Title	Modelling the conference paper assignment problem
Author(s)	de Givry, S.; Schiex, T.; Schutt, A.; Simonis, Helmut
Publication date	2020-09
Original citation	de Givry, S., Schiex, T., Schutt, A. and Simonis, H. (2020) 'Modelling the conference paper assignment problem', ModRef 2020 - 19th workshop on Constraint Modelling and Reformulation, Louvain-la-Neuve, Belgium, 7 September, pp. 1-17.
Type of publication	Conference item
Rights	**© 2020, the Authors.** https://modref.github.io/ModRef2020.html
Item downloaded from	http://hdl.handle.net/10468/11447

Modelling the Conference Paper Assignment Problem

S. de Givry[3], T. Schiex[3], A. Schutt[2], and H. Simonis[1]

[1] Insight Cente for Data Analytics, School of Computer Science and Information Technology, University College Cork, Ireland helmut.simonis@insight-centre.org
[2] Data61 and CSIRO, Melbourne, Australia
[3] INRA, MIA Toulouse, UR-875, 31320 Castanet-Tolosan, France

Abstract. In this paper[4] we describe different constraints and models for the conference paper assignment problem. While the core problem is a simple flow problem, additional constraints often arise to tailor a solution to specific wishes, or to increase perceived fairness for reviewers and/or submissions. We show some results from actual conferences paper assignments, and also investigate scalability of the method for large-scale events.

Keywords: Conference Paper Assignment,reviewer assignment,Modelling

1 Introduction

Assigning reviewers to submitted papers can be seen as the main task for a program chair of a scientific conference, with good assignments leading to insightful reviews and a high quality of the resulting paper selection. While in the past, this assignment was largely done by hand, more recently tools to automate this assignment have become popular. There are broadly speaking two types of systems, one approach uses one-sided bids by reviewers for papers, and tries to maximize the overall sum of achieved preferences, the other tries to find a matching value between reviewers and submissions by analyzing the submissions and the published work of the reviewers. The latter is very useful for large conferences, where expressing preferences for thousands of submissions would create a very high workload for the potential reviewers. In this paper we assume that the preference values between reviewers and submissions are given as input, focus on constraints beyond the matching of reviewers to papers, and describe a system that has been used for the last two CP conferences. We first describe the constraints handled by the system informally, in order to encourage discussion about which problem we should be solving. For the selected constraints, we then show a relatively simple MIP model, and its integration with the easychair conference management system. This allows a program chair to build and control the paper assignment interactively, adding or disabling constraints as required.

[4] This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 12/RC/2289-P2 which is co-funded under the European Regional Development Fund.

1.1 Related Work

The Conference Paper Assignment Problem (CPAP) has been studied by many scientific communities including AI [3,6,4]. Its simplest formulation, having a set of submissions to be assigned to a set of reviewers with a fixed number of reviews per paper and a limit on the number of reviews per reviewer, and, given a list of preferences or bids, maximizing the total value of the bids, corresponds to a minimum cost maximum network flow problem that can be solved in polynomial time [3]. More complex formulations with respect to the objective criterion, by introducing fairness [1,4] and/or topic or geographical coverage aspects [5,2], yield to NP-hard problems.

In this paper, we do not consider fairness criteria, but rely on hard constraints for that. We consider several additional constraints to meet the program chair preferences.

1.2 Structure of Paper

In Section 2 we give an informal description of the problem addressed, detailing the rationale for different constraints beyond the straight-forward paper assignment. In Section 3 we present a MIP model that covers the constraints described informally before. This is followed in Section 4 by a discussion of the integration with the easychair conference management system. Section 5 presents the two real-world datasets from the CP 2019 and CP 2020 conferences, and shows an analysis of the generated paper assignments.

2 Informal Problem Description

The Conference Paper Assignment Problem (CPAP) arises for most academic conferences, where submitted papers (submissions) must be reviewed to decide which papers should be accepted at the conference. This is different from reviewing decisions for journal publication, as hundreds (sometimes thousands) of submissions must be reviewed at the same time. This requires a correspondingly large set of reviewers, especially if the number of assigned papers per reviewers should be kept low. The assignment problem arises as the program chair must decide which reviewers deal with which papers, so that papers are competently reviewed, and reviewers are assigned papers that match their interest. While in a journal these decisions can be done by hand for each individual submission as they arrive one by one, the large submission set for conferences requires some form of automation. A link between reviewers and submissions is generated either by a bidding process, or by matching areas of interest. The bidding is one sided, reviewers bid for papers, based on abstracts or the full submission, but submissions (authors) do not bid for reviewers. In addition, some form of conflict recognition is needed to ensure that conflicts of interests are avoided in the assignment. In this paper we do not consider how the bid preferences and conflicts are derived, we assume that they are given as input data.

There are three basic variants of the review assignment problem:

fixed number of reviews per paper In this model, each paper must be reviewed by the the same number of reviewers, typically three. The number of papers for each reviewer may vary.

fixed number of reviews per reviewer To set the same workload for each reviewer, an identical number of papers is assigned to each reviewer. This means that some papers have more, and other fewer reviews.

Senior PC Committee In this model, the papers are partitioned between the members of the senior program committee, only one senior PC member is assigned to each paper. In some conferences, the assignment is fixed, given by the track structure, in others papers are assigned based on bids.

We denote as $|S|$ the number of submissions, $|P|$ the number of reviewers, r the number of reviews required per paper, and a the number of papers per reviewer.

2.1 Fixed Number of Reviews per Paper

This is probably the most common scenario, every paper is assigned three reviews. The number of available reviewers determines how many papers each reviewer should handle. As the number of reviewers in each track varies, different lower and upper bounds on the number of assigned papers can be specified for each track. In our model we generalize this to allow individual lower and upper bounds for each reviewer. The most balanced case assumes that the number of assigned papers per reviewer only varies by one, and the lower and upper bounds are given by

$$\underline{a} \leq \frac{|S|r}{|P|} \leq \overline{a}$$

Note that we may be able to satisfy more preferences if we vary the number of assigned papers by more than one. Whether this improvement in preference value justifies the more imbalanced workload is not obvious.

2.2 Fixed Number of Assigned Papers

If the aim is to assign the same workload to each reviewer, then we can choose the number of papers per reviewer in such a way that at least r reviews are produced for each paper, but some papers may be assigned more. It seems unlikely that from the start one would plan for more than $r + 1$ reviews, although more controversial papers may require additional reviews at a later stage to reach a decision. In this scenario, the number of reviews per reviewer is given by

$$a = \lceil \frac{|S|r}{|P|} \rceil$$

resulting in $|P|a$ overall reviews.

2.3 Reviewer and Chair Bids

In easychair, reviewers can express preferences as "yes" or "maybe" and a default "neutral" value, we translate those choices into numeric value via a configuration.

We allow the PC chair to express preferences on top of the reviewer defined bid structure. This is useful in three scenarios:

- If not enough reviewers have bid on some paper, the solver must find other reviewers for these papers. Typically, the solver will choose reviewers that have few bids or only bids that are shared with many other reviewers. The chair can express some preferences by manually matching reviewers to papers, even if they have not bid for the paper.
- The chair may consider some reviewers more qualified than others for certain papers. By expressing an extra preference value the model can be asked to take these considerations into account. This is very useful for popular papers with many bids, where the program chair may want to guide the assignment towards the most qualified reviewers. Note that this acts as a form of soft constraint which only affects the objective.
- Conversely, the chair may reduce the preference value for some assignment if he considers the assignment as a weaker alternative, this again is a soft version of a hard pre-assignment.

2.4 Multi-Track and Shared Allocation

In a large conference, submitting and reviewing all papers in one pool quickly becomes unmanageable. Instead, thematic tracks are used to partition the papers at submission time. Each track has its own program committee, knowledgable in the topic, and all papers in the track are reviewed only by reviewers from that track. This partitions the paper assignment problem into more easily solved disjoint sub-problems. Only at the final decision time there is some interaction between tracks, to make sure that acceptance across all tracks is done fairly.

While this model works reasonably well for large conferences, there is an issue with a smaller conference like CP. Many reviewers are expert in more than one area, and should be assigned papers from multiple tracks. By treating the tracks independently, the workload for such reviewers may be excessive, as each track chair assigns the maximum number of papers, multiplying the workload for multi-track reviewers. Previously, these conflicts were resolved by hand, negotiating assignment limits between track chairs. In our model, we can set global lower and upper bounds per reviewer, and the model decides where the reviewers are needed most.

2.5 Pre-assignment

For some papers, the program chair may decide on some reviewers based on their unique qualification in the area, disregarding bids and forcing the assignment. In the same way, the chair may decide that some papers must not be assigned to certain reviewers, either because of some undeclared perceived conflict, or because the chair does not trust the expertise of the reviewer for this specific field, even if the reviewer has bid on the paper.

2.6 Shared Reviews

A very common problem is that some submissions are quite similar, either since they are written by the same or nearly same set of authors, or since they address the same problem. In this case, it is important that at least some reviewers see both these papers, and can discuss the merits of the papers in the discussion phase.

2.7 Organization Related Constraints

It is commonly accepted that reviewers should not review papers from authors in the same organization, and reviewers should declare conflicts with such authors. At the same time, it is also important that not all reviewers of a paper are from the same organization, even if they are not linked to any of the authors. This avoids an institutional bias, where colleagues form the same opinions, even if they do not explicitly discuss the paper during the reviewing. One may extend this to co-authors, stating that long-time collaborators should not be assigned to the same paper. This constraint may be too strong, excluding very qualified reviewers from some paper, and can therefore not be enforced between all reviewers that are in conflict with each other.

2.8 Geographical Region Based Constraints

A slightly weaker version of the constraint states that not all reviewers should come from the same geographical region, and that reviewers from countries different from the author's country should be considered. This improves the diversity of the review assignment, and may help to overcome national biases in the reviewing.

2.9 Preferential Treatment

The program chair may need some mechanism to guide the assignment process, to implement specific wishes for some reviewers. A reviewer may state at invitation time that he can only review a limited number of papers because of other workload, and the chair may wish to enforce this limit in order to satisfy this hard reviewer constraint. The chair may also decide that some PC members (for example relatively junior members) should only be asigned papers they have bid for, while more experienced PC members may be asked to review papers outside their area of expertise, if no other reviewers can be found. The chair may want to decide these preferences on a case by case basis.

2.10 Regret

Ideally, all reviewers would be assigned papers that they strongly want to review. In most cases, this will not be possible, and a regret constraint may be required. The actual assignment of papers for each reviewer should not be much worse than the best possible assignment.

This is not handled at the moment. It is straightforward to express the constraint if the number of assigned papers is fixed and known, but requires additional variables, if the number of papers for the reviewer is not known a priori.

2.11 Easy Papers

The program chair may decide that certain papers are much easier to review than others, and therefore do not contribute as much to the workload of the reviewer. It would be unfair if a reviewer with few reviews is assigned to such a paper, these should go to reviewers with a high workload, instead.

2.12 Fairness

If we only try to optimize the overall sum of preferences, then some reviewers may be assigned all their preferred bids, while others only get papers they have not bid for. This is unfair to both the reviewers and the submissions, as papers where none of the reviewers wanted to review the paper operate under a heavy disadvantage. One way of improving fairness is to state that all reviewers should be assigned at most k papers they have not bid for, and any submission should not have more than l reviewers that have not bid for the paper. Enforcing these constraints with strong limit values may make the problem infeasible, requiring either adding more diverse reviewers, or allowing a relaxation of the constraint.

2.13 Senior Program Committee Assignment

In the senior PC assignment, only one committee member needs to be assigned to each paper. This typically is done after the initial paper assignment is done. If we want to enforce some constraint that reviewers and senior PC members of the same submission should not come from the same organization, it may be required to link the two problems together to find a feasible or optimal solution.

3 MIP Model

In this section we describe a mixed integer linear programming model for the problem. The main decisions taken are whether a specific reviewer is assigned to a specific paper, expressed with 0/1 integer variables. But first we need to formalize some of the concepts and data required by the application.

3.1 Constants and Indices

We use the following sets of entities:

P set of persons that are assigned, indexed with p
S set of submissions to be assigned, indexed by s
T set of tracks considered at the same time, indexed by t
R set of shared review constraints (see Section 2.6), indexed by r
G set of countries for all authors and reviewers, indexed by g
G_s set of countries of the authors of submission s

The following constants are used in the model. To increase flexibility, we allow separate limits for each person or submission, whereever possible.

b_{ps} bid, non-negative preference of person p for submission s, larger values indicate higher preference
r_s non-negative integer, number of reviews required for submission s
l_p non-negative integer, lower bound on number of reviews assigned to person p
u_p non-negative integer, upper bound on number of reviews assigned to person p
c_{ps} boolean, true value indicates that there is a conflict between person p and submission s, and the paper cannot be assigned
f_{ps} boolean, value true indicates that submission s must be assigned to person p
e_{ps} boolean, value true indicates that submission s must not be assigned to person p
w_{pt} boolean, true value indicates that person p can be assigned to papers of track t
t_s track of submission s
g_p the geographical area (country) of reviewer p
s_{r1} first submission of shared review constraint r
s_{r2} second submission of shared review constraint r
q_r number of reviewers sharing review constraint r
easy(s) submission s is considered an easy paper to review
onlybid(p) person p should only be assigned papers which have a positive bid
rev global value for relaxing the bid for reviewers, states how many non-bid papers can be assigned to a reviewer
sev global value for relaxing the bid for submissions, states how many non-bid reviewers can be assigned to a submission
glv each submission must have reviewers from at least this many countries
gsc at most this many reviewers of a submission can come from the same country as an author for the submission

3.2 Variables

The main decision variables x_{ps} indicate if person p is assigned to review submission s.

x_{ps} 0/1 integer variable, denotes if person p is assigned submission s
y_{pr} 0/1 integer variable, indicates if shared review r is performed by person p
z_p non-negative continuous variable, states how many papers are assigned to person p
v_{gs} 0/1 integer, at least one reviewer from country g is assigned to submission s

The 0/1 variables y_{pr} are used to express shared review constraints. They state if reviewer p is assigned both papers mentioned in constraint r. The non-negative z_p variables count how many papers reviewer p is assigned to. They are used to deal with the easy paper constraints.

3.3 Constraints

The number of assigned papers is computed from the individual assignment decision variables.

$$\forall_{p\in P}: \quad z_p = \sum_{s\in S} x_{ps} \tag{1}$$

Each submission gets the correct number of reviews, that number can vary between submissions.

$$\forall_{s\in S}: \quad \sum_{p\in P} x_{ps} = r_s \tag{2}$$

The number of papers assigned to a reviewer lies between the given lower and upper bound.

$$\forall_{p\in P}: \quad l_p \leq z_p \leq u_p \tag{3}$$

If there is a conflict between reviewer p and submission s, then that paper can not be assigned.

$$\forall_{p\in P}\forall_{s\in S}: \quad c_{ps} \Rightarrow x_{ps} = 0 \tag{4}$$

If the reviewer p does not review papers for track t, then all papers in that track cannot be assigned.

$$\forall_{p\in P}\forall_{s\in S}: \quad \neg w_{pt_s} \Rightarrow x_{ps} = 0 \tag{5}$$

If the assignment of reviewer p to submission s is given as input, then the decision variable is set to one.

$$\forall_{p\in P}\forall_{s\in S}: \quad f_{ps} \Rightarrow x_{ps} = 1 \tag{6}$$

If the assignment of reviewer p to submission s is excluded as input, then the decision variable is set to zero.

$$\forall_{p\in P}\forall_{s\in S}: \quad e_{ps} \Rightarrow x_{ps} = 0 \tag{7}$$

Easy papers should only be given to reviewers who are assigned the largest possible number of papers.

$$\forall_{p\in P}\forall_{s\in S|\text{easy(s)}}: \quad z_p \geq l_p + (u_p - l_p)x_{ps} \tag{8}$$

Reviewers who should only be assigned to papers in their bid are excluded from being assigned to other papers.

$$\forall_{s\in S}\forall_{p\in P|\text{onlybid(p)}}: \quad b_{ps}=0 \Rightarrow x_{ps}=0 \tag{9}$$

The number of non-bid papers assigned to any reviewer is limited by the constant *rev*.

$$\forall_{p\in P}: \quad \sum_{s\in S|b_{ps}=0} x_{ps} \leq \text{rev} \tag{10}$$

For any submission, the number of reviewers that have not bid on the paper, but are assigned, is limited by the constant *sev*.

$$\forall_{s\in S}: \quad \sum_{p\in P|b_{ps}=0} x_{ps} \leq \text{sev} \tag{11}$$

The following constraint (12) deals with shared review requirements, expressed as constraints r in the set R. For a specific shared constraint r between two papers s_{r1} and s_{r2}, the number of reviewers assigned both papers must be greater or equal to q_r.

$$\forall_{r\in R}: \quad q_r \leq \sum_{p\in P} x_{ps_{r1}} x_{ps_{r2}} \tag{12}$$

As this is non-linear, we have to linearize the constraint by introducing new 0/1 variables y_{pr} and the constraints (13) to (16).

$$\forall_{r\in R}: \quad q_r \leq \sum_{p\in P} y_{pr} \tag{13}$$

$$\forall_{r\in R}: \quad y_{pr} \leq x_{ps_{r1}} \tag{14}$$

$$\forall_{r\in R}: \quad y_{pr} \leq x_{ps_{r2}} \tag{15}$$

$$\forall_{r\in R}: \quad x_{ps_{r2}} + x_{ps_{r2}} \leq 1 + y_{pr} \tag{16}$$

The next constraints deal with the geographical regions. The first one states that if we assign reviewer p to paper s, then we have a reviewer from country g_P

$$\forall_{p\in P}\forall_{s\in S}: \quad x_{ps} \leq v_{g_p s} \tag{17}$$

The next constraint states that each submission must have reviewers from at least *glv* countries.

$$\forall_{s\in S}: \quad \sum_{g\in G} v_{gs} \geq \text{glv} \tag{18}$$

Finally, we state that for each paper, there can be at most *gsc* reviewers that come from the countries where the authors of the paper reside.

$$\forall_{s\in S}: \quad \sum_{p\in P|g_p\in G_s} x_{ps} \leq \text{gsc} \tag{19}$$

3.4 Objective

We want to maximize the total value of preferences

$$\max \sum_{p \in P} \sum_{s \in S} b_{ps} x_{ps} \tag{20}$$

under the existing constraints. The objective maximizes the preferences over all review assignments, but some of the side constraints can be seen as optional, and the "best" solution is the one which finds the best compromise between total preference costs, and side constraints that are satisfied. We do not try to find this compromise automatically, but provide an interactive environment where the user can experiment with different variants of the problem. In the end, the user decides which solution best satisfies their wishes.

4 Integration with easychair

The CP conference series has been using the conference management system (CMS) easychair (`https://easychair.org`) for many years. easychair allows for manual paper assignment, and also has an automated assignment tool, which based on the bid preferences, assigns reviewers to papers. Unfortunately, easychair does not describe how the assignment is produced, and in the past, clearly sub-optimal solutions were returned. This motivated the program chairs for the CP conference to manage the assignment process by an automated tool outside the CMS. In 2019, and then again in 2020, easychair's data export functionality was used to automate the paper assignment problem for the CP conference. There are two main APIs for data export. In the professional version, a complete dataset for the conference, including submission data, program committee data, bids and conflicts can be exported in csv or Spreadsheet form. In the Assignment subsystem, more limited data about submissions, and bids can be exported as well, in addition a generated assignment can be uploaded into the easychair platform. Tables 1 and 2 describe the files provided in the two interfaces, while Table 3 describes the upload format[5] . All information about the API is current as of June 2020.

One of the challenges of the data export is that names and email addresses are not normalized across submissions and tracks, the same person can appear with different email addresses if the submissions or reviewer invitations are using different formats. In the data download there is a personId number that addresses this problem, but which is not persistent, as it may change over multiple exports of the data. This means that any additional data we define cannot use this id, and must rely on the name/email pair to identify persons. Care must be taken to generate the match between persons and names correctly.

[5] The reviewerId required for the upload is not present in the data export files, therefore the `reviewer.csv` files for all tracks must be used in the assignment system

Table 1. Relevant Data Export Tables of easychair

Name	Description
author.csv	submission author information
bidding.csv	bid information for all tracks
committee.csv	program committee members for all tracks, including chairs and super chairs
conflict.csv	user generated and auto-generated conflicts of reviewers and submissions
submission.csv	lists all submissions to the conference, including all tracks
track.csv	lists all tracks of the conference

Table 2. Assignment Export Tables of easychair, one set of files per track

Name	Description
reviewer.csv	Description of program committee, reviewerId is the first column
conflict.csv	list of conflicts by authors and reviewers, also includes auto-detected conflicts
bid.csv	current bids for papers by reviewers
assignment.csv	current assignment in easychair system

Table 3. Data Format for Assignment Upload (no headers), one file per track, separate file for SPC

Column	Type	Description
reviewerId	int	id number of reviewer from reviewer file
submission	int	submission id number

Another challenge is the form of the organization field, which does not uniquely identify the organization of an author or reviewer, as different people may use different forms of identification. As an example, people from the same research group in Cork have identified their organization as

- 4C
- Insight Centre for Data Analytics
- Computer Science Department, UCC
- University College Cork
- Insight Centre for Data Analytics, Schools of Computer Science and Information Technology, University College Cork

This variation makes it impossible to automatically use the organization field to identify conflicts, and assignment preferences. On the other hand, a large manual effort would be required to properly identify all organization equivalence classes by hand.

Fig. 1. System Architecture

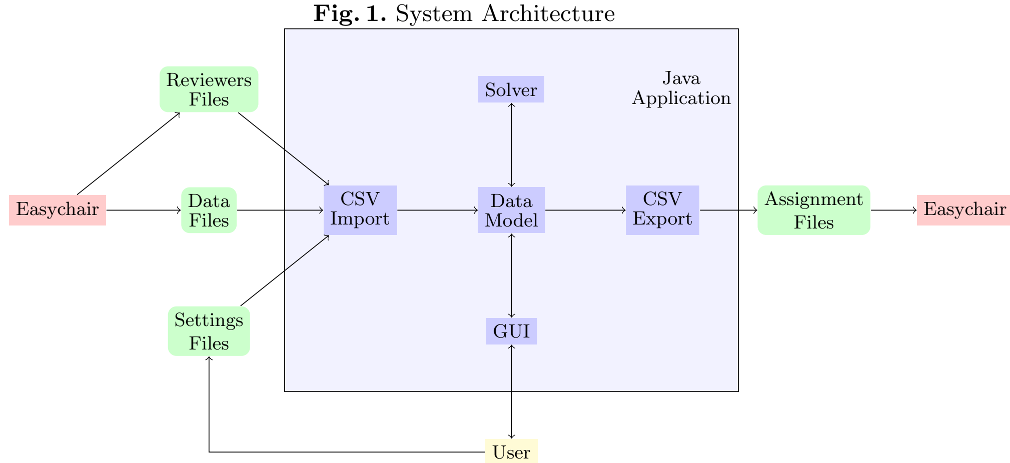


Figure 1 shows the overall architecture of the application. The input data is loaded into an interactive Java application, and held in an internal data model. The user interacts with the system through a user interface, where he/she can change constraints, and explore different settings of the solver. The solver is run from within the application, possibly many times with different settings. Parameters, settings and KPIs obtained are stored in the system for each solver run. When the user is satisfied with the solution, the assignment is exported as csv files (one for each track), which can then be uploaded to easychair.

5 Data and Results

5.1 Characterising Bid Distributions

We use bid data from CP 2019 and CP 2020 to characterize the distributions of bids for reviewers and submissions. This ignores the bids of the senior program committee, which can be easily handled by a simpler solver instance. The 2019 data only cover the technical track (63 papers and 85 reviewers, 539 yes bids, 1256 bids total), while the 2020 data cover all tracks (121 papers, 99 reviewers[6], 470 yes bids, 1049 bids total). We note that even though fewer reviewers bid for papers in 2019, there are more yes and more overall bids in 2019 compared to 2020. We can only speculate why this happened, in both cases reviewers were asked to bid for at least 10 papers.

Figure 2 shows the 2019 data in the first row, the 2020 data in the second row. It displays the distribution of the bids from reviewers, on the left only counting the yes bids, on the right counting all positive (yes + maybe) bids. The y-axis is expressed as a fraction of all reviewers. We compare the observed distributions with the values of a Poisson distribution, which approximates the observed distribution, but not very well. Note that in both years there are a number of reviewers without any yes bids, and in 2019 some reviewers did not enter any positive bids.

Figure 3 shows the distribution of bids per submission, again yes bids on the left, and all positive bids on the right. For both 2019 and 2020, there is one very popular paper each, attracting 22 yes bids (2019) and 16 yes bids (2020). Note that in both 2019 and 2020 there are a number of papers without any yes bid, this may either indicate a gap in program committee diversity, or indicate papers that were not attractive to any (even qualified) reviewers. In 2020, there is one paper without any positive bids at all, this paper will require three non-bid reviewers.

5.2 Constraints Used

The 2019 dataset considers a single track, the main technical track of the conference, and assigns reviewers for that track only. While an upper bound of three papers is enforced for all reviewers, the lower bound is set to zero, as there is a surplus of reviewers. A penalty for not assigning any papers to a reviewer in the objective function tries to balance the overall workload per reviewer. For the actual conference, the interaction of the assignments of multiple tracks was handled by hand, by enforcing reviewer specific upper bounds in discussion between the track chairs. In the 2020 model, all six tracks were assigned by the program at the same time, using possibly overlapping program committees for each track. This was done to balance the workload fairly, and reduce the need for multiple iterations of the assignment. Track specific lower and upper bounds are enforced, varying the number of assigned papers between tracks, but only allowing a difference of one within each track. The additional constraints for program

[6] Reviewers in multiple tracks counted once.

Fig. 2. Reviewer Bid Distribution; Observed distribution as bars, Poisson Distribution with same average as line for comparison; first row 2019, second row 2020

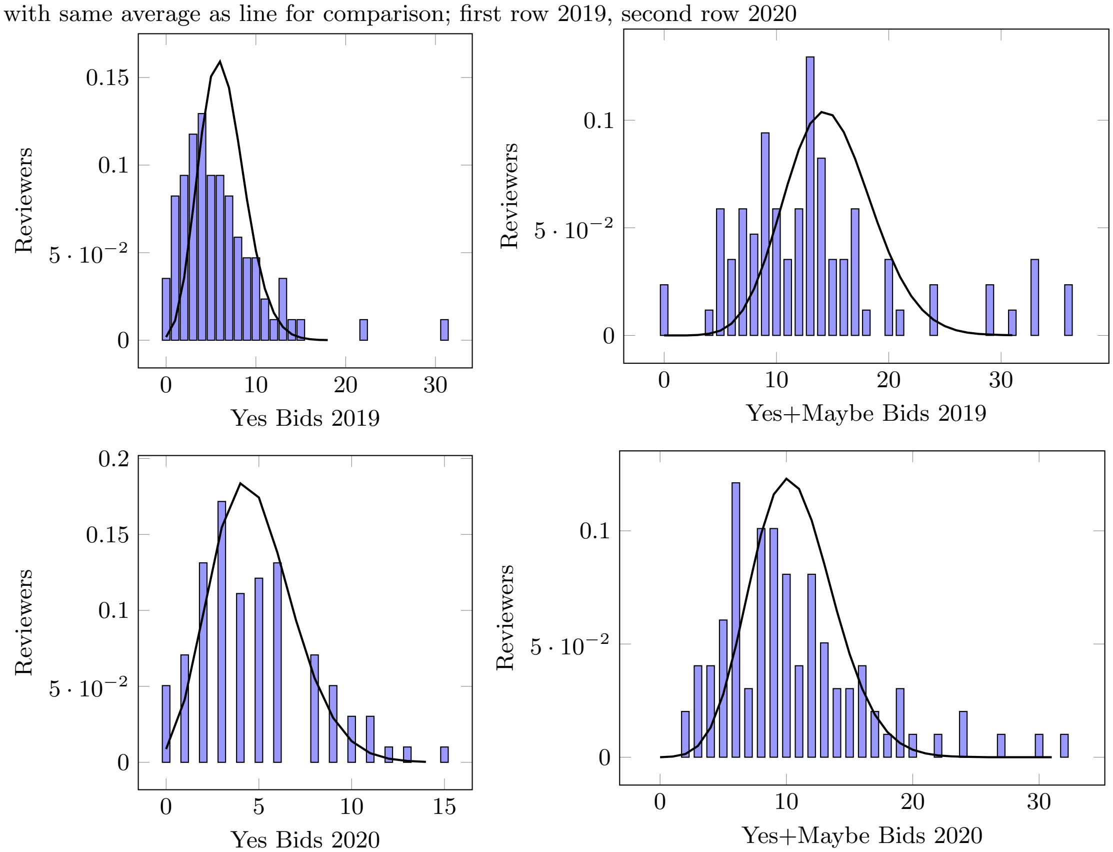
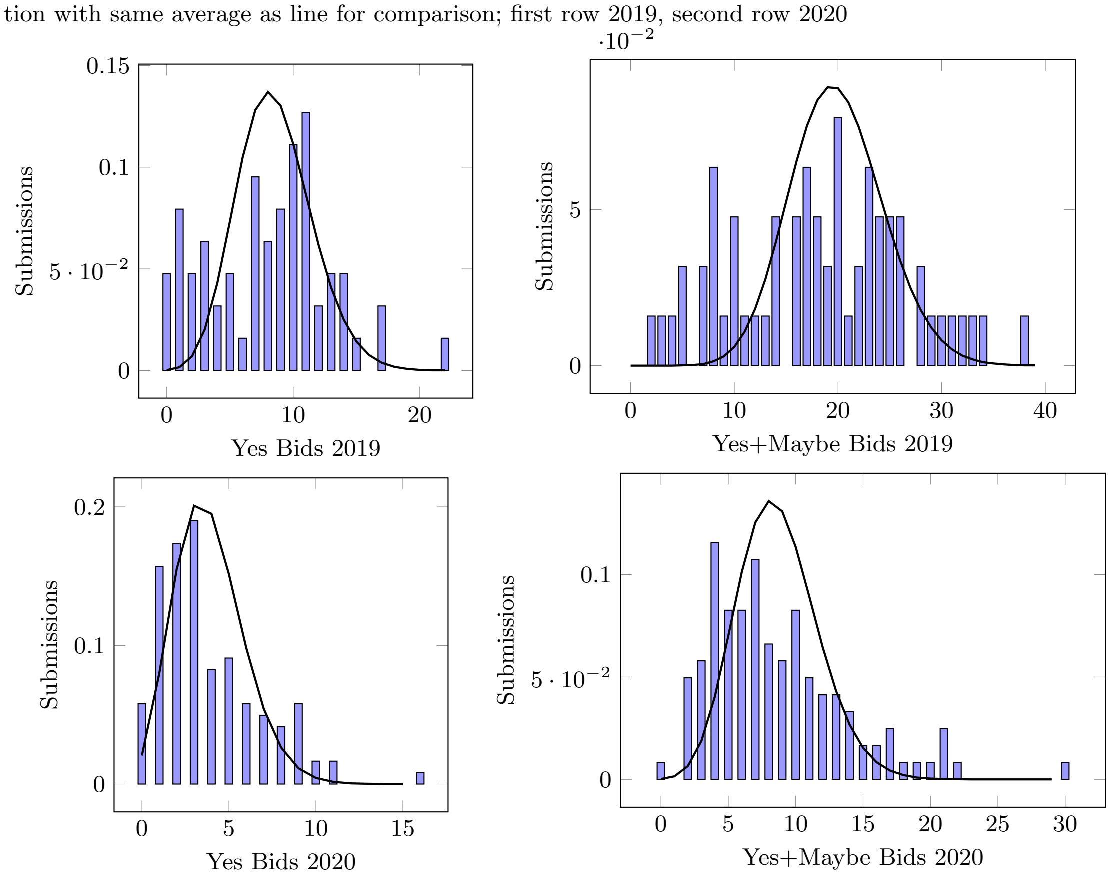


Fig. 3. Submission Bid Distribution; Observed distribution as bars, Poission Distribution with same average as line for comparison; first row 2019, second row 2020



chair preferences, forced assignments and exclusion, easy paper assignment and shared reviewers for related papers were all introduced in the 2020 model.

5.3 Results

Table 4 show the results for the assignment for CP 2019 and CP 2020, while also recalling some of the main metrics for the problems. For 2019, only one assignment (0.53 % of all 189 assignments) was not preferred, while 9 of the available reviewers were not assigned any papers at all. In 2020, the solution used all available reviewers, while 17 (5.68 %) of the assignments were not preferred. Note that the program finds optimal solutions in seconds.

Table 4. Results for years 2019 and 2020; Percentage of assignment are preferred (yes), maybe, or unpreferred, number of unused reviewers

			Bids		Percentage of Assignments			Unassigned
Year	Submissions	Reviewers	Yes	All	Yes	Maybe	Non Preferred	Reviewers
2019	63	85	539	1256	84.13	15.34	0.53	9
2020	121	99	470	1049	72.83	21.49	5.68	0

The two scenarios show clearly different behaviour. In 2019, with a surplus of reviewers, and more bids per reviewer, the assignment can be done nearly without any unpreferred assignment, and some reviewers were not assigned any papers at all. In 2020, with fewer bids for papers, more unpreferred choices needed to be assigned, to find the required number (three) of reviewers for all papers. The reduced review time window for paper bidding and assignment in 2020, due to the COVID-19 pandemic and the postponed submission deadline for CP 2020, precluded adding more reviewers and their bids to the system, with around 10% of program committee invitations being rejected or ignored. This points to an issue in the workflow for paper assignment. Solving a realistic version of the paper assignment problem is only possible when all bids for papers have been entered, only at this time will gaps in the submission coverage be discovered. But recruiting more PC members at this stage would incur further delays in the overall assignment, and reduce the time available for writing reviews.

6 Conclusions and Future Work

In this paper we have presented an extended version of the conference paper assignment problem, describing some constraints extending the standard bid-based paper assignment. We have split the presentation into an informal, natural language part and a second part describing a MIP model for the problem, in order to promote more discussion about which problem we want to solve, rather than the technical details on how to solve a known, well-specified problem. Results from two, quite different, uses of the system for the CP conferences of 2019 and 2020 were presented. One finds a solution nearly without non-preferred assignments, while the other, with fewer positive bids, but more papers, requires more use of non-preferred choices. The model and constraints used should be usable for many small to medium sized conferences, giving program chairs an opportunity to express additional constraints on top of the bids by the program committee.

References

1. Garg, N., Kavitha, T., Kumar, A., Mehlhorn, K., Mestre, J.: Assigning papers to referees. Algorithmica **58**(1), 119–136 (2010)
2. de Givry, S., Lee, J.H., Leung, K.L., Shum, Y.W.: Solving a judge assignment problem using conjunctions of global cost functions. In: O'Sullivan, B. (ed.) Principles and Practice of Constraint Programming - 20th International Conference, CP 2014, Lyon, France, September 8-12, 2014. Proceedings. Lecture Notes in Computer Science, vol. 8656, pp. 797–812. Springer (2014). https://doi.org/10.1007/978-3-319-10428-7_57, https://doi.org/10.1007/978-3-319-10428-7_57
3. Goldsmith, J., Sloan, R.: The AI onference paper assignment problem. In: Proceedings of the 22nd AAAI ConferenceWorkshop on Preference Handling for Artificial Intelligence (MPREF) (2007)
4. Lian, J.W., Mattei, N., Noble, R., Walsh, T.: The conference paper assignment problem: Using order weighted averages to assign indivisible goods. In: McIlraith, S.A., Weinberger, K.Q. (eds.) Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018. pp. 1138–1145. AAAI Press (2018), https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17396
5. Long, C., Wong, R.C.W., Peng, Y., Ye, L.: On good and fair paper-reviewer assignment. In: 2013 IEEE 13th International Conference on Data Mining. pp. 1145–1150 (2013)
6. Price, S., Flach, P.A.: Computational support for academic peer review: a perspective from artificial intelligence. Commun. ACM **60**(3), 70–79 (2017). https://doi.org/10.1145/2979672, https://doi.org/10.1145/2979672