Advanced Logistic Systems, Vol. 7, No. 2 (2013), pp. 27–34.

CONCEPTUALIZATION OF FLUID FLOWS OF LOGISTIFICATED PROCESSES

ÁKOS GUBÁN[1]–RICHÁRD KÁSA[2]

Abstract: The amelioration of dysfunctional processes can be approached by several points of view. On the one hand customers' perceptions are analysed in our researches and on the other hand we perform the research from the view of the analyses of flowing documents, resources, information, data – which called fluids – which also lead to process amelioration. Our goal during this research project is to modelling real service processes through the detection of flowing elements (fluids) in the process flow and building a fluid-process flow. This requires the comparison and the determination of the fluid-coordination and the development of "flow-distance" concept. Using these definitions it will be possible to classify and thus to determine the real processes of the system. In this paper only the framework of real process design is given.
Keywords: logistification, fluid, fluid flow, weak similarity, strong similarity, general similarity, clustering.

1. Introduction

It is clearly justified in technical logistic literature that well-defined technical-mathematical models can be made to logistical and supply chain processes. The same unfortunately cannot be said of the service processes as they are much more stochastic. Most important characteristics of logistics processes that material, information, resources, human resources, data, documents, etc. flow occurs through them and these flows can be observed and analysed and in many cases even organized into processes depending on the various elements. In this sense even a service process can be considered as a kind of logistic process if the flown elements of the system are determined. Thus it is not sure, that a service system requires different process analysis method than a logistic system. Our investigations are led by this consideration in order to create a "techno-mathematical" model to analyse and reorganize dysfunctional service processes.

Logistification is a previously shown [2, 3] business process amelioration [2] method. The analysis of fluid flow offers a very particular opportunity for analysing system of processes. In this recent paper we would like to present the similarity of flow processes in the fluid-flow system analysis. This requires the comparison and the determination of the fluid-coordination and the development of "flow-distance" concept. Using these definitions it will be possible to classify and thus to determine the real processes of the system. A brief summary of the necessary preliminaries to this paper is presented in the next section.

2. Research background

In term of flow the fluid-flow can be divided into two groups: it can be either nodal flow or continuous flow. In case of nodal flow the fluid transformation is visible/measureable and

[1] PhD, Budapest Business School
guban.akos@pszfb.bgf.hu
[2] MSc, Budapest Business School
kasa.richard@ pszfb.bgf.hu
H-1149 Budapest, Hungary

have affect only on process nodes, while in case of continuous flow the effect of fluid transformation can be realized and measured at any point of the process. In terms of our investigation – as we are about to carry out simulation of service processes – nodal flows will be important and give an overview of this kind of flow.

Let d $\in D$ be a fluid (where D is a finite set), and let P_0 be the process at t_0 initial time of fluid analysis (observation) which or whose input involves the fluid and be the initial type of the fluid T_0.

Then $(d, \tau_0, t_0) \in P_0(c_i)$ marking means that the observed (d) fluid has arrived to c_i node of P_0 process at t_0 point in time notably in τ_0 type.

The flow of this fluid can be described at $[t_s; t_f]$ period of time with the following sequence:

$$F(d;\tau;w)_{[t_s;t_f]} = \langle (c_{i_0 j_0};\tau_0;w_0;t_{s0};t_{o0}); (c_{i_1 j_1};\tau_1;w_1;t_{s1};t_{o1}); \dots; (c_{i_m j_m};\tau_m;w_m t_{s1};t_{o1}) \rangle, \quad (1)$$

where

$$(c_{i_l j_l};t_{sl};t_{ol}): l = 1;\dots;m+1$$

means that the fluid entered into the $c_{i_l j_l}$ node of a certain process with τ and τ_{l-1} status and with the type's w and w_{l-1} weight value. The outgoing fluid has τ_l type and w_l weigh. The other two parameters represents the time of node entry and exit time, where $t_{sl} \leq t_{ol} \in [t_s; t_f]$ as well as $t_{ol} \leq t_{sl+1}$. Obviously in sequence

$$\check{C}(F) = \langle c_{i_0 j_0}; c_{i_1 j_1}; \dots; c_{i_m j_m} \rangle$$

a certain node may appear several times, however the sequence of nodes is finite. Thus $|\check{C}(F)|$ refers to the number of nodes in the sequence.

At this point the similarity of flows must be defined. The uncovering of real service processes is started by the analysis of the previous system. That is the pre-condition analysis. Every systems in its basic model is predefined and partly or fully described by its basic processes.

This context is used as a basis in analysing and modelling of the system. To uncover real processes of the system fluid-flows must be explored as shown on Figure 1.

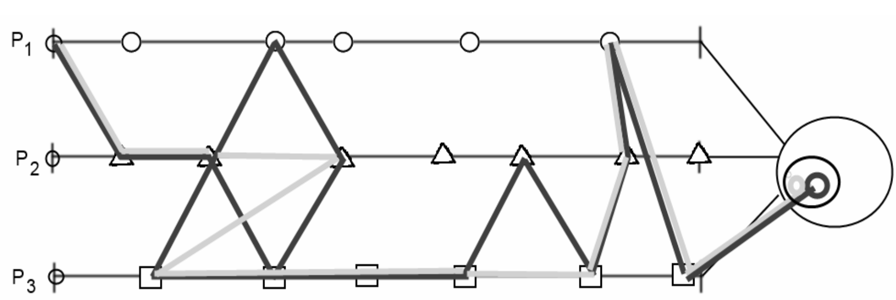

Figure 1. Fluid flow in processes

3. Similarity of fluid flows

A brief explanation is given here for the determination of similarities of fluid flows. At initial stage of the examination of similarities the fluid weights and types are disregarded.

The fluid weights and types may have role in classification, if when only those nodes are examined which are involved in fluid flow and the interesting is which nodes are involved in the fluid flow as the path of the fluid flow is set as a baseline.

The first method is a more permissive approach: we are focusing during the similarity analysis only on how each processes' nodes relates to other process's nodes. For demonstration consider two fluid flow in the same period of time (disregarded to type and weigh parameters at this time):

$$F_1(d) = F_1(d)_{[t_s;t_f]} = \langle (c_{i_0j_0}; t_{s0}; t_{o0}); (c_{i_1j_1}; t_{s1}; t_{o1}); \dots; (c_{i_{m_1}j_{m_1}}; t_{sm_1}; t_{om_1}) \rangle \text{ and}$$

$$F_2(d) = F_2(d)_{[t_s;t_f]} = \langle (c_{l_0k_0}; t_{r0}; t_{p0}); (c_{l_1k_1}; t_{r1}; t_{p1}); \dots; (c_{l_{m_2}k_{m_2}}; t_{rm_2}; t_{pm_2}) \rangle$$

Let $\check{C}(F_1) = \langle c_{i_0j_0}; c_{i_1j_1}; \dots; c_{i_{m_1}j_{m_1}} \rangle$ and $\check{C}(F_2) = \langle c_{l_0jk_0}; c_{l_1k_1}; \dots; c_{l_{m_2}k_{m_2}} \rangle$ be node sequences of two flows.

Let $\tilde{C}(F_1) = \langle c_{i_0j_0}; c_{11}; \dots c_{1k_1}; c_{i_1j_1}; c_{21;\dots}c_{1k_2}; \dots; c_{i_{m_1}j_{m_1}} \rangle$ be an extended sequence of $\check{C}(F_1)$ where $c_{11}; \dots c_{1k_1}; c_{21}; \dots c_{2k_2}; \dots; c_{m_11}; \dots c_{m_1k_{m_1}}$ are nodes and $k_1; k_2; \dots; k_{m_1} \in \mathcal{N}$.

In such case when $\check{C}(F_1)$ subsequence is fully incorporated by $\check{C}(F_2)$, than $\check{C}(F_1)$ is an extended subsequence of $\check{C}(F_2)$: $(\widetilde{F_1})_{F_2}$.

(Notation: $F_1 \sqsubseteq F_2$ means, that F_1 is a subsequence of F_2; and $F_1 \sqsubset F_2$ means, that F_1 is an extended subsequence of F_2.) This is illustrated in Figure 2.

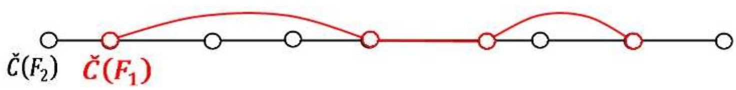

Figure 2. Extended subsequence

Two sequences are called δ weakly similar, if and only if when the ratio of the length of their maximum extended subsequences is $\delta > 0$ i.e. F is a subsequent of F_1 and F_2 at the same time,

$$\max\left(\frac{|\tilde{C}(F)_{F_2}|}{|\check{C}(F_2)|}; \frac{|\tilde{C}(F)_{F_1}|}{|\check{C}(F_1)|} \middle| \text{for all mutually extended subsequences} \right) \geq \delta$$

where $\delta \in]0; 1]$ represents the similarity of the two sequences. (Notation: $F_1 \sim_\delta F_2$.) The similarity measure of two sequences is δ, if

$$\delta = \max\left(\frac{|\tilde{C}(F)_{F_2}|}{|\check{C}(F_2)|}; \frac{|\tilde{C}(F)_{F_1}|}{|\check{C}(F_1)|} \middle| \text{for all mutually extended subsequences} \right).$$

In Figure 3: $\delta = \max\left(\frac{3}{7}; \frac{3}{10}\right) = \frac{3}{7}$ and $\delta = \frac{6}{11}$.

Let $C_1 \subseteq \check{C}(F_1)$ and $C_1 \subseteq \check{C}(F_2)$ subsequences of our sequences, than two sequences called δ strictly similar if and only if when

$$\max\left(\frac{|C_1|}{|\check{C}(F_2)|}; \frac{|C_1|}{|\check{C}(F_1)|} \middle| \text{for all subsequences} \right) \geq \delta$$

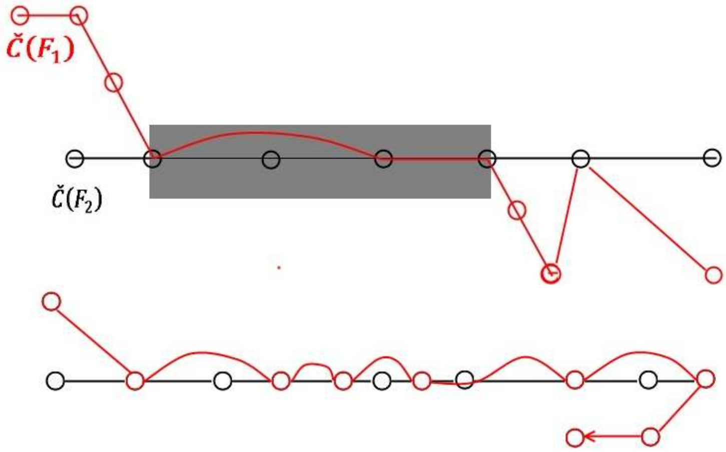

Figure 3. Weakly similar sequences

In such case, the mutual containment of the maximal subsequences of a sequence is assumed. (Notation: $F_1 \approx_\delta F_2$.) The similarity measure of these two sequences is

$$\delta = \max\left(\frac{|C_1|}{\left|\check{C}(F_2)\right|};\frac{|C_1|}{\left|\check{C}(F_1)\right|}\middle|\text{for all subsequences}\right)$$

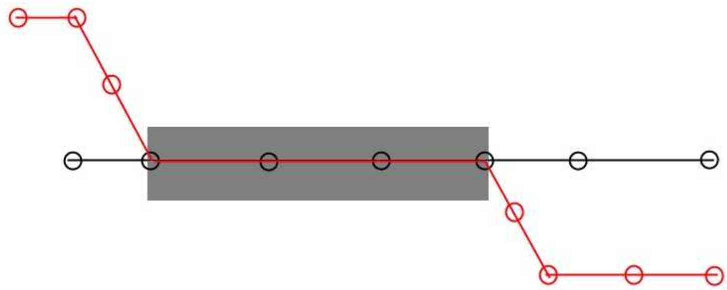

Figure 4. Strictly similar sequences

Additional sequence similarities also can be introduced. Maximal ordinal node matching is called generalized similarity, i. e. the extension is used for both sequences. This reasoning is shown in Figure 5.

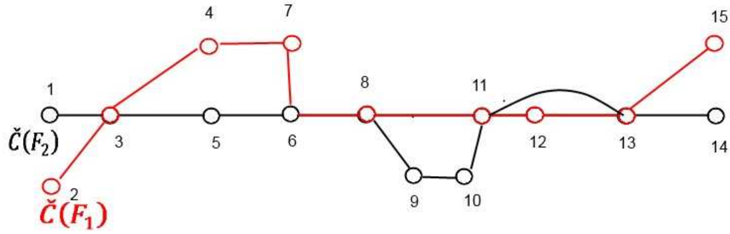

Figure 5. Generalized similarity

Both flows of Figure 5 have sequence $\langle$ 3; 6; 8; 11; 13 $\rangle$ in this order. The measure of similarity should be appropriately defined by the ratio of the shorter sequences, i.e. as Figure 5. shows: $\delta(F_1;F_2)=0,5$ (Notation: $F_1 \simeq_{0,5} F_2$.)

In such case, when $F_1;F_2$ has no common nodes, then the measure of their similarity is $\delta(F_1;F_2)=0$ and can be concluded as they are not similar.
These similarities are satisfying some important characteristics:

1. The self-similarity measure of a sequence (in all three above defined cases)

$$F_1 \sim_1 F_1;\ F_1 \approx_1 F_1;\ F_1 \simeq_1 F_1.$$

Consequently all three relation is reflexive by $\delta \in\ ]0;1]$ value, thus:

$$F_1 \sim_\delta F_1;\ F_1 \approx_\delta F_1;\ F_1 \simeq_\delta F_1.$$

2. All three similarities are symmetric, thus:

$$\text{if } F_1 \sim_\delta F_2, \text{ then } F_2 \sim_\delta F_1;$$
$$\text{if } F_1 \approx_\delta F_2, \text{ then } F_2 \approx_\delta F_1;$$
$$\text{if } F_1 \simeq_\delta F_2, \text{ then } F_2 \simeq_\delta F_1.$$

   It follows the maximization of the definition in the first two cases and from the reference to the minimal sequence in the third case.

3. If $F_1 \sqsubseteq F_2$ or $F_1 \sqsubset F_2$ is satisfied, then:

$$F_1 \sim_1 F_2;\ F_1 \approx_1 F_2;\ F_1 \simeq_1 F_2.$$

4. Let $\delta_1; \delta_2 \in\ ]0;1]; \delta_1 \leq \delta_2$

$$\text{if } F_1 \sim_{\delta_1} F_2, \text{ then } F_1 \sim_{\delta_2} F_2;$$
$$\text{if } F_1 \approx_{\delta_1} F_2, \text{ then } F_1 \approx_{\delta_2} F_2;$$
$$\text{if } F_1 \simeq_{\delta_1} F_2, \text{ then } F_1 \simeq_{\delta_2} F_2.$$

   Unfortunately transitivity is not satisfied. For example let

$$\check{C}(F_1) = \langle 1;2;3;4;5;6 \rangle\quad \check{C}(F_2) = \langle 1;2;3;4 \rangle\quad \check{C}(F_3) = \langle 3;4;5;6 \rangle$$

$$F_1 \approx_{\frac{2}{3}} F_2 \quad F_1 \approx_{\frac{2}{3}} F_3 \quad \text{but} \quad F_1 \approx_{\frac{1}{2}} F_3$$

   This is a problem because direct classification with the measure of similarity is not possible.

4. Classification and similarity of flows

If and when the types and weights of flowing fluids in the process flow are disregarded then flows can be classified by their measure of similarity. As it was mentioned above unfortunately, the similarity is not equivalence relation as transitivity is not satisfied, thus it is not suitable directly of classification – but it is not necessarily important. For example, if we take the above mentioned example, the three processes can be brought to a group without any aversion.

1. Hierarchical clustering: This method offers the benefit that known centroids are not needed during the iteration which is desirable in respect to fluid flow classification. However out of clustering methods ab initio only single linkage and complete linkage methods could be applied, as other methods assume such distances relative to derivative value which were not interpreted so far and even not possible without some particular excessive force.
2. Non-hierarchical clustering: For this method we would need to know the number of core processes and to identify the established centres of classes. It would be easier to apply than a hierarchical clustering, but it still requires a constrained solution.
3. Heuristic solution: Due to its advantages a heuristic solution will be used, but a centre selection method will be requisite.

5. Weighted fluid flows

In this section of the paper the previously outlined problem is approached from another point of view. A well-defined system of weights for the fluids and fluid flows will be introduced. It is also not easy to construct a well-designed system of weights as the fluids have different types, relevance and importance and even their quantity is different.

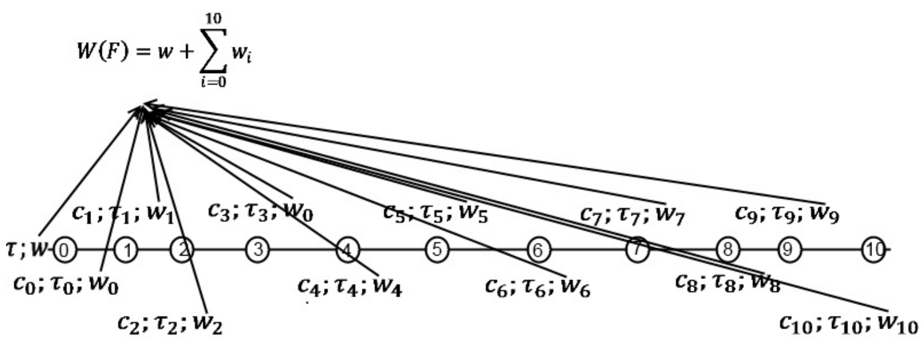

Figure 6. Weights of fluid flows

Let $\tau \in T$ be the type of a fluid, let c be a node, this fluid may appear, and let q be the quantity of this fluid appeared on c. Thus $w: (\tau; c; q) \mapsto \mathcal{R}^+ \cup \{0\}$ function is the weight of the fluid going out of node c. (Note: The specific definition of the value of the weight relies both on the perception of the process/service user and on the certain system and service. It will be defined during a simulation.)

Let $F(d;\tau;w)_{[t_s;t_f]}$
$$= \langle (c_{i_0j_0};\tau_0;w_0;t_{s0};t_{o0}); (c_{i_1j_1};\tau_1;w_1;t_{s1};t_{o1}); \dots; (c_{i_mj_m};\tau_m;w_m t_{s1};t_{o1}) \rangle$$
according to (1), where $w_l = w\left(\tau_l; c_{i_lj_l}; q(\tau_l; t_{ol})\right): l = 0; \dots; m.$

Thus $W(F) = w + \sum_{i=0}^{m} w_i$ is called the weight of the fluid flow.

After that the generous steps of building the classification method is presented, which will be the basis of the heuristic methodology:

1. ** Definition of fluid-flow nodes **
2. Location of the known processes of the system and determine the nodes.
3. Determination of the fluids in the system. If too many is detected consider disregarding the less important ones or integrate them into relevant fluids. (This step requires a lot of attention, and must be carried out with the help of experts.)
4. Determination of the fluid flows, and the dynamical assignation to nodes.
5. Determination of the node sequences in the resulting flows.
6. Examination of the fluid transformations in every node (if there is any transformation in a certain node). Thus the types and quantities of the outgoing fluids are derived. (And these will be the parameters of the fluid entering into the next node of the sequence.)
7. Formation of the nodes-leaving weights of the fluids. (The rules of weight formation should be determined prior to the modelling.)
8. Determination of fluid weights.
9. Evaluation of the initial centroids of the fluid clusters.
   a) Arrangement of the fluid flows in descending order according to their weights
   b) Adjustment of *P(F)* to 0 for each process.
   c) Let $\delta = 0$ and define Δ step-intervals and the *M* maximum number of processes ($M \leq$ number of fluid flows)!
   d) Take the first stream to the list!
      i. Let (*N(F)=0*) (flow counter)
      ii. Take the first not equal process from the list
         1. If its similarity to *F* is $\leq \delta$ (δ away from each other), then *N(F)* will be incremented.
         2. Take the next process to be compared
      iii. After surpassing the last flow, we need to move on to the next activity.
      iv. If $N(F) \geq M - 1$, then the flow must be marked as *P(F)* = 1. (That means that it has $M - 1$ neighbours in δ distance.
      v. Move forward to the next process to be examined.
   e) If we do not go beyond the last process should be repeated.
   f) Examine how many flows satisfy *P(F)* = 1.
   g) If it is less than *M*, then:
      i. δ shall be increased by the value of Δ step-interval: $\delta := \delta + \Delta$
      ii. Go back to step d.

      Otherwise:

      iii. Examine if there is so many *M* flows, which are mutually δ distance to each other.
         1. If there is no M pieces, then
            a) $\delta := \delta + \Delta$
            b) Go back to step d.

2. Otherwise:
   a) Select those M pieces out of them, which have the greatest weights
   b) Let $\phi\,(\left|\phi\right| \geq M)$ be this list

10. ** Any additional flows are assigned to the centroid flows of the list. **
11. The assignment is made on the basis of similarity. (Procedure is not detailed here, as it is a current task of the LOST in Services Modelling Working Group.)

6. Summary

In this paper the determination of the similarity degree of fluid-flows have been presented. The significance of this issue lies in the creation of real processes via the designation of similarity and weights of fluids. The exploration of centroids is performed by a heuristic method, whose procedures have been defined. It requires further studies to determine the effectiveness and sensitivity of the centroid selection. Further analysis should be implemented to decide whether the centroids have enough distance from each other and whether they have enough number of similar flows. After a proper refinement of the method it is necessary to build a procedure for the flow assignment to centroids, which will be based on the Shopping Basket Theory (Demetrovics, Hua, Guban [4]). This procedure will form the basis of other fields of our research, such as service user segmentation.

Acknowledgements

The research is supported by the EMMI-26130-2/2013/TUDPOL research grant of the Ministry of National Resources at Budapest Business School.

Literature

[1] Cua, K. O.–McKone, K. E.–Schroeder, R. G. (2001): *Relationships between implementation of TQM, JIT, and TPM and manufacturing performance*. Journal of Operations Management, Vol. 19 (6), pp. 675–694.

[2] Gubán, Á.–Kása, R. (2013): *A Literature Based Review of Business Process Amelioration Methods and Techniques Regarding Service Orientation*. Journal of Advanced Management Science, Vol. 1 (2), pp. 230–235.

[3] Kása, R.–Gubán, Á. (2013): *Evolution of methods and technics of process improvement and their customization for services. (in Hungarian)* In 3. Vezetéstudományi Konferencia: Vezetés és szervezetek Taylor után 102 évvel. Szeged.

[4] Demetrovics, J.–Hua, N. S.–Gubán A. (2011): *An algebric approach to market basket model: explicit representation of frequent market basket and association rules*. Proceedings of the 8th International Conference on Computer Science and Information Technologies (CSIT 2011) Yerevan: National Academy of Sciences of Armenia, pp. 170–173.

[5] Demetrovics, J.–Rónyai L.–Hua N. S. (2003): *Dependencies and generalized relations*. Mathematical And Computer Modelling, Vol. 38, pp. 773–782.

[6] Telek, P.–Bányai, T. (2013): *Complex design of integrated material flow systems*. Advanced Logistic Systems. Theory and Practice, Vol. 7, No. 1, pp. 105–110.