Enumeration of Minimal Control Sets of Vertices in Oriented Graph

G.Sh. Tsitsiashvili, Yu.N. Kharchenko, A.S. Losev and M.A. Osipova

Institute of Applied Mathematics Far Eastern Branch of the RAS,
690041 Vladivostok, Radio street 7, Russia

Far Eastern Federal University, 690950, Vladivostok, Sukhanov street 8, Russia

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
In this paper new algorithm of oriented graph factorization with square complexity is constructed. This algorithm allows to enumerate all control sets of graph vertices. Appropriate numerical experiment is made.

Mathematics Subject Classification: 605C80

Keywords: a cluster, a factor graph, a partial order, a control set

Introduction
In this paper we consider oriented graph with finite sets of vertices and edges. A sequential algorithm of graph factorization with square number of arithmetical operations by a number of graph vertices is constructed. A decrease of calculation complexity in a comparison with boolean multiplication of contiguity matrixes is connected with an introduction of partial order matrix which is defined completely by assignment operations.

Using this algorithm a problem of an enumeration of minimal by a volume vertices sets controlling all graph vertices is analyzed. In other words for any
graph vertex there is a vertex from control set so that in the graph there exists a way from control vertex to separated graph vertex. Such concept of control set of vertices is close to analogous concept [1] but it allows to enumerate all control sets of vertices as in initial graph so in subgraphs with fixed minimal vertices. The procedure may be considered as calculation development of [2] approach.

A numerical experiment demonstrated possibilities of suggested algorithm is made.

1 Preliminaries

Consider oriented graph $G$ with finite sets of vertices $V$ and edges $E$. Say that two vertices $v_1, v_2 \in V$ of oriented graph belong to binary relation $v_1 \sim v_2$ if in this graph there is a cycle which contains both vertices. In [3], [4] the binary relation ” $\sim$ ” of oriented graph vertices is called the strong connectivity which is the equivalence relation [5, § 3]. Divide the set $V$ into classes of the equivalence and denote $[v]$ the cluster which contains the vertex $v \in V$, $[V]$ is the set of these clusters. Call the oriented graph $[G]$ with the set of vertices $[V]$ and the set of edges $[E] = \{(v_1, v_2) : \exists v'_1 \in [v_1], v'_2 \in [v_2], (v'_1, v'_2) \in E\}$ a factor graph. It is obvious that $[G]$ is acyclic graph.

On the set $[V]$ of the graph $[G]$ vertices define the binary relation ” $\succeq$ ”: $[v_1] \succeq [v_2]$, if in the graph $[G]$ there is a way from the cluster $[v_1]$ to the cluster $[v_2]$. By the definition of the graph $[G]$ the binary relation ” $\succeq$ ” is the partial order relation [5, § 4]. It is obvious that if $[v_1] \succeq [v_2]$ then in the graph $G$ there is a way from any vertex of the cluster $[v_1]$ to any vertex of the cluster $[v_2]$.

Separate in the set $[V]$ the subset $[V]'$ of maximal by the relation ” $\succeq$ ” elements. By construction for any cluster $[v] \in [V]$ there is a cluster $[v]' \in [V]'$ so that $[v]' \succeq [v]$. Take in each cluster $[v]' \in [V]'$ an element $v''$ and compose from these elements the controlling set $V' = \{v'' : [v]' \in [V]'\}$. In such a way it is possible to enumerate all minimal (by a number of elements) control sets in the graph $G$.

Remark. Assume that the subset $\tilde{V} \subseteq V$. All ways in the graph $G$ coming into the vertices $\tilde{v} \in \tilde{V}$ create the subgraph $\tilde{G}$ of the graph $G$. Enumerate all minimal sets $\tilde{V}'$ of the graph $G$ vertices which satisfy the condition: for any $\tilde{v} \in \tilde{V}$ there is $\tilde{v}' \in \tilde{V}'$ and a way from $\tilde{v}'$ to $\tilde{v}$. To solve this problem it is necessary to construct the set of clusters $[\tilde{V}]$ containing vertices from $\tilde{V}$ and find the set $[\tilde{V}']$ of all maximal elements of the graph $[G]$, exceeding some elements from $[V]$. Then each set $\tilde{V}'$ may be constructed by a separation in each cluster from $[\tilde{V}]'$ a single element and their aggregation into $\tilde{V}'$. 
2 Algorithm of factorization and partial order construction by contiguity matrix

Describe now an algorithm of a factorization of the graph $G$ vertices and a construction of partial order $\succeq$. Assume that the set $V = \{1, \ldots, n\}$ consists of $n$ vertices and denote $t = \min(f : 2^f \geq n) = \lceil \log_2 n \rceil + 1$ and $A = |a_{ij}|_{i,j=1}^n$ the contiguity matrix of the graph $G$. Construct a sequence of zero-one matrixes $A(k) = ||a_{jr}(k)||_{j,r=1}^n$ by recurrent relation

$$a_{jr}(1) = \max_{1 \leq s \leq n} \min(a_{js}, a_{sr}), \quad a_{jr}(k + 1) = \max_{1 \leq s \leq n} \min(a_{js}(k), a_{sr}(k)), \quad 1 \leq k < t.$$

It is obvious that if $a_{jr}(t) = 1$ then in the graph $G$ there is a way from the vertex $j$ to the vertex $r$ and $j' \in [j] \iff \exists j'' \in [i], j' \in [j] : a_{j''r}(t) = 1$.

3 Sequential algorithm of factorization and partial order construction

Cubic complexity of previous section algorithm is strong restriction for numerical experiments with oriented graphs occurred in an analysis of protein networks. So it is naturally to decrease calculation complexity.

Describe sequential algorithm of a construction of clusters and the matrix of partial order between clusters. On the step 1 there is the single vertex 1 which creates the cluster $[1]$ and the set of clusters $K = \{[1]\}$. Introduce the matrix $a = |a([p],[q])|_{[p],[q] \in K}$ which characterizes the partial order $\geq$ : $a([p],[q]) = 1$, if $[p] \geq [q]$ and in opposite case $a([p],[q]) = 0$. On the step 1 $a([1],[1]) = 1$.

Assume that on the step $t$ there is the clusters set $K$ and the matrix $a$. These clusters create a division of the set $V = \{1, \ldots, n\}$ into non intersected subsets, each cluster $[k] \in K$ is indexed by maximal number $k \in V$ of its vertices.

On the step $t + 1$ new vertex $t + 1$ appears. It is connected with edges which come into vertices from the set $P \subseteq V$ and with vertices which come into the vertex $t + 1$ from the vertices from the set $Q \subseteq V$. Each vertex from the set $P$ (the set $Q$) comes into some cluster. Denote $[P]$, $[Q]$ the sets of clusters corresponding the sets of vertices from $P$, $Q$ and define

$$K_{[p]} = \{[k] \in K : a([p],[k]) = 1\}, \quad [p] \in [P],$$

$$K_{[q]} = \{[k] \in K : a([k],[q]) = 1\}, \quad [q] \in [Q], \quad A = \left( \bigcup_{[p] \in [P]} K_{[p]} \right) \cap \left( \bigcup_{[q] \in [Q]} K_{[q]} \right).$$
\[ A_1 = \left( \bigcup_{[p] \in [P]} K_{[p]} \right) \setminus A, \ A_2 = \left( \bigcup_{[q] \in [Q]} K_{[q]} \right) \setminus A, \ B = K \setminus (A \cup A_1 \cup A_2). \]

New vertex \( t + 1 \) and clusters from the set \( A \) create new cluster
\[ [t + 1] = \{ t + 1 \} \cup A, \ K = (K \setminus A) \cup \{ [t + 1] \}. \]

A recalculation of the matrix \( a \) on renewed set of clusters \( K \) is following \([8]\):
\[ a([t+1], [i]) = 1, \ [i] \in A_1 \cup \{ [t+1] \}, \ a([i], [j]) = 1, \ [i] \in A_2, \ [j] \in A_1 \cup \{ [t+1] \}, \]
\[ a([i], [j]) = 0, \ [i] \in A_1, \ [j] \in A_2 \cup \{ [t+1] \} \cup B, \]
\[ a([i], [j]) = 0, \ [j] \in A_2, \ [i] \in B \cup \{ [t+1] \}, \ a([t+1], [i]) = a([i], [t+1]) = 0, \ [i] \in B. \]

All other meanings of the matrix \( a \) elements coincide with previous ones on the step \( t \). This algorithm has square complexity (a number of arithmetical operations) by a number of graph vertices. All other operations are assignment operations.

To define maximal (minimal) by the partial order ” \( \succeq \) ” clusters it is necessary in the matrix \( a \) calculated on last step to separate columns with the unit only at the diagonal (rows with the unit only at the diagonal).

To test suggested factorization algorithm we consider the human protein network Interactome with 7520 vertices and 22737 edges \([2]\). Results of the factorization of this network vertices are represented in the table. The table 1 allows to represent this protein network essentially more compact and obtain very contrast distribution by clusters by numbers of their vertices.

<table>
<thead>
<tr>
<th>numbers of cluster vertices</th>
<th>numbers of clusters</th>
<th>numbers of maximal clusters</th>
<th>numbers of minimal clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4526</td>
<td>0</td>
<td>2290</td>
</tr>
<tr>
<td>2</td>
<td>63</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2800</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>4608</td>
<td>29</td>
<td>2322</td>
</tr>
</tbody>
</table>

Table.
References


Received: February 1, 2014