Estimation of job-site work progress through on-site monitoring

Alberto Giretti [1]*, Alessandro Carbonari [1], Gabriele Novembri [2], Federico Robuffo [1]

[1] Università Politecnica delle Marche, Department of Civil and Building Engineering and Architecture, "Building Construction and Automation" research Team, Ancona, Italy
[2] University "La Sapienza", Department of Architecture and Town Planning for Engineering, Rome, Italy
* Corresponding author (a.giretti@univpm.it)

Purpose This paper reports on the development of intelligent probabilistic models for real-time estimation of construction progress, which operate on the basis of a continuous data flow collected by monitoring networks deployed on-site. Several authors listed the advantages that would be provided by the availability of such models, like project performance and quality control, timely onsite inspections, better control of health and safety prescriptions against job injuries and fatalities. The findings reported in this paper represent a feasibility study and preliminary examples of Bayesian Networks, which are able to infer the work progress attained at every step, starting from real-time tracks of the construction site activities. Activity tracks are represented as a set of state variables figuring out workers' effort, equipment and materials usage rates and other knowledge about the context. **Method** As estimations are always related to dynamic processes, Dynamic Object Oriented Bayesian Networks have been used to develop a set of first order Hidden Markov Models. Hence, the models are arranged as a sequence of time steps, where each time step propagates evidences collected by the site monitoring sensor network along the time line. The actual cumulative progress is computed as a function of the progress achieved in each time step. Models representing a number of typical tasks (external piping, on-site cast of reinforced concrete floor slab, walls erection, ceiling installation) for a real case of a construction site have been developed. Their structure has been designed as part of a general monitoring framework, covering all the phases from design to execution, where BIM design, monitoring systems, methodological process innovations, intelligent inferences and advanced visualization are combined. **Results & Discussion** The networks have been developed and validated through data collected from a real case, and they have been shown to be able to infer work progress, the accuracy of which depends on the resolution and quality of the collected data.

Keywords: automation, Bayesian networks, work progress, construction sites

INTRODUCTION

Among the outstanding benefits that would derive from automated on-site workers, equipment and material tracking, real-time work progress estimation is considered as one of the most critical. Most of the ongoing research on this field, which is summarized in the next paragraph, is targeted to the following three main goals:

- setting up a cost-effective construction project management, featuring real-time deployment of information, including material and equipment inventory and their traceability;
- performing intelligent waste management, which makes its delivery to the appropriate facility for reuse, recycling, recovery or disposal feasible;
- providing a safe working environments to employees, whose translation into practical situations may be in the form of automated control of proper wearing of safety gears, signaling hazards in real-time, automated predictive collision detection and fall hazards warning in crowded site's areas, and so on depending on the particular kind of work to be performed.

The main focus of this concerns a methodology, based on the use of advanced probabilistic models (Bayesian Networks) and real-time and low-invasive monitoring networks, to automatically estimate the work progress at the execution phase. In particular, the construction of a shopping mall made up of a precast concrete technology has been considered, and the feasibility of a monitoring approach for work progress estimation and based on the automated detection of those resources present on site has been shown.

Moreover, this research step is part of a wider approach aiming at the development of an integrated framework[1] for advanced construction management. Effective monitoring is conditioned upon embedding non-invasive sensors within resources which are expected to operate on-site. To this aim, BIM-based engineering design constitutes a fundamental support, because it allows the disaggregation of any building's design into its elements and phases. In the proposed example a kind of tasks list has been pre-

liminary arranged through BIM, making the assignment of one or more sensor kit to each phase of the construction plan easier and faster (Fig. 1). Once resources and relevant variables are tracked, data collected can be filtered and processed in real-time by advanced probabilistic models, in order to infer the work progress, so that higher reliability and efficiency in Project Control reporting (e.g. Daily Site Reports) and quality inspection management can be achieved.
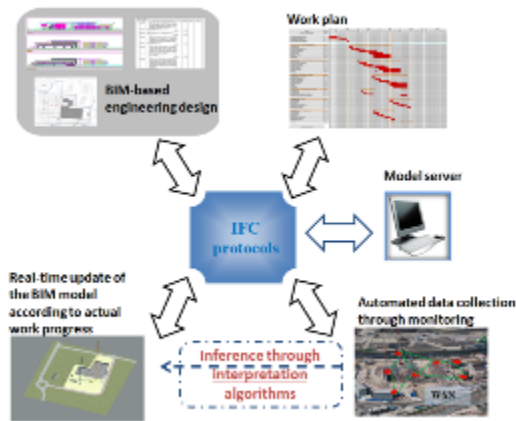

Fig.1. Interoperability for work progress management.

The next paragraph reports about the most relevant research in this field. The following paragraph discusses the conceptualization of the whole framework. Following the details of a case study will be described. Conclusions close the paper, before the list of references.

BRIEF OVERVIEW OF THE SCIENTIFIC BACKGROUND

On-site tracking is seen as much challenging as promising, thanks to the many advantages that would be brought into the construction process[2]: material management and inventory traceability, automated waste management (i.e. automatic recognition of the waste destination and its cycle), safety support through real-time laborers' warning in case of imminent and probable risk occurrence (e.g. collision avoidance, fall from heights etc…). Assisted design can help avert accidents at the design phase[3], but a real-time warning system is anyway useful to avoid unexpected or difficult to model occurrences[4,5].

Automated progress monitoring would reduce the burden of work usually required to produce project reports editing[6,7], and would make easier communication through automated visualization of construction data[8]. Information awareness is an undisputable excellent tool to manage machines, as it can record operators and operational times in a central server, planning and ensuring inspection and maintenance. However, general economic efficiency is conditioned upon the development of a system for automated project performance control, where indirect data would be intelligently converted into performance measurements[9]. The availability of such an approach would pave the way to many innovations, and several contributions in literature stressed those ones related to quality management: automated remote inspection of infrastructures[10]; control of the construction quality of details that are no more visible after the task is accomplished (e.g. to validate the depth of foundation piers)[11]; tracking and understanding the context for automated relevant information retrieval during site inspection[12].

This paper contributes to the use of advanced probabilistic models (namely Dynamic Bayesian Networks) to perform inference from real-time collected data about the estimation of on-site work progress.

THE CONCEPT

The work progress estimation process, we are going to discuss, is part of a more comprehensive framework (Fig. 1) made up of interrelated technologies, which all together figure out an automated construction management system. Within this framework, the work progress monitoring function is the result of a management procedure that encompasses many different technologies. As depicted in Fig. 2, initially the construction process is broken down into basic activities, so that the resources used in each of them can be unambiguously identified. BIM-based engineering design can be used to support this task, since BIM object oriented modeling allows activities and resources to be clearly identified and assigned to design objects, such as building components and technical systems. A set of sensors is then assigned to each resource so that real-time data about the resource usage can be gathered. Given the research results and experiences about on-site materials and laborers tracking reported in the previous paragraph, and given the ongoing technology innovations[13], at present data about on-site resource usage, gathered in real time by embedded monitoring networks, can be considered sufficiently reliable to support management tasks.

Hence resource usage data is then processed by intelligent algorithms, capable estimating the progress of the ongoing activities. The results are finally arranged in a management database so that further elaborations, including scheduling of real-time inspection and quality assessment, as labeled in the fourth phase of Fig. 2, can be performed.

In the next paragraphs we will discuss the feasibility of automated progress monitoring by means of Dynamic Bayesian Networks, which is the key algorithmic step of this procedure, through a case study.
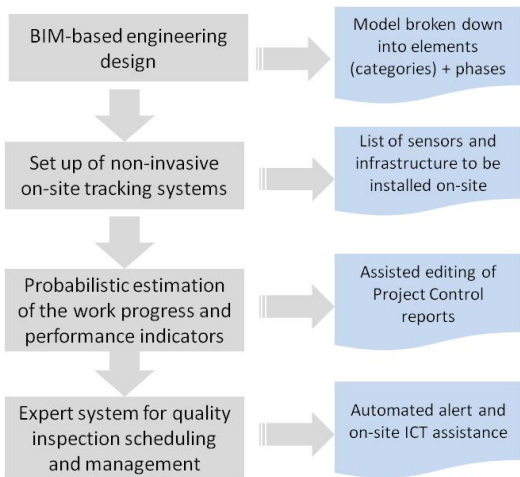
Fig.2. A possible procedure to perform automated work progress management.

CASE STUDY

The construction site and monitored tasks

The resource usage data, the will be feed to the progress estimation algorithm, have been collected by monitoring the execution of a shopping mall in the village Cerreto d'Esi (AN), built by the company Torelli & Dottori SpA. The whole project includes three areas: the shopping mall, the office building and the parking lot. Our monitoring was relative to the erection of the shopping mall, which is on the left side of Fig. 3.
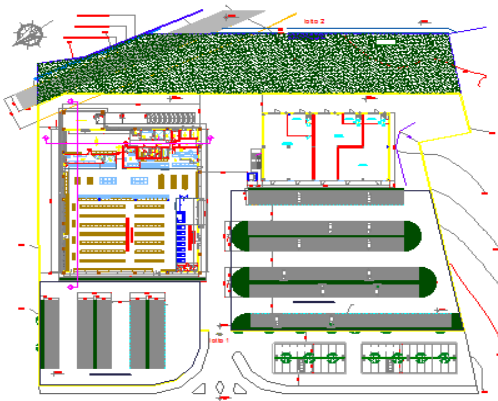

Fig.3. Whole picture of the monitored site.

The squared shaped mall has a total surface exceeding 9,000 m², external height 6.90m and internal net floor-to-ceiling height 3.50m. It is made of precast concrete bearing frames; the roof is made of precast concrete floor slabs, walls made of concrete panels and aluminum framed windows. It is provided with a conditioning system, a solar heating system, a fire protection system, electric and lighting systems.

Three tasks have been monitored during the execution:

- excavation and pipelines laying (sewage, electrical power line, water supply), both external and internal;
- site cast ground floor concrete slab;
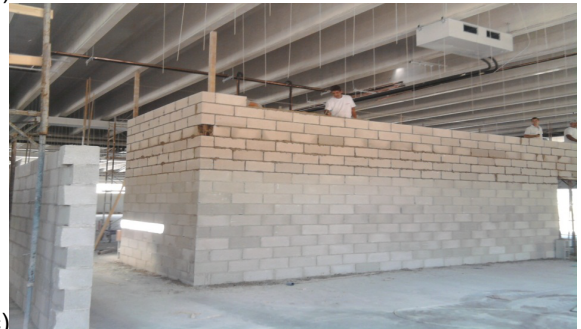- fabrication of internal hollow brick partitioning walls.

During excavation (Fig. 4a) the following state variables have been collected: rate of usage and position of excavators; usage and positions of dumpers; number of man-hours used to operate equipment, to lay pipes underground and to fill in trenches; number of pits laid underground. In addition, every half a day the actual work progress for the task was measured by our operators in charge of monitoring. Collected data extend from Monday 2011, July 18th to Thursday 2011, July 21st, when the task was accomplished. The site cast ground floor concrete slab execution took place between Tuesday and Wednesday 2011, July 26th–27th (Fig. 4b): first of all reinforced bars have been put in place (through the use of dumpers to transfer them from storage areas) and laborers and formworks were installed along the boundaries; then concrete was brought to the site by concrete truck mixers and poured in the formwork through the use of one concrete pumping truck setup on-site. So the usage of all the equipment and man-hours were monitored and recorded, together with work progress.

a)

b)

c)

Fig.4. Excavation phase (a), site cast concrete ground floor (b), hollow brick internal walls (c).

The hollow brick partitions were executed indoor (Fig. 4c) during 9 working days split into two phases: from Monday 2011, August 1st to Thursday 2011, August 4th and then again (after summer holidays) between Tuesday 2011, August 30th and Tuesday 2011, September 6th. In this case equipment (small dampers to move bricks from storage to the interior and concrete mixers to produce mortar) and number amount of bricks employed constituted relevant indicators for work progress estimation, together with employed man-hours.

Data gathering

As relevant variables for work progress monitoring were not know *a-priori* before the development of probabilistic models, monitoring was performed by operators observing the work progress and writing down records every 5*min* relative to the use of the resources listed above. This approach allowed redundancies and boost awareness on the procedures performed on site. So for each of the monitored tasks the following documents have been produced:

- a database comparing the work progress with the amount of resources employed during a time duration of 5*min*;
- reports with description of the work performed and monitored data;
- a photographic survey of the activities.

As the Dynamic Bayesian Network presented in the next paragraph is relative to the excavation and pipe laying task, Table 1 shows the data collected and reworked after such monitoring.

Id	excavator (%/60 min)	excavation man-hours	excavation progress	pipelines man-hours	dumper (%/60min)	pipelines progress	truck (%/60 min)	dumper (%/60 min)	trench filling progress
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.77	0.58	5.00	1.50	0.00	5.00	0.08	0.00	0.00
3	1.00	0.17	15.00	2.17	0.00	10.00	0.08	0.17	10.00
4	1.00	0.25	10.00	2.92	0.00	10.00	0.00	0.25	5.00
5	0.57	0.00	20.00	1.42	0.00	10.00	0.29	0.00	0.00
6	1.00	0.00	5.00	1.00	0.00	5.00	0.00	0.00	0.00
7	0.00	0.50	0.00	2.58	0.00	10.00	0.00	0.92	20.00
8	0.50	1.00	0.00	1.00	0.00	0.00	0.08	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.85	0.10	20.00	0.17	0.54	20.00	0.00	0.00	0.00
11	1.00	0.08	25.00	2.83	0.54	30.00	0.00	0.00	0.00
12	0.92	0.15	25.00	4.25	0.58	40.00	0.00	0.00	0.00
13	0.42	0.10	25.00	2.67	0.08	20.00	0.17	0.00	20.00
14	1.00	0.15	25.00	0.67	0.00	20.00	1.00	0.00	30.00
15	1.00	0.07	20.00	0.17	0.00	12.00	1.00	0.00	30.00
16	1.00	0.08	25.00	2.08	0.00	20.00	0.67	0.00	30.00
17	1.00	0.08	10.00	1.33	0.00	20.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.75	1.00	0.00	1.00	0.00	0.00
19	0.00	1.08	0.00	4.58	0.00	10.00	1.00	1.00	10.00
20	0.00	1.08	0.00	3.25	0.00	5.00	1.00	1.00	10.00
21	0.00	1.00	0.00	4.08	0.00	5.00	1.00	1.00	10.00
22	0.00	1.00	0.00	7.75	0.00	15.00	1.00	1.00	10.00
23	0.00	0.58	0.00	3.83	0.00	10.00	1.00	1.00	10.00
24	0.00	0.50	0.00	2.92	0.00	10.00	1.00	0.83	5.00
25	0.00	1.00	0.00	6.75	0.42	20.00	1.00	0.00	5.00
26	0.00	1.00	0.00	6.58	0.00	10.00	1.00	0.42	10.00
27	0.00	1.00	0.00	6.83	0.00	5.00	1.00	0.25	10.00
28	0.00	1.00	0.00	1.25	0.00	0.00	0.25	0.00	5.00
29	1.00	1.58	0.00	0.17	0.23	0.00	0.00	0.00	0.00
30	1.00	1.58	30.00	0.00	0.00	0.00	0.23	0.00	35.00
31	1.00	1.58	30.00	2.42	0.00	20.00	0.00	0.00	0.00
32	1.00	1.17	30.00	1.75	0.00	20.00	0.43	0.00	30.00
33	1.00	1.08	30.00	1.50	0.00	20.00	0.00	0.00	0.00
34	1.00	2.67	30.00	2.58	0.00	30.00	0.75	0.00	20.00
35	0.83	1.42	20.00	2.83	0.00	30.00	1.00	0.00	45.00
36	1.00	0.42	20.00	1.00	0.00	0.00	0.00	0.00	0.00
37	0.75	0.06	15.00	2.50	0.25	10.00	0.42	0.00	20.00
38	0.00	0.08	0.00	0.33	0.00	0.00	0.00	0.00	0.00
39	1.00	1.58	0.00	0.17	0.23	0.00	0.00	0.00	0.00
40	0.58	1.08	25.00	2.67	0.00	20.00	0.00	0.00	0.00
41	0.67	0.67	10.00	3.00	0.00	15.00	0.33	0.00	10.00

Table 1. Hourly dataset worked out from the database collected during the execution of excavation and pipelines.

A pre-processing phase showed that a sampling time of one hour is short enough to capture all the relevant dynamics of the monitored activities. In addition the task under analysis has been subdivided into the three sub-tasks:

- trench excavation;
- laying of pipelines on a concrete bed;
- trench filling with sand and previously excavated ground.

That's because the overall work progress of the main task resulted in fact from the composition of the work progress of these sub-activities. Therefore the data initially recorded have been resampled with a time scale of 1 hour (i.e. the time of equipment usage and the man-hours has been summed). Photos, site reports and measurements allowed to assign the actual work progress to each working hour. The data in Table 1 have been consequently organized through 41 rows (one for each hour) and 9 columns: three devoted to the work progress of each sub-task and other six devoted to resource usage during their execution. These data have been used to train the Dynamic Bayesian Network Model, as it will be discussed in the next section.

DEVELOPMENT OF THE PROBABILISTIC MODELS
Dynamic Bayesian Networks

Bayesian Networks (BN) have the unique capability to provide both intuitive and scientifically rigorous representations of complex systems. In addition, after validation, they can be used for performing both scenario analyses, through inference propagation algorithms, and diagnostic reasoning, through backward propagation based on the inversion rule[14].

These networks also have the advantage of enabling qualitative and explicit representation, where nodes represent variables and arcs represent quantitative relationships among the same, worked out through parametric probabilistic models[15].

When the domains to be modeled are very complex, Object Oriented Bayesian Networks (OOBN) are usually used: they are made up of several elementary networks, sharing some of the variables, which constitute the links between the networks. Each elementary network is generally developed separately (and models one of the involved many physical phenomena) but the inference algorithms are propagated over the whole set of elementary networks.

Dynamic Bayesian Networks (DBN) are used to represent statistical models that depends on time, usually called stochastic processes. DBN are based on a discretized time line, and are made up of several time slices, each representing a snapshot of the state of the system at a particular moment in time. Transition relationships among different state variables in different time slices capture the system temporal dynamic. The application of BNs to model the

evolution of processes that have temporal dynamics requires, in its simplest formulation[15]:

- an initial instance of the Bayesian network that contains the formulation of the problem at time t=0, that is the set of random variables $X_{i,0}$ and the related conditional probability distributions: $P(X_{i,0}|X_{i-1,0})$, $P(X_{i-1,0}|X_{i-2,0})$, etc.;
- one or more transition networks that correlate the variables of the BN instance at t=0 with the variables of the BN instance at t=1.

Fig. 5 shows a graphical representation of three time slices of a DBN.
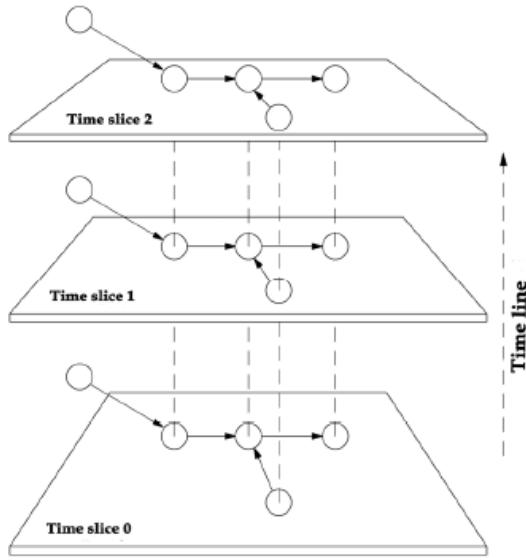
Fig.5. Graphical representation of a general Dynamic Bayesian Network: it is made up of three instances of the same BN.

Two assumptions are usually made about the physical processes at hand:

- all the information needed to predict the state of the process at time t+1 is contained in the description of the process state at time t. No information about earlier time is needed. These kinds of processes are called Markov processes of order one;
- the process is steady, that is, the transition networks remain the same for any ti → ti+1.

The BN for work progress estimation

Fig. 6 depicts a plot of the work progress versus man-hours in the case of excavation sub-task. Similarly to the other sub-tasks (laying of pipelines and trench filling) it's clear that no functional dependence can be defined established by the two variables, Hence the processes must have be modeled as stochastic Markov processes. Both first-order and second-order Markov networks have been tested, showing that the second worked better.

Networks development followed three steps:

- the qualitative network relative to a second-order Markov process for each of the sub-tasks has been developed;
- the dataset has been rearranged to train the network according to the above defined variables and the NPC structural learning performed to find out other hidden causal relationships;
- the EM learning process has been performed by means of the Hugin Expert[TM] software and demonstration of reliability carried out.
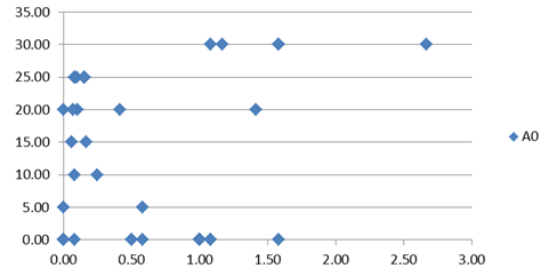

Fig.6. work progress plotted vs man-hours relative to the excavation sub-task.

Fig. 7 depicts the structure of the DBN relative to the excavation sub-task; the meaning of the nodes being as follows: A = work progress; B = equipment usage; L = amount of man-hours. The second digits (i.e 0, 1, 2 and 3) are referred to the time slice: for example L0 is the amount of man-hours used the hour before L1 and two hours before L2. It can be noticed that in each time slice the work progress (A) has been assumed as conditionally dependent upon man-hours (L) and equipment usage (B). In addition every variable in each time slice is influenced by the value of the same variable in the two previous time slices (e.g. A2 is dependent on A1 and A0): this is the translation of a second-order Markovian process into a Dynamic Bayesian Network.
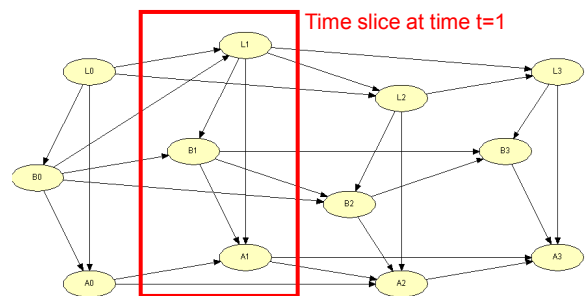

Fig.7. Initial qualitative representation of a sub-network.

Then the dataset in Table 1 has been rearranged to replicate each variable in the four time slices (e.g. L0, L1, L2, L3) with a shift of 1 hour each time slice and NPC structural learning performed. This algorithm allowed us to test whether other conditional dependence relationships occurred among each pair of variables, through an independence test, which is performed according to equations[16]:

$$E(x_{ij}) = \frac{x_{i+} \cdot x_{+j}}{n} \qquad (1)$$

$$E(x_{ijk}) = \frac{x_{i+k} \cdot x_{+jk}}{x_{++k}} \qquad (2)$$

Referring to eq. (1), it is run on a dataset of n records, where x_{ij} is any record. x_{i+} is every observation where x is found with a given value and the same holds for x_{+j}. If the two variables are conditionally independent, their mean value for the generic x_{ij} must be given by equation (1). Eq. (2) holds on the same bunch of theory, but considers marginal probability distributions where dependence of the variable x_i from x_j is surveyed once the value of x_k is given. The use of eqs. (1) or (2) depends on the kind of qualitative relationships assumed among the variables. The statistics is distributed as a "chi-squared".

After performing the test with the help of Hugin Expert[TM] software, in the network of Fig. 7 more causal relationships (given by arrows) as in Fig. 8-a were added. Similarly, it was done for the other sub-networks: pipelines laying (Fig. 8b) and trench filling (Fig. 8c).
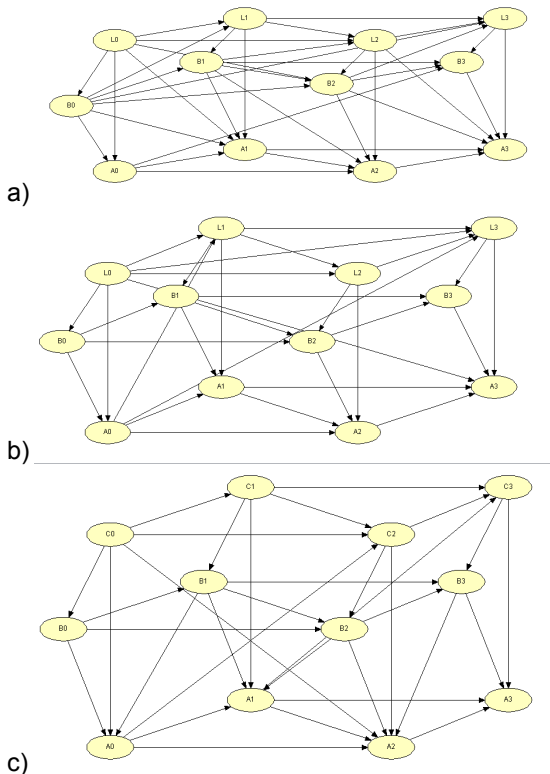

a)

b)

c)

Fig.8. Final qualitative representations of the three sub-networks relative to: excavation (a), pipeline laying (b) and trench filling (c).

At this juncture the dataset was used to perform EM learning[16] for each sub-network: this kind of learning is capable of estimating multivariate "Dirichlet" distributions describing every conditional relationships in the network, the order of the distribution of each variable (or node) being as high as the number of parents (i.e. incoming arrows) pertinent to any varia-

ble. The final qualitative structure of the OOBN capable of mixing all the inputs and computing the expected overall progress in given in Fig. 9a: the three sub-networks discussed above give as outputs the single work progress, which is passed through the OOBN in Fig. 9-a to the sub-network in Fig. 9-b, where the overall work progress is then computed as a weighted input from each of the sub-networks.
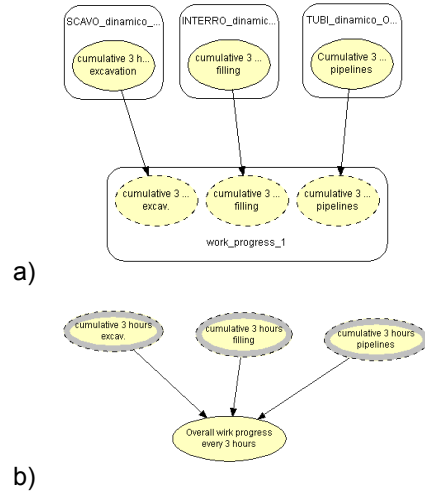

a)

b)

Fig.9. Qualitative structure of the overall work progress network (a) and sub-network for the computation of the work progress (b).

Demonstrations

In this paragraph it is shown that the three sub-networks developed in the previous sub-section are able to model the processes under analysis. To this aim the networks have been used in the running modes and their inferences compared to the dataset. Fig. 10 shows how validation was performed: the variables on the left with red rows are those ones where evidences have been inserted, that is to say the state of the variables have been observed and fixed in the networks.
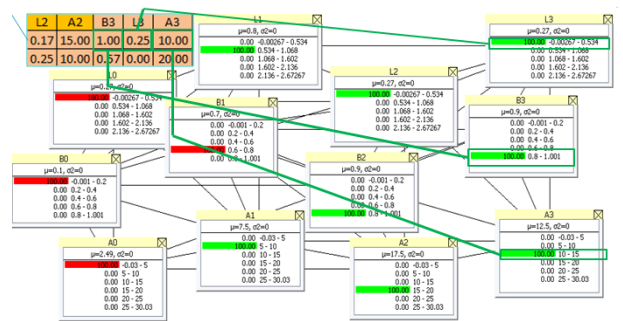

Fig.10. One of the validation cases for the sub-network relative to the excavation phase.

Then the network performs inferences on the future states (time slice no. 3) and gives back the intervals depicted in Fig. 10, which are in accordance with the dataset. In Fig. 11a and 11b two other validation cases for the same sub-network are shown. It came out that the three sub-networks are capable of repre-

senting the process and restituting the non-linear relationships holding between the variables and regarding the work progress on the considered construction site.
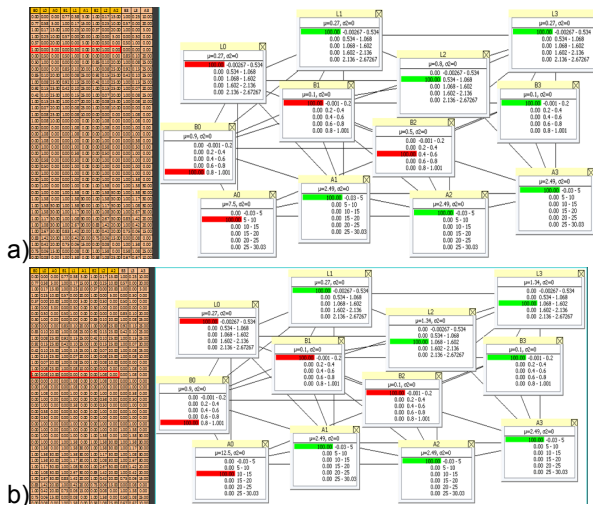

Fig.11. Two validations where estimations by the sub-network have been compared to the database.

CONCLUSIONS AND FUTURE DEVELOPMENTS

In this paper we have faced the problem of real-time work progress estimation. One of the difficulties lays in the need to decompose the tasks into sub-tasks and find out causal relationships among the involved variables so the whole progress may be estimated. In the case of construction sites there is no linear dependence (or functional relationship) between the resources employed at every hour and the work progress. So it was necessary to work out Bayesian networks representing second-order Markovian processes, whose causal relationships have been modeled through the EM learning algorithm, based on the use of Dirichlet probability functions. As it was demonstrated, the relative sub-networks well represent the several processes as probabilistic inferences are within the observations recorded in the dataset. Finally the qualitative structure of the overall network, working out the progress estimation has been proposed. Future steps will be relative to validations of the overall network and inferences performed with respect to tasks spanning longer time periods.

References

1. Giretti, A., Carbonari, A., Vaccarini, M., Robuffo, F., and Naticchia, B., "Interoperable approach in support of semi-automated", 2011 Proceedings of the 28th ISARC, Seoul, Korea, 2011.
2. Lu, W., Huang, G.Q., Li, H., "Scenarios for applying RFID technology in construction project management", *Automation in Construction*, Vol. 20, pp. 101-106, 2011.
3. Zhang, J.P., Hu, Z.Z., "BIM- and 4D-based integrated solution of analysis and management for conflicts and structural safety problems during construction: 1. Principles and methodologies", *Automation in Construction*, Vol. 20, pp. 155-166, 2011.
4. Woo, S., Jeong, S., Mok, E., Xia, L., Choi, C., Pyeon, M., Heo, J., "Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guanzhou MTR", *Automation in Construction*, Vol. 20, pp. 3-13, 2011.
5. Lee, U-K., Kim, J-H., Cho, H., Kang, K-I., "Development of a mobile safety monitoring system for construction sites", *Automation in Construction*, Vol. 18, pp. 258-264, 2009.
6. Navon, R., and Goldschmidt, E., "Can Labor Inputs be Measured and Controlled Automatically?", *Journal of Construction Engineering and Management*, ASCE, July/August 2003.
7. Roh, S., Aziz, Z., Pena-Mora, F., "An object-based 3D walk-through model for interior construction progress monitoring", *Automation in Construction*, Vol. 20, pp. 66-75, 2011.
8. Chiu, C-Y., Russel, A.D., "Design of a construction management data visualization environment: A top-down approach", *Automation in Construction*, Vol. 20, pp. 399-417, 2011.
9. Navon, R., "Automated project performance control of construction projects", *Automation in Construction*, Vol. 14, pp. 467-476, 2005.
10. Zhu, Z., German, S., Brilakis, I., "Detection of large-scale concrete columns for automated bridge inspection" *Automation in Construction*, Vol. 19 pp. 1047-1055, 2010.
11. Wang, L-C., "Enhancing construction quality inspection and management using RFID technology", *Automation in Construction*, Vol. 17, pp. 467-479, 2008.
12. Bezhadan, A.H., Aziz, Z., Anumba, C.J., Kamat, V.R., "Ubiquitous location tracking for context-specific information delivery on construction sites", *Automation in Construction*, Vol. 17, pp. 737-748, 2008.
13. Cheng, T., Venugopal, M., Teizer, J., Vela, P.A., "Performance evaluation of ultra wideband technology for construction resource location tracking in harsh environments", *Automation in Construction*, Vol. 20, pp. 1173–1184, 2011.
14. Pearl, J., Probabilistic reasoning in intelligent systems: networks of plausible inferences, 2nd edition. San Mateo, California: Morgan Kaufmann, 1988.
15. Neapolitan, R.E., Learning Bayesian Networks. Upper Saddle River: Prentice Hall, 2004.
16. Korb, K.B., Nicholson, A.E., Bayesian Artificial Intelligence, 2nd edition. London: Chapman & Hall, 2004.