Constraint Optimization on Privacy Preserved Meeting Organization

By

M.M.A.S.T. Akmeemana - Index No: 20020015
A.I. Vidanage - Index No: 20021089
K.P.G.K. Jayathilake - Index No: 20020521

Supervisor: Dr. C.I. Kappetiyagama

Co-supervisor: Mr. Tharindu Wijethilaka

IS 4101 – Final Year Project in Informatioon Systems

Degree of Bachelor of Science Honours in
Information Systems

University of Colombo School of Computing
35, Reid Avenue, Colombo 07,
Sri Lanka

April 2024

Table of Contents

Contents

Table of Contents

1 Project profile

1.1 Revised introduction

A meeting is a synchronous communication, based on one or more documents, occurring at one or more locations. In today's interconnected world, meetings are held across various contexts and serve a multitude of purposes, from business discussions to academic collaborations. The organization of these meetings is increasingly supported by a range of platforms offered by different vendors, each designed to streamline the scheduling process and enhance user experience. Substantial research has focused on improving the quality and efficiency of meeting organization, addressing challenges such as preventing the overlap of participants in concurrent meetings or avoiding scheduling multiple meetings at the same location simultaneously. But, optimization of privacy concerning meeting content remains an unexplored area. Current research has not adequately addressed the constraints necessary to safeguard meeting privacy, during the meeting organization process.

1.2 Motivation

It's observed that the privacy of meetings is violated when they are organized in an adhoc manner. Because, when there are various topics to discuss, it's difficult to decide on participants relevant to each topic, without the help of a formal constraint optimization process. Privacy violation mainly takes place in the following manners.

- Adding unauthorized participants to meetings, such that they get access to confidential data restricted from them.
- Selecting privacy-violating meeting modes, in which meeting content can leak to unauthorized people.

1.3 Scope

- Identifying constraints related to the privacy of meetings. (Currently, we have identified the following factors as constraints affecting the privacy of meetings)
  - Meeting content
  - Participants
  - Meeting mode
- Optimizing identified constraints in a viable manner, producing an effective solution for supporting privacy-preserved meeting organization.
- Evaluating the solution to prove the validity of the solution.

1.4 Aims

- Selecting meeting participants, and ensuring the privacy of meeting content discussed in the meeting.
- Selecting a privacy-preserving meeting mode, for conducting the particular meeting.

- Evaluating the solution in an acceptable, credible approach for proving its validity.

1.5 Objectives

- Based on the assumption that every meeting has at least one document associated with it (Meeting agenda document), selecting meeting participants depending on Access Control Lists (ACL) of documents associated with the meeting.
- Selecting a privacy-preserving meeting mode, based on availability and locations of selected meeting participants.
- Evaluating the solution by mapping it to an already existing well-known constraint optimization solution, using a standard mapping notation.

2 Comprehensive literature survey and identified research gap

2.1 Comprehensive Literature survey

1. **Introduction to Scheduling Problems in Various Domains**
   - **Business-to-Business (B2B) Meeting Scheduling Problem (Bofill et al. 2022):** This problem addresses optimizing the scheduling of meetings in B2B contexts by minimizing gaps in participants' schedules and adhering to their availability constraints. This is particularly relevant for events where limited availability and high participant turnover complicate scheduling. Various solving techniques like Constraint Programming (CP), Mixed-Integer Programming (MIP), and Maximum Satisfiability (MaxSAT) have been examined, with MaxSAT showing the highest effectiveness in managing the scheduling complexities of B2B meetings.
   - **Multi-Mode Resource-Constrained Project Scheduling (MRCPSP) with Preemptive Extensions (Peteghem and Vanhoucke 2010):** This study explores the multi-mode scheduling problem with a focus on minimizing project duration under resource constraints, incorporating both renewable and nonrenewable resources. It introduces a genetic algorithm with a preemptive scheduling feature, allowing for splitting activities and facilitating dynamic adaptability. This approach is crucial for managing project tasks in resource-limited settings like construction or manufacturing, where project delays and resource conflicts need to be minimized.
   - **Job-Shop Scheduling under the Theory of Constraints (TOC) (Golmohammadi 2015):** Job-shop scheduling focuses on maximizing production efficiency in complex, resource-intensive environments, like the automotive industry. This paper applies the Theory of Constraints (TOC) to optimize throughput by identifying system bottlenecks and adjusting scheduling to improve resource allocation and task flow. TOC's effectiveness in job-shops lies in its adaptability to different production systems and complexity levels, ensuring high efficiency despite the variability of production demands.

- **Distributed constraint optimization with MULBS (Enembreck and André Barthès 2012):** The paper introduces a new distributed constraint optimization problem (DCOP) algorithm called MULBS, which stands for Multiple Local Bounded Search. This algorithm is designed to enhance the efficiency of solving complex optimization problems, particularly in the context of collaborative meeting scheduling. MULBS utilizes a pseudo depth-search tree to prioritize agents, allowing for asynchronous problem-solving and minimizing communication overhead. The study demonstrates that MULBS can quickly find high-quality solutions for loosely connected graphs, although it may struggle to converge optimally in over-constrained scenarios. The empirical results indicate that MULBS outperforms existing algorithms like ADOPT and DPOP in terms of communication efficiency and runtime, making it a promising approach for distributed scheduling tasks in collaborative environments.
- **Work-Package Planning and Schedule Optimization for Projects with Evolving Constraints (Abuwarda and Hegazy 2016):** Traditional schedule optimization models often struggle with predefined activities and fixed constraints, which limits their effectiveness in dynamic environments. To address these challenges, this paper introduces a flexible constraint programming (CP) framework. This framework is designed to optimize schedules both during the early planning stage and just before construction begins. This dual focus allows for better adaptability to evolving project needs

2. **Approaches and Techniques for Scheduling**

   - **Constraint Programming (CP), Mixed-Integer Programming (MIP), and Maximum Satisfiability (MaxSAT) in B2B Scheduling (Bofill et al. 2022):** CP provides a robust way of managing constraints like availability and capacity, while MIP optimizes for overall schedule feasibility. MaxSAT, however, has emerged as the preferred technique for B2B scheduling due to its high solving speed and flexibility in handling participant preferences. In cases where a feasible solution isn't possible, MIP's effectiveness in quickly identifying unsolvable instances proves advantageous.
   - **Bi-Population Genetic Algorithm in Project Scheduling (Peteghem and Vanhoucke 2010):** This paper leverages a genetic algorithm with two populations—left-justified and right-justified schedules—for optimizing the resource-constrained project scheduling problem. This approach enhances flexibility by accommodating preemptive scheduling, where tasks can be interrupted and resumed, which is crucial for project environments with varying resource demands. The genetic algorithm's comparative advantage lies in its adaptability, outperforming existing methods under certain preemptive scheduling conditions.
   - **Theory of Constraints (TOC) in Job-Shop Scheduling (Golmohammadi 2015):** TOC-based scheduling focuses on identifying and elevating bottlenecks in complex job-shop systems to maximize output. Through simulation and analysis, this method proves effective in managing both free and non-free product flows, adjusting batch sizes, and re-prioritizing tasks based on throughput requirements. The TOC approach aligns well with real-world

production systems, where resource constraints and unpredictable bottlenecks challenge optimal scheduling.

- **The MULBS Scheduling Approach (Enembreck and André Barthès 2012):** MULBS, or Multiple Local Bounded Search, is a specific algorithm designed to enhance the DCOP framework for scheduling tasks. It employs a pseudotree structure to organize agents, which helps prevent infinite loops and allows for asynchronous problem-solving. The algorithm begins with a bottom-up process to generate an initial solution using a minimal conflict strategy, followed by a top-down refinement phase where agents asynchronously propose better solutions. This dual-phase approach not only reduces communication overhead but also ensures that high-quality solutions are consistently found, even in loosely connected graphs. While MULBS may struggle to find the optimal solution in over-constrained scenarios, it generally outperforms other algorithms like ADOPT and DPOP, making it a robust choice for distributed meeting scheduling tasks.
- **Schedule Optimization Framework (Abuwarda and Hegazy 2016):** The primary method employed in this research is constraint programming (CP). The research methodology consists of developing two mathematical models. The first model focuses on schedule optimization during the early planning stage, considering all necessary features to address initial constraints. The second model is designed for schedule optimization immediately before construction, ensuring adaptability to any last-minute changes or constraints. The proposed CP formulation was implemented using the IBM ILOG CPLEX Optimization Studio 12.6. This software is linked to Microsoft Excel and Microsoft Project to facilitate input and output processes, making it user-friendly for project managers. Overall, these methods provide a robust framework for optimizing project schedules in dynamic environments, allowing for effective decision-making throughout the project lifecycle.

3. **Challenges and Limitations in Scheduling Models**

   - **B2B Meeting Scheduling (Bofill et al. 2022):** One limitation lies in managing fixed meetings and precedence relationships, which restricts flexibility. Additionally, adapting these models for real-time scheduling changes remains challenging. The dynamic nature of B2B events, where availability can shift unexpectedly, demands adaptive solutions that current models struggle to provide.
   - **Resource-Constrained Project Scheduling (MRCPSP) (Peteghem and Vanhoucke 2010):** In project scheduling, handling infeasible schedules within the genetic algorithm and managing the complexities of preemptive scheduling constraints presents significant challenges. The difficulty in finding an optimal solution increases with project size, and additional constraints for activity splitting further complicate the model's applicability.
   - **Job-Shop Scheduling (TOC) (Golmohammadi 2015):** The main challenges are adjusting TOC rules for complex, real-world job-shop systems and managing setup times and batch sizes that can impact scheduling accuracy. Real-world job-shop environments also introduce interdependent bottlenecks,

making it difficult to apply TOC rules without modifications for specific operational contexts.

- **MULBS (Enembreck and André Barthès 2012):** *Performance in Over-Constrained Graphs:* MULBS may struggle to converge to the optimal solution in over-constrained scenarios, particularly in almost-complete graphs where tight constraints hinder the effective exploration of solutions Reliance on Local Search Strategies: The algorithm's two-phase process can lead to suboptimal solutions if the initial candidate is weak. Limited exploration during local searches may prevent significant improvements. *Computational Resource Requirements:* While reducing communication overhead, the need for agents to manage local solutions can increase computational demands, especially in larger networks. This trade-off can be challenging in resource-constrained environments.
- **Schedule Optimization Framework (Abuwarda and Hegazy 2016):** The crashing strategy currently treats overtime and overstaffing as mutually exclusive, which may not reflect real-world practices, Additionally, the model does not permit overlapping activities, which could help reduce project duration. The model also assumes that durations, costs, and resources are deterministic, overlooking uncertainties like weather and site conditions, which could be addressed through probabilistic scheduling in future work. Furthermore, the model lacks the capability for activity splitting, which could optimize resource allocation.

4. **Performance Analysis and Results**

- **B2B Scheduling with MaxSAT (Bofill et al. 2022):** MaxSAT consistently outperforms other approaches in solving speed and satisfaction of participant preferences. Experiments show that fairness constraints, which reduce idle times for participants, lead to more balanced schedules. However, MIP proves more efficient in detecting unsolvable instances, providing valuable feedback in the early stages of scheduling.
- **Genetic Algorithm for MRCPSP (Peteghem and Vanhoucke 2010):** The genetic algorithm demonstrates competitive performance, especially when employing a bi-population strategy. Preemptive scheduling has been shown to improve project makespan under specific conditions, making it effective for tasks that require frequent rescheduling due to resource constraints. The study shows that adjusting population size based on activity count enhances the efficiency of the genetic algorithm in handling complex projects.
- **TOC in Job-Shop Scheduling (Golmohammadi 2015):** The TOC-based scheduling methods effectively address bottlenecks and prioritize tasks, yielding better throughput outcomes than standard scheduling plans. Simulation results underscore the need to prioritize non-free products and manage batch sizes to optimize production flow. Despite these successes, modifications to TOC rules are essential to adapt them for broader real-world applications.
- **MULBS (Enembreck and André Barthès 2012):** MULBS demonstrated performance comparable to DPOP in scenarios with 10 variables and varying densities, while ADOPT struggled due to its search strategy and local information exchange. The algorithm significantly reduces communication needs,

as MULBS messages maintain a constant size, unlike DPOP messages, which grow exponentially, resulting in better throughput for MULBS. In experiments with higher variable counts (25 and 30) at a density of 3, MULBS achieved 100% completeness, consistently identifying the best solution. The algorithm also scales effectively with increasing problem complexity, outperforming both ADOPT and DPOP in throughput and runtime.

- **Schedule Optimization Framework (Abuwarda and Hegazy 2016):** Presents a work-package planning and schedule optimization model that effectively manages projects with evolving constraints. Key findings include the model's flexibility, demonstrated by consistent performance across various objective functions at different project stages. During the early planning phase, it utilized a discrete time-cost tradeoff optimization to select suitable network paths and construction methods, achieving a project cost of $1.467 million while meeting a 165-day deadline. As projects evolved, the model adapted to new constraints using linear time-cost tradeoff optimization, allowing for fine-tuning of execution schedules. Overall, the results indicate the model's potential for integration into standard project management software, making optimization a more mainstream tool in construction management.

5. **Future Research Directions**

- **B2B Meeting Scheduling (Bofill et al. 2022):** Future research could focus on integrating machine learning for real-time preference prediction, enhancing adaptive scheduling for dynamic events. Adding constraints for multi-participant meetings and minimizing travel or location changes are additional areas for development. Real-time data usage for dynamic adaptability would further improve practical applicability in B2B contexts.
- **MRCPSP with Genetic Algorithms (Peteghem and Vanhoucke 2010):** Exploring adaptive algorithms that dynamically adjust based on project changes and expanding preemptive features to other scheduling types could advance project scheduling solutions. Additionally, adapting the genetic algorithm to other scheduling contexts and further refining its performance with varying population sizes and fitness functions could yield broader applicability.
- **TOC for Job-Shop Scheduling (Golmohammadi 2015):** Future work could investigate more complex job-shop settings, incorporating batch adjustments and adapting TOC rules to evolving bottlenecks and resource dependencies. Expanding TOC's applicability to diverse industries and refining the algorithms for fluctuating production schedules would enhance TOC's utility in job-shop environments.
- **MULBS:** Future work suggests conducting more evaluations on complex problems, particularly those involving over-constrained graphs, to assess the algorithm's effectiveness in challenging scenarios. Additionally, involves comparing MULBS with other almost-complete DCOP algorithms, aiming to position it within the broader landscape of distributed constraint optimization techniques and identify its relative strengths and weaknesses. Lastly, the scalability of MULBS in various problem settings, especially in loosely connected dependency graphs, will be further explored to understand its performance under different conditions.

- **Schedule Optimization Framework (Abuwarda and Hegazy 2016):** Future work will focus on allowing a combination of crashing strategies, enabling overlapping activities to reduce project duration, and incorporating uncertainties related to factors like weather. Additionally, the model may include activity splitting to improve resource allocation and adapt to the construction phase by integrating progress events and practical constraints. Finally, there is an emphasis on examining the model's performance in large-scale projects to validate its effectiveness and flexibility in diverse environments.

6. **Conclusion**

   - This literature review highlights the strengths and limitations of CP, MIP, MaxSAT, genetic algorithms, and TOC across different scheduling challenges. Each approach shows promise within its specific context, yet the need for adaptability and real-time solutions remains prominent. Future advancements will likely focus on integrating predictive analytics, refining preemptive scheduling, and customizing TOC rules to address evolving real-world scheduling demands. Together, these improvements aim to bridge the gap between theoretical models and practical scheduling needs, ultimately leading to more robust and versatile scheduling frameworks.
   - The paper presents the MULBS algorithm for distributed constraint optimization problems (DCOP), focusing on collaborative meeting scheduling. It highlights the algorithm's efficiency in generating solutions through a local search strategy, which minimizes communication overhead. Experimental results indicate that MULBS outperforms traditional algorithms like ADOPT and DPOP, particularly in terms of scalability and resource usage, although it may not always guarantee optimal solutions in high-density graphs. The authors suggest further research to compare MULBS with other algorithms and explore its adaptability to complex problems.

2.2 Research gap

- **Constraint optimization related to meeting privacy** remains a critical research gap. Existing meeting constraint optimization research has not adequately incorporated specific privacy constraints, limiting their effectiveness in the domain of meeting privacy.

3 Revised research question

3.1 Hypothesis

Meeting participant selection is primarily dominated by the access control lists of the **documents** presented, and the choice of **meeting mode** depends on the participants' locations.

3.2 Revised research question

How to optimize identified privacy-preserving meeting constraints (meeting content represented by documents, meeting participants, meeting mode) to validate the hypothesis

?

4 Research approach

1. Identify constraints associated with privacy-preserved meeting organization.
2. Develop a basic solution that includes the identified constraints.
3. Fine tune the solution to optimize the identified constraints, using an already existing well-known constraint optimization solution, which is proven to be correct.
4. Evaluate the finalized solution by mapping it formally to the well-known solution utilized, for proving the validity of the finalized solution.

5 Experiments and preliminary results

5.1 Experiments

- We have started to define constraints associated with privacy-preserved meeting organizations. The latest version of currently identified constraints and defined concepts are included in the following document. This content is to be further fine-tuned in the future, depending on our findings in this research.
  - `https://github.com/Thejana-A/IS-4101-git/blob/master/latex/definitions_Sep_20/definition.pdf`
- We have developed a basic workflow in Python, including steps for,
  - Meeting participant selection, based on Access Control Lists of documents related to the meeting.
  - The meeting time slot selection is based on the availability of selected participants.
  - Meeting mode selection is based on the locations of participants, during the eligible time slot.
- This basic workflow was implemented using a set of functions present in Python language. Its latest version is available at the following link.
  - `https://github.com/Thejana-A/IS-4101-git/blob/master/python_code/intersection_nov04.py`
- For constraint optimization purposes, we need to replace Python functions fulfilling above mentioned 3 steps, with eligible well-known constraint optimization algorithms. Because, there is no data-driven or experimental evaluation method to evaluate a program developed by us from scratch, regarding our constraint optimization problem.
- Meeting constraint optimization research conducted by Bofill et al concluded that MaxSAT was the best algorithm for their research problem (Bofill et al. 2022). Therefore, we decided to analyze the MaxSAT algorithm to check whether it is eligible for our meeting constraint optimization problem as well.

- Currently, we are analyzing the following MaxSAT implementations simultaneously, to see whether they can satisfy our constraint optimization requirement regarding **meeting participant selection, based on Access Control Lists of documents related to the meeting**.
  - Pure MaxSAT algorithm:
    - https://github.com/sukrutrao/MaxSAT-Solver
  - Partial weighted MaxSAT algorithm:
    - https://github.com/FlorentAvellaneda/EvalMaxSAT
  - Both Pure and partial weighted algorithms:
    - https://github.com/pysathq/pysat/blob/master/solvers/glucose30.tar.gz
    - https://github.com/pysathq/pysat/blob/master/solvers/glucose421.tar.gz
- We have used inputs created by ourselves and exhaustive input sets available at the following link, for analyzing the above MaxSAT solvers, by observing the output delivered for various inputs.
  - https://github.com/harshad-p/max-sat-solver/tree/master/input

5.2 Current progress

- For now, none of the above MaxSAT solvers were capable of providing promising support for constraint optimization related to **meeting participant selection**. We hope to analyze them further, before shifting to a different algorithm, to check whether any of these MaxSAT algorithms can be mapped to our constraint optimization requirement successfully.

6 Finalized evaluation plan

- Since there is an uncountable number of possibilities for meetings, it is impractical to evaluate the finalized solution (proof of concept) in a data-driven or experimental approach.
- Therefore, it is difficult to evaluate a solution developed from scratch, regarding our constraint optimization problem.
- Accordingly, we are planning to develop our solution based on a well-known constraint optimization solution, which is already proven to be correct. Then, we will formally map our finalized solution to that well-known constraint optimization solution.
- In addition, invariant rules supporting that mapping will be proven by using a standard proof assistant, (Ex: Coq, Isabelle, Lean) to reinforce the credibility.

List of References

References

Abuwarda, Zinab and Hegazy, Tarek (2016). "Work-Package Planning and Schedule Optimization for Projects with Evolving Constraints". In: *Journal of Computing in Civil Engineering* 30.6, p. 04016022. DOI: `10.1061/(ASCE)CP.1943-5487.0000587`. URL: `https://ascelibrary.org/doi/abs/10.1061/%5C%28ASCE%5C%29CP.1943-5487.0000587`.

Bofill, M., Coll, C., Garcia, M., Giraldez-Cru, J., Pesant, G., Suy, J., and Villaret, M. (2022). "Constraint Solving Approaches to the Business-to-Business Meeting Scheduling Problem". In: *Journal of Artificial Intelligence Research* 74, pp. 263–301. DOI: `https://doi.org/10.1613/jair.1.12670`. URL: `https://dl.acm.org/doi/pdf/10.1613/jair.1.12670`.

Enembreck, Fabrício and André Barthès, Jean-Paul (2012). "Distributed constraint optimization with MULBS: A case study on collaborative meeting scheduling". In: *Journal of Network and Computer Applications* 35.1. Collaborative Computing and Applications, pp. 164–175. ISSN: 1084-8045. DOI: `https://doi.org/10.1016/j.jnca.2011.02.016`. URL: `https://www.sciencedirect.com/science/article/pii/S1084804511000543`.

Golmohammadi, Davood (2015). "A study of scheduling under the theory of constraints". In: *International Journal of Production Economics* 165, pp. 38–50. ISSN: 0925-5273. DOI: `https://doi.org/10.1016/j.ijpe.2015.03.015`. URL: `https://www.sciencedirect.com/science/article/pii/S0925527315000833`.

Peteghem, Vincent Van and Vanhoucke, Mario (2010). "A genetic algorithm for the preemptive and non-preemptive multi-mode resource-constrained project scheduling problem". In: *European Journal of Operational Research* 201.2, pp. 409–418. ISSN: 0377-2217. DOI: `https://doi.org/10.1016/j.ejor.2009.03.034`. URL: `https://www.sciencedirect.com/science/article/pii/S037722170900191X`.