Research Article

Multilayer Network-Based Production Flow Analysis

Tamás Ruppert, Gergely Honti, and János Abonyi

MTA-PE "Lendület" Complex Systems Monitoring Research Group, Department of Process Engineering, University of Pannonia, Veszprém, Hungary

Correspondence should be addressed to János Abonyi; janos@abonyilab.com

Received 24 February 2018; Revised 18 May 2018; Accepted 27 May 2018; Published 16 July 2018

Academic Editor: Dimitris Mourtzis

Copyright © 2018 Tamás Ruppert et al. 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.

A multilayer network model for the exploratory analysis of production technologies is proposed. To represent the relationship between products, parts, machines, resources, operators, and skills, standardized production and product-relevant data are transformed into a set of bi- and multipartite networks. This representation is beneficial in production flow analysis (PFA) that is used to identify improvement opportunities by grouping similar groups of products, components, and machines. It is demonstrated that the goal-oriented mapping and modularity-based clustering of multilayer networks can serve as a readily applicable and interpretable decision support tool for PFA, and the analysis of the degrees and correlations of a node can identify critically important skills and resources. The applicability of the proposed methodology is demonstrated by a well-documented benchmark problem of a wire-harness production process. The results confirm that the proposed multilayer network can support the standardized integration of production-relevant data and exploratory analysis of strongly interconnected production systems.

1. Introduction

Industry 4.0 is a strategic approach to design optimal production flows by integrating flexible and agile manufacturing systems with Industrial Internet of Things (IIoT) technology [1] enabling communication between people, products, and complex systems [2–4]. The integration of manufacturing and information systems is, however, a challenging task [5]. Horizontal and intercompany integration should connect the elements of the supply chain [6], while vertical integration should connect information related to the entire product life cycle [7]. According to this new concept, the improvement and optimization of production technologies based on cyber-physical systems (CPS) are realized by the simultaneous utilization of information related to production systems [8], products, models [9], simulators, and process data [10, 11].

CPS- and Industry 4.0-type solutions also enable the compositions of smaller cells providing more flexibility with regard to production [12]. This idea leads to decentralized manufacturing [13] and emerging next generation machine systems [14]. This trend highlights the importance of the relationship between flexibility and complexity [15].

The complexity of production systems can be divided into the physical and functional domains [16]. To analyze this aspect, our focus is on the production flow analysis of production systems as production analysis has multiple perspectives according to the hierarchical decomposition of the production system: (1) production flow analysis studies the activities needed to make each part and machines to be used to simplify the material flow. (2) Company flow analysis studies the flow of materials between different factories to develop an efficient system in which each facility completes all the parts it makes. (3) Factory flow analysis plans the division of the factory into groups or departments each of which manufactures all the parts it makes and plans a simple unidirectional flow system by joining these departments. (4) Group analysis divides each department into groups, each of which completes all the parts it makes—groups which complete parts with no backflow, crossflow (between groups), and no need to buy any additional equipment. (5) Line analysis analyzes the flow of materials between the
m副主任 each group to identify shortcuts in the plant layout, and (6) tooling analysis tries to minimize setup time by finding sequences that minimize the required additional tooling for the following job [17].

Production flow analysis (PFA) is a technique to identify both groups and their associated "families" by analyzing the information in component process routes which show the activities (often referred as operations) needed to make each activity [18, 19]. Every production flow analysis begins with data gathering during which nonvalue adding activity should be optimized [20]. When dealing with large quantities of manufacturing data, a representational schema that can efficiently represent structurally diverse and dynamical system have to be taken into consideration. Standards like ISO 18629, 10303 (STEP), and 15531 (MANDATE) support information flow by standardizing the description of production processes [21]. Based on these standards and web semantics, a manufacturing system engineering (MSE) knowledge representation scheme, called an MSE ontology model, was developed as a modeling tool for production [22]. The MSE ontology model by its very nature can be interpreted as a labeled network.

A simple multidimensional representation is proposed that can unfold the complex relationships of production systems. Network models are ideal to represent connections between objects and properties [23]. However, as a multidimensional problem that requires flexibility due to the continuously growing amount of information is in question and a new multidimensional approach in the form of a multilayer network [24] is presented.

For the analysis of the resultant ontology-driven labeled multilayer network, techniques to facilitate cell formation and competency assignment for operators were developed.

Manufacturing cell formation aims to create manufacturing cells from a given number of machines and products by partitioning similar machines which produce similar products. Standard cell formation problems handle products and machines while their connections are represented by two-layered bipartite graphs or machines-products incidence matrices. Classical algorithms are based on clustering and seriation of the incidence matrices. Recently, various alternative algorithms have been developed, for example, self-organizing maps [25] of fuzzy clustering-based methods [26]. What is common in most of these approaches is that they only take two variables into account [27]. However, complex manufacturing processes should be characterized by numerous properties, like the type of products and resources, and the required skills of operators should be also taken into account at successful line balancing since the skills of the operators are influencing the speed of the conveyor belt [28]. Dynamic job rotation [29] also requires efficient allocation of the assembly tasks while taking into account the constraints related to the available skills of the operators.

To handle these elements of the production line, the traditional cell formation problem was extended into a multidimensional one. The main idea is to represent these problems by multilayered graphs and apply modularity analysis to identify the groups of items that could be handled together to improve the production process.

An entirely reproducible benchmark problem was designed to demonstrate our methodology. As an example, the problem of process flow analysis of wire-harness production was selected as this product is complex and varies significantly [30] as the geometries and components of the harness vary depending on the final products [31]. Since there are challenges in the selection of the cost-effective design [32] and the demand for flexibility and a short delivery time urge the definition of product families produced from the submodules [33], the problem requires the advanced integration of process- and product-relevant information.

The remaining part of the paper is structured as follows. In Section 2, a multilayer network model is formalized that was developed to represent production systems. In Section 3, how production flow analysis problems can be interpreted as network analysis tasks is discussed. Section 3.2 formalizes the projection of the multilayer networks and studies how conditional connections can be defined, while Section 3.3 applies this projection to calculate the node similarities. The group formation task is described in Section 3.4, where the results of this approach on benchmark examples are also presented. The detailed case study starts in Section 4 with the definition of the wire-harness production use case. The details of the problem are given in the Appendix. Section 4.1 demonstrates the applicability of similarity and modularity analysis. The workload analysis is given in Section 4.2, while interesting applications related to the evaluation of the flexibility of operator-task assignment problems are discussed in Section 4.3. Finally, conclusions are drawn in Section 5.

2. Multilayer-Network Representation of Production Systems

Essential information about the products to be assembled, parts to be manufactured, materials to be used, methods and techniques to convert the material to the required finished components, and manpower to operate the plant is usually available to a company, but rarely in an appropriate form for ease of digestion by the manager [34]. In this section, we propose a network-based model to study the relationship between these elements.

As can be seen in Figure 1, the proposed network consists of a set of bipartite graphs representing connections between the sets of products \( p = \{ p_1, \ldots, p_{N_p} \} \), machines/workstations \( w = \{ w_1, \ldots, w_{N_w} \} \), parts/components \( c = \{ c_1, \ldots, c_{N_c} \} \), activities (operations) \( a = \{ a_1, \ldots, a_{N_a} \} \), and their categorical properties (referred as activity types) \( t = \{ t_1, \ldots, t_{N_t} \} \) and skills of the operators needed to perform the given activity \( s = \{ s_1, \ldots, s_{N_s} \} \).

The relationships among these sets are defined by bipartite graphs \( G_{ij} = (O_i, O_j, E_{ij}) \) represented by an adjacency matrices, where \( O_i \) and \( O_j \) are used as a general representation of a set of objects, as \( O_i, O_j \in \{ p, w, c, a, t, s \} \).

The edges of these bipartite networks can represent material, energy or information flows, structural relationships, assignments, attributes, and preferences, and the edge weights
of production management and transform them into a multidimensional network model. The model is capable of representing information at different levels, so it can support factory flow analysis and departmental flow analysis, or, according to the concept of Industry 4.0, it can also integrate interorganizational supply chains. The development of organizational models is also supported, for this purpose, solutions following the standard of UN/EDIFACT (the United Nations rules for Electronic Data Interchange for Administration, Commerce and Transport) could be used.

The extracted models lend themselves to be handled in the databases of graphs [37, 38] or RDF-based ontologies [39]. In our work, the related technical details of building and storing graph-based decision systems are not the focus; rather, how information from this model can be extracted to support production flow analysis is of concern. In the next section, such techniques are presented.

3. Production Flow Analysis Relevant Operations on Networks

3.1. From Problems of Production Analysis to Tools of Network Science. The main benefit of the multidimensional network model is that it provides a transparent and easily interpretable integration of process- and product-relevant information and as well as facilitating the tools of network science for production flow analysis.

The aim of production flow analysis (PFA) is to identify bottlenecks and groups in products, components, and machines to highlight possible improvements by redesigning the layout, forming manufacturing cells, scheduling the activities, or identifying line families of products based on clustering the sequences of machine usage.

Modules/part families are sets of machines and parts that are highly likely to work together in one group or be processed in a similar order. Since this definition is similar to the concept of modules in networks, it is assumed that fining modules in (multidimensional) networks can be considered as a useful heuristical approach of PFA.

The application of heuristics in PFA is a well-accepted approach since in most cases, the economic benefits are complicated and time-consuming to calculate, and the resultant complex optimization problems are not easy to solve with classical optimization algorithms/operation research tools. In this paper, we suggest that the following network analysis tools should serve as a good heuristic solutions for specific PFA problems:

1. Calculation of the loads and usage frequencies—identification of the bottlenecks
   (i) Calculation of unknown dependencies
   (ii) Analysis of node and edge centralities
2. Group formation—clustering nodes and identifying communities
   (i) Rank-order-based clustering
Table 1: Definition of the biadjacency matrices of the bipartite networks used to illustrate how a production system can be represented by a multidimensional network.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Nodes</th>
<th>Description</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Product (p)-activity (a)</td>
<td>Activity required to produce a product</td>
<td>$N_p \times N_a$</td>
</tr>
<tr>
<td>W</td>
<td>Activity (a)-workstation/machine (w)</td>
<td>Workstation assigned for the activity</td>
<td>$N_a \times N_w$</td>
</tr>
<tr>
<td>A'</td>
<td>Activity (a)-activity (a')</td>
<td>Precedence constraint between activities</td>
<td>$N_a \times N_a$</td>
</tr>
<tr>
<td>B</td>
<td>Product (p)-component/part (c)</td>
<td>Component/part required to produce a product</td>
<td>$N_p \times N_c$</td>
</tr>
<tr>
<td>P</td>
<td>Product (p)-module (m)</td>
<td>Module/part family required to produce a product</td>
<td>$N_p \times N_p$</td>
</tr>
<tr>
<td>C</td>
<td>Activity (a)-component (c)</td>
<td>Component/part built in or processed in an activity</td>
<td>$N_a \times N_c$</td>
</tr>
<tr>
<td>M</td>
<td>Activity (a)-module (m)</td>
<td>Activity required to produce a module</td>
<td>$N_a \times N_m$</td>
</tr>
<tr>
<td>T</td>
<td>Activity (a)-activity type (t)</td>
<td>Category of the activity</td>
<td>$N_a \times N_t$</td>
</tr>
<tr>
<td>S</td>
<td>Activity type (t)-skill (s)</td>
<td>Skill/education required for an activity category</td>
<td>$N_t \times N_s$</td>
</tr>
<tr>
<td>O</td>
<td>Skill (s)-operator (o)</td>
<td>Skills of the operators</td>
<td>$N_s \times N_o$</td>
</tr>
</tbody>
</table>

Table 2: The edge types of the proposed multilayer network.

<table>
<thead>
<tr>
<th>Flow type</th>
<th>Attribute type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>Material, energy, or information flow between the nodes</td>
</tr>
<tr>
<td>Edge weight</td>
<td>Physical attributes of the flow, like quantity, or during discrete events, the frequency of the flow, like the number of hours between events</td>
</tr>
<tr>
<td>Self-loop</td>
<td>Inner activities</td>
</tr>
<tr>
<td>Parallel edges</td>
<td>Multiple flows can be represented by multilayer/multidimensional networks</td>
</tr>
<tr>
<td>Serial connections</td>
<td>Paths of the flow of different entities</td>
</tr>
<tr>
<td>Modularity</td>
<td>Highly cooperative nodes</td>
</tr>
</tbody>
</table>

(ii) Similarity-based clustering
(a) Calculation of node similarities of (projected) networks
(b) Clustering nodes and edges based on the calculated similarities
(c) Joining of clusters of different objects to form modules

(iii) Finding modules in the (multilayer) network

(3) Line formation—ordering modules to minimize sequential transfers
(i) Ordering based on the ratio of in/out degrees—Hollier’s method [40]
(ii) Application of graph layout techniques
3.2. Projections of the Multilayer Network and Calculation of Undefined Connections. As Figure 3 illustrates, when relationships among the $O_i$ and $O_j$ sets are not directly defined, it is possible to evaluate the relationship between its $O_{ik}$ and $O_{jl}$ elements as the number of possible paths or the length of the shortest path between these nodes.

In the case of connected unweighted multipartite graphs, the number of paths intersecting the $O_0$ set can be easily calculated based on the connected pairs of bipartite graphs as

$$A_{O_0}[O_0, O_i] = A[O_0, O_i]^T \times A[O_0, O_j].$$

(1)

Conditional connections could also provide useful information in terms of PFA. To demonstrate the problem, let us have a look at Figure 4 which shows the network defined in (2). In this example, although operators $o_1$ and $o_3$ do not share any machines, the fact that machines $m_1$ and $m_2$ produce identical products results in the $A[O_2|O_1(O_0, O_j)]$ projection operators defining a connection between these operators.

<table>
<thead>
<tr>
<th>Table 3: Node types of the proposed network.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event type</td>
</tr>
<tr>
<td>Fundamental properties</td>
</tr>
<tr>
<td>Node degree</td>
</tr>
<tr>
<td>Modularity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4: Node edge matchings in the proposed network.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow type (edges)</td>
</tr>
<tr>
<td>Event type (nodes)</td>
</tr>
<tr>
<td>Resource type (nodes)</td>
</tr>
<tr>
<td>Competency type (nodes)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5: The ADACOR predicates can be directly applied to define layers of the network [36] (please note that we use the term activity to refer to operations).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicates</td>
</tr>
<tr>
<td>ComponentOf$(x,y)$</td>
</tr>
<tr>
<td>Allocated$(x,y,t)$</td>
</tr>
<tr>
<td>Available$(x,y,t)$</td>
</tr>
<tr>
<td>RequiresTool$(x,y)$</td>
</tr>
<tr>
<td>HasTool$(x,y,t)$</td>
</tr>
<tr>
<td>HasSkill$(x,y)$</td>
</tr>
<tr>
<td>HasFailure$(x,y,t)$</td>
</tr>
<tr>
<td>Precedence$(x,y)$</td>
</tr>
<tr>
<td>UsesRawMaterial$(x,y)$</td>
</tr>
<tr>
<td>RequestSetup$(x,y)$</td>
</tr>
<tr>
<td>HasProcessPlan$(x,y)$</td>
</tr>
<tr>
<td>OrderExecution$(u,x,w,y)$</td>
</tr>
<tr>
<td>HasRequirement$(x,y)$</td>
</tr>
<tr>
<td>HasGripper$(x,y,t)$</td>
</tr>
<tr>
<td>ExecutesOperation$(x,y)$</td>
</tr>
</tbody>
</table>

3.2. Projections of the Multilayer Network and Calculation of Undefined Connections. As Figure 3 illustrates, when relationships among the $O_i$ and $O_j$ sets are not directly
Formally, in some cases, the $A[O_i|O_j(O_p, O_q)]$ conditional projections might be of interest defined by

$$A[O_i|O_j(O_p, O_q)] = A[(O_j, O_k)] \times (A[(O_p, O_q)])^T \times A[(O_j, O_k)]^T,$$

where the resultant $A[O_i|O_j(O_p, O_q)]$ network states that the $i$th property set is analyzed based on the $A[O_j|O_i(O_p, O_q)]$ inner network defined by the inner projection of the objects to the $j$th set.

The projections are not applicable for all types of edges (e.g., the projection with precedence constraints does not result in interpretable networks). Generally, the projections calculate the number of paths between the nodes which number is directly interpretable (e.g., it can reflect the number of assignable operators for a given workstation).

To support these calculations, it is beneficial to utilise the adjacency matrix of the whole multiplex network obtained by flattening or matricization:

$$A_{M} = \begin{bmatrix}
A_{1,2} & \cdots & A_{1,N} \\
A_{2,1} & \cdots & A_{2,N} \\
\vdots & \ddots & \vdots \\
A_{N,1} & \cdots & A_{N,2} & \cdots & 0_N
\end{bmatrix},$$

where $A_{i,j}$ is used to represent the $A[O_p, O_q]$ biadjacency matrices of the $G_{i,j}$ bipartite graphs.

3.3. Calculation of Node Similarities. Node similarities can reveal useful information with regard to PFA, for example, if the similarities of the machines need to be defined based on how many common parts they are processing. When the machines are denoted as $k$ and $j$, and $S_k$ and $S_j$ as the sets of parts that are connected to these machines, the similarities of the machines can be evaluated according to the Jaccard similarity index [41]:

$$\text{sim}(k, j) = \frac{|S_k \cap S_j|}{|S_k| + |S_j| - |S_k \cap S_j|}. \quad (5)$$
The proposed network-based representation is also beneficial in similarity analysis. When \( O_o = w \) represents the set of machines/workstations and \( O_i = c \) represents the set of components, the \( a_{ij} \) is an edge weight stored at the intersection of the \( j \)-th row and \( i \)-th column of the \( A[O_o, O_i] \) biadjacency matrix. This represents that the \( i \)-th type of component is built in at the \( j \)-th workstation and the degree of the \( j \)-th node, \( k_j = \sum a_{ij} \), is identical to the cardinality of the \( |S_j| \) set, which means how many component types are built in at the \( j \)-th workstation.

We can generate two projections for each bipartite network. The first projection connects two \( O_i \) nodes (in our case, two workstations) by a link if they are linked to the same \( O_o \) node (same components). As Figure 5 illustrates, the \( |S_i \cap S_j| \) cardinality is identical to the \( j-k \) edge weight of the projected network which represents how many identical components are built in at the \( k \)-th and \( j \)-th workstation:

\[
A_{O_o, [O_o, O_i]} = A[O_o, O_i]^T \times A[O_o, O_i]. \tag{6}
\]

The second projection connects the \( O_i \) nodes (in our case, two components/parts) by a link if they connect to the same \( O_o \) node (workstations), which projection represents how parts are connected by the machines:

\[
A_{O_i, [O_o, O_i]} = A[O_o, O_i] \times A[O_o, O_i]^T. \tag{7}
\]

When the similarities of more layers are taken into account, multiple projections on the same machines can be defined by the weighted sum of their projections:

\[
A[O_0, O_o] = \sum_{\Gamma} \omega_{\Gamma} A[O_0, O_i] \times A[O_0, O_i]^T. \tag{8}
\]

### 3.4. Identifying Modules for Group Formation.

Communities are locally dense connected subgraphs in a network, so nodes that belong to a community have a higher probability to link to the other members of that community than to nodes that do not belong to the same community. Our key idea is that finding communities in (multilayer) networks of the proposed models can be used to solve group/cell formation problems of PFA. To formalize the cell formation problem, we utilized the modularity measure introduced by Newman [42] and improved for bipartite graphs by Barber [43].

A module of the network consists of a subgraph whose vertices are more likely to be connected to one another than to the vertices outside the subgraph. Modularity reflects the extent, relative to a random configuration network, to which edges are formed within modules instead of between modules. The modularity can be determined for each community of a network (in PFA, this means the modularity of each production cell can be calculated). For a network with \( n_c \) communities, the following modularity value is used to determine the modularity value of community \( C_c \) in terms of each \( C_c \) community with \( N_c \) nodes connected by \( L_c \) links, \( c = 1, \ldots, n_c \):

\[
Q_c = \frac{1}{L_c} \sum_{(i,j) \in C_c} \left( a_{ij} - \frac{k_i k_j}{L_c}\right) = \frac{L_c - k_i k_j}{L_c}. \tag{9}
\]

If the \( Q_c \) modularity value of a cluster is a positive value, then the subgraph \( C_c \) tends to be a community. The modularity of the full network can be evaluated by summing \( Q_c \) over all \( n_c \) communities, \( Q = \sum Q_c \).

As can be seen, the definition of modularity perfectly fits the problem of manufacturing cell formation. Therefore, we propose a graph modularity maximization-based approach for this purpose. In this study, we adapt the Newman [42], LP-BRIM [44], and adaptive BRIM [43] algorithms available in the BiMAT MATLAB toolbox [45].

To illustrate the applicability of this approach, Figure 6 visualizes a cell formation problem and how the extracted modules can be assigned as manufacturing cells.

The efficiency of the formation of the cell can be evaluated based on \( e \), the total number of activities; \( e_o \), the number of exceptional elements that are excluded from the cells; and \( e_v \), the number of zeros in the cells [46]:

\[
\Gamma = \frac{e - e_o}{e + e_v}. \tag{10}
\]

Table 6 compares the efficiencies of cell formation achieved by the proposed clustering and the modularity-based algorithms of cell formation with recently developed advanced goal-oriented optimization results in several benchmark problems of [46]. As can be seen, modularity-based algorithms perform surprisingly well, the \( \Gamma \) values (given as rounded parentages) are near to the optimized performances, and most importantly, the number of machine-part matchings outside of the modules (\( e_o \) values) and the number of modules are much smaller in almost all cases than the optimized reference solutions.
Based on this success, several modularity optimization algorithms were applied. As will be demonstrated in the following section, the approach is also applicable when searching for modules in multiple layers by the multilayer InfoMap algorithm \[47, 48\].

## 4. Application to the Analysis of Wire-Harness Production

To provide a detailed and reproducible case study for production flow analysis, an open-source benchmark model of modular wire-harness production was developed. The details of the model are given in the Appendix. The multilayer network model of the production flow analysis problem is formed and analyzed in the MuxViz framework developed for the interactive visualization and exploration of multilayer networks \[49\]. The established network is depicted in Figure 2.

### 4.1. Similarity and Modularity Analysis

Analysis of the reducibility of a multilayer network provides useful information about the similarities of the layers \[50, 51\]. To demonstrate the applicability of this metric, the C, Z, S, O, and T layers were analyzed (see Figure 7).

As can be seen in Figure 8, based on the reducibility of the network two clusters were formed. The first cluster is related to product-process (Z-T-C) layers, while the second collects the operator-skills- (O-S-) relevant information. The importance of the definition of the activity types (layer T) is also highlighted.

Although our network defines part families indirectly in layer M and also groups of these activities (in layer T), it is interesting to observe how the multilayer network is structured and how the analysis of the modularity of the network can form part and activity groups. For this purpose, a multilayer InfoMap algorithm was applied \[47, 48\].

![Figure 6: Modularity analysis of the 30 × 41 machine-part benchmark example.](image-url)

### Table 6: Cell formation efficiency of bipartite modularity optimization algorithms. The Θ values are given as rounded percentages.

<table>
<thead>
<tr>
<th>Problem size</th>
<th>Optimization [46] (Number of c, Θ [%])</th>
<th>Newman (Number of c, Θ [%])</th>
<th>LP-BRIM (Number of c, Θ [%])</th>
<th>Adaptive BRIM (Number of c, Θ [%])</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 × 24</td>
<td>7, 72</td>
<td>4, 67</td>
<td>4, 67</td>
<td>2, 67</td>
</tr>
<tr>
<td>20 × 20</td>
<td>5, 43</td>
<td>4, 41</td>
<td>4, 40</td>
<td>4, 40</td>
</tr>
<tr>
<td>24 × 40</td>
<td>11, 53</td>
<td>7, 41</td>
<td>7, 40</td>
<td>8, 43</td>
</tr>
<tr>
<td>28 × 46</td>
<td>10, 45</td>
<td>4, 37</td>
<td>3, 33</td>
<td>5, 39</td>
</tr>
<tr>
<td>30 × 41</td>
<td>10, 59</td>
<td>6, 45</td>
<td>7, 51</td>
<td>8, 52</td>
</tr>
<tr>
<td>30 × 50</td>
<td>12, 60</td>
<td>9, 44</td>
<td>10, 47</td>
<td>9, 44</td>
</tr>
<tr>
<td>37 × 53</td>
<td>3, 59</td>
<td>4, 49</td>
<td>3, 53</td>
<td>2, 53</td>
</tr>
</tbody>
</table>

Based on this success, several modularity optimization algorithms were applied. As will be demonstrated in the following section, the approach is also applicable when searching for modules in multiple layers by the multilayer InfoMap algorithm \[47, 48\].
The analysis yielded useful and informative results. 26 modules were identified. Although layer M which represents how the activities are grouped according to different products, this analysis was able to detect the modules of the products \((m_1, \ldots, m_7)\) in terms of the types of the activities \((t_1, \ldots, t_{16})\). This result confirms that the analysis of the modularity of the proposed multilayer network model is useful in fine-tuning the existing part families based on multiple aspects representing the layers of the model.

To demonstrate how such information is useful in the early process-design phase to define technical modules, layer T of the C-Z-S-O-T multilayer network is shown in Figure 9. As can be seen, the most significant module is separated into...
six smaller groups by following the structure of layer $Z$ that defines in which zone the activities occur. The central role of the most frequent and widely distributed $t_{10}$ type of activity (wire-terminal attachment) is also highlighted.

4.2. Workload Analysis. The balancing of modular production is challenging due to the great diversity of products [52]. Besides group formation, the analysis of the workloads is also an important task in production flow analysis. The proposed bipartite network-based model can be directly applied for this purpose as the biadjacency matrices of the layers result in simple calculations. To illustrate this applicability, let us consider the analysis of how well the production line is balanced. The equation $L_w = \text{MP}_p$ represents the activities of the production of the $p$th product (where $P_p$ represents the $p$th column of the $P$ product-module matrix). As these activities are assigned to the workstations as $L_w = \text{diag}(L_w)W$ and $T^\top L_w$, represents the number of activities grouped by activity types and $T^\top CC^\top L_{w_p}$ is the number of built-in components at a given workstation. The central role of the most frequent and widely distributed $t_{10}$ type of activity (wire-terminal attachment) is also highlighted.

As Figure 10 illustrates, the calculations above can be used to check how the process is balanced and how the complexity of the product influences the workloads of the workstations.

Although the presented workload analysis is not unique to the proposed model, we believe that the results demonstrated the rich information content and broad applicability of multilayer networks which can also be interpreted as a linear algebraic approach model of the system.

4.3. Analysis of the Flexibility of Operator Assignment. In the early 80s, [53] suggested that organisational research should incorporate network perspective. In the early 90s, six themes (turnover/absenteeism, power, work attitudes, job design, leadership, and motivation) dominated the research of microorganisational behaviour [54]. Recently, multilayer networks are becoming widely used in the analysis of social networks where people interact with each other in multiple ways like via mobile phone and emails [55–59]. In this paper, we make the first attempt to integrate such analysis to the modelling and optimisation of production process.

For successful line balancing of wire-harness production, the skills of the operators influencing the speed of the conveyor belt should also be studied [28] and handled [60]. Dynamic job rotation [29] requires efficient allocation of the assembly tasks while taking into account the constraints related to the available skills of the operators. Figure 11 shows the distribution of the required skills as a function of different product modules, $M^TSO$. As can be seen, the most in demand is the $s_3$ terminal-attaching skill, while $s_6$ is the visual testing skill which is required only once during production. The abilities of the operators can also be calculated, for example, $W'TSO^T$ yields how many activities can be performed at a given operator-workstation assignment (see Figure 11(a)).

The presented analysis can be useful in designing the sessions of the operators by determining the components of critical skills and knowledge. Figure 12 shows the layers $S$ and $O$ of the network. Five groups of activity, skill, and operator nodes were identified with the help of multilayer modularity analysis. The smallest module contains the $t_{15}$ clip installation activity type which requires specialist skills.

As can be seen, the skill $s_5$ can be considered a key piece of knowledge, because it is related to five types of activities. Operators $o_9$ and $o_{10}$ possess specialist knowledge, while $s_1$ consists of group-wise knowledge because it is the most related to the operators.

**Figure 10**: The workloads (number of activities, built-in components, and total activity times) can be easily calculated based on the biadjacency matrices of the proposed model, which supports the balancing of the conveyor belt.

![Figure 10](image-url)
The presented analysis demonstrated that the analysis of the node degrees can identify the critically essential skills and resources. Skills that have small degrees in the O layer can be considered as the knowledge of specialists, while skills with large degrees are quantified as group-wise knowledge. Skills that have no links at the S layer are useless, while skills that have a small degree at the O layer and high degree at the S layer are critical, as this reflects that a small number of operators can be assigned to a large number of tasks which requires this knowledge.

5. Conclusions

A multilayer network model was developed for production flow analysis to represent the physical and functional domains of production systems by taking into account the aspects of the structure of the system, the variety of machines, products, components, and operators and their interdependencies.

Most of the layers of the model are represented by a bipartite graph, where edges represent material, energy, or information flows and attributes of the objects represented by the nodes of the graph. It was highlighted that the nodes and connections could be easily defined based on standards of process management. As the layers of the network represent different aspects of the production system, the proposed model is flexible and easily extendable.

Following the introduction of the new modeling concept, it was demonstrated how the tools of network science should be used to support production flow analysis. Firstly, it was shown that the analysis of the paths in the network provides...
Useful information about hidden, previously undefined connections. It was recognized that modularity analysis of the network is a promising tool for forming groups in PFA, and the performances of advanced (bipartite and multilayer) network modularity algorithms (like InfoMap) are comparable to the most advanced optimization algorithm tailored to the problem of cell formation.

A detailed benchmark problem was developed to make the research of multivariable algorithms of production flow analysis reproducible. With the help of the studied wire-harness process, the benefits of the modularity analysis of problem-specific sets of layers were demonstrated. The results confirm that the detected groups of activities are useful in terms of fine-tuning of modules (part families). Workload and capability-related network measures were developed. Along with analysis of the node degrees and their correlations, individual-, key-, and group-wise skills could be identified. The biadjacency matrices of the network lead to the calculation of workloads, and the investigation of how the production line is balanced. Besides the numerical analysis, visualizations were presented to demonstrate how multilayer networks provide insights into the critical factors of interconnected production systems, and the results of which confirm that multilayer networks can support the integration of production-relevant data and decision-making related to complex production systems.

Since the handling of the time-varying behaviour of process systems is becoming ever more critical in the field of cyber-physical systems, our future work will focus on the integration of historical process data to define networks of sequential procedures and temporal connections.

Appendix

Details of the Wire-Harness Production Technology

To support the reproducible development of production flow analysis and optimization algorithms, an open-source benchmark problem of a modular wire-harness production system was developed. The core of the system is a paced conveyor shown in Figure 13. Based on data published in [32, 61], $N_p$ was based on 64 products and defined $N_m$ as a combination of 7 modules: $m_1$ base module, $m_2$ as left- or right-hand drive, $m_3$ normal/hybrid, $m_4$ halogen/LED lights, $m_5$ petrol/diesel engine, $m_6$ 4 doors/5 doors, and $m_7$ manual or automatic gearbox. $N_a$ was defined 654 activities/tasks categorized into $N_i$, which consisted of 16 activity types with well-modeled activity times (see Table 7). In these activities, $N_i$ was equal to 64 different built-in part families (component types), and $N_a$ was defined as 63 bandages, $C_1 = 180$ terminals, $C_2 = 63$ wires 4.6, $C_3 = 25$ clips, and $C_4 = 90$ wires). The conveyor $N_w$ consisted of 10 workstations (tables). For every table (workstation), one operator is assigned, $N_o = 10$. The required $N_r$ was also defined as 6 skills of the operators, namely, $s_1$—laying cable, $s_2$—spot-tying, $s_3$—terminal attaching, $s_4$—connector installing, $s_5$—clip installing, and $s_6$—visual testing. $N_z$ was also defined as 6 zones for the workstations (see Figure 14) to study the distribution of

**Table 7: Types of activities and the related activity times [61].** The activity times are calculated based on fixed and proportional values, for example, when an operator is laying four wires over one foot, according to the $t_4$ model, the activity time will be $1\times 6.9\ s + 4\times 4.2 = 23.7\ s$.

<table>
<thead>
<tr>
<th>ID</th>
<th>Activity</th>
<th>Remark</th>
<th>Unit</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>Point-to-point wiring on chassis</td>
<td>Direct wiring</td>
<td>Number of wires</td>
<td>4.6</td>
</tr>
<tr>
<td>$t_2$</td>
<td>Laying in U-channel</td>
<td></td>
<td></td>
<td>4.4</td>
</tr>
<tr>
<td>$t_3$</td>
<td>Laying flat cable</td>
<td></td>
<td></td>
<td>7.7</td>
</tr>
<tr>
<td>$t_4$</td>
<td>Laying wire(s) onto harness jig</td>
<td>Laying flat cable</td>
<td>Base time</td>
<td>6.9</td>
</tr>
<tr>
<td>$t_5$</td>
<td>Laying cable connector (one end) onto harness jig</td>
<td>To the same breakout</td>
<td>Per wire</td>
<td>7.4</td>
</tr>
<tr>
<td>$t_6$</td>
<td>Spot-tying onto cable and cutting it with a pair of scissors</td>
<td></td>
<td></td>
<td>16.6</td>
</tr>
<tr>
<td>$t_7$</td>
<td>Lacing activity</td>
<td></td>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td>$t_8$</td>
<td>Taping activity</td>
<td></td>
<td></td>
<td>3.6</td>
</tr>
<tr>
<td>$t_9$</td>
<td>Inserting into tube or sleeve</td>
<td></td>
<td></td>
<td>1.8</td>
</tr>
<tr>
<td>$t_{10}$</td>
<td>Attachment of wire terminal</td>
<td>Terminal-block fastening (fork lug)</td>
<td></td>
<td>22.8</td>
</tr>
<tr>
<td>$t_{11}$</td>
<td>Screw fastening of terminal</td>
<td></td>
<td></td>
<td>17.1</td>
</tr>
<tr>
<td>$t_{12}$</td>
<td>Screw-and-nut fastening of terminal</td>
<td></td>
<td></td>
<td>24.7</td>
</tr>
<tr>
<td>$t_{13}$</td>
<td>Circular connector</td>
<td>Installation only</td>
<td></td>
<td>11.3</td>
</tr>
<tr>
<td>$t_{14}$</td>
<td>Rectangular connector</td>
<td>Latch or snap-on</td>
<td></td>
<td>24.0</td>
</tr>
<tr>
<td>$t_{15}$</td>
<td>Clip installation</td>
<td></td>
<td></td>
<td>8.0</td>
</tr>
<tr>
<td>$t_{16}$</td>
<td>Visual testing</td>
<td></td>
<td></td>
<td>120.0</td>
</tr>
</tbody>
</table>
the fixtures on the tables. The related Z matrix is defined based on the layout of the table and shows the relationship between the activities and zones of the workstation, which facilitates a detailed analysis of the workload in the workstations. All of this information is represented by a set of bipartite graphs defined in Table 1 and depicted in Figure 2. The related dataset is freely and fully available on the website of the authors: https://www.abonyilab.com.

Data Availability
The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
The authors declare that no conflict of interest exists with regard to the publication of this paper.

Acknowledgments
This research was supported by the National Research, Development and Innovation Office (NKFIH), through Project OTKA-116674 (process mining and deep learning in the natural sciences and process development) and the EFOP-3.6.116-2016-00015 Smart Specialization Strategy (S3) Comprehensive Institutional Development Program.

References


