An extrapolation of Fleury's Algorithm for determining the longest path in an undirected graph

F. Ghionea, M. Pirvan

Abstract: Existing solutions for determining the longest path in an undirected graph are indirect or heuristically, since the existence of the cycles is discouraging. The"exfoliation" method is the solution proposed to avoid the problem of cycles: iteratively, the peripheral structure of edges and nodes of the graph is removed; concatenating then the sub-graphs resulted after the exfoliation procedure. If a graph admits an elementary chain containing all its nodes, then the concatenation of successive sub-graphs is possible and leads to the determination of the longest path in the undirected graph. This statement is proved adequately adjusting Fleury's algorithm for Eulerian paths , not in the analyzed graph, but in a matagraph(an auxiliary graph which, instead of nodes, has the sub-graphs resulted after the "exfoliation" procedure is applied).

Keywords: directed graph, undirected graph, incident nodes, maximal chain, edges, odd-degree nodes, Eulerian cycle.

I. INTRODUCTION

One of the first categories of graphs can be made upon the possibility of getting over of a conjunction between two nodes in one or both senses [1-15]; an undirected graph permits the transit of any existing conjunction in both senses. Once it is established the affiliation of the analyzed graph to one of the two categories, the differences between the concepts used are multiplied [1,4]:

F.Ghionea is Conf..PHD. Eng. at University Politehnica of Bucharest, Independence Street, nr. 313, Bucharest,Romania, fghionea@yahoo.com,
M. Pirvan is Math PHD Lect at University Politehnica of Bucharest, Independence Street, nr. 313, Bucharest, Romania, monicapirvan@yahoo.com.

- What is a path in an oriented graph becomes a chain in an undirected graph;
- What is a circuit in an oriented graph becomes a cycle in an undirected graph;
- The length of a path – measured by the number of the component links – becomes the length of the chain – measured by the number of the component edges;
- etc., and not at least,
- the minimum path between two corners in an oriented graph –characterized by the smallest number of kilometers between the two corners, becomes the chain of minimum value between two nodes in an undirected graph – characterized by the smallest number of kilometers between the two nodes;
- the path of maximal value between two corners in an oriented graph – characterized by the biggest number of kilometers between the two corners - way that exists only if there isn't any possible variant of repeating its own steps inside the graph (it means that the graph doesn't have circuits) - *doesn't have correspondence in an undirected graph* because any imagined algorithm for the determination of this maximal chain would barge into the problem of cycles, which, repetitively got over, may lead to the getting of any value, obviously of the infinite value for the searched length [1,3, 7-10]. Example :

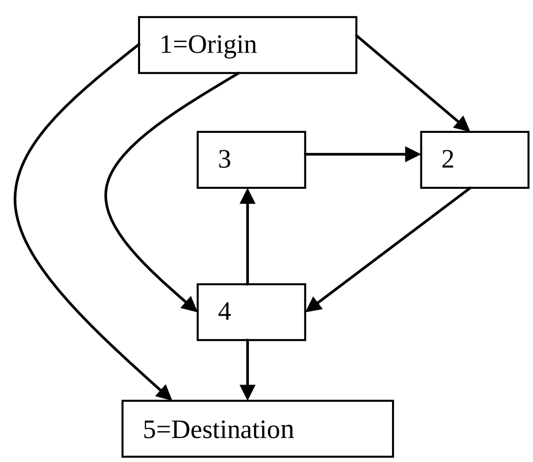

Fig. 1 Oriented graph (which permits the repetition of the own steps inside the graph)

It may find out that the way 1-2-4-3-2-4-3-… has the infinite value (as number of links as well as number of kilometers) and the algorithms that are known (Ford, Bellman-Kalaba, etc. [1,2,3]) are inefficient. Analogously**,** if in the graph in fig. 1 the edges aren't taken into consideration – but the connections still remain– it appears again the same problem of infinite chain, and the algorithms are unknown, because the cycling has probably discouraged.

II. THE PROBLEM STATEMENT

In the following, the authors have as a goal the identification of an algorithm to determine the longest chain in a graph, under the circumstances in which the analysis is based only upon undirected, connected[1] and unarticulated[2] graphs.

To exemplify the procedure which leads to the searched result, it will initially consider the graph in fig. 2.

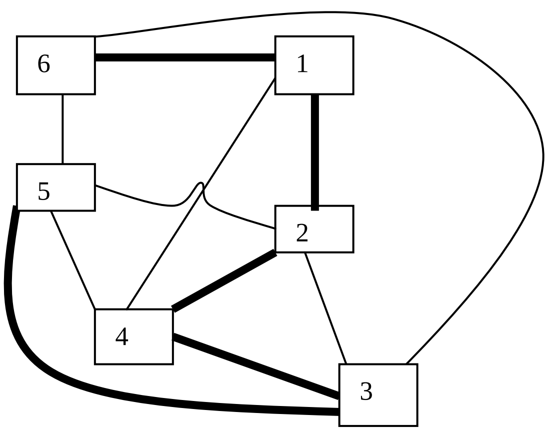

Fig. 2 Undirected, connected and unarticulated graph

It makes the initial choice: for example: node 6 as the origin of the chain. It realizes the lot of incidence nodes for node 6 : = {1,3,5}

- it casts one of the continuances; may be node 1: the result is the chain 6-1
- it follows some steps in which :
  - each time it analyzes the last and the penultimate incidence lot, from which it excludes the nodes that already compose the chain (marked by an asterisk);
  - it prefers to continue the chain by nodes with double representation – and in the incidence lot of the last component (noted U) and in the incidence lot of the penultimate component (noted P); *otherwise, it may choose any continuance.*

Step 1

$U(1) = \{2,4,6*\}$, $P(6) = \{1,3,5*\}$

$[U(1) \setminus \{6\}] \cap [P(6) \setminus \{1\}] = \emptyset$

- any continuance from the lot U(1) \ {6}
- it choose node 2, resulting the chain 6-1-2

Step 2

$U(2) = \{1*,3,4,5\}$, $P(1) = \{2*,4,6*\}$

$[U(2) \setminus \{1\}] \cap [P(1) \setminus \{1, 6\}] = \{4\}$

- it is recommended node 4 resulting the chain 6-1-2-4

Step 3

$U(4) = \{1*,2*,3,5\}$, $P(2) = \{1*,3,4*,5\}$

$[U(4) \setminus \{1, 2\}] \cap [P(2) \setminus \{1, 4\}] = \{3, 5\}$

- it choose node 3 resulting the chain 6-1-2-4-3

Step 4

$U(3) = \{2*,4*,5,6*\}$, $P(4) = \{1*,2*,3*,5\}$

$[U(3) \setminus \{2, 4, 6\}] \cap [P(4) \setminus \{1, 2, 3\}] = \{5\}$

- it is recommended node 5 resulting the chain 6-1-2-4-3-5

Because all the nodes are contained, the succession 6-1-2-4-3-5 represents the longest chain (with thickened line in fig. 2). For the same graph: if the initiation of the calculations would be done in node 2, it would result the longest chain represented by the succession 2-3-4-5-6-1.

III. THE ISOLATION OF THE TRIANGLE IN THE NONTRIVIAL GRAPH

Definition: A connected, undirected and unarticulated graph will be called nontrivial if in the graph there are at least 3 nodes {i,j,k} so that the edges i-j, j-k and k-i make a cycle.

Let's consider an nontrivial graph in which the triangle (i,j,k) has been isolated - Fig. 3. The set of the nodes that are incident to the node i, others than {j,k}, would be noted $\{X_m\}$ with m = 1…q, the set of the nodes that are incident to the node j, others than {i,k}, would be noted $\{Y_n\}$ with n = 1…r and the set of the nodes that are incident to the node k, others than {i,j}, would be noted $\{Z_p\}$ with p = 1…t. As an origin node, it may choose any of the nodes {i,j,k}, the procedure being similar to any choice.

[1] Connected graph = a graph in which there is an elementary or composed chain between any two joints.

[2] Unarticulated graph =a graph in which the elimination of a certain joint leads to the apparition of two sub graphs (independent one of another).

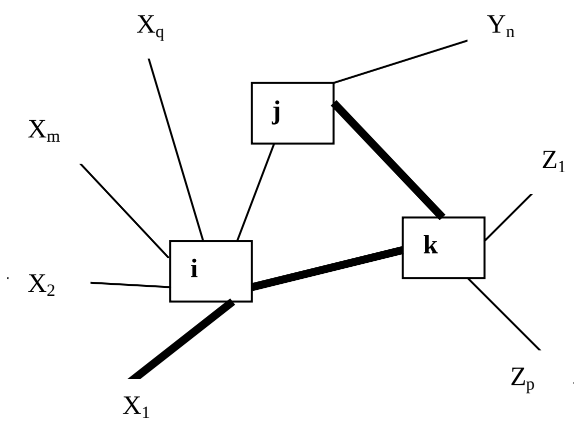

Fig. 3 The isolation of the triangle in the nontrivial graph

Be k the first chosen node. It realizes an initial chain, for example k-i. It appeals to the sets type U and P in this manner:

$U(i) = \{j,k,X_1, \ldots X_q\}$, $P(k) = \{i,j, Z_1, \ldots Z_t\}$

Case I. If $\{X_m\} \cap \{Z_p\} = \Phi$ then:

$$[U(i) \setminus \{k\}] \cap [P(k) \setminus \{i\}] = \{j\}$$

which permits the taking into consideration, for the continuation, of j-k-i or k-i-j. Be one of the alternatives j-k-i. Now the last node is i again and the penultimate is k so it must again take into consideration the sets U(i) and P(k) above, under the following circumstances :

$$[U(i) \setminus \{j,k\}] \cap [P(k) \setminus \{i,j\}] = \acute{\O}$$

and, according to the procedure, it permits the choice, for the co-optation in the chain, of the node X_1 (as a matter of fact any of the junction nodes from the incidence matrix of the node i) with the mention that in the graph there is compulsory a polygon instead of a triangle in order to assure another junction between the node X_1 and any of the nodes $\{Z_p\}$ because, *if there isn't this polygon, then the node i or the node k would be articulation points* (which might be contrary to the working hypothesis) [4]. Consequently the chain has become j-k-i-X_1

Case II. If $\{X_m\} \cap \{Z_p\} \neq \Phi$ then there is necessarily in the graph another triangle than (i,j,k) and, according to the procedure, this node X_m (identical to one of the nodes $\{Z_p\}$) is recommended to be co-opted in the structure of the chain. Consequently, the chain has become k-i-X_m when k, i and X_m form a triangle.

In both cases, a certain triangle could be excluded from the graph, because no incidence lot examination will include in the searching horizon:

- Either the nodes i, j, k in the first case;
- Or the nodes k, i, X_m in the second case.

In other words, if the graph is unarticulated, there is the possibility to reach to the nodes which develop from the nodes i, j, k, respectively from the nodes k, i, X_m on other paths. By developing this logic: If the nodes i, j, k for example, had been articulation ones, then the procedure wouldn't have been able to contain in the chain the nodes $\{Z_p\}$ and it would have become inconsistent.

To apply again the procedure there mustn't be any articulation point in the sub application of the graph (in "the rest of the graph"). For conviction, it makes the remark that the exclusion of the (i,j,k) triangle nodes from the graph automatically leads to the exclusion of other arcs, thus– fig. 3 :

- all the edges k-Z_p an i-X_m which can't be co-opted because the chain structure would be changed, plus:
- j-k, k-i and i-X_m which form the chain, plus :
- j-i which would turn the chain into a cycle.

As a conclusion: if Z_p aren't articulation points, the exclusion of the edges and nodes above specified doesn't lead to the forming of artificial articulation points.

Example. The railway Chinese constructors had to face the very difficult soil conditions from Tibet: apart from the permafrost bedding– inappropriate to the infrastructure – had to exceed marshy areas which block the normal routing of the railway bed, though under the attendance obligation of all the localities on the platform. Obviously the transport service is done in both senses of the railway conjunctions and the application of the algorithms which determine the Hamilton route faces difficulties. By the present procedure it is possible to easily determine the longest chain in the following graph which represents the situation on the ground:

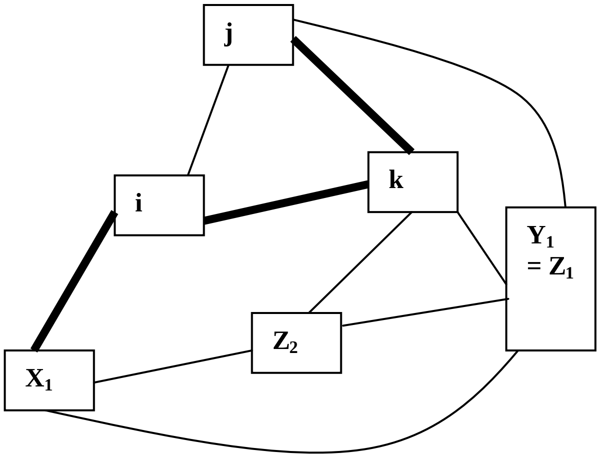

Fig. 4 The demonstration exemplification

In the variant j-k-i-X_1 it obtains the following sub-application of the graph:

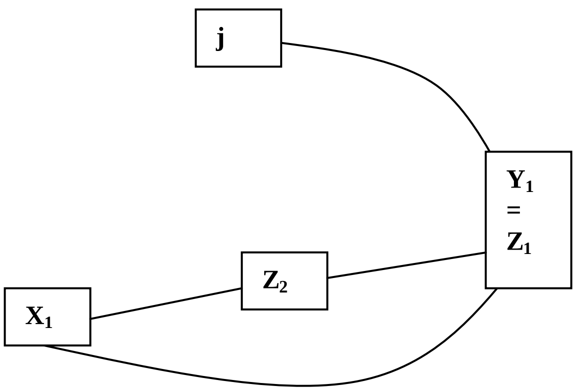

Fig. 5 Sub-application of the graph in figure 4

Because the chain continuation will be done from the node j or from the node X1 the node Y_1 can't be considered an artificial articulation point, because:

- either it continues from the node j and the edge i-Y_1 pieces together the partial graphs "excluded" (j-k-i-X_1) and "left";
- or it continues from the node X_1 and the edge i-Y_1 won't be part anymore from the chain because the procedure goal is to determine a chain, not a circuit;

Furthermore, if the node i had been also articulated by other edge to one of the other nodes, then it couldn't have raised a question about the node Y_1 as an articulation one (namely the problem is reduced, in all the complex cases, to the considerations above).

III. THE ALGORITHM OF THE LONGEST CHAIN IN A GRAPH

There are also graphs for which, *by choosing at random the continuation* nodes from the step 2 specified in paragraph, it may reach to the finalization of the chain before co-opting all the nodes, although this can be possible through other choice. For the graph in fig. 6 (the chain being "blocked" in node 7 before the finalization): if, after the node 8 it had chosen the continuation, the complete chain would have been realized, namely the chain which would have contained all the nodes.

To "unblock" the situation, *the procedure offers the solution:*

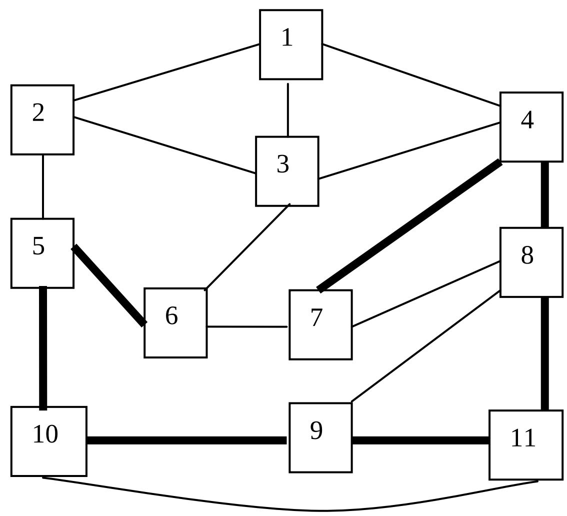

Fig. 6 The premature "blocking" of the chain

- the chain continues, reaching even the admissible limit, namely the chain 6-5-10-9-11-8-7-6 rounded, getting a circuit;
- another chain or, possibly, another circuit is formed by the nodes that have left – fig.7.

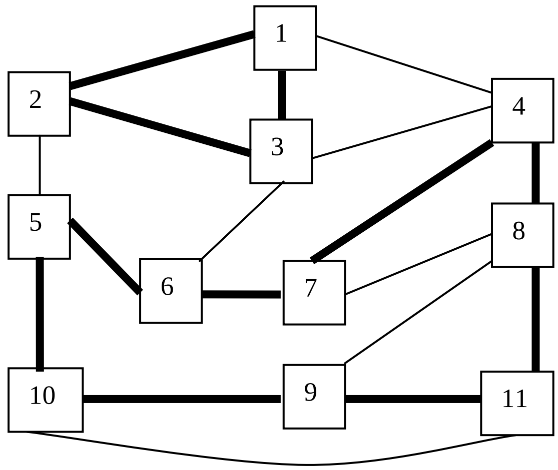

Fig. 7 The multiplication of the circuits

- in order to get the longest possible chain, 3-1-2-5-10-9-11-8-4-7-6, by the concatenation of the two circuits-fig. 8.

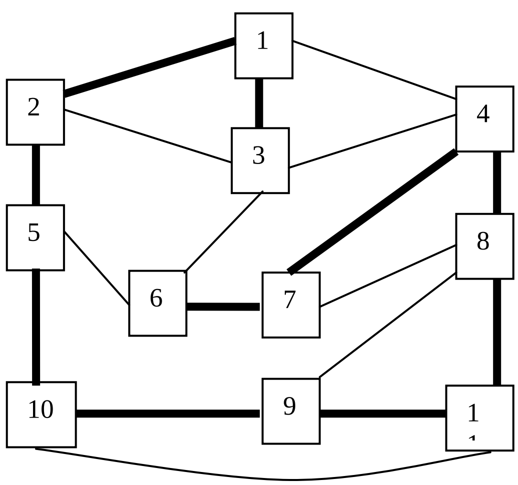

Fig. 8 The maximal chain (with 10 edges)

This type of solution – in two phases – inevitably leads to the practical modality to solve the problem of the longest chain: namely by EXFOLIATING a graph., which means:

- for all graphs it can easily determine what is its peripheral structure (that combination of angles and nodes which separate the loops of the graph from the space which is exterior to the graph – in fact one of the loops of the graph, the exterior loop);
- the removal of the peripheral structure reduces the initial graph to one of its sub applications :

a. either a graph to which it can repeat the exfoliating procedure;
b. or one or more trees;
c. or one or more isolated nodes;
d. or a combination of these cases.

- joining reasonably the “coatings” resulted from the exfoliation, it can determine the longest chain (searched).

Concretely, for the graph in fig. 8, in a uniform representation (without bold lines), it must exfoliate the first coating formed by the nodes 1-2-5-10-11-8-4-1 and the corresponding angles which permit these nodes to become a circuit, the application left of the initial graph being of type d) above (a tree 3-6-7 and an isolated node – 9).

In the next section, will demonstrate that – if the graph admits a chain in the structure of which all its nodes can enter – then the concatenation of the succeeding coatings in order to compose the chain is possible. So the solution is easy obtained-fig.9.

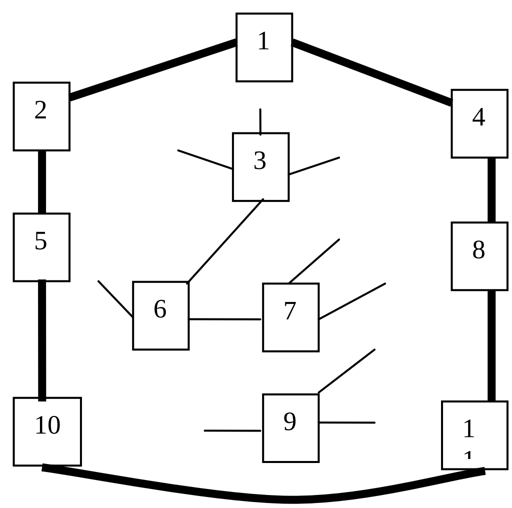

Fig. 9 The first coating – the thickened line – plus the sub-application left after the exfoliation

In the concrete case above: because the node 9 is isolated, it must be considered a bottom of the chain, and the link with the first coating is reasonably realized through the node 11, which call off the continuity of the first coating and it permits, on the other hand, the integration of the tree 3-6-7 in the searched chain.(fig. 10).

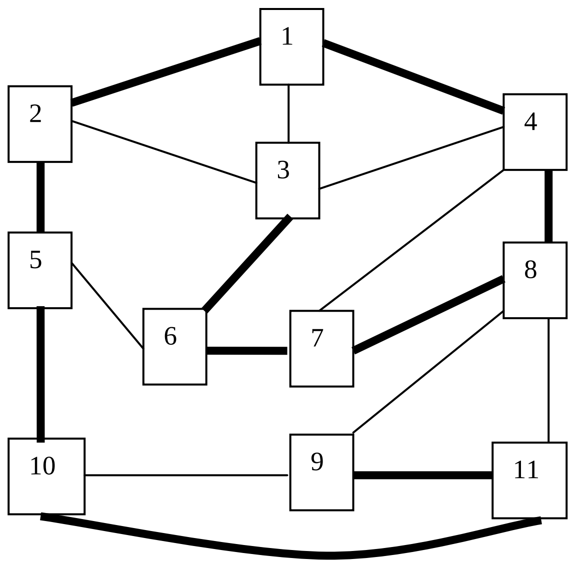

Fig. 10 The chain obtained after the exfoliation procedure.

The procedure that allowed obtaining the complete chain was heuristically exposed. In the next section we offer a mathematically proof for this procedure.

IV. CONCATENATION OF SUB-GRAPHS RESULTED AFTER THE EXFOLIATION PROCEDURE.

Lemma : If a graph admits an elementary chain containing all its nodes, then the concatenation of successive sub-graphs is possible and leads to the determination of the searched chain.

To justify this lemma, the consequences of two main results from graphs theory are used:

Theorem 1. [6] : In every graph exists an even number of odd -degree nodes.(a node has an odd -degree if the number of incident edges is odd);

Theorem 2. (Fleury's Algorithm) [2]: An Eulerian cycle exists and can be determined -if the graph hasn't odd -degree nodes, or an Eulerian path exists and can be determined-if the graph has exactly two odd -degree nodes, respecting the following rules:

- Starting from every node, each time an edge is covered, this is marked not to be considered again
- During the formation procedure of the cycle will never be used an edge which in that moment is an isthmus(the removal of the edge would create two components having only a connection edge between its)

Combining the results of the previous two theorems, we identify two possible cases in the proof of lemma:

Case 1. : The graph hasn't any odd- degree nodes or has exactly two odd- degree nodes.

In this situation, the proof of lemma's statement is a direct consequence of the structure of Eulerian cycle (or path) determined with Theorem 2: *if all the edges are covered, implicitly all the nodes are considered, even repeatedly.*

So, all the sub applications a), b), c) and possibly d) mentioned in the previous section will be part of the Eulerian cycle (or path) as residues , after the removal of the peripheral structure of the studied graph.

By selecting a minimum expanse succession of nodes, but containing all the sub-applications a), b), c) and possibly d), the searched chain is obtained, together with the description of the trail that produced it. This is possible due to the remark that *any sub-application of the analyzed graph must be connected through at least two edges with the peripheral structure or with one of the other structures***,** otherwise the graph being articulated and outside the scope of our search established by the initial hypothesis:

- For every " entrance" in a sub-application of the graph it exists an "exit";
- If all the connection edges between the peripheral structure and the sub-application are covered, then the "entrance" in the sub-application becomes impossible.
- If the procedure reached the peripheral structure:
  - either exists a way to continue through a sub-application,
  - either there are no more sub-application to reach.

Remark: All the previous considerations are referring to a " metagraph" (the graph construction determined from the peripheral structure – represented by a single node and sub-applications a), b), c) and possibly d)).

Exemplifying this technique for the graph presented in fig 2, we obtain the complete chain after the removal of the peripheral structure of the respective graph (edges 3-5, 5-6 and 6-3) and the remained sub-component is presented in fig. 11.

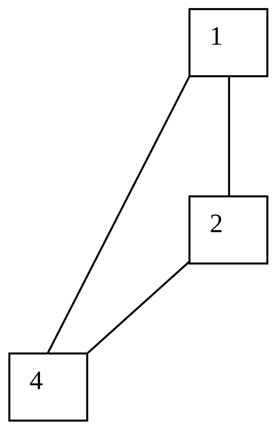

Fig. 11 The remained “inside “graph after the removal of the peripheral structure.

Concatenation of the two components- the peripheral structure and a)-type sub-component is presented in fig 12.

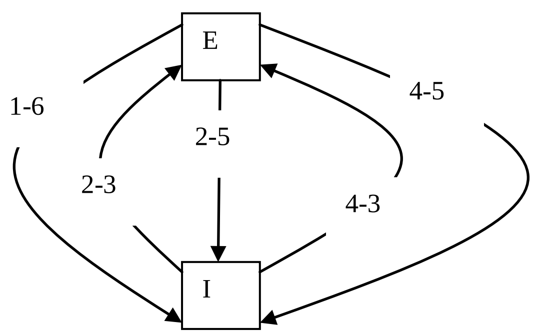

Fig. 12 “Metagraph” obtained by concatenation (with two odd -degree nodes)

Fleury’s Algorithm gives the succession of nodes E – I – E – I – E – I , which stays at the base of the procedure:

- one of the succession (E – I or I – E) is selected;
- give up of the rest of the Eulerian trail;
- obtain the complete chain (fig 13).

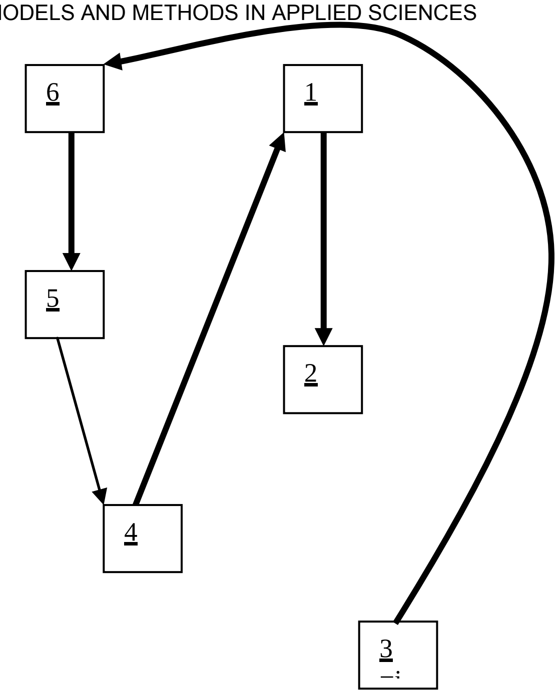

Fig. 13. A version of the chain: peripheral structure edges 3-6 and 6-5; sub-component edges 4-1 and 1-2; concatenation edge 5-4.

Case 2. : The graph has four or more odd -degree nodes.

The basic idea for solving this more complicated situation is to introduce artificial (temporarily) an additional connection which reduces the “metagraph” to a simpler one, with exactly two odd -degree nodes.

To exemplify this procedure, we consider the following graph (fig.14), where E is the notation for the nodes of peripheral structure, Z1, Z2, …, are the isolated nodes of the graph and I is the notation for the “inside” graph we analyze it.

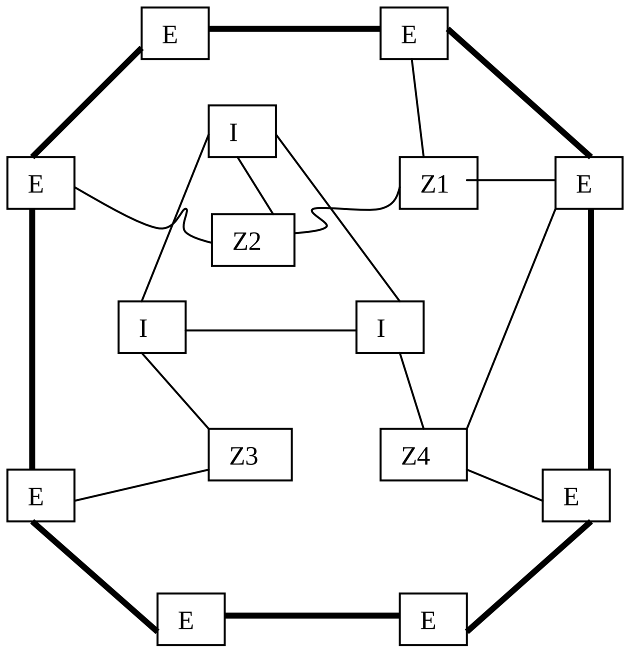

Fig.14. The analyzing graph with peripheral structure, isolated nodes and “inside” graph .

After the removal of the peripheral structure, we obtain a sub-component of the analyzed graph, presented in fig.15.

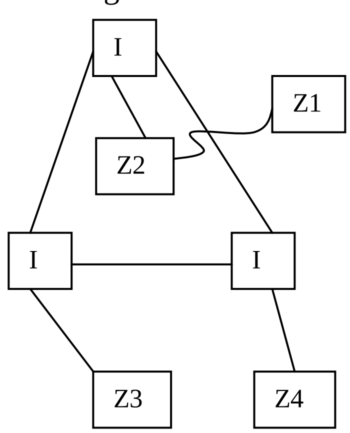

Fig. 15 Sub-component of the analyzed graph, after the removal of the peripheral structure.

The “artificial” connection is created between the nodes Z2 and Z4(fig.16) and must be considered in the sequences of components Z1 – Z2 … … Z4 – E – **Z4 – I – Z2 – E – Z1 – E – Z3 – I** to obtain the complete chain .(fig.17)

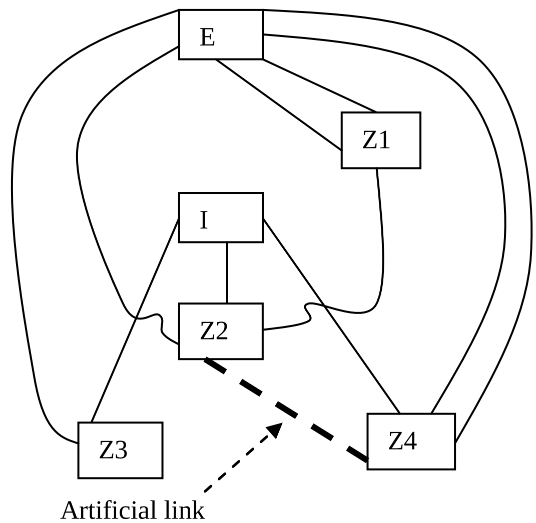

Fig. 16 The “ metagraph”

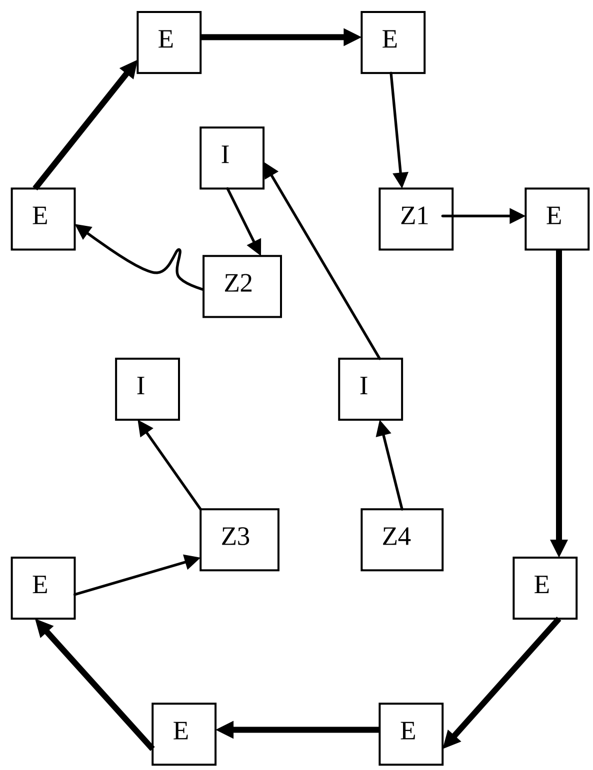

Fig. 17 The complete chain.

The graph-analytic operations previous detailed simplify the analyzed graph, removing only the useless elements of it, because, from our hypothesis, the searched chain exists.

V.CONCLUSIONS

In the theoretical physics it discussed about the problem of the "natural form" of a liquid. The experience proved that, under the circumstances of different liquids difficult to be mixed, but of the same density, the liquid to which they search its form self disposed– after its pouring out without disorders – in the substance of the witness liquid, having the form of a sphere even under the circumstances of the earthly gravitation.

The first problem solved by the determination of the maximal chain of a graph is the one of getting "the natural form" of a graph. For the graph in fig. 2 the longest graph, given by the succession already determined 6-1-2-4-3-5, permits the following representation without a doubt:

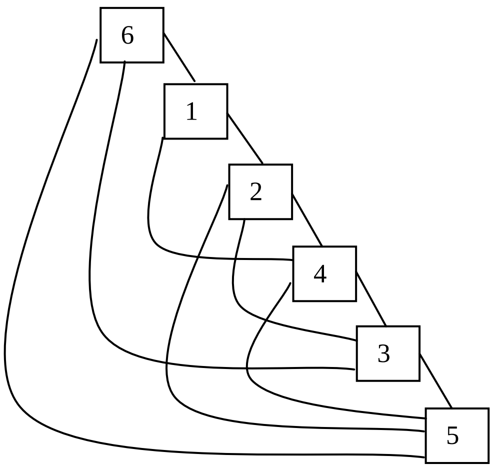

Fig. 18 The "natural" form of a graph

namely : giving as information the number of nodes, their correlations and the structure of the maximal length chain, two analysts can figure the same representation of a graph without image differences – fig. 18 – in opposition with the representation in fig. 2 also correct, but "another" (the number of correspondent graphs being very big). Only the representation based upon the maximal length chain can be identically reproduced by any analyst.

More, the previous representation can be used, for example, to analyze a utilities disposal, in a logistic terminal, so the number of mobile resources fluxes intersections on the terminal area to be minimum possible. The advantages are obvious for the freedom of movement of its, with benefits in time measurements. The method is an alternative to the heuristically methods used in the present for solving the planning problem for terminal areas.

The second problem that the maximal chain can solve is the one of the "Hamilton way". There are methods which permit the determination of the way which passes through all the nodes once, but both the method of the Latin multiplication and the Little method are extremely laborious [2,4]. On the contrary, the procedure based upon the EXFOLIATION of a graph is simple.

From engineering perspective, the benefits of "exfoliation" method proposed are sustained by the following reasons:

- Distribution problems, especially in overbuilding, solved using "Litte & Co" algorithm [14], are modeled with directed graphs theory and even so, the number of successive operations is very large.. Or this hypothesis is artificial since urban roads need both ways movements. By the contrary, the "exfoliation" method followed by the concatenation procedure solves this problem in an elegant way.
- In a urban area there are working places that need to be supplied and for which the space location and the moment represent main objectives, taking in consideration that some of its are in the center of the city or in the suburbs and the crowding moments are different (fig 14 can be a representation). "Exfoliation" method permits to establish an optimal working-program for supply vehicles since, initiating the chain from the center towards suburbs, we obtain a minimum overlap of the congestion traffic. Open problem: can the mileage element be also included in this algorithm?
- The connection between "sales man "problem and longest path in a graph is highly revealed by the method exposed in this article; Salah Elmaghraby [5] describes the significance of determining the longest path in an undirected graph in solving "sales man "problem .

REFERENCES :

[1] C. Berge "Graph theory and its application"(in Romanian), Ed. Tehnica, Bucharest,1969;

[2] A. Kaufmann, "Methods and models of operational research" (in Romanian) Ed. Stiintifica. si Enciclopedica, Bucharest, 1967;

[3] O. Oysten," Graphs and applications "(in Romanian), Ed. Stiintifica. si Enciclopedica, , Bucharest, 1968 ;

[4] I. Tomescu, " Combinations Problems and Graph Theory "(in Romanian),Editura didactica si pededagogica, Bucharest, 1986;

[5] S. E. Elmaghraby, "Some network models in managements": *Lecture notes in operations research*, Ed. Springer-Verlag, New York, 1970, vol. 29;

[6] H. B Maynard, "Industrial Engineering" (in Romanian), Ed. tehnica, Bucharest, 1975;

[7] M.Pirvan, F.Ghionea, C. Flaut , "A mathematical model of the relevant request determination in the transport problem", *Proceeding of the 9th WSEAS International Conference on Mathematics and Computers in Business and Economics (MCBE'08)*, Bucharest, June 24-26, 2008, pg. 105-110;

[8] C. Flaut, F.Ghionea, M. Pirvan , "Optimization of urban transportation network", *Proceeding of the 9th WSEAS International Conference on Mathematics and Computers in Business and Economics (MCBE'08),* Bucharest, June 24-26, 2008, pg. 111-115;

[9] F. Ghionea, M.Pirvan, "Unique Links in Graphs", *Proceeding of the 13th WSEAS International Conference on Mathematics and Computers in Business and Economics(MCBE'12*), Iasi, Romania, June 13-15, 2012, pg.146-151;

[10] M. J.Attalah, D.Z.Chen, "Computing the all-pairs longest chain in the plane", *International Journal of Computational Geometry and Applications*,World Scientific Publishing Company, 1994;

[11] M. H.Albert,"Random graph orders", *Research report* No 88-3, 1988.

[12] S. Encheva, "Connected Graphs Associated with Comparison Matrices", *Management, Marketing and Finances, (MMF'08),* Harvard University, Cambridge, Massachusetts, USA, March 24-26, 2008**,** pg.59-64;

[13] B. Brtnik, "Solving Circuits by Multiple Transformation Graphs", *Proceedings of the 11th Int. Conf. on Mathematical Methods and Computational Techniques in Electrical Engeneerings , (MMACTEE '09*), Vouliagmeni Beach, Athens, Greece, September 28-30, 2009, pg. 63-65;

[14] R. S. Ackoff, "Bases of Operational Research"(in Romanian),, Ed. tehnica, Bucharest, 1975;

[15] A. Spiteri Staines," Rewriting Petri Nets as Directed Graphs", *International Journal of Computers*, Issue 2, Volume 5, 2011, pg. 289-297;