ADVANCED SCHEDULING DATA PREPARATION, REPRESENTATION, AND ANALYSIS IN SUPPORT OF CONSTRUCTION PLANNING AND SCHEDULING

Jianfeng Wu[1] and Lucio Soibelman[2]

ABSTRACT

Considerable knowledge about construction methods, processes, performance, and risks is essential for decision making in project planning and scheduling. Currently, most companies are still relying on human planners to make important decisions in generating, reviewing, and modifying project schedules manually and in a case-by-case manner, while large volumes of computerized schedules from previous projects are not analyzed for lessons learning and knowledge discovery after projects are finished.

This paper introduces a research effort on preprocessing, representing, and analyzing historical scheduling data to discover more comprehensive, objective and explicit knowledge in support of decision making in project planning and scheduling. The motivations, related work, and a methodology with detailed steps are provided in this paper. Preliminary results from a case study applying this methodology are also presented to show its feasibility in identifying possible correlations between architectures of scheduling networks and their performance during implementation.

KEY WORDS

construction planning and scheduling, knowledge discovery, network analysis

INTRODUCTION

Planning and scheduling is a knowledge-intensive job that is critical to successful completion of construction projects. Most companies rely on human experts to make important decisions in various scheduling tasks. Novice planners either learn lessons from experienced peers, or accumulate their own knowledge by trials and errors. Such practices are usually error-prone, human-dependent, and time-consuming. As a result, poor quality of scheduling work, which is largely due to lack of necessary knowledge, is one of the major causes for a low workflow reliability of 30~60% (Ballard, 1999) that results in costly delays in the $800+ billion dollar construction industry (US Census data quoted by Macomber, 2002).

At the same time that we are starving for knowledge we are drowning in information. Critical Path Method (CPM) based scheduling programs for project planning and scheduling, such as Primavera and MS Project, have been widely applied in the construction industry for many years. Planning and construction history of previous projects, from which lessons were

---

[1]   PhD Student, Dept. of Civil and Envir. Engrg., PH 118G, Carnegie Mellon Univ., Pittsburgh, PA 15213, Phone +1 412/519-7563, jianfen1@andrew.cmu.edu

[2]   Associate Professor, Dept. of Civil and Envir. Engrg., PH 118N, Carnegie Mellon Univ., Pittsburgh, PA 15213, Phone +1 412/268-2952, FAX +1 412/268-7813, lucio@andrew.cmu.edu

learned by human experts, are increasingly available in computerized scheduling databases. However, most of such information-intensive historical data are left intact after projects are finished, without being analyzed to improve scheduling practices in future projects.

To address such a problem, we have initiated a research effort for reusing and analyzing historical scheduling data in support of decision making for project planning and scheduling. Upon its completion, two specific goals are expected to be achieved: 1) development of an explicit mechanism to retrieve and reorganize scheduling data from previous projects into analysis-friendly data representations supporting interactive exploration of project history; 2) development of new applications of graphical analysis tools on CPM-based schedules for pattern recognition and knowledge discovery. This paper presents our initial work in this research effort, including our literature survey, a developed methodology for domain-specific scheduling knowledge discovery, and a case study validating its feasibility.

RELATED WORK IN CONSTRUCTION PLANNING AND SCHEDULING

Due to the critical role that planning and scheduling plays in construction management, many research efforts have been put to improve current practices. Three major types of such efforts were identified in our literature survey:

RESEARCH IN SCHEDULE ANALYSIS AND OPTIMIZATION

Schedule analysis tools, most of which are based on CPM networks, were applied in previous research, including the work of Bubshait and Cunningham (1998) to compare as-planned, as-built, and modified as-built method for their performance and applicability on delay analysis; window analysis techniques employed by Finke (1999) and Hegazy and Zhang (2005); and a project delay computation method developed by Shi, Cheung and Arditi (2001).

Machine Learning (ML) based scheduling tools have also been introduced and developed for construction scheduling tasks, with genetic algorithms (GA) being the most popular tool in recent years. Many researchers have applied GA to explore optimal/near-optimal solutions for scheduling problems by searching only a small part of large and complicated sample spaces of construction alternatives (Feng, Liu and Burns, 1997; Hegazy and Kassab 2003; El-Rayes and Kandil, 2005). Other machine learning tools are also suggested, such as neural networks (Senouci and Adeli, 2001) and case-based reasoning (Dzeng and Tommelein 1997).

RESEARCH IN CONSTRUCTION PROCESS MODELS

Construction process models are collections of information that describe, abstract, or present AEC (Architect/Engineering/Construction) projects (Fischer and Froese, 1996) from the perspective of resource and information flows within activities. Many such models have been proposed and applied in support of scheduling work. Two IDEF (Information Definition) modeling languages, IDEF0 (IDEF, 1993) for function modeling and IDEF3 (IDEF, 1995) for process description have been widely used by construction researchers. Much work has been done to detail the modeling languages into conceptual process models, such as MoPo for construction process analysis and planning (Karhu, 2003); Petri net for construction modeling and simulation (Wakefield and Sears, 1997); WorkPlan scheduling based on lean construction principles (Choo et al., 1999); IDEF3-based ontology development for AEC

interoperability (Tesfagaber et al., 2003). The conceptual models can be integrated with other project information into specific information models for visualization and simulation tasks, as shown in the research by Halpin (1976), Ioannou and Martinez (1996), Lu and AbouRizk (2000), Zhang, Shi and Tam (2002), Kamat and Martinez (2005), among many others.

RESEARCH IN KNOWLEDGE-BASED SCHEDULING SYSTEMS

Solutions based on knowledge-based systems have also been developed in previous research (Zozaya-Gorostiza, Hendrickson, and Rehak, 1989; De La Garza and Ibbs, 1990; Dzeng and Tommelein, 1997; Dzeng and Lee, 2004), in which scheduling knowledge is collected from paper-based sources and human experts, processed and represented by researchers, and integrated into expert systems for automated generation and reviewing of project schedules.

COMPARISON BETWEEN EXISTING RESEARCH AND THIS RESEARCH

Research efforts like the ones described above proved to be helpful in improving efficiency and quality of scheduling work from various perspectives. However, they addressed different research issues from what we are working on. CPM-based and GA-based schedule analysis tools are focused on providing optimal solutions for given problems, without employing or identifying general and explicit scheduling knowledge; most knowledge-based solutions rely on capabilities of researchers and planners involved to collect and represent valid scheduling knowledge for decision support; and related process modeling work is also dependent on domain knowledge and modeling skills of researchers/developers to abstract and present scheduling knowledge at different levels of details. Overall, none of the above research projects were intended to learn explicit and objective knowledge from companies' planning and construction history, which is the major issue that this research is trying to address. A research map below (Fig 1) shows how this research is related to other research efforts and what's new in this research.
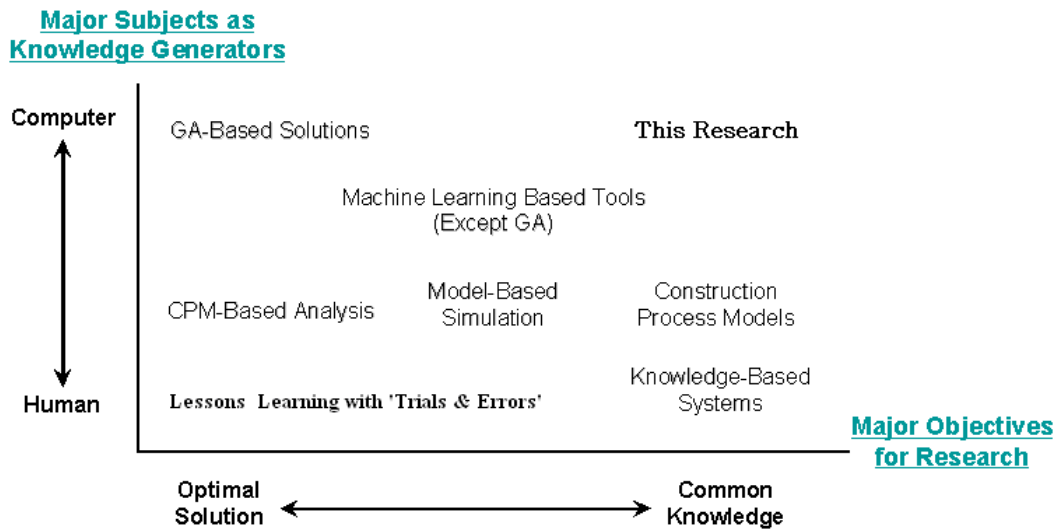

Figure 1: Research Map for This Research and Existing Research

GRAPHICAL ANALYSIS & SCHEDULING KNOWLEDGE DISCOVERY

RECENT DEVELOPMENTS IN GRAPH MINING RESEARCH

Traditionally knowledge discovery research has been done on transactional databases, in which each instance is represented by one row in a data table. With developments of data analysis techniques, objects for data analysis have been extended to more generic graph-based data, so that not only instances (nodes), but also connections (edges) between them are included for pattern recognition. Recently, graph mining tools has received extensive attention from researchers trying to address problems in bioinformatics, social networks, web services, and workflow management. Currently, many graph mining research projects are focused on identifying frequent substructures in large graph databases (Cook and Holder, 2000; Palmer, Gibbon and Faloutsos, 2002; Yan and Han, 2003). Other research efforts include learning process models from work-flow logs (Agrawal, Gunopulos and Leymann, 1998) and graphical data generation (Chakrabarti, Zhan and Faloutos, 2004). Getoor (2003) suggests that other interesting patterns could also be identified on links in graphs, such as link-based classification/clustering and predictions of link type, strength, and cardinality.

SCHEDULING KNOWLEDGE DISCOVERY

Construction schedules can be viewed as a special type of "directed acyclic graph" (DAG), with activities as nodes and directed dependencies between them as edges. However, graphs in schedules are more complicated than current objects in graph mining (e.g., web links and protein molecules) with additional information on activity nodes and dependencies, such as construction products, activity durations, resource constraints, etc. The additional complexity requires that existing graph mining tools be adapted to learn domain-specific knowledge from historical scheduling data, which we are working on in this research.

   With the development of appropriate data exploration and analysis tools, accumulated schedules could provide planners with a wealth of embedded construction knowledge that is critical for their decision making, such as mutual interdependencies between activities as observed from previous projects (e.g. "Pouring of concrete must be followed by at least 2 days of curing"); sequences for finishing a given job as frequently chosen by planners (e.g. "Installing of concrete columns is executed in the order of installing reinforcement→erecting formwork→ pouring concrete→curing→removing formwork); and potential problems in specific sub-networks as implied by repetitive discrepancies between as-planned and as-built schedules (e.g., "In all schedules, 80% of the subsections with 4 or more parallel sequences ended up with at least one activity sequences being delayed").

   Similar to the processes of knowledge discovery in transactional databases, we developed a methodology as a detailed guideline for scheduling knowledge discovery in graphical and CPM-based project schedules. For better understanding, a case study applying graphical analysis tools on a project control database for scheduling knowledge discovery is presented in the next section, with detailed steps including data acquisition, preparation, representation, and analysis.

CASE STUDY

SCHEDULING DATA ACQUISITION

The data for this case study was collected from a large capital facility project. The data spans four months of excavation and foundation work. Roughly 21,000 tasks were recorded with dependencies between them. Such tasks were planned separately into about 2,600 networks composed of 3~80 tasks. Also, data regarding the reasons for non-completions of individual tasks during implementation have been collected, so that project managers could have an overview of the work flows, their variability, and causes impacting non-completions. Such information can be helpful in guiding project managers to focus on specific causes for certain tasks. However, little is known about other crucial scheduling knowledge, such as influence of the original design of scheduling networks on non-completions of tasks in these networks.

Oliveira, Soibelman and Choo (2004) worked on this data to identify a specific graphical pattern, frequent sequences for failures to complete planned tasks, i.e., how the failure of one task may contribute to other failures of its downstream tasks. Different from their work, this study was intended to find more general graphical patterns from the same set of scheduling networks by: 1) creating generic and concise type descriptions for networks with varied size and complicated architectures; 2) identifying correlations between network types and their probabilities of non-conformances in implementation; 3) analyzing influences of positions of one task within a network on its probability of being not completed as planned.

SCHEDULING DATA PREPARATION

In addition to other general data preparation operations as studied in previous construction KDD research (Soibelman, Kim and Wu, 2005), a specific and necessary data preprocessing operation in this case study was the removal of redundant dependencies in networks. According to Kolisch, Sprecher, and Drexel (1995), a dependency from activity X to Y is redundant if there is another activity Z, such that Z is succeeding to X and Y is succeeding to Z, directly or indirectly. As illustrated in Fig 2, a direct dependency from X to Y is unnecessary because Y can not be started immediately after X anyway.
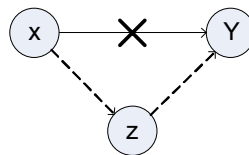

Figure 2: An Example for Redundant Dependencies

DATA REPRESENTATION WITH NETWORK TYPE DESCRIPTION

The architectures of scheduling networks may influence their risks and performance during implementation to a great extent. A challenge here was to describe the networks in a concise and precise way, so that networks having similar architectures and performance could be abstracted into same or close descriptions. A novel network type description addressing this challenge is detailed below using a network in the following figure (Fig 3) as an example.

N1

N2

Further
Subdivision
of N2

(a) N21 (Upper Sub-Network)

B

C

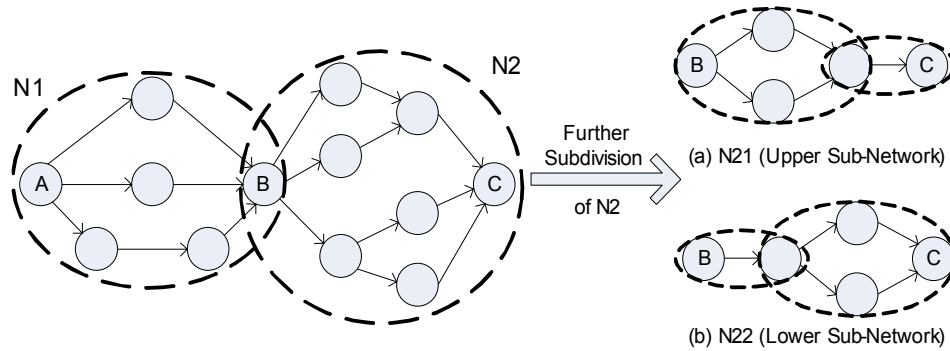(b) N22 (Lower Sub-Network)

Figure 3: An Example for Network Type Description

Three major observations could be seen from Fig 3: 1) how a network is split into branches or converged from branches is a good indicator of its complexity since such splits/converges generate or eliminate parallelisms that result in variability; 2) there exist multi-hierarchical sequential/parallel sub-networks within a large and complicated network that are interrelated only through very few connecting activities – for instance, a network in Fig 3 can be divided into two sequential sub-networks N1 and N2 so that N1 is preceding to N2 only through activity B, and N2 could be further divided into two parallel sub-networks which have only two common activities, the starting activity B and ending activity C; 3) when subdivisions continues, some sub-networks would eventually consist of 'atomic' sequences of activities without any branches (e.g., all the three sequences from A to B), which could be simplified into paths from the starting task to the end.

Based on these observations, a scheduling network could be described in an abstract way focusing on how a work flow goes through it by splitting/converging, and finally arriving at its end. Taking the same network as an example, the abstraction comprises two major tasks:

1) Recursive divisions of the network and sub-networks: sequential/parallel divisions are alternately applied in a top-down manner, until the network is eventually composed of only 'atomic' sequences. In Fig 3, N1 could be divided into 3 parallel 'atomic' sequences, while N2 could be divided into two parallel sub-networks that could be further subdivided.

2) Type description in a bottom-up manner: in this stage, activity identifications were removed since we only concerned the abstract description of the network architecture. In a reverse direction to recursive divisions, the type description for a network/sub-network could be obtained by following the basic rules as below:

- Basic descriptions: if a sub-network is composed of multiple parallel 'atomic' paths from a starting activity to the end, it would be represented as $1{\rightarrow}n{\rightarrow}1$, where n is the number of paths. For example, N1 in Fig 3 could be represented as $1{\rightarrow}3{\rightarrow}1$.

- Sequential combinations: the type description of a sub-network like N21 could be got by combining the descriptions of its two sequential components as shown in Fig 3(a), $1{\rightarrow}2{\rightarrow}1$ and $1{\rightarrow}1$, into $1{\rightarrow}2{\rightarrow}\mathbf{1}{\rightarrow}1$, where the middle $\mathbf{1}$ is shared by both parts. Similarly, the sub-network N22 in Fig 3(b) could be represented as $1{\rightarrow}\mathbf{1}{\rightarrow}2{\rightarrow}1$.

- Parallel combinations: the type description of a sub-network like N2 could be obtained by combining the descriptions of N21 and N22 into $\mathbf{1}{\rightarrow}(2{\rightarrow}1,1{\rightarrow}2){\rightarrow}\mathbf{1}$, in

which *1*'s at both ends are shared by the two components (N21/N22), while the paratheses enclose detailed structures inside N21 and N22 separated by comma(s).

As for the network in Fig 3, its final type description would be the sequential combination of descriptions of N1 and N2 as 1→3→1→(2→1,1→2)→1. A special case here is that if a network is composed of one single sequence of activities, it could be simply described as '1'.

SCHEDULING DATA ANALYSIS FOR NON-CONFORMANCES

Type descriptions for all 2,600 scheduling networks in this case study were generated using a computer program developed on a Matlab® 7.0 platform. When the abstract type descriptions of these networks were compared with their probabilities of non-conformances during implementation, some interesting patterns were identified as follow.

First of all, among the identified 64 distinct types for all 2,600 networks (except ~5% of them with exceptional network structures on which our current solution was not applicable, which we are still studying on and will present in later works), some types of network were apparently less likely to have non-conformances than others. To simplify the analysis, these types were categorized into 4 groups based on their common graphical representations: single sequence, single start to multiple branches, multiple starts to single sequence, and multi-hierarchical starts. The following table (Table 1) shows the general description, the non-conformance rate, and the number of networks for each group, together with examples of generated type descriptions for networks in those groups.

Table 1: Groups of Network Types and their Non-Conformance Rates

Group No.	Group	Examples of Type Descriptions	General Description	# of Networks	# of Non-Conf.	Non-Conf. Rate (%)
1	Single Sequence	1	1	1851	493	26.6
2	Single Start to Branches	1→2, 1→3→1, 1→2→1→2→1, …	1→n→…	51	12	23.5
3	Multiple Starts to Single Sequence	2→1, 3→1→2→1, 4→1…	n→1→…	365	159	43.6
4	Multi-Hierarchical Starts	(1→6,1)→1, (2→1,2→1)→1, ...	Others	167	98	58.7

These non-conformances rates were then compared in a pair-wise manner using a statistical hypothesis testing equation below (Hogg and Tanis, 2001):

$$z_{\alpha/2} = \frac{|p_a - p_b|}{\sqrt{\dfrac{p_a(1-p_a)}{n_a} + \dfrac{p_b(1-p_b)}{n_b}}} \qquad (1)$$

in which 2 groups, A and B, with different non-conformances rates, were tested with a null hypothesis $p_a=p_b$ vs. an alternative hypothesis $p_a \neq p_b$. In such tests, $z_{\alpha/2} \geq 1.960$ means that the null hypothesis was rejected with a 95% confidence, i.e., that groups A and B had dissimilar

non-conformance rates by statistics. In comparisons using equation (1), groups 1 and 2 had significantly lower non-conformance rates than group 3 (with $z_{\alpha/2}$=6.06 for group 1 vs. group 3, and 3.09 for group 2 vs. group 3, respectively), and non-conformance rates of groups 3 and 4 were significantly different as well (with $z_{\alpha/2}$=3.28).

Also, comparisons between non-completion rates of activities at different positions turned out some meaningful results. The non-completion rate of tasks connecting sequential or parallel sub-networks, e.g. those with more than one preceding and/or succeeding tasks, was significantly larger than that of tasks among 'atomic' sequences (15.6% vs. 8.4% with a significant value $z_{\alpha/2}$=8.94). When looked into more closely, non-completion rates of tasks in 'atomic' sequences varied with their positions in their corresponding sequences as well: obviously, the closer that a task is to the starting/ending tasks, the higher non-completion rate it would have during implementation.

EVALUATION AND FUTURE WORK

With these results, we may conclude that in this construction project, a scheduling network could be more reliable if it was designed to be a single sequence, or to start with just one task. This makes sense because when a schedule is started, the management team may not have all resources well prepared; but after the jobs are started, usually there will be more necessary resources ready to support more parallel implementations. If the network architectures could not be changed, project managers should focus more managerial efforts to control networks with type descriptions from group 3 and 4 to prevent possible non-completions. Also, in this case study, tasks at the splitting or converging points within a network, and those close to these tasks, seemed to have higher probabilities of non-completions. This is a reasonable discovery too, considering that conflicts between tasks, and thus non-completions of related tasks, are most likely to occur when parallel jobs are started or ended. Project managers should be better prepared for such tasks by double-checking available resources, prerequisites, and other conditions.

The methodology applied in this case study proved to be both feasible and valid since it enables automated characterization of networks with similar architectures and performance into same or close type descriptions. Also, this study is an important initial step for further and more in-depth developments in many aspects: 1) the recursive division method used in this case study made it possible to decompose large and complicated scheduling networks into multiple levels of sub-networks and sequences, which is required to build the analysis-friendly data representation for project planning and construction history; 2) the generated type descriptions could be viewed as extracted graphical features of scheduling networks, which could be integrated with other features of activities and dependencies for identifying other general and useful patterns in scheduling data; 3) the complete process of scheduling data preparation, representation, and analysis in this case study provided primary insights and experiences for the following development of this research on a larger scale.

CONCLUSIONS

Lack of appropriate data representation and analysis tools for large and complex CPM-based networks is an obstacle for lessons learning from computerized project schedules that have

been widely available in the construction industry. This paper introduces a methodology developed to preprocess, represent, and analyze historical project schedules for knowledge discovery in support of construction planning and scheduling. Preliminary results from a case study applying the process of scheduling data preparation, representation, and analysis are also presented, with interesting and valid graphical patterns discovered in a project planning and control database. Research efforts is under way to extend the current research with the objective of retrieving and reorganizing scheduling data from previous projects into analysis-friendly data representation, and to allow the application of latest data analysis techniques including graph mining tools for scheduling knowledge discovery.

ACKNOWLEDGEMENT

REFERENCES

Agrawal, R., Gunopulos, D. & Leymann, F. (1998) "Mining Process Models from Workflow Logs." *Proc. of 6th Intl. Conf. on Extending Database Technology,* Valencia, Spain.

Ballard, G. (1999). "Improving Work Flow Reliability". P*roc. of the 7th Conf. of the Intl. Group for Lean Construction*, Berkeley, California, USA

Bubshait, A. and Cunningham, M. (1998). "Comparison of Delay Analysis Methodologies". ASCE, *J. of Constr. Engrg and Mgmt*, 124(4), 315-322

Choo, H. et al. (1999). "WorkPlan: Constraint-Based Database for Work Package Scheduling." ASCE, *J. of Constr Engrg & Mgmt*, 125(3), 151-160.

Chakrabarti, D., Zhan. Y. and Faloutsos, C. (2004) "R-MAT: A Recursive Model for Graph Mining". *Proc. of SIAM Intl. Conf. on Data Mining*, Lake Buena Vista, Florida, USA

Cook, D. and Holder, L. (2000) "Graph-Based Data Mining". *IEEE Intelligent System*, 15(2), pp 32-41

De La Garza, J. and Ibbs, C. (1990). "Knowledge-elicitation study in construction scheduling domain". ASCE, *J. of Comp. in Civil Engrg.*, 4(2), 135-153

Dzeng, R. & Lee, H. (2004). "Critiquing Contractors' Scheduling by Integrating Rule-Based and Case-Based Reasoning". Elsevier Science, *Auto. in Constr.*, 13(5), 665-678

Dzeng, R. and Tommelain, I. (1997). "Boiler Erection Scheduling Using Product Models and Case-Based Reasoning". ASCE, *J. of Constr. Engrg and Mgmt*, 123(3), 338-347

El-Rayes, K. and Kandil, A. (2005). "Time-Cost-Quality Trade-Off Analysis for Highway Construction". ASCE, *J. of Constr. Engrg and Mgmt*, 131(4), 477-486

Feng, C., Liu, L. and Burns, S. (1997). "Using Genetic Algorithms to Solve Construction Time-Cost Trade-Off Problems". ASCE, *J. of Comp. in Civil Engrg.*, 11(3), 184-189

Finke, M. (1999). "Window Analyses of Compensable Delays". ASCE, *J. of Constr. Engrg and Mgmt*, 125(2), pp. 96-100

Fischer, M. and Froese, T. (1996). "Examples and Characteristics of Shared Project Models." ASCE, *J. of Comp. in Civil Engrg.*, 10(3), 174-182.

Froese, T. et al. (1999). "Industry Foundation Classes for Project Management - A Trial Implementation." *ITcon*, Vol. 4, 17-36, http://www.itcon.org/1999/2

Getoor, L. (2003). "Link mining: A New Data Mining Challenge". *ACM SIGKDD Explorations Newsletter*, 5(1), 84-89

Halpin, D. (1976). "CONSTRUCTO: An Interactive Gaming Environment" ASCE, *J. of the Constru. Div.*, 102(1), 145-196

Hegazy, T. and Kassab, M. (2003) "Resource Optimization Using Combined Simulation and Genetic Algorithms". ASCE, *J. of Constr. Engrg and Mgmt*, 129(6), 698-705

Hegazy, T. and Zhang, K. (2005) "Daily Windows Delay Analysis". ASCE, *J. of Constr. Engrg and Mgmt*, 131(5), pp. 505-512

IDEF0: tegration Definition for Function Modeling. (1993). http://www.idef.com/idef0.html

IDEF3 – Process Description Capture Method. (1995). http://www.idef.com/idef0.html

Ioannou, P. & Martinez, J. (1996). "Comparison of Construction Alternatives Using Matched Simulation Experiments". *J. of Constr. Engrg. and Mgmt.*, 122(3), 231-241

Karhu, V. (2003). "A View-Based Approach for Construction Process Modeling". *J. of Computer-Aided Civil & Infrastructure Engrg.*, Issue 18, 275-285

Kolisch R., Sprecher A. and Drexl A. (1995). "Characterization and Generation of a General Class of Resource-Constrained Project Scheduling Problems." *Management Science*, Vol. 41, 1693-1703

Lu, M. and AbouRizk, S. (2000). "Simplified CPM/PERT Simulation Model". ASCE, *J. of Constr. Engrg and Mgmt*, 126(3), 219-226

Macomber, J. (2002). "US Construction Demographics: Some Relevant Statistics to Consider". [online] (accessed September 20, 2005) http://ocw.mit.edu/NR/rdonlyres/Civil-and-Environmental-Engineering/1-464Spring2004/

Palmer, C., Gibbons, P. and Faloutsos, C. (2002). "ANF: A Fast and Scalable Tool for Data Mining in Massive Graphs". *ACM SIGKDD '02*, Edmonton, Alberta, Canada

Senouci, A. and Adeli, H. (2001) "Resource Scheduling Using Neural Dynamics Model". ASCE, *J. of Constr. Engrg. and Mgmt.*, 127(1), pp 28-34.

Shi, J., Cheung, S. and Arditi, D. (2001). "Construction Delay Computation Method". ASCE, *J. of Comp. in Civil Engrg.*, 127(1), 60-65

Soibelman, L, Kim, H., and Wu, J. (2005) "Knowledge Discovery for Project Delay Analysis" *Bauingenieur*, Springer VDI Verlag, Feburuary 2005.

Tesfagaber, G. et al. (2003). "Ontology Modeling of IDEF3 Process Description for AEC Applications Interoperability". *Proc. of Constr. Research Congress*, Honolulu, HA, USA

Wakefield, R.R. and Sears, G.A. (1997). "Petri Nets for Simulation and Modeling of Construction Systems." ASCE, *J. of Constr. Engrg and Mgmt*, 123(2), 105-112

Yan, X. and Han, J. (2003) "CloseGraph: Mining Closed Frequent Graph Patterns", *Proc. of ACM SIGKDD Intl. Conf. on Knowledge Disovery & Data Mining*, Washington, DC

Zhang, H., Shi, J. and Tam, C. (2002). "Iconic animation for activity-based construction simulation". ASCE, *J. of Comp. in Civil Engrg.*, 16(3), 157-164

Zozaya-Gorostiza, C., Hedrickson, C. and Rehak, D. (1989). *Knowledge-Based Process Planning for Construction and Manufacturing.* Academic Press, Boston, MA, USA