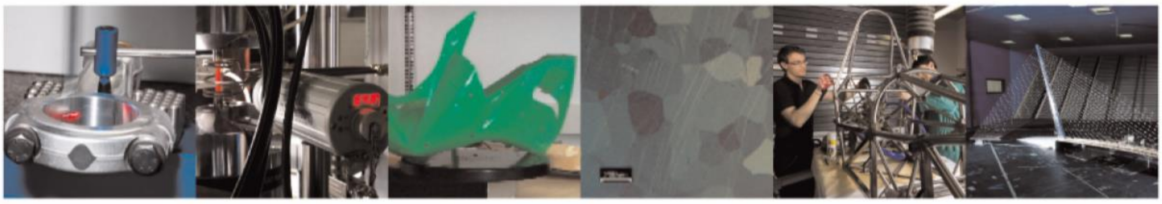

Integrated production and reconfiguration planning in modular plug-and-produce production systems

Colledani, Marcello; Angius, Alessio

This is a post-peer-review, pre-copyedit version of an article published in CIRP Annals. The final authenticated version is available online at: http://dx.doi.org/10.1016/j.cirp.2019.04.087

Integrated production and reconfiguration planning in modular plug-and-produce production systems

Marcello Colledani[1](2), Alessio Angius[1]

1 Politecnico di Milano, Department of Mechanical Engineering, Via la Masa, 1, 20156, Milan, Italy.

Modular plug-and-produce production systems have been proposed as promising architectures to face the challenge of evolving market requirements, large product variety and small lots. These systems enable fast reconfiguration through on-line production modules substitution. However, such capability poses challenges at planning level, as the sequencing of lots and the selection of production modules need to be performed simultaneously. This paper proposes an integrated method for production and reconfiguration planning combining stochastic lot completion time distribution analysis and lot sequence optimization to maximize the system service level. The approach is validated in a real industrial system producing hydraulic valves.

reconfiguration; production planning, optimization.

1. Introduction, motivation and objectives

Manufacturing companies are facing the challenge of delivering high quality products in large number of variants, with customized product features and in very small lots [1]. In reply to these market requirements, highly reconfigurable and evolvable manufacturing systems, endowed with modularity and scalability technical enablers, have been proposed [2], [3]. In these architectures, the system configuration and processing capabilities can easily evolve depending on the dynamic evolution of product feature and volume requirements [4].

Recently, the advent of industry 4.0 technologies and advanced distributed manufacturing control systems has provided additional plug-and-produce capabilities to these systems, further reducing reconfiguration times and efforts. According to the concept of modular plug-and-produce production systems, production modules can be re-arranged in-line, during the manufacturing system operations. Indeed, specific connectivity hubs are integrated in the system where mechatronic modules can be added and removed with limited effort. Such hubs are endowed with standardized power, compressed air supply, and data connectors, which support automatic module recognition within the control system and self-configuration of the production modules. These features highly reduce the system reconfiguration time as well as the skills required during reconfiguration, making it possible to perform multiple system architecture adjustments during the production shift.

In spite of their benefits, modular plug-and-produce production systems also bring additional challenges at system design and management level, which bound their diffusion in real manufacturing contexts [5]. At design level, the most suitable mechatronic modules need to be selected, considering the expected evolution of product feature and volume requirements [6]. At management level, the available modules need to be properly exploited in order to meet the target service level for the demanded production lots. While the design problem has been widely analyzed in the literature [7], only few approaches have been recently proposed to address system management challenges. In [8] a method to allocate products to reconfigurable lines and to optimally manage production modules is developed, neglecting reconfiguration times and lot sequencing. In [9] a stochastic programming approach has been proposed for designing a modular production system in the automotive industry. The production modules are selected by considering demand scenarios. Once fixed the configuration, the lot sizes are optimized in order to meet the demand. In [10] a method to dynamically re-arrange available production modules to meet the evolving demand is proposed. Although the mentioned papers contributed to the formalization of the problem, the joint definition of production and reconfiguration plans remains a challenging issue for practitioners, limiting the applications of plug-and-produce systems in industry. Indeed, the plug-and-produce capabilities bring the reconfiguration planning problem at the same operational level of the short-term production planning problem, thus requiring an integrated approach to properly capture this interaction.

This paper proposes for the first time a modeling framework and a methodology to jointly plan the production lots and the reconfiguration actions in modular plug-and-produce production systems, given a constraint on the target service level. The benefits of the proposed approach towards less integrated methods are discussed. The application to a real industrial case in the production of hydraulic valves shows that significant benefits can be achieved by properly manage the modular system with the proposed method.

2. System description and modeling assumptions

The considered modular system is formed by N workstations and $N-1$ buffers of finite capacity, configured in serial layout, (Figure 1). Workstations are denoted as w_i, $i=1,..,N$ and buffers (yellow circles) are denoted as b_i with $i=1,..,N-1$. The capacity of each buffer is B_i. Depending on the technological content of each workstation, V different product types can be produced in the system. Each product type v requires T_v different tasks to be processed, $v=1,..,V$.

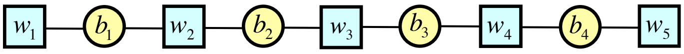

Figure 1. Representation of the analyzed manufacturing system.

A workstation is a standard physical platform, or hub, which is capable to dynamically host mechatronic modules by providing power and compressed air supply, and data flow connection. A mechatronic module is a composition of production units, or mechatronic objects, which has the capabilities to satisfy the requirements of a specific task on a product. For example, an automatic press is a mechatronic module for performing o-ring insertion tasks, composed of several mechatronic objects, among which the pneumatic module, the table, the fixture, and the fixing tool. Let $m_{y,v,j}$ indicate the *y*-th mechatronic module able to perform task $t_{v,j}$, with $j=1,..,T_v$, and $O = \{o_1, o_2, \dots o_K\}$ be the set of mechatronic objects that can be dynamically arranged to compose mechatronic modules. Then, $m_{y,v,j}$ can be defined as a set of couples $\{(x_1, z_1), (x_2, z_2), \dots\}$, such that (x, z) exists only if the mechatronic object $x \in O$ is physically connected to the mechatronic object $z \in (O \cup \{\emptyset\})$, where $\emptyset$ is a placeholder to handle the case in which *x* is not connected to other units.

A given allocation of mechatronic modules on workstations that is able to satisfy the full production requirements of at least one product type is defined as a system configuration, c_l, $l=1,..,L$. Each product type can be produced with at least one system configuration. The set of mechatronic modules populating workstation *i*, with $i=1,..,K$, in configuration *c* is referred as $M_{c,i}$.

Therefore, the considered system is dynamically reconfigurable in the sense that production units can be installed, removed and recombined into modules in order to provide the capabilities to produce different products, by exploiting the plug-and-produce functionalities of standardized workstations.

2.1. Single workstation modeling

Within a configuration, the dynamics of each workstation is modeled by a discrete-time and discrete-state Markov chain of general complexity. This setup allows to analyze a wide set of different workstation models within a unique framework. For example, workstations integrating unreliable mechatronic modules characterized by generally distributed up and down times and also multiple modules with non-identical processing times can be considered within the same framework, thus making the proposed approach applicable to a wide set of real modular production systems.

In detail, each workstation w_i in configuration *c* is represented by $I_{i,c}$ states. The dynamics of each workstation in visiting its states in configuration *c* is captured by the transition probability matrix $\lambda_{i,c}$, that is a square matrix of size $I_{i,c}$. Moreover, a binary quantity reward vector $\mu_{i,c}$ is considered, with $I_{i,c}$ entries. Therefore, state s with $\mu_{i,c}(s)=1$ can be considered as an operational state for workstation *i* in configuration *c*, while state *s* with $\mu_{i,c}(s)= 0$ is a down state for workstation *i*.

The transition probability matrix $\lambda_{i,c}$ and the rewards vector $\mu_{i,c}$ can be directly constructed as the Kronecker product of the transition probability matrices and the reward vectors characterizing the dynamics of the composing mechatronic modules contained in the set $M_{c,i}$ [9].

2.2. System dynamics

A discrete flow of parts is considered in the system. Workstation w_i is blocked if the buffer B_i is full. Workstation w_i is starved if the buffer B_{i-1} is empty. Workstation w_N is never blocked and workstation w_1 starves only when it completes the processing of the last part of the lot.

2.3. Production and reconfiguration plan

Within a planning period τ, a production plan $\{l_1, l_2, .., l_L\}$, composed of production lots of different sizes and types, need to be produced. Thus, $l_i = (a_i, v_i)$ where $a_i > 0$ defines the amount of parts that compose the lot *i*, and v_i determines the product type.

When a lot is completed, the system may experience a down-time due to reconfiguration before switching to the production of the next lot. This stochastic reconfiguration time, $r(c, c')$, is required to implement all the mechatronic object replacement actions that modify the system from configuration *c* to configuration *c'*. The production of the next lot starts when all the parts of the previous lot have been completed and reconfiguration of all the workstations ended.

An integrated production and reconfiguration plan is then characterized by the double (ω,φ), where ω is a vector that provides the sequence of lots and φ is a vector that defines the configurations used for producing the lots, in such a way that if $\omega_j = l_i$ and $\varphi_j = c$, then lot l_i will be produced as *j-th* in the sequence, with the system arranged in configuration *c*.

2.4. Performance measures

Let $A_i(\mathrm{t})$ be the random variable describing the number of parts of the lot l_i completed by the system at the time instant *t*. Then, the completion time of a single lot in isolation corresponds to a random variable $C_i(a_i) = min_t(P(A_i(t) = a_i)$ whereas the completion time of the whole production plan corresponds to $C = \sum_{i=1}^{L} C_i(a_i) + S$, where *S* is the random variable representing the total reconfiguration time.

Then, the main performance measures of interest are:

- $SL_{l_i} = P(C_i(a_i) \leq \tau)$, the lot service level, i.e. the probability that the lot is completed in τ time units;
- $CT_{l_i} = min_{t'}(P(C_i(a_i) \leq t') \geq \varepsilon)$, i.e. the minimal time required to complete the lot with a service level equal to ε;
- *SL*, the service level for the whole production plan, i.e. $P(C \leq t)$.
- *CT*, i.e. minimal time to complete the whole production plan, with a service level equal to ε, i.e. $min_{t'}(P(C \leq t') \geq \varepsilon)$;

2.5. Integrated production and reconfiguration planning problem

The integrated production and reconfiguration planning problem, in a planning period τ, can be stated as one of the following dual formulations:

Problem I: find the plan (ω,φ) that minimizes the completion time *CT* under a target service level SL^{Obj}.

Problem II: find the plan (ω,φ) that maximizes the service level *SL* under a target completion time CT^{Obj}.

3. Description of the method

With the objective to solve the problems stated in Section 2, an innovative method has been developed which is based on the analytic derivation of the lot completion time distribution. This result is exploited to compute the completion time distribution of a given production and reconfiguration plan. Then, a search algorithm is adopted to explore the space of feasible plans in order to find the optimal solution.

3.1. Lot completion time distribution analysis

In this paragraph, the exact analytical method used to compute the completion time of a single lot, l_i, under a given system configuration *c* is described. The rationale of the approach, inspired by the method proposed in [11], is explained in the following. A discrete time, discrete state Markov chain, denoted by Θ, is built that characterizes the overall behavior of the system, in a given configuration *c*. A state of Θ is described by a vector $\boldsymbol{s}=(n_1,n_2,..,n_{K-1},\alpha_1,\alpha_2,..,\alpha_K)$ where n_k is the number of parts in the buffer b_k, and α_k is the state of workstation w_k. The transition probability matrix Q is built as the sum of two matrices, i.e. $Q = \widehat{D} + D$, where $\widehat{D}$ collects the transitions that do not lead to the completion of a part and D contains all the other transitions. These two matrices are used to create an absorbing process that

counts the number of parts left to complete the lot of size a_i. The absorbing process is composed of $(a_i + 1)$ states. The process dynamics is described by a matrix P which is a block matrix that has the matrix $\widehat{D}$ on the diagonal block k and matrix D in the positions $(k, k+1)$, $1 \leq k \leq a_i$. The only not null block at $k = a_i + 1$ is an identity matrix of the same size of D. Level $a_i + 1$ corresponds to the absorbing state in which the lot is completed.
Let $\pi(t)$ be the distribution vector at time t where $\pi_{k,j}(t)$ corresponds to a state j in the block k, then $P(C_i(t) = a_i) = \sum_{j \in |D|} \pi_{a+1,j}(t)$, from which the measure of interest can be derived. The vector $\pi(t)$ is recursively calculated as:

$$\pi(t) = \begin{cases} \pi_{k-1}(t-1)D + \pi_k(t-1)\widehat{D} & 2 \leq k \leq a_i \\ \pi_1(t-1)\widehat{D} & k = 1 \\ \pi_{k-1}(t-1)D & k = a_i + 1 \end{cases} \quad (1)$$

3.2. Production and reconfiguration plan completion time distribution analysis

The completion time distribution for a given production and reconfiguration plan is determined by using the convolution operator (*) and including the reconfiguration times. Formally:

$$P(C \leq t) = P(r(\omega_0, \omega_1) = t) * P(C_1(a_1) \leq t) * P(r(\omega_1, \omega_2) = t) * P(C_2(a_2) = t) \dots * P(r(\omega_{L-1}, \omega_L) = t) * P(C_L(a_L) = t). \quad (2)$$

3.3. Search algorithm

Since the completion time distribution of each lot l_i with configuration c is computed in isolation, the exhaustive exploration of all the possible production and reconfiguration plans reduces to the analysis of each lot with all the possible configurations that can be used to produce it. As a consequence, if $\beta(l_i)$ corresponds to the set of configurations that can produce lot l_i, $\prod_{i=1}^{L} |\beta(l_i)|$ small models instead of $L! \cdot \prod_{i=1}^{L} |\beta(l_i)|$ large models, considering the entire production and reconfiguration plan, are evaluated. The small models are then aggregated by equation (2) and used to evaluate all the possible production and reconfiguration plan combinations. Given its reduced complexity, this algorithm supports the implementation of the developed method within industry relevant application cases.

4. Real case study

The proposed method has been validated through the application to a real industrial case producing hydraulic valves at Bosch Rexroth. The system produces five product families featuring significant differences in terms of processing requirements. Within a family, several product variants are available, originating about 200 different product types. Each product type is produced by performing a sequence of tasks, including mounting, joining, screwing, pressing, stamping, gluing, riveting, and leakage testing. Examples of mechatronic modules adopted in the real case to perform these tasks are reported in Figure 2.


(a) (b) (c)

Figure 2. Example of mechatronic modules adopted in the real case: (a) Automatic press, (b) Screwing spindle, (c) Riveting unit.

In order to analyze the behavior of the system and the benefits of the proposed approach towards less integrated methods, in the first experiment the analysis is limited to a planning period where only four part types are produced, with four feasible system configurations. Each configuration is formed by three workstations, separated by buffers of capacity equal to 3. The association between configurations and product types is reported in Table 1, together with the corresponding total ideal lead time, computed as the sum of the processing times of workstations. Part types 1 and 4 can be produced only with dedicated configurations because their tasks require the installation of specific tools for hammering, riveting and mechanical joining. Configurations 2 and 3 can be instead used to produce part type 2 and 3, with different ideal lead-times. Failure and repair times of the mechatronic modules populating the workstations are collected by analyzing the records of the production monitoring system. The time required to install, uninstall and move mechatronic objects among reconfigurations and starting from an empty configuration, considered as an unbiased initial condition for the planning, is reported in Table 2.

Table 1. Ideal system lead-time for the considered system configurations (the time unit is omitted for confidentiality reasons).

Configuration	**Product Type**			
	1	**2**	**3**	**4**
1	0.3	n.a.	n.a.	n.a.
2	n.a.	5.04	20.25	n.a.
3	n.a.	18.24	5.04	n.a.
4	n.a.	n.a.	n.a.	10.8

Table 2. Reconfiguration times between configurations.

Configuration	**1**	**2**	**3**	**4**
1	0	336.02	336.02	690.08
2	156.02	0	396.00	726.07
3	180.02	60.03	0	750.07
4	378.68	378.67	378.67	0
Empty	4.0135	336.02	336.02	690.09

Table 3. Completion time under variation of the lot size and the service level, for different production and reconfiguration plans (the optimal plans are highlighted in bold).

Lot size	**SL**	**Completion Time**			
		Plan 1 $\omega = \lvert 1,3,2,4 \rvert$ $\varphi = \lvert 1,3,3,4 \rvert$	*Plan 2* $\omega = \lvert 1,3,2,4 \rvert$ $\varphi = \lvert 1,2,3,4 \rvert$	*Plan 3* $\omega = \lvert 1,3,2,4 \rvert$ $\varphi = \lvert 1,2,2,4 \rvert$	*Plan 4* $\omega = \lvert 1,2,3,4 \rvert$ $\varphi = \lvert 1,2,3,4 \rvert$
5	0.5	**1291.1**	1769.6	1312.8	1627.2
	0.9	**1388.5**	1870.8	1414.1	1724.9
	0.99	**1431.1**	1911.3	1454.6	1768.3
10	0.5	**1412.8**	1958.6	1420.8	1681.5
	0.9	**1510.1**	2053.1	1515.3	1779.2
	0.99	**1558.8**	2100.3	1562.6	1826.2
25	0.5	1796.1	2478.3	**1758.3**	1862.4
	0.9	1905.6	2593.1	**1859.6**	1970.9
	0.99	1412.8	1958.6	**1420.8**	1681.5
50	0.5	2428.7	3376.1	2311.8	**2188.0**
	0.9	2538.2	3484.1	2419.8	**2289.3**
	0.99	1455.4	1999.1	1461.3	**1724.9**

Under the aforementioned problem settings, the integrated production and reconfiguration planning problem I is solved. In total $4! = 24$ lot sequences are possible in four different configuration sequences, resulting in 96 combinations to be analyzed. The total completion times for the optimal sequences, ω, in each of the four configuration sequences, φ, are reported in Table 3, under variation of the lot sizes and service levels.

The following considerations hold:

- The optimal integrated production and reconfiguration plan is not necessarily the one including configurations with the lowest ideal lead-time for each individual product (plan 4). For example, focusing on the case with lot size 5 and service level 0.9, plan 1 provides the shortest completion time, with a difference of about 25% with respect to plan 4. This is due to the fact that the additional processing time caused by producing part type 2 with configuration 3 instead of configuration 2 is compensated by a reduction in the reconfiguration time. This result proves that if the production and reconfiguration planning problems are decoupled and the optimal configuration for each product is fixed before planning the production sequence, a significant loss in performance can be observed.

- By increasing the lot size, from 5 to 50, the optimal production sequence and reconfiguration plan changes. While for lot sizes equal to 5 and 10, plan 1 is optimal, for lot size 25, plan 3 becomes optimal and for lot size 50, plan 4 becomes optimal. Indeed, for larger lot sizes, the additional processing time of suboptimal configurations becomes predominant and it is not compensated by the reduction of reconfiguration time. As a consequence, as the lot sizes increase, the optimal planning solution tends to the one obtained by decoupling the production and reconfiguration planning problems.
- The target service level imposed on the system affects the attainable completion time but does not impact on the optimal production and reconfiguration plan. In particular, it is remarkable to observe that by considering only the average completion time of the plan (*SL=0.5*), the completion time is considerably underestimated with respect to the industry relevant case of *SL=0.99*. For example, for the case of lot size equal to 5, a difference of 10% is observed.

These results show that the benefits of implementing the proposed integrated production and reconfiguration planning method are particularly evident in the case of small lot productions, where the short-term behaviour of the system is predominant and longer processing times can be compensated by shorter reconfiguration times.

In the second experiment, the performance of the method is investigated when the number of lots and possible system configurations increase. Two additional product types and configurations are considered. In particular product type 5 can be produced in configuration 1 and 5, and product type 6 can be produced with configurations 2, 3 and 6, with different processing times. In this case, the number of possible configuration sequences increases up to $3^3 \times 2^2 = 108$ and the number of possible production sequences corresponds to $6! = 720$. Thus, the overall number of possible reconfiguration and production plans is equal to 77760.

Identical lot sizes of 50 parts are selected. The completion time of the whole plan is computed with a service level equal to 0.95. The method, implemented in JAVA and executed on an off-the-shelf hardware, took ~504 seconds to evaluate all the possible combinations and to provide the optimal configuration. By comparing the worst and best plans, a plan completion time of 3606.4 time units, against some 6494.9 time units, can be observed, leading to a completion time reduction of about 50%. The whole production plan completion time distributions for the best and the worst cases are reported in Figure 2. It is possible to notice that the difference between the two distributions is remarkable and that the major portions of the two probability masses do not overlap. As a final remark, also in this experiment the best production and reconfiguration plan does not exploit the configurations with shorter ideal lead-time for each product type, as observed in the previous experiment.

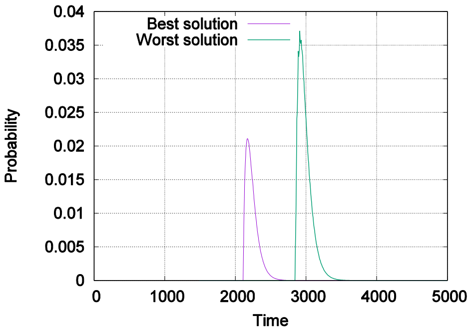

Figure 3 Probability mass function of the completion time for the best and worst integrated production and reconfiguration plans, in the case of six product types and six feasible configurations.

6. Conclusions

This paper proposes a modeling framework and a methodology to jointly plan the production of lots and the system reconfigurations in the context of modular, plug-and-produce production systems. The method grounds on the analytical derivation of the production plan completion time distribution, thus entailing low computational effort, making it applicable in real industrial settings, even in high planning dynamics contexts. The real case analysis shows that high benefits can be achieved by implementing the developed approach with respect to less integrated methods, especially for small production lots.

The theoretical work developed in this paper paves the way to further extensions to more complex problem settings. For example, the case where system reconfigurations entail modifications of the buffer capacities can be considered by extending the concept of mechatronic objects also to transportation units. Moreover, the case where multiple production lines can share mechatronic objects is relevant and will be subject of future investigations.

References

[1] ElMaraghy H., Schuh G., ElMaraghy W., Piller F., Schönsleben P., Tseng M., Bernard A., 2013, Product Variety Management, CIRP Annals–Manufacturing Technology, 62(2): 629-652.

[2] H.-P. Wiendahl (1), H.A. ElMaraghy (1), P. Nyhuis, M.F. Zaeh (2),H.-H. Wiendahl, N. Duffie (1), M. Brieke, 2007, Changeable Manufacturing - Classification, Design and Operation, CIRP Annals–Manufacturing Technology, 56(2): 783-809.

[3] Putnik, G., Sluga, A., ElMaraghy, H., Teti, R., Koren, Y., Tolio, T., Hon, K.K.B., 2013, Scalability in Manufacturing Systems Design and Operation: state-of-the-art and future development roadmap, CIRP Annals–Manufacturing Technology, 62(2): 751-774.

[4] Tolio, T., Ceglarek, D., ElMaraghy, H.A., Fischer, A., Hu, S.J., Laperriere, L., Newman S.T., Váncza J., 2010, SPECIES Co-evolution of products, processes and production systems, CIRP Annals–Manufacturing Technology, 59(2): 672-694.

[5] Hu, S.,J., Ko, J., Weyand, L., ElMaraghy, H.A., Lien, T.K., Koren, Y., Bley, H., Chryssolouris, G., Nasr, N., Shpitalni, M., 2011, Assembly System Design and Operations for Product Variety, CIRP Annals–Manufacturing Technology, 60(2): 715-735.

[6] Tolio, T., 2009, Design of Flexible Production Systems, Methodologies and Tools, Springer Berlin Heidelberg.

[7] Koren Y, Gu X, Guo W, 2018, Reconfigurable Manufacturing Systems: Principles, Design, and Future Trends. Frontiers of Mechanical Engineering 13(2):121– 136.

[8] Gyulai D, Kádár B, Kovács A, Monostori L, 2014, Capacity management for assembly systems with dedicated and reconfigurable resources. CIRP Annals–Manufacturing Technology 63(1):457-461.

[9] Colledani, M., Gyulai, D., Monostori, L., Urgo, M., Unglert, J., Van Houten, F., 2016, Design and Management of Reconfigurable Assembly lines in the Automotive Industry, CIRP Annals - Manufacturing Technology, 65 (1): 441-446.

[10] Li, X., Bayrak, A., E., Epureanu, B., I., Koren, Y., 2018, Real-time teaming of multiple reconfigurable manufacturing systems, CIRP Annals - Manufacturing Technology, 67 (1): 437–440.

[11] Angius, A., Colledani, M., Horvath, A. 2014, Moments of accumulated reward and completion time in Markovian models with application to unreliable manufacturing systems, Performance Evaluation, 75: 69-88.