



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

Betweenness Centrality of Event Graph Application to Metabolic Network Modelled by Elementary Net System

H. Raaf and B. Messabih

Département d'informatique faculté des sciences université des sciences et de la technologie Mohamed Boudiaf Oran Algérie

Abstract: The ENS (Elementary Net Systems) is derived from low Petri nets Conditions/events systems. We present in this study an approach to compute the betweenness centrality index on a depending graph of an ENS. This measure can be used not only to predict the most influential event in ENS also to detect community structure (sub-network structure). So we firstly give a new strategy to represent ENS network as an event depending graph. Then we give some definitions to compute betweenness centrality indices on this depending graph. Pathway analysis of large metabolic networks meets with the problem of combinatorial explosion of pathways. To help us better understand the organization principle of complex metabolic network, we propose to present a metabolic network with ENS. This aims to deconstruct metabolic networks into sub-networks based on the global network structure by using the betweenness index. This strategy of modeling is most adequate to apply the betweenness algorithm.

Key words: Petri net, shortest path, depending graph, betweenness algorithm, biochemical network, metabolic pathways

INTRODUCTION

Elementary Net System (ENS) known as Conditions/events systems and have been defined by Thiagarajan (1986), Rozenberg (1986) and Andre (1989) to describe easily intrinsic characteristics of real-time systems such as high degree of parallelism and synchronisation. This model is derived from low level Petri net model.

Although, Petri nets and ENS where been defined several years ago, nothing was known about their decomposition into sub-networks or the most influential event or the notion of centrality. The idea of centrality as applied to human communication was introduced by Bavelas (1948). It is defined on social network represented as simple graph. We propose to use the notion of betweenness centrality used for detecting communities in social large network (Girvan and Newman, 2002) to detect the most influential event in ENS. To find community (sub-network) structures in graphs, the interest in the domain has started with a new divisive approach proposed by Girvan and Newman (2002), the edges with the largest betweenness (number of shortest paths passing through an edge) are removed one by one in order to split hierarchically the graph into communities.

The behavior of Metabolic network is well modelled and analyzed by Petri Nets (Reddy *et al.*, 1993). The

introduction of the Petri nets to represent the biochemical reactions was carried by many researchers (Hofestaedt and Telen, 1998; Goss and Peccoud, 1998; Matsuno *et al.*, 2000; Chen, 2004) they showed that the graphic structure (static) of the Petri nets easily (qualitatively) models the interactions in a cell: the places represent the substances, the transitions represent the reactions, the arrows the tokens and the weights of the arrows model the quantity of the substances which is subjected to a reaction to produce a quantity of products ensuring the expression of temporal evolution of these networks (dynamics). The execution of Petri nets allows the quantitative analysis (kinetic) (Hofestaedt and Telen, 1998) of these networks. Structural analysis of large metabolic network (Schuster *et al.*, 2002) meets with the problem of combinatorial explosion. A first work to decompose a metabolic network into sub-network was proposed by Schuster *et al.* (2002). It considers the metabolites degree of network connectivity. Then others works followed (Holme *et al.*, 2003; Ma *et al.*, 2004; Ravasz *et al.*, 2002). All these works represent the metabolic network as a simple graph.

In this study, we investigate an approach to compute betweenness centrality measure on the model of ENS. As the behavior of metabolic network is well represented by Petri nets, we propose to use ENS to model it (substances/enzymatic reactions). This aims to

deconstruct the metabolic network into sub-networks based on the global network structure by using the centrality index. Identifying the modular organization of metabolic network by certain network decomposition methods can help us better understand the organization principle of complex metabolic network.

An essential tool for the analysis of social networks are centrality indices defined on the vertices of the graph (Freeman, 1979; Newman and Girvan, 2003). They are designed to rank the actors according to their position in the network and interpreted as the prominence of actors embedded in a social structure. Many centrality indices are based on shortest paths linking pairs of actors (Brandes, 2001; Girvan and Newman, 2002), measuring, e.g., the average distance from other actors, or the ratio of shortest paths an actor lies on. Many network-analytic studies rely at least in part on an evaluation of these indices. All of them are based on the same idea. If two communities are joined by only a few inter-community edges, then all paths through the network from vertices in one community to vertices in the other must pass along one of those few edges. Given a suitable set of paths, one can count how many go along each edge in the graph and this number we then expect to be largest for the inter-community edges, thus providing a method for identifying them.

To apply this measure, we first represent ENS as a graph of events in which events are defined as nodes and conditions as depending connections between the nodes. Secondly we adapt the definition of shortest path to compute centrality indices on this graph and we propose to use Brandes (2001) algorithm.

MATERIALS AND METHODS

Elementary net systems: Elementary net systems are based on low level Petri nets Known as conditions/events systems. Places also called conditions are allowed to contain at most one mark. Consequently, transitions, also called events, may fire not only if input places contain one mark but also if output places are empty. Obviously, such nets are bounded. Note that the firing takes no time, as for reactive systems.

Definition 1: An ENS is a tuple $E = (G, M_0)$ where G and M_0 are the Petri part:

$G = (S, R, Prec, Post)$ where S is the set of conditions and R is the set of events, the set $Prec$ of arrows represents directed link from pre-conditions to event and the set $Post$ of arrows represents directed link from event to post-conditions.

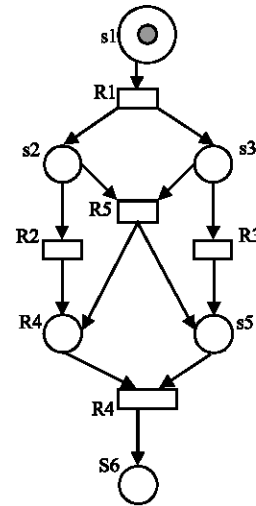


Fig. 1: An elementary net system

M_0 is the initial marking $M: S \rightarrow \{0,1\}$ (number of marks in a condition of set S is zero or one).

- $Prec(r) = \circ r$ where $\circ r$ represent the set of preconditions needed by an event r , $Prec \subseteq R \times partyset(S)$.

- $Post(r) = r^\circ$ where r° represent the set of post-conditions of an event r , $Post \subseteq R \times partyset(S)$.

Example 1: Figure 1 shows an elementary net system with R_1, R_2, R_3, R_4, R_5 are events presented by rectangles. $s_1, s_2, s_3, s_4, s_5, s_6$ are conditions presented by circles. And one mark in s_1 (Fig. 1).

Definition 2: An event r is enabled iff all its pre-conditions are enabled

$$\forall s \in \circ r \quad M(s) = 1 \text{ and } \forall s \in r^\circ M(s) = 0$$

The firing of an enabled event r from marking M results in marking M' according to the following:

$$\forall s \in \circ r \quad M(s) = 0, \forall s \in r^\circ M'(s) = 1, \forall s \in S - (\circ r \cup r^\circ) \\ M'(s) = M(s)$$

Events graph: A directed depending events graph is defined as follows:

- A directed link from event node A to event node B is added iff all pre-conditions of B composed with post-condition of the event A . We say an event B can occur (to fire) only after A (if B is connected only with A) or B depend on A
- An event B can occur only after the events A and C iff pre-conditions of B composed with post-conditions of A and post-conditions of C

- An even B can occur only after A1, A2... and An iff pre-conditions of B are composed with post conditions of A1 and A2 ... An

Definition 3: Let $G = (S, R, \text{Prec}, \text{Post})$ be an ENS. $\text{Gr} = (R \cup T, P)$ is the directed event graph of G , if the following conditions hold for all z and r in R and all t in T :

$$\begin{aligned}
 &-(r1, r2) \in P \text{ iff } r1^\circ = \circ r2 \\
 &-(z_i, t) \in P \text{ and } (t, r) \in P \text{ iff } z_i \circ \cap \circ r \neq \emptyset \text{ and} \\
 &z^* \cap \circ r = r. \text{ With } z^* = z_i^\circ \text{ and} \\
 &\quad \quad \quad i=1,2..
 \end{aligned}$$

In an events graph $\text{Gr} = (R \cup T, P)$, we denote by Pre and Succ the following sets:

$$\begin{aligned}
 \text{-Pre}(r) &= \{u \in R \cup T / (u, r) \in P\} \\
 \text{-Succ}(r) &= \{u \in R \cup T / (r, u) \in P\}
 \end{aligned}$$

Example 2: Figure 2 shows the events graph of Fig. 1.

Shortest path and betweenness

Shortest path: A quantity of interest in many social network (modelled with a graph) studies is the betweenness of an actor i , which is defined as the total number of shortest paths between pairs of actors that pass through i . This quantity is an indicator of who the most influential people in the network are, the ones who control the flow of information between most others. We are interesting to do this idea for detecting the most influential event in the events graph of an ENS.

As the events graph is composed with two types of nodes (R or T nodes) an adequate shortest path definition is necessary.

Let w be a weight function on the arrow. We define:

$$\begin{aligned}
 w((u, r)) &= 1 \quad (u, r) \in P, u \in R \cup T, r \in R \text{ and} \\
 w((r, t)) &= 0 \quad (r, t) \in P, r \in R, t \in T
 \end{aligned}$$

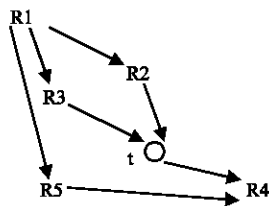


Fig. 2: The event graph

Weights are used to measure, the strength of a link.

Define a path from $s \in R$ to $r \in R$ as an alternating sequence of vertices and edges, beginning with s (firing the event s) and ending with r (r is enabled), such that each edge connects its preceding with its succeeding vertex. We use $d(s, r)$ to denote the distance between vertices s and r , i.e., the minimum length of any path connecting s and r in Gr .

By definition, $d(s, s) = 0$ for every $s \in R$ and $d(s, r) = d(r, s)$ for $s, r \in R$. We assume familiarity with standard algorithms for shortest-paths problems.

The minimal path of an event $t \in T$ to occur after firing an event s is the maximal value of shortest paths of a succeeding events such that pre-conditions of t are taken on and all events preceding t can be enabled with event s . We consider here the shortest-path as the longest succeeding shortest path of events which makes some pre-conditions of t taken on. In other words:

$$\begin{aligned}
 d(s, t) &= \text{Max}_r d(s, r) \text{ and for all event } r \text{ preceding} \\
 &\quad \quad \quad t \quad d(s, r) > 0
 \end{aligned}$$

Definition 4:

$$\text{for every } s \in R \cup T \quad d(s, s) = 0 \tag{1}$$

$$d(s, t) = \text{Min}_v (d(s, v)) + 1 \tag{2}$$

with $(v, t) \in P, v \in R \cup T, s \in R \cup T, t \in R$

$$d(s, t) = \text{Max}_r d(s, r) \tag{3}$$

with $(r, t) \in P \quad s \in R \cup T, r \in R, t \in T$

Centrality indices based on shortest paths: Several measures capture variations on the notion of a vertex's importance in a graph. Let σ_{st} denote the number of shortest paths from $s \in R \cup T$ to $t \in R \cup T$, where $\sigma_{ss} = 1$ by convention. Let $\sigma_{st}(v)$ denote the number of shortest paths from s to t , that some $v \in R \cup T$ lies on.

Given pairwise distances and shortest paths counts, the pair-dependency

$$\delta_{st}(v) = \sigma(v) / \sigma_{st} \tag{4}$$

of a pair $s, t \in R \cup T$ on an intermediary $v \in R \cup T$, i.e. the ratio of shortest paths between s and t that v lies on, is given by:

$$\delta_{st}(v) = 0 \quad \text{if } d(s, t) < d(s, v) + d(v, t)$$

$$\delta_{st}(v) = \sigma_{sv} \cdot \sigma_{vt} \quad \text{otherwise}$$

To obtain the betweenness centrality index of a vertex v , we simply have to sum the pair-dependencies of all pairs on that vertex,

$$CB(v) = \sum_{s \neq v \in R \cup T} \delta_{st}(v) \quad (5)$$

Therefore, betweenness centrality is determined (Brandes, 2001) in two steps:

- Compute the length and number of shortest paths between all pairs
- sum all pair-dependencies

RESULTS

We adapt the faster Algorithm defined by Brandes according with our definition of shortest path (definition 4). In Brandes algorithm, the betweenness centrality can be computed in $O(nm + n^2 \log n)$ time and $O(n+m)$ space for weighted graphs. Where n is the number of vertices and m the number of links. For unweighted graphs, run-ning time reduces to $O(nm)$. Brandes count shortest paths using traversal algorithms. Both breadth-First Search (BFS) for unweighted and Dijkstra's algorithm for weighted graphs start with a specified source vertice and, at each step, add a closest vertex the set of already discovered vertices in order to find shortest paths from the source to all other vertices. He determines the betweenness centrality index by solving one single source shortest-paths problem for each vertex. At the end of each iteration, the dependencies of the source on each other vertex are added to the centrality score of that vertex. Our adaptation doesn't modify Brandes's algorithm complexity. It takes in count tow types of nodes ($s \in R \cup T$) and applies definition 4 to determine the length of shortest paths between nodes.

Example 3: Figure 3 shows the betweenness index of Fig. 2.

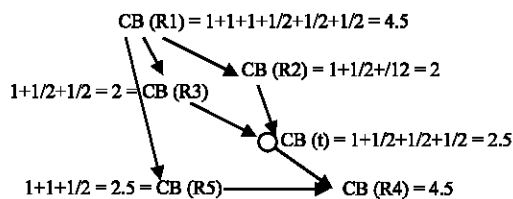


Fig. 3: The betweenness indices

Application to metabolic network: The metabolic networks are represented by the biochemists as complex graphs whose nodes constitute the metabolites and the arrows represent the reactions which are catalysed by specific molecules called enzymes. The structural (topological) analysis of metabolic pathways only uses information about stoichiometric structure and the reversibility or irreversibility of enzymatic reactions. It consists in computing a minimal enzymes set that describe the conical steady state solution space for flux distribution in a metabolic network.

In the modelling of metabolic systems, usually a distinction is made between external and internal metabolites. External metabolites are the sources and sinks of the network; their concentrations are assumed to be buffered. Internal metabolites (intermediates) have to be balanced with respect to production and consumption at steady state. In many cases, there are specific biochemical reasons to treat a metabolite as internal (for example, when it is known to be merely an intermediate) or as external (e.g., nutrient or excreted product).

Classification of metabolites is not clear: an external metabolite is considered to be present in large excess so that its concentration is virtually unaffected by the reactions under study. An internal metabolite, however, must fulfil a balance equation implying that production is equals consumption. A first work to classify metabolites was proposed by Schuster *et al.* (2002). It considers the metabolite degree of network connectivity. Above a given connectivity threshold, some metabolites are considered as external. Then others works followed to decompose metabolic networks (Holme *et al.*, 2003; Ma *et al.*, 2004; Ravasz *et al.*, 2002). Identifying the modular organization of metabolic network by certain network decomposition methods can help us better understand the organization principle of complex systems. Here we represent a metabolic network as an ENS and than represent it as a reaction depending graph by definition 3. The connectivity in this graph describes the sequences and the dependency of enzymatic reactions and not only connectivity.

Definition 5: A metabolic network is represented as $ENS = (S, R, Prec, Post)$ where S is the set of metabolites and R is the set of enzymatic reactions, the set $Prec$ of arrows represents directed link from metabolite substrates to reactions and the set $Post$ of arrows represents directed link from reactions to products.

$-R = R_{irev} \cup R_{rev}$ (irreversibles and reversibles enzymatics reactions)

-Prec(r) = r° where r° represent the set of substances needed by a reaction r, $\text{Prec} \subseteq \text{Rx party set}(S)$.

-Post(r) = r° where r° represent the set of products of a reaction r, $\text{Post} \subseteq \text{Rx party set}(S)$

Example 4: The pentose-phosphate-pathway (PPP) (Fig. 4, 5). In this example we represent the PPP by ENS and give the corresponding reaction depending graph.

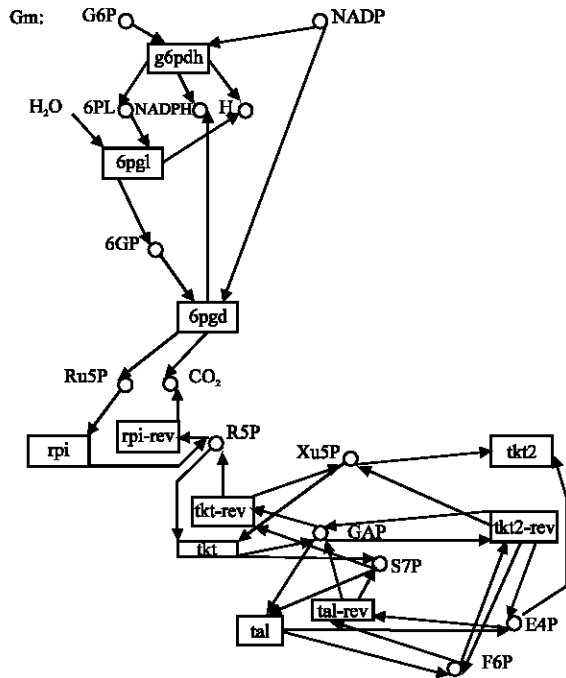


Fig. 4: Metabolic network with ENS

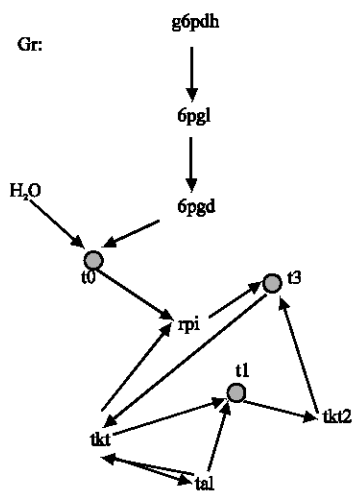


Fig. 5: The reaction depending graph of metabolic network of Fig. 4

$S = \{ G6P, NADP, 6PL, NADPH, H, H_2O, 6GP, CO_2, Ru5P, Ru5P, R5P, Xu5P, R5P, GAP, S7P, F6P, E4P \}$

$R = \{ g6pdh, 6pgl, 6pgd, rpi, rpe, tkt, tal \}$

$g6pdh: G6P + NADP \rightarrow 6PL + NADPH + H$

$6pgl: 6PL + H_2O \rightarrow 6GP + H$

$6pgd: 6GP + NADP \rightarrow NADPH + CO_2 + Ru5P$

$rpi: Ru5P \rightleftharpoons R5P$ rpi is reversible reaction we can represent it as rpi and rpi_rev as follow:

rpi: $Ru5P \rightarrow R5P$

rpi_rev: $R5P \rightarrow Ru5P$

tkt: $Xu5P + R5P \rightarrow GAP + S7P$

tkt_rev: $GAP + S7P \rightarrow Xu5P + R5P$

tal: $GAP + S7P \rightarrow F6P + E4P$

tal_rev: $F6P + E4P \rightarrow GAP + S7P$

tkt2: $Xu5P + E4P \rightarrow GAP + F6P$

tkt2_rev: $GAP + F6P \rightarrow Xu5P + E4P$

DISCUSSION

This investigation of modelling the metabolic network as a reaction depending graph and the application of the betweenness centrality algorithm helps to determine the most influential reaction and eventually to decompose the metabolic network into subnetworks. The strategy of representing the metabolic network as a reaction depending graph (event graph definition 3) gives a graph which has only reactions types of nodes (reactions without metabolites). This is important to apply the Brandes algorithm. What makes the difference with the Holme *et al.* (2003) work. We preferred to compute the betweenness centrality on a graph which is primarily composed with only reaction nodes. Secondly, the connectivity in this graph describes the sequences and the dependency of enzymatic reactions (definition 3).

CONCLUSION

In this study we have investigated a centrality structure in ENS. This betweenness centrality is based on compute of the length and number of shortest paths between all pairs events and sum all pair-dependencies. Our approach permits firstly to describe an ENS network as an event depending graph, secondly, to give a new definition of shortest path on graph which defined with tow reaction types of nodes and we have adapt the betweenness algorithm for our definitions. The use of ENS to model metabolic network permits to represent the metabolic network as a reaction (event) depending graph.

REFERENCES

Andre, C., 1989. Synchronized Elementary Net Systems. In: Advances in Petri Nets 1989, Rozenberg, G. (Ed.). Springer-Verlag, Berlin, Heidelberg, ISBN-13: 978-3-540-52494-6, pp: 51-76.

- Bavelas, A., 1948. A mathematical model for group structure. *Hum. Organisations*, 7: 16-30.
- Brandes, U., 2001. A faster algorithm for betweenness centrality. *J. Mathematical Sociol.*, 25: 163-177.
- Chen, M., 2004. *In silico* systems analysis of biopathways. <http://bieson.ub.uni-bielefeld.de/volltexte/2004/556/index.html>.
- Freeman, L.C., 1979. Centrality in social networks: Conceptual clarification. *Social Networks*, 1: 215-239.
- Girvan, M. and M.E.J. Newman, 2002. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA.*, 99: 7821-7826.
- Goss, P.J.E. and J. Peccoud, 1998. Quantitative modelling of stochastic system in molecular biology by using stochastic Petri nets. *Proc. Natl. Acad. Sci. USA.*, 95: 6750-6755.
- Hofstaedt, R. and S. Telen, 1998. Quantitative modelling of biochemical networks. *In Silico Biol.*, 1: 39-53.
- Holme, P., M. Huss and H. Jeong, 2003. Subnetwork hierarchies of biochemical pathways. *Bioinformatics*, 19: 532-538.
- Ma, H.W., X.M. Zhao, Y.J. Yuan and A.P. Zeng, 2004. Decomposition of metabolic network into functional modules based on the global connectivity structure of reaction graph. *Bioinformatics*, 20: 1870-1876.
- Matsuno, H., A. Doi, M. Nagasaki and S. Miyano, 2000. Hybrid Petri Net representation of Gene regulatory network. *Proc. 5th Pacific Symp. Biocomput.*, 5: 338-349.
- Newman, M.E.J. and M. Girvan, 2003. Finding and evaluating community structure in networks. <http://www.citeulike.org/user/salvoscellato/article/154>.
- Ravasz, E., A.L. Somera, D.A. Mongru, Z.N. Oltvai and A.L. Barabasi, 2002. Hierarchical organization of modularity in metabolic networks. *Science*, 297: 1551-1555.
- Reddy, V.N., M.L. Mavrouniotis and M.N. Liebman, 1993. Petri net representation in metabolic pathways. *Proc. Int. Conf. Intel. Syst. Mol. Biol.*, 1: 328-336.
- Rozenberg, G., 1986. *Behavior of Elementary Net Systems*. Springer-Verlag, Berlin, pp: 60-94.
- Schuster, S., T. Pfeiffer, I. Koch, F. Moldenhauer and T. Dandekar, 2002. Exploring the pathway structure of metabolism: Decomposition into subnetworks and application to *Mycoplasma pneumoniae*. *Bioinformatics*, 18: 351-361.
- Thiagarajan, P.S., 1986. *Elementary Net Systems*. In: *Petri Nets: Central Models and Their Properties*, Brauer, W., W. Reisig and G. Rozenberg (Eds.). Springer-Verlag, Berlin, Heidelberg, ISBN-13: 978-3-540-17905-4, pp: 26-59.