

# 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	$\mathbf{Pro}$	ject profile	1
	1.1	Introduction	1
	1.2	Motivation	1
	1.3	Scope	1
	1.4	Basic definitions	2
		1.4.1 Notations	2
		1.4.2 Access Control List	2
		1.4.3 Privacy-preserved meeting	3
	1.5	Aims	3
	1.6	Objectives	3
2	$\operatorname{Lit}_{oldsymbol{\epsilon}}$	erature survey and identified research gap	3
	2.1	Introduction to literature survey	3
	2.2	Literature survey	4
		2.2.1 Meeting Scheduling Challenges	4
		2.2.2 Constraint-Based Optimization Techniques	4
		2.2.3 Genetic Algorithms and Adaptive Scheduling	5
		2.2.4 Theory of Constraints (TOC) and Bottleneck Management	6
		2.2.5 Machine Learning and Hybrid Approaches in Constraint Solving .	6
		2.2.6 Privacy-Centric Distributed Models for Scheduling	6
	2.3	Research gap and conclusion	7
3	Res	earch question	7
	3.1	Hypothesis	7
	3.2	Research question statement	7
4	Res	earch approach	7
5	Experiments and preliminary results		
	5.1	Experiments	8
	5.2	Current progress	9
6	Fina	alized evaluation plan	9
List of References			10

# 1 Project profile

#### 1.1 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, organizing meetings preserving privacy of meeting data remains an unexplored area. Current research has not adequately addressed the constraints necessary to safeguard meeting privacy, during the meeting scheduling process.

#### 1.2 Motivation

It is observed that the privacy of meetings is violated when they are organized in an ad-hoc manner. Because, when there are various topics to discuss, it is difficult to decide on participants relevant to each topic, without the help of a formal constraint resolution 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

- During identification of constraints related to the privacy of meetings, we have identified the following factors as constraints affecting the privacy of meetings.
  - Meeting content
  - Participants
  - Meeting mode
- Resolving identified constraints in a viable manner, producing an effective solution for supporting privacy-preserved meeting organization.
- Evaluating the solution to ensure the validity of the solution.

#### 1.4 Basic definitions

#### 1.4.1 Notations

Following finite sets are defined:

- $\mathcal{D}$ : The set of all documents.
- $\mathcal{R}$ : The set of all roles.
- $\mathcal{I}$ : The set of all individuals
- $\mathcal{L}$ : The set of all locations.
- $\mathcal{T}$ : The set of all time slots.

Following functions are also defined:

$$access: \mathcal{D} \mapsto 2^{\mathcal{R}}(2^{\mathcal{R}} = \text{power set of } \mathcal{R})$$

$$access(d) = \{r \in \mathcal{R} \mid r \text{ has access to } d\}$$

$$transform: \mathcal{I} \times \mathcal{L} \times \mathcal{T} \mapsto \mathcal{R}$$

$$transform(i, l, t) = r : r \text{ is role of } i \text{ at location } l \text{ at time slot } t$$

$$location: \mathcal{I} \times \mathcal{T} \mapsto \mathcal{L}$$

$$location(i, t) = l : i \text{ is at } l \text{ at } t$$

A meeting M is a 4-tupple,

$$M = \langle D, I, L, t \rangle$$

such that,

$$D \subseteq \mathcal{D}$$

$$L \subseteq \mathcal{L}$$

$$I \subseteq \mathcal{I}$$

$$t \in \mathcal{T}$$

#### 1.4.2 Access Control List

We define following 2 relationships.

$$d = \{d \in \mathcal{D} \mid \text{d is a document}\}$$
 
$$i = \{i \in \mathcal{I} \mid i \text{ is an individual }\}$$
 
$$g = \{g \subseteq \mathcal{I} \mid g \text{ is a subset of one or more individuals in } \mathcal{I}\}$$

Above relationships mean that d is an element of set  $\mathcal{D}$ , and i is an element of set  $\mathcal{I}$ . Further, g is a group of one or more individuals (i), where  $i \in \mathcal{I}$ , such that  $g \neq \emptyset$ .

Consider that following finite set is also defined:

•  $\mathcal{G}$ : Set of all possible not-null subsets of  $\mathcal{I}$ 

Based on above all sets, we define following relationship.

$$access(d) = \{g \in \mathcal{G} \mid g \text{ has access to } d\}$$

Above relationship means that g is an element of set  $\mathcal{G}$ , and that access(d) is the set of groups (g) having access permission to document d.

#### 1.4.3 Privacy-preserved meeting

We assume that every meeting has an agenda document associated with it, to define the set of groups(g) required to attend the meeting

A privacy-preserved meeting is a meeting in which individual i representing role r has no access to the meeting, when  $r \in g$  and  $g \notin access(agenda)$ .

#### 1.5 Aims

- Resolving constraints associated with privacy-preserved meeting scheduling.
- Deciding the possibility of a privacy-preserved meeting to ensure the privacy of meeting content discussed, concerning the selected meeting participants.
- Selecting a privacy-preserving meeting mode, for conducting the particular meeting.
- Evaluating the solution in an acceptable, credible approach for ensuring its validity.

### 1.6 Objectives

- To map the identified constraints to a well-known constraint optimization problem.
- To determine whether a privacy-preserved meeting is possible depending on Access Control Lists (ACL) of documents associated with the meeting, based on the assumption that every meeting has at least one document associated with it (Meeting agenda document).
- To select a privacy-preserving meeting mode, based on availability and locations of selected meeting participants.
- To evaluate the solution by defining an invariant set of rules, using a standard proof assistant (Ex: Coq, Isabelle, Lean)

# 2 Literature survey and identified research gap

## 2.1 Introduction to literature survey

Meeting scheduling becomes increasingly complex when incorporating privacy controls to restrict meeting access to authorized participants only. While many scheduling models address efficiency and availability, fewer approaches ensure access is controlled

to prevent unauthorized participants from joining. This type of access-controlled privacy preservation is critical in sensitive or confidential meeting environments, such as corporate or client meetings, where unauthorized access could lead to security breaches. Despite notable advances, the field lacks a comprehensive, privacy-preserved scheduling model that fully addresses these needs. This literature review critically examines existing approaches, compares their methodologies and limitations, and identifies the gap in privacy-preserved meeting scheduling.

### 2.2 Literature survey

#### 2.2.1 Meeting Scheduling Challenges

Meeting scheduling that integrates access control requires decentralized management of permissions, which can be achieved through distributed models like Distributed Valued Constraint Satisfaction Problem (DVCSP) and Distributed Constraint Optimization Problem (DCOP). In these models, agents independently manage access for each participant, minimizing data sharing and reducing the risk of unauthorized access (Tsuruta and Shintani n.d.; Enembreck and André Barthès 2012).

DVCSP and DCOP offer semi-optimal solutions with limited data exchange between agents, making them suitable for environments where meeting access needs to be restricted. However, while DVCSP allows for flexible constraint relaxation, it may fall short in settings where strict participant authorization is necessary. In contrast, the MULBS algorithm within DCOP provides a more secure option by focusing on minimal communication, reducing the risk of unauthorized access through limited information exposure.

DVCSP's main advantage is its flexibility, allowing agents to relax lower-priority constraints to achieve a balance between ideal scheduling and participant privacy. In contrast, DCOP uses independent agents to manage individual schedules with minimal data sharing, achieving semi-optimal solutions with lower computational costs. MULBS, a specific DCOP algorithm, emphasizes minimal communication requirements and privacy preservation, making it highly suitable for privacy-sensitive environments. Another distinct approach comes from GeoInformatica, which frames the scheduling problem within computational geometry, focusing on optimizing both time and spatial constraints through algorithms with O(n log n) complexity (Berger et al. 2009).

While DVCSP and DCOP both excel in preserving privacy through decentralized scheduling, they differ significantly in their computational approaches and the quality of solutions. DVCSP's semi-optimal solutions are sufficient for low-density environments but may be inadequate in high-density scheduling contexts where optimization is critical. DCOP, particularly the MULBS algorithm, achieves greater computational efficiency by minimizing communication, yet it also sacrifices some scheduling accuracy due to its semi-optimal focus. GeoInformatica's computational geometry-based model achieves excellent efficiency for time-space scheduling but does not address privacy directly, making it limited in privacy-sensitive scheduling environments.

#### 2.2.2 Constraint-Based Optimization Techniques

Constraint-based optimization techniques like Constraint Programming (CP), Mixed-Integer Programming (MIP), and Maximum Satisfiability (MaxSAT) have been applied to scheduling, focusing on managing structured constraints around availability, priority,

and location. CP allows for precise encoding of scheduling constraints, yet it lacks scalability in larger, more dynamic settings (Bofill et al. 2022). MIP, with its ability to rapidly identify infeasible solutions, is effective for structured scheduling tasks with rigid constraints but is too inflexible for applications with dynamic privacy needs (Peteghem and Vanhoucke 2010).

Recent advances in MaxSAT, such as the Soft Conflict Pseudo Boolean (SPB) constraints, have introduced adaptive clause weighting strategies to improve local search efficiency in scheduling contexts. The SPB-MaxSAT algorithm, specifically, provides a high-performing solution for complex scheduling problems by adjusting clause weights to satisfy "soft" constraints, which are more flexible and adaptable (Zheng et al. 2024).

CP and MIP serve well in traditional scheduling scenarios, but both fall short in contexts requiring real-time adaptability and privacy prioritization. While CP offers flexibility with strict constraints, its lack of scalability makes it unsuitable for larger, privacy-sensitive scheduling environments. MIP provides robust solutions for structured schedules but lacks the adaptability needed to handle frequent privacy-related adjustments. In comparison, MaxSAT—particularly the SPB-MaxSAT algorithm—introduces soft constraints that enable adaptability, positioning it as a more flexible alternative to CP and MIP. However, despite these improvements, MaxSAT lacks explicit privacy-focused mechanisms, meaning it primarily supports flexibility rather than privacy preservation.

Overall, constraint-based techniques each contribute unique strengths to the scheduling field. CP's precise encoding, MIP's efficiency in identifying infeasible constraints, and MaxSAT's adaptive clause weighting all provide valuable solutions to specific scheduling problems. Yet, the lack of real-time adaptability and privacy-preserving mechanisms across these methods underscores their limitations in handling complex, privacy-focused meeting scheduling scenarios. MaxSAT's introduction of soft constraints offers the most promise for dynamic scheduling, but further integration with privacy-preservation techniques is essential to address privacy-sensitive needs fully.

#### 2.2.3 Genetic Algorithms and Adaptive Scheduling

Genetic algorithms (GAs) introduce adaptability through iterative optimization, allowing for schedules to evolve based on changing constraints or resource availability. Applied extensively in resource-constrained scheduling, GAs use a bi-population model that supports continuous adjustments, which is advantageous in settings where participant availability and privacy needs fluctuate frequently (Peteghem and Vanhoucke 2010).

GAs are more adaptable than constraint-based methods like CP and MIP, due to their iterative nature and ability to explore multiple solutions. This adaptability makes GAs suitable for environments with dynamic constraints, where immediate schedule adjustments are required. However, GAs' computational intensity poses a challenge for real-time applications, particularly in privacy-sensitive contexts where participant selection may require immediate recalibration based on privacy needs. Compared to MaxSAT's adaptive clause weighting, GAs offer greater flexibility but are less efficient computationally, requiring careful tuning of parameters to achieve optimal performance.

The bi-population structure of GAs provides an advantage in dynamic environments, offering an adaptable alternative to more rigid constraint-based models. However, GAs are hindered by computational demands, which can limit their real-time responsiveness in privacy-sensitive scheduling. This indicates a need for hybrid approaches that incorporate GA adaptability with computational efficiency to better serve privacy-preserved meeting

scheduling needs.

#### 2.2.4 Theory of Constraints (TOC) and Bottleneck Management

The Theory of Constraints (TOC) focuses on identifying and alleviating bottlenecks within scheduling systems, improving throughput in resource-constrained environments. TOC has proven valuable in production-oriented scheduling, where specific bottlenecks often restrict scheduling efficiency and resource allocation (Golmohammadi 2015).

While TOC is effective in managing resources and prioritizing critical constraints, it does not incorporate privacy considerations or adaptability for dynamic participant needs. In comparison, models like DVCSP and DCOP offer more delicate scheduling management by focusing on privacy through decentralized data management. While TOC targets throughput and bottleneck alleviation, its limited adaptability and lack of privacy mechanisms restrict its usefulness in privacy-sensitive scheduling environments.

TOC's bottleneck prioritization method is beneficial for environments where resource constraints drive scheduling but falls short in privacy-sensitive contexts. The rigid focus on throughput limits TOC's applicability to meeting scheduling, which often requires flexible participant selection based on confidentiality needs. TOC's strength in optimizing resource allocation could be enhanced with additional privacy and adaptability mechanisms to better address modern scheduling requirements.

#### 2.2.5 Machine Learning and Hybrid Approaches in Constraint Solving

Machine learning (ML) enhances adaptive scheduling by using supervised learning to predict participant preferences and optimize scheduling algorithms based on past data. Hybrid models, such as Cube-and-Conquer, combine ML with traditional SAT solvers, improving scalability and efficiency in complex scheduling tasks (Popescu et al. 2022; Heule et al. 2012).

ML's ability to predict participant preferences makes it a valuable addition to adaptive scheduling, offering enhanced flexibility over CP, MIP, and TOC. In contrast to traditional models, ML enables real-time responsiveness to changing conditions, aligning well with the demands of privacy-preserved meeting scheduling. However, ML's applications in privacy-sensitive contexts remain limited due to the absence of standardized privacy metrics and evaluation tools, which restricts its ability to fully manage confidentiality alongside scheduling.

ML and hybrid models present significant opportunities for improving adaptability in scheduling through predictive capabilities. However, the lack of privacy-specific applications and standard metrics in ML means that its current role in privacy-preserved scheduling is limited. Developing standardized evaluation metrics and incorporating privacy mechanisms would enhance ML's suitability for privacy-sensitive scheduling scenarios.

#### 2.2.6 Privacy-Centric Distributed Models for Scheduling

Privacy-preserving scheduling has advanced through distributed models such as DCOP and DVCSP, which decentralize data management to protect participant confidentiality. The MULBS algorithm exemplifies the DCOP approach by minimizing communication among agents, which helps maintain privacy (Enembreck and André Barthès 2012). GeoInformatica's computational geometry-based approach addresses scheduling efficiency

by using LP-type problems and graph models, but it lacks privacy-preserving mechanisms (Berger et al. 2009).

DCOP and DVCSP models emphasize privacy through minimal data exchange, contrasting with centralized approaches that are more susceptible to data exposure. MULBS's minimal communication requirements make it highly suitable for privacy-sensitive contexts, while GeoInformatica's time-space optimization algorithms provide efficiency but lack explicit privacy management. Compared to other scheduling models, these distributed approaches offer a balanced solution for privacy-sensitive environments but achieve only semi-optimal results due to the trade-off with computational efficiency.

Distributed models effectively protect privacy but sacrifice some optimization to achieve this. DCOP and DVCSP's decentralized approach addresses privacy needs better than centralized methods but falls short of delivering fully optimized solutions. There remains a gap in achieving optimal scheduling without compromising privacy, highlighting the need for models that can balance privacy, efficiency, and scheduling quality.

### 2.3 Research gap and conclusion

Despite significant advancements, privacy-preserving meeting scheduling remains largely unaddressed. Constraint-based methods excel in managing structured constraints but lack real-time adaptability and privacy-preserving mechanisms. Adaptive scheduling techniques, such as GAs and TOC, offer flexibility but fall short of the privacy requirements essential for modern scheduling contexts. Privacy-centric distributed models partially address privacy needs but achieve only semi-optimal outcomes, highlighting a clear research gap.

Future research must focus on developing hybrid scheduling models that integrate constraint-based optimization, machine learning, and distributed frameworks to create secure, responsive, and privacy-focused scheduling solutions. Addressing this gap would advance the field, providing compliant, adaptable scheduling frameworks tailored to modern privacy-sensitive environments.

# 3 Research question

# 3.1 Hypothesis

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

## 3.2 Research question statement

What is the algorithm to validate the above hypothesis?

# 4 Research approach

1. Identify and define constraints associated with privacy-preserved meeting organization.

- 2. Map the identified constraints to a well-known constraint optimization problem.
- 3. Evaluate the finalized solution by defining an invariant set of rules, using a standard proof assistant (Ex: Coq, Isabelle, Lean).

# 5 Experiments and preliminary results

### 5.1 Experiments

- We have started to define constraints associated with privacy-preserved meeting organization. The latest version of currently identified constraints and defined concepts is 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 validation, based on Access Control Lists of documents related to the meeting, including meeting agenda.
  - The meeting time slot selection based on the availability of validated participants.
  - Meeting mode selection 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 resolving privacy-preseved meeting constraints in more optimum manner, we need to map above mentioned 3 steps of workflow to eligible well-known constraint optimization algorithms. Because, it is not feasible to evaluate all possibilities in problem space for a solution developed from scratch, regarding our constraint resolving 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 resolving problem as well.
- Currently, we are analyzing the following MaxSAT implementations simultaneously, to see whether they can satisfy our constraint resolving requirement regarding meeting participant validation, 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 resolving constraints related to **meeting participant validation**. 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 resolving requirement successfully.

# 6 Finalized evaluation plan

If a solution is developed from scratch for our constraint resolving problem, it is difficult to evaluate the finalized solution (proof of concept) experimentally, since there are extremely large number of possibilities for meetings in problem space.

Accordingly, we are planning to formally map our constraint resolving problem to a well-known constraint optimization problem, which is already addressed. We hope to develop an invariant set of rules for ensuring validity of that mapping. In addition, those invariant rules will be proven by using a standard proof assistant (Ex: Coq, Isabelle, Lean), to reinforce the credibility.

# List of References

### References

- Berger, Florian, Klein, Rolf, Nussbaum, Doron, Sack, Jörg-Rüdiger, and Yi, Jiehua (2009). "A meeting scheduling problem respecting time and space". In: *Geoinformatica* 13.4, pp. 453–481. ISSN: 1384-6175. DOI: 10.1007/s10707-008-0053-4. URL: https://doi.org/10.1007/s10707-008-0053-4.
- 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.
- Heule, Marijn J. H., Kullmann, Oliver, Wieringa, Siert, and Biere, Armin (2012). "Cube and Conquer: Guiding CDCL SAT Solvers by Lookaheads". In: *Hardware and Software: Verification and Testing*. Ed. by Kerstin Eder, João Lourenço, and Onn Shehory. Vol. 7261. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 50–65. ISBN: 978-3-642-34187-8. DOI: 10.1007/978-3-642-34188-5\_8.
- 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.
- Popescu, Andrei, Polat Erdeniz, Seda, Felfernig, Alexander, Uta, Mathias, Atas, Müslüm, Le, Viet Man, Pilsl, Klaus, Enzelsberger, Martin, and Tran, Trang (2022). "An overview of machine learning techniques in constraint solving". In: *Journal of Intelligent Information Systems* 58.1, pp. 91–118. ISSN: 0925-9902. DOI: 10.1007/s10844-021-00666-5.
- Tsuruta, Takuo and Shintani, Toramatsu (n.d.). "Scheduling Meetings using Distributed Valued Constraint Satisfaction Algorithm". In: ().
- Zheng, Jiongzhi, Chen, Zhuo, Li, Chu-Min, and He, Kun (2024). Rethinking the Soft Conflict Pseudo Boolean Constraint on MaxSAT Local Search Solvers. arXiv: 2401. 10589 [cs.AI]. URL: https://arxiv.org/abs/2401.10589.