.

1.1 Security of Big Data in Cloud Storage

The correctness of the data is an important factor to ensure in almost any data and computation related activities. It serves as an important part of data security and privacy than on quality of service. With the rapid interest on cloud computing and the increasing needs in big data analytics, verification of correctness of data is becoming essentially important, particularly on outsourced data.[2]

Big data is the massive amounts of data which implies performing computation and database operations remotely from the owner’s data enterprise.[3] Since in big data a key value proposition is access to data from various different domains, security and privacy will play a very important role in big data research and technology. Big data is attracting more and more interests from numerous industries. A few examples are oil and gas mining, scientific research (biology, chemistry and physics), online social networks (Twitter, Facebook), multimedia

* Author for correspondence
data and business transactions. With mountains of data collected from increasingly efficient data collecting devices as well as stored on fast-growing storage hardware, people are keen to find solutions to store and process the data more efficiently and to discover more values from the mass at the same time. When referring to big data research problems, people often bring the 4 v’s - volume, velocity, variety and veracity. This poses various brand-new challenges to computer scientists nowadays.

The recently emerged cloud computing, known to be the latest development and the most promising technological backbone for solving big data problems in data center technology especially in parallel distributed systems and service computing is widely considered. In many practical applications involving fast-updating dynamic data, cloud also offers elasticity and scalability which can result in further saving of costs. In present situation, large amounts of business data of numerous big companies such as Amazon AWS, IBM SmartCloud and Microsoft Azure have been moved into and managed by clouds. Despite those interesting advantages of cloud, there are still strong concerns regarding service qualities, particularly on data security. In fact, data security has been frequently raised as one of the top concerns in using cloud.

1.2 Broadcasting Scheme
In the broadcasting scenario, all the nodes in the network receive the same message from the source node. In one-to-all model, transmission by each node can reach all nodes that are within radius distance from it, while in the one-to-one model, each transmission is directed towards only one neighbor (using, for instance, directional antennas or separate frequencies for each node). In this paper we shall use one-to-all model and broadcasting is been studied mainly on that model. Broadcasting is often termed as flooding. This term is used to denote to the broadcasting scheme wherein the identical message is sent through all nodes receiving it. Conventionally Flooding has been used for broadcasting. In broadcasting they are sending an alarm signal or paging a particular host. Additionally, for route discovery in a source-initiated on-demand routing, flooding/broadcasting is applied. Similarly, broadcasting can be used in the context of an effective location-aware routing algorithm like this: the source S might start the destination search process through broadcasting a short message containing the location of S, id of destination D and some control bits. D applies any location-based routing algorithm, when the destination search message reaches successfully D, and reports back to S with a short message containing the location. The source S can then apply again the same routing algorithm to send the full message toward D. We argued that flooding can be a viable candidate for multicast and routing protocols in very dynamic ad hoc networks.

Flooding has been replaced by a method where each Cluster Head and gateway (or border, as renamed in this paper) node in a clustered wireless network forwards the message exactly once. The maintenance of cluster structure, however, requires excessive communication overhead due to “chain effect” caused by mobility. The concept of dominating nodes is proposed. We use this method in this paper in which the Dominating Nodes (DN) broadcast to all the nodes within its two-hop neighbor.

Unplanned networks are modeled in the best way by the unit graphs that are constructed in the subsequent way: The nodes of A and B are neighbors in the network (hence joined by an edge) if the Euclidean distance between them is at most R, that R is the transmission radius equal for all nodes. Random unit graphs are used in our experiments.

Applying the broadcasting algorithms in a study, there has been an investigation to achieve a high ratio of nodes to receive the message of a reduced amount of rebroadcasting. All of the methods from study have a parameter that its best value might depend upon network conditions, that is a global information. Furthermore, their methods are not reliable. Assuming, there are no message collision, if the method guarantees message delivery to all nodes connected to the source, a broadcasting method can be referred as reliable. It means there will be the guaranty of delivery of each message to all intended neighboring nodes without collisions at any of these nodes, if an ideal medium access scheme. Such scheme is time division multiple accesses wherein each node is allocated a time slot different from time slots allocated to any of its two hop neighbors. Reducing the broadcast redundancy the packet loss can be reduced because of contention or collision and potentially improve the reliability of broadcasting of our reliable proposed algorithm.

In this paper, the main contribution is reliable dominating-node-based broadcasting algorithms in cloud structure of capturing and transmitting large data,
improved by highest degree priority in selecting the dominating nodes, cooperative jamming scheme and retransmission after negative acknowledgement scheme. If GPS or another location method is available to all the nodes in the network, dominating node maintenance is incorporated into location updates between neighbouring nodes. Through communication with a satellite network, Global Positioning System (GPS) provides location information (latitude, longitude and possibly the height) to host in a wireless network. Instead, through exchanging signal strength information with their neighbours nodes may measure signal strengths of incoming messages and determine the location of its neighbors. The notion in our algorithms is proposed according to the fact that each node needs only the location of its neighbors or alternatively, the list of neighbors of each neighbor. They are also degree independent in two senses: Firstly, there is no parameter in the algorithms set based on the network average degree and secondly the performance of proposed algorithms appears to be relatively stable with respect to degree.

1.3 Cooperative Jamming in Preventing Eavesdroppers

Cooperative communication helps in exploiting spatial diversity to enhance the quality of wireless links. Security can be improved by cooperative networks by having the information content minimum to the eavesdropper nodes of the expected destination and having maximum to the relay nodes of the expected destination. The recently proposed cooperative network technique is cooperative jamming to improve physical layer security in the presence of eavesdroppers. The Figure 1 illustrates this technique.

**Figure 1.** Co-operative jamming of location unknown.

In wireless communication, occurrence of interference is considered redundant. This fetches the work of cooperative jamming for flexible and efficient wireless network technique to confuse the eavesdroppers and making the source message uncertain by generating friendly jamming signal to the eavesdroppers. In this, if the data has to be transmitted from source S to destination D, jamming signal will be emitted by the relay nodes to have the secure communication and to prevent the eavesdroppers of location unknown from capturing the data. In our novel algorithm, cooperative jamming scheme is considered.

### 2. Novel Deg-Greedy Algorithm for Secure Transmission of Big Data

To capture and transmit the big data securely from the source node to the destination node using the Novel deg-Greedy algorithm from a given cloud network G = (V, E). All nodes in the network are of three classes namely, Dominating Nodes (DN), Adjacent to Dominating Nodes (ADN), Next to Adjacent Dominating Nodes (NADN).

#### Table 1. Nomenclature

<table>
<thead>
<tr>
<th>Shapes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>🍀</td>
<td>Dominating Nodes (DN)</td>
</tr>
<tr>
<td>🍀</td>
<td>Adjacent to Dominating Nodes (ADN)</td>
</tr>
<tr>
<td>🍀</td>
<td>Next to Adjacent Dominating Nodes (NADN)</td>
</tr>
<tr>
<td>🍀</td>
<td>Number of Dominating Nodes</td>
</tr>
</tbody>
</table>

#### 2.1 Selection of DN

In the given cloud network, the first process is to cover the entire network in the two-hop neighbors of the Dominating Nodes (DN). The DN is measured by the degree based calculation process. The process is similar to the graph theory, where the degree of a vertex of a graph is the number of edges incident to the corresponding vertex with loops counted twice. Based on the degree of all the nodes, the nodes with maximum degree will be selected as Dominating Node (DN).
2.2 Selection of ADN
The nodes which are adjacent to the dominating node are selected and are marked as ADN (Adjacent to Dominating Node).

2.3 Selection of NADN
The node neither belongs to dominating node nor adjacent to dominating node are selected as NADN (i.e. two-hop neighbor of Dominating Node). This process is repeated until all the nodes in the cloud are covered within the two-hop neighbor to the Dominating Nodes (DN).

2.4 Source Request
After covering the entire network within the two-hop neighbors of Dominating Nodes (DN) the source is requesting to capture the big data that it is in need along with the destination id to which the data is to be transmitted. The source request is send to all the dominating nodes in the given cloud network. During this flow of transmission there may be presence of eavesdropper nodes. To prevent the eavesdropper nodes from capturing the big data cooperative jamming scheme is used, through which jamming signal will be emitted by the adjacent nodes in the path to confuse the eavesdroppers.

2.5 Capturing and Transmitting the Big Data Content in the Cloud Network (CTBDC in CN)
The received Dominating Nodes (DN) broadcast the source request to all the nodes within its two-hop neighbors in the corresponding cloud along with the destination id. The node which contains the big data content requested by the source node will send the data to the corresponding destination node with the corresponding id number.

If a node in ADN or NADN is in the path of more than one DN, based on the probability of the DN, a node will be selected for the particular ADN or NADN. The probability value will be equally distributed for each DN participating in this scenario.

To guaranty delivery of each packet of data to all intended neighboring nodes without collisions at any of these nodes, medium access scheme called time division multiple accesses, where each node is assigned a time slot which is different from time slots assigned to any of its two hop neighbors. The proposed algorithm along with jamming signals exhibits the efficient and secure transmission of big data in the wireless networks.

3. Evaluation of Proposed Algorithm

The performance of the cloud network is classified as follows, namely, Strongly Perfect Network (SPN), Strongly Weakly Perfect Network (SWPN) and Weakly Strongly Perfect Network (WSPN).

3.1 Strongly Perfect Network (SPN)
The network is considered as SPN, if the number of DN equals to 1 (i.e. in the given network, all the nodes are directly connected to the DN). The Figure 2 represents the examples of some SPN’s.

3.2 Strongly Weakly Perfect Network (SWPN)
The network is considered as SWPN, if there is more than one DN without the presence of NADN to cover the entire network. The Figure 3 exhibits the various network formats for SWPN.
4. Weakly Strongly Perfect Network (WSPN)

The network is considered as WSPN, if there is more than one DN with the presence of NADN to cover the entire network. So in this scenario, speed of transmission of data is quite slow when compared with the above two scenario. The Figure 4 illustrates the diagrammatic representation of WSPN.

![Figure 4. Weekly Strongly Perfect Network (WSPN).]

4. Analytical Approach of Proposed Algorithm

4.1 Theorem

The countable union of countable Strongly Perfect Network (SPN) need not be a Strongly Perfect network (SPN).

**Proof.** Let us consider two different strongly perfect networks. We can prove the above theorem by 3 cases.

4.1.1 Case 1.1 (a) \( D_1 (SPN_1) \rightarrow D_2 (SPN_2) \)

If the dominating nodes of the two networks are connected by an edge (i.e.) dominating node \( D_1 \) from SPN\(_1\) network is merged by an edge to the dominating node \( D_2 \) of SPN\(_2\) network. The Resultant Network (RNW) will have the same number of dominating nodes but the network will emerge as the Strongly Weekly Perfect Network (SWPN) satisfies the following,

\[ \chi (RNW) = \chi (SPN_1) + \chi (SPN_2) = 2 \text{ DN} \]

Thus the theorem is verified.

4.1.2 Case 1.1(b) \( D_1 (SPN_1) \leftrightarrow D_2 (SPN_2) \)

Suppose the dominating nodes of two Strongly Perfect Network (SPN) coincides with each other (i.e.) if dominating node \( D_1 \) from SPN\(_1\) network and dominating node \( D_2 \) from SPN\(_2\) network are coincides and form a single dominating node for the whole network, the resultant network remains as a Strongly Perfect Network (SPN) follows the rule

\[ \chi (RNW) = \chi (SPN_1) + \chi (SPN_2) - 1 = 1 \text{ DN} \]

Thus the theorem is proved.

4.1.3 Case 1.2 \( D_1 (SPN_1) \rightarrow ND_2 (SPN_2) \)

If we connect \( D_1 \) from SPN\(_1\) network and ND\(_2\) from SPN\(_2\) network by an edge, the network will be resulted as Strongly Weekly Perfect network (SWPN). In this Scenario the number of dominating nodes in the resultant network (RNW) will be equals to 2 that is expressed as \( \Psi (RNW) = 2 \). Suppose \( \Psi (RNW) = n \) then \( n \) numbers of networks are connected by the above mentioned manner.

Thus proves the theorem.

4.1.4 Case 1.3. \( ND_1 (SPN_1) \rightarrow ND_2 (SPN_2) \)

The connection of non-dominating node ND\(_1\) of SPN\(_1\) network with the non-dominating node of SPN\(_2\) network by an edge will result as same as above case 1.2.

4.2 Theorem

The countable union of countable Strongly Weekly Perfect Network (SWPN) is always the Strongly Weekly Perfect Network (SWPN).

**Proof.** The above stated theorem is proved by 3 cases by considering two different Strongly Weekly Perfect Network (SWPN).

4.2.1 Case 2.1(a). \( D_1 (SWPN_1) \rightarrow D_2 (SWPN_2) \)

If the dominating nodes of the two Strongly Weekly Perfect Networks (SWPN) are connected by an edge (i.e.) dominating node \( D_1 \) from SWPN\(_1\) network is connected by an edge to the dominating node \( D_2 \) of SWPN\(_2\) network. Then the Resultant Network (RNW) will have the number of dominating node of SWPN\(_1\) network added to the number of dominating node of SWPN\(_2\) network, thus the theorem is verified by retaining the network as Strongly Weekly Perfect Network (SWPN).

4.2.2 Case 2.1(b). \( D_1 (SWPN_1) \leftrightarrow D_2 (SWPN_2) \)

Similarly as above case, if any one of the dominating node from each of the network coincides with each other (i.e.) if dominating node \( D_1 \) from SWPN\(_1\) and dominating node \( D_2 \) from SWPN\(_2\) are coincides then the Resultant
Network (RNW) having the number of dominating node is expressed as follows:

\[ \chi (RNW) = \chi (SWPN_1) + \chi (SWPN_2) - 1 = 1 \text{ DN} \]

Thus the theorem is proved.

4.2.3 Case 2.2. \( D_1 (SWPN_1) \rightarrow ND_2 (SWPN_2) \)

If we connect \( D_1 \) from \( SWPN_1 \) network to \( ND_2 \) from \( SWPN_2 \) network by an edge, the network will be resulted as Strongly Weekly Perfect Network (SWPN) having the property,

\[ \chi (RNW) = \chi (SWPN_1) + \chi (SWPN_2) = 2 \text{ DN} \]

This completes the proof of the theorem.

4.2.4 Case 2.3. \( ND_1 (SWPN_1) \rightarrow ND_2 (SWPN_2) \)

The connection between non- dominating node \( ND_1 \) of \( SWPN_1 \) network with the non-dominating node \( ND_2 \) of \( SWPN_2 \) network by an edge will result as same as above case 2.2.

4.3 Theorem

The countable union of countable Weekly Strongly Perfect Network (WSPN) is always the Weekly Strongly Perfect Network (WSPN).

**Proof.** The above said theorem is proved by 3 cases by considering different Weekly Strongly Perfect Networks (WSPN).

4.3.1 Case 3.1(a). \( D_1 (WSPN_1) \rightarrow D_2 (WSPN_2) \)

If the dominating node of the two Weekly Strongly Perfect Network (WSPN) is connected by an edge (i.e.) dominating node \( D_1 \) from \( WSPN_1 \) network is merged by an edge to the dominating node \( D_2 \) of \( WSPN_2 \) network. Then the Resultant Network (RNW) will be obeying the following Equation:

\[ \chi (RNW) = \chi (WSPN_1) + \chi (WSPN_2) = 2 \text{ DN} \]

This completes the proof.

4.3.2 Case 3.1(b). \( D_1 (WSPN_1) \leftrightarrow D_2 (WSPN_2) \)

Similarly as above case, if the dominating nodes of the two different network coincides with each other (i.e.) if dominating node \( D_1 \) from \( WSPN_1 \) network and dominating node \( D_2 \) from \( WSPN_2 \) network are coincides then the Resultant Network (RNW) will have the following property:

\[ \chi (RNW) = \chi (WSPN_1) + \chi (WSPN_2) - 1 = 1 \text{ DN} \]

Thus the theorem is proved.

Table 2. Analysis approach for proposed algorithm

<table>
<thead>
<tr>
<th>Cases specification</th>
<th>Figures specification</th>
<th>Features specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1.1 (a)</td>
<td>APPENDIX</td>
<td>1. The dominating nodes are connected with one another from different SPN’s, here the cooperation between these network will be very highly satisfied level, since the security is too strong and prevention of attacker namely eavesdroppers can be easily identified than the other cases.</td>
</tr>
<tr>
<td>Case 2.1 (a)</td>
<td>APPENDIX</td>
<td>2. If the dominating node in RNW transmit the recognized packet of data immediately through the path with the connected neighbor network. This scenario will yield to high speed of transmitting the dataset along with reduced cost.</td>
</tr>
<tr>
<td>Case 3.1 (a)</td>
<td>APPENDIX</td>
<td>In these cases, we are merging the dominating node of the network with one another therefore the efficiency of the network is reduced as well as the speed of transmission of data packet between the node will be comparatively slow with the cases i.j.(a).</td>
</tr>
<tr>
<td>Case 1.1 (b)</td>
<td>APPENDIX</td>
<td>In these cases there will not be any major changes as previously mentioned cases (1.2, 2.2, 3.2), in addition the transmitting speed will be more and information can be easily exchanged between the networks involved in this scenario.</td>
</tr>
<tr>
<td>Case 2.1 (b)</td>
<td>APPENDIX</td>
<td>In these scenarios the resultant network will be strong, since the dominating node of any one of the two network’s performance capacity is increased and it receives the data packet and gains the knowledge of the neighbor node through the connected non- dominant node.</td>
</tr>
<tr>
<td>Case 3.1 (b)</td>
<td>APPENDIX</td>
<td>In this case the efficiency will be improved between the networks; in addition the transmission of data will be increased in a significant ratio.</td>
</tr>
</tbody>
</table>
4.3.3 Case 3.2. $D_1 (\text{WSPN}_1) \rightarrow ND_2 (\text{WSPN}_2)$

If we connect $D_1$ from WSPN$_1$ network and ND$_2$ from WSPN$_2$ network by an edge, the network will be resulted as Weekly Strongly Perfect Network (WSPN) having the property:

$$\chi(\text{RNW}) = \chi(\text{WSPN}_1) + \chi(\text{WSPN}_2) = 2 \text{DN}$$

This completes the proof.

4.3.4 Case 3.3. $ND_1 (\text{WSPN}_1) \rightarrow ND_2 (\text{WSPN}_2)$

The connection of non-dominating node ND$_1$ of WSPN$_1$ network with the non-dominating node ND$_2$ of WSPN$_2$ network by an edge will result as same as above case 3.2.

4.3.5 Case 3.4. $D_1 (\text{WSPN}_1) \rightarrow NADN_1 (\text{WSPN}_2)$

The connection of dominating node $D_1$ to the non-dominating node NADN$_1$ by an edge will result in the creation of ND$_1$ to the node in which the $D_1$ is connected. Further the efficiency gets increased.

Appendix

Case 1.1(a)
$D_1 (\text{SPN}_1) \rightarrow D_2 (\text{SPN}_2)$

Case 1.1(b)
$D_1 (\text{SPN}_1) \leftrightarrow D_2 (\text{SPN}_2)$

Case 1.2
$D_1 (\text{SPN}_1) \rightarrow ND_2 (\text{SPN}_2)$

Case 1.3
$ND_1 (\text{SPN}_1) \rightarrow ND_2 (\text{SPN}_2)$

Case 2.1(a)
$D_1 (\text{SWPN}_1) \rightarrow D_2 (\text{SWPN}_2)$

Case 2.1(b)
$D_1 (\text{SWPN}_1) \leftrightarrow D_2 (\text{SWPN}_2)$

Case 2.1(b)
$D_1 (\text{SWPN}_1) \leftrightarrow D_2 (\text{SWPN}_2)$
Case 2.2
\[ D_1 (SWPN_1) \rightarrow ND_1 (SWPN_2) \]

Case 2.3
\[ ND_1 (SWPN_1) \rightarrow ND_2 (SWPN_2) \]

Case 3.1(a)
\[ D_1 (WSPN_1) \rightarrow D_2 (WSPN_2) \]

Case 3.1(b)
\[ D_1 (WSPN_1) \leftrightarrow D_2 (WSPN_2) \]

Case 3.2
\[ D_1 (WSPN_1) \rightarrow ND_2 (WSPN_2) \]

Case 3.3
\[ ND_1 (WSPN_1) \rightarrow ND_2 (WSPN_2) \]

Case 3.4:
\[ D_1 (WSPN_1) \rightarrow NADN_1 (WSPN_2) \]
5. Conclusion and Future Work

The paper has demonstrated the secure and efficient big data transmission in cloud structure using the proposed broadcasting algorithm (CTBDC in CN). The broadcasting reliability is achieved with significant reduction in the number of rebroadcasting messages, resulting in reduced contention and collision problems in the network. In addition, dominating nodes concept has reduced the maintenance communication cost compared to cluster structure. Another issue in wireless networks is the presence of unidirectional links. Different transmission ranges of the nodes or hidden terminal problem can cause unidirectional links. The performance of proposed scheme in the presence of unidirectional scheme is left for future study.

6. Acknowledgment

The authors would like to thank Dr. Ponnammal Natarajan worked as Director–Research, Anna University-Chennai, India and Dr. K. Sarukesi former vice chancellor in Hindustan University-Chennai, India for their cognitive ideas and dynamic discussions with respect to the paper’s contribution.

7. References