On Using Real-Time and Post-Contest Data to Improve the Contest Organization, Technical/Scientific Procedures and Build an Efficient Contestant Preparation Strategy

Jamaladdin HASANOV, Habil GADIRLI, Aydin BAGIYEV
ADA University, School of IT and Engineering
Ahmadbey Aghaoglu str. 61, 1008 Baku, Azerbaijan
Email: jhasanov@ada.edu.az; hgadirli2019@ada.edu.az; abaghiyev@ada.edu.az

Abstract. Nowadays, the coaches of various sports disciplines use analytical tools to process the game data, analyze the behavior of their team, individual players, and build efficient strategies based on the strength and weaknesses of the competitors. Considering its global scope, time limitation, complexity of tasks, ranking, and medals, IOI can also be considered as an intellectual sports contest. With the direct relation to Information Technology and sports, the statistics of the IOI contests need to be analyzed to deliver better results, initate changes in the preparation strategy and increase the quality of the event as well. This paper provides a statistical analysis based on the last on-site IOI and shares insights on each of them.

Keywords: CMS, data analysis, contestant, organization.

1. Introduction

International Olympiad in Informatics is a global algorithmic programming contest that involves school-aged contestants worldwide, organized as teams, representing their countries. According to the statistics of the last 5 years (IOI Statistics, 2020a), more than 300 contestants from at least 80 countries have joined the contest each year. The final ranking of the contestants after a 2-day competition is highlighted with the corresponding medal's color code on the results page of every contest year (IOI Statistics, 2020b). The number of countries, contestants, team members, guests, and standings are the basic statistics shared with the public by the International Committee of the IOI.

More detailed statistics prepared by the Technical and Scientific committees are shared with the internal community:

- The International Technical Committee shares a summary on the technical issues, common problems and provides statistics on the language usage, task completion, and other contest-related information that is fetched from the contest database. These inputs help improve the infrastructure-related planning (contest environment, contestant laptop parameters, and so on), optimize procedures (of printing, translation, contests start/stop, extension rule, backup rules, etc.), and help Scientific Committee make adjustments to the rules (on adding/removing programming languages, modifying limitations, tasks types, changes in regulations and so on) if necessary.
- The International Scientific Committee runs a survey to get contestant and team leader responses. This survey usually covers the tasks (statements, difficulty of tasks, and subtasks), contest environment, online judge systems, programming language preference, and organizational questions. The survey result may also affect the procedures and regulations – during its presentation in the joint meeting of the committees, some common concerns and points may raise an internal discussion.

As stated previously, these inputs help better understand the strong/weak points and continuously improve the quality of the contest. According to the reports of the last years, more common and basic organizational problems are not encountered anymore, the contest environment, judging system, and procedures have been improved to eliminate the usage and capacity-related problems.

Despite its immense value on organizational planning, the mentioned reports do not cover the information that describes the contestant behavior, which might be interesting to the contestants and their team leaders for retrospective analysis and also provide Technical and Scientific teams with valuable information right on time. In recent years, there has been valuable research and applications of analytical methods and Machine Learning techniques to predict the contestant performance (Alnahhas and Mourtada, 2020) and analysis of the task difficulty (Fantozzi and Laura, 2020; Vegt and Schrijvers, 2019; Pankov and Kenzhaliyev, 2020). Some specialists have provided valuable pedagogical analysis of the key factors that may lead to contestant's success (Tsvetkova and Kiryukhin, 2020; Lodi, 2020). There have also been ideas on platform-based automatic analysis of the contestant activities in competitions and e-learning systems (Kostadinov *et al.*, 2018).

This paper introduces a report, based on the IOI 2019 data, which aims at understanding of the contestant behavior during the contest days. This report is the first phase of the "IOI data analysis" project supported by the grant, which provides an analysis of the Contest Management System (CMS, n. d.) database of the corresponding IOI archive.

The following chapter shares various statistical reports, provides authors' inference based on the data, and some recommendations derived from them.

2. Insights

The Contest Management System (CMS) is an open-source distributed contest manage-
ment and grading system that was developed for the IOI 2012, hosted in Italy. Starting
from 2017, the CMS system has been used consecutively in the 5 last IOI contests,
which made it a potential nominee for the de-facto standard for the IOI. CMS stores all
the task and contest data in a compact Postgres DB structure. CMS has decent documen-
tation on the structure and installation of the system but the structure of the tables and
their relationship are not fully described. As a part of this report, the structure of the DB
and the purpose of the tables have been documented and shared in a public repository
(GitHub, 2021a).

   This chapter introduces the insights on IOI 2019 data, that start as general statistics
and then specific targeted analysis, accompanied with the findings.

2.1. *Submission Dynamics*

On each day of the contest, the contestants receive 3 tasks and 5 hours of time to
submit their solutions. The order in which they solve the tasks and frequency of their
submissions might be different (which is analyzed in 2.3). The following report shows
the number of submissions in a 30-minute interval (Fig. 1). Each diagram shows the
number of total (blue line) and successful (dashed orange line) submissions. Any sub-
mission scored more than 0 is considered as successful. So the gap between the lines
shows the submissions that did not score anything.

   As seen from Day 1 results, "Shoes" was a comparably easy task based on a greedy
algorithm and therefore gets a lot of submissions from the start. "Rect" has some in-
creasing dynamics but "Split" seems challenging till the end.

   Before the contest, the Scientific Committee (SC) of the IOI evaluates the complex-
ity of the tasks and makes assumptions on solvability by the contestants (like 65% of
the contestants may partially or completely solve it). Day 2 tasks have the same flavor –
"Line" is easy, almost all the attempts have been graded (only less than 10% of the total
contestants couldn't score anything), but the other two tasks were quite challenging for
the contestants.

   The mentioned insight might be useful for the real-time and post-contest analysis by
the SC, and add value to the task selection process.

2.2. *Contestant Requests*

The IOI's scope is not limited only to the technical procedures – the administrative,
logistic, and support services are also primary responsibilities of the host committees.
During their 5-hour work, contestants use their contest environment to ask various ques-
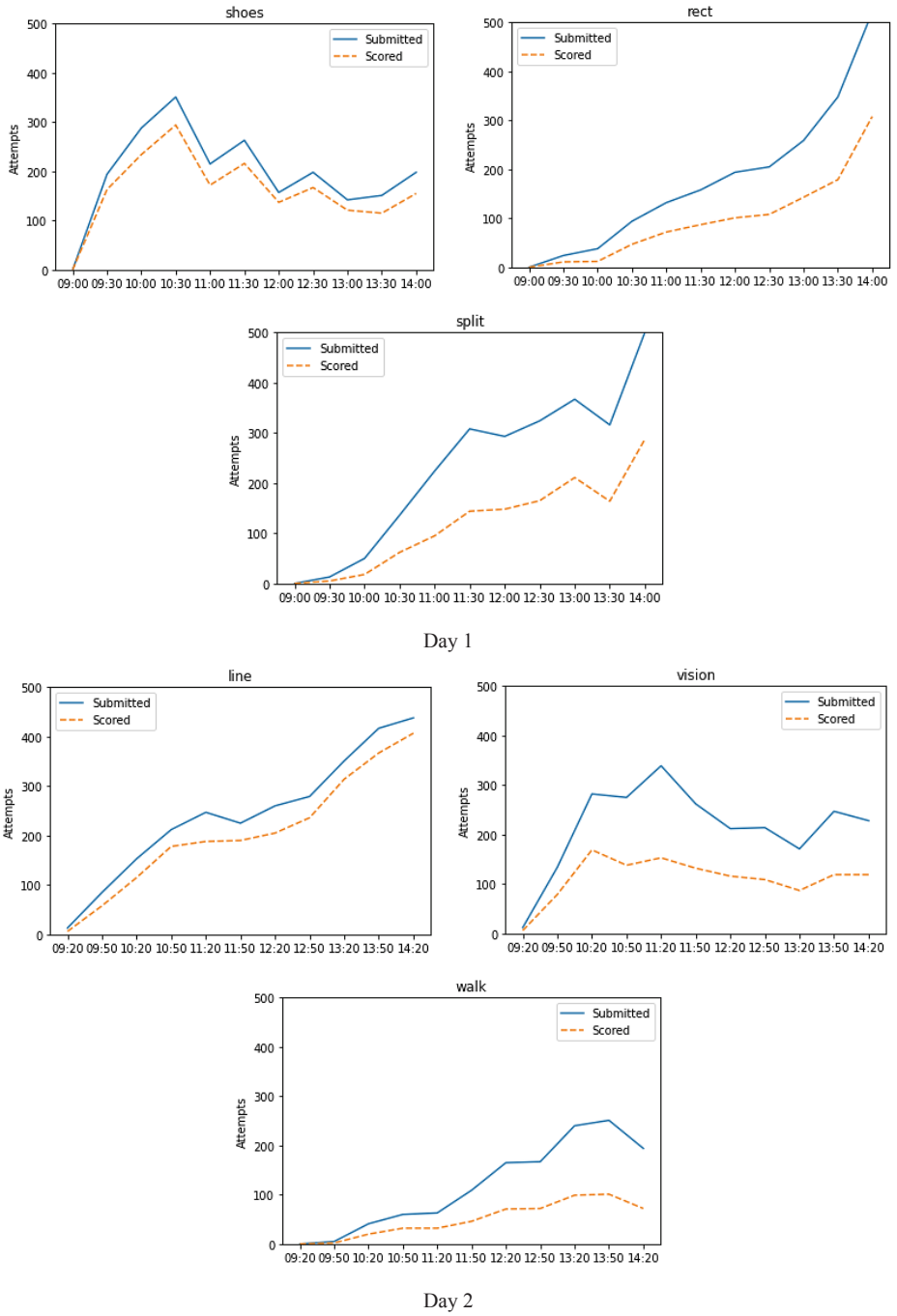
Day 1

Day 2

Fig. 1. The number of total and successful submissions grouped by tasks.

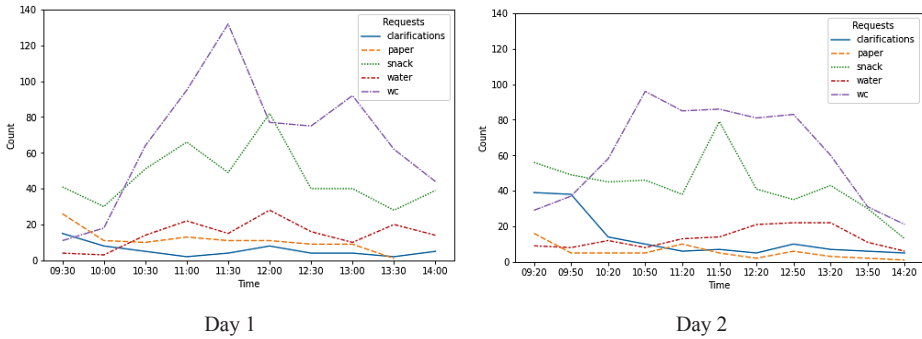Day 1   Day 2

Fig. 2. Contestant requests and questions.

tions, starting from the clarifications to the request for food, beverage, and other sup-
plies. During the IOI, this statistic is usually shared in a summarized way (what type of
requests in total we had). Fig. 2. shows the number of the requests made by contestants
every 30-minute, grouped by the request type.

In the IOI2019 version of the ContestWebServer, each question had a correspond-
ing icon that helped contestants register their needs/questions/inquiries just by clicking
on items and adding a small note. As a result, it was quite easy to group the results by
categories. The "snack" category includes fruits (banana and apple), chocolate, and
cupcake. The increase of the "snack" and "WC" requests at the beginning of day 2 is
connected to the late start of the contest (in contrast with Day 1, the Day 2 statistics in-
cludes the contest's start time, which was 9:20. Due to delay on the 2nd day, the requests
have started before the contest time. Therefore 9:20 shows the accumulation of the
requests till that time.). As seen from the graphs of both days, the clarification requests
are naturally high at the beginning of the contest – all technical and rule-related ques-
tions are clarified in the first 30–40 minutes of the contest. Also, there is an increase
at the start point of the "clarifications", which was due to technical problems. As seen
from the graphs, for both days, after 45 minutes from the start, the "clarifications" count
starts to decline.

As seen from the graph, putting water bottles on the desks before the contest and add-
ing a few empty A4 papers in the task envelopes reduced the request on these tasks. This
improvement had been noted from the previous IOI observations – a systematic review
of the request fulfillment statistics might increase the efficiency of the process.

During the contest time, HTC team members are the only staff that has a chance
to closely observe the situation, see the patterns and share their thoughts and forecasts
immediately with their coordinators or HTC chair. At the 2nd hour of the contest day
1 of IOI 2019, HTC team members shared with the HTC coordinators that there is an
increasing queue to the toilets (as might be seen from Fig. 2. Day 1). In the beginning, it
was thought that it might be normal (usually this peak is common between the 2nd and
3rd hours of the contest) but considering the number of allocated WC cabins in National
Gymnastic Hall (which hosted events with 10,000 visitors), it was decided to double-
check the case. According to the observation of the HTC members, the reason could

Table 1

Sequence analysis between the snack and WC requests

Original requests	The number of WC requests	
	Day 1	Day 2
Banana	70	-
Apple	-	34
Chocolate bar	58	50
Cupcake	22	16
Water	98	77

be the order of bananas and drinking water over it. As a result, it was decided to stop delivering bananas (replaced them with chocolate bars) and to replace the banana with the apple on day 2. As seen from Fig. 2. Day 2, the "WC requests" do not cross the 100 requests and distribution seems quite normal.

Now in the retrospective analysis, we can see the proof of the behavior. Table 1 shows the number of WC requests made within the 45 minutes after each snack and water-related request. As seen from the Day 1 numbers, the order of 'Banana' and 'Water' seem to have a correlation with the 'WC' requests. There is a high number for 'Chocolate bar' as well, but when analyzed, it was realized that mainly they are either ordered or consumed with water.

If such a report had been used in a real contest time, the mentioned incident could have been detected as a pattern way before its peak time. Similar to technical alarm and monitoring systems, HTC could easily use pre-defined thresholds and rules to see the irregular or abnormal patterns and involve HTC members in the detailed investigation.

2.3. *Switching Between the Tasks*

Knowledge for sure plays an important role in a contestant's success. But problem-solving and evaluation skills definitely cannot be ignored as well. A few scenarios with the various outcomes might be possible:

- A contestant may start with the most difficult task, spend the majority of time on it and exhaust himself/herself without leaving no chance to solve the others.
- A contestant may decide to spend a fixed amount of time (small, just enough to analyze the problem) on each of the problems in the beginning, evaluate the complexity, select the easiest one and solve one at a time.
- A contestant may decide to work on the hardest task first, such that it requires more concentration, and then spend the remaining part of the contest to the simple tasks.
- Same as in the previous option, but a contestant may decide to use an iterative approach to spend a fixed amount of time on each task and this way focus on the incremental total score, rather than the finalization of the particular task.

The selection of the right strategy for the contestant is a primary job of the coach and the decision on the particular strategy shall consider many factors, including:

- The continuous concentration time of the contestant.
- The total amount of time, a contestant can be productive (it is in no way all 5 hours of the contest day for everybody – there have been contestants who decided to take a nap during the contest).
- The knowledge of the contestant on different problem types.
- The cost (of time and attention) of switching from one task to another.

The teams that regularly deliver good results definitely have well-rehearsed strategies. The teams that usually have lower standings or are aiming to achieve even better results need to analyze the behavior of the successful contestants and infer knowledge from it. The ranking page of CMS allows to see the dynamics of the submissions for the particular task of each contestant but the analysis of each might be quite a tedious job. In this paper, we consolidate this data and deliver statistics on task solving. To demonstrate a successful model, we provide the list of IOI 2019 gold medalists (on the Y-axis) with their work progress throughout the contest (X-axis) (Fig. 3). Each color in the graph corresponds to the particular task and the number on it shows the minutes spent on that task during that work. Each bar in this graph may include several submission attempts. As seen from the graph, some contestants haven't made that many task switches (On day 1, Ildar, Ziqian, and Eric have worked on one task at a time, but Benjamin and Zixiang have returned to the previous tasks twice), whereas others had the mentioned iterative approach. The same approach is followed the next day, which means the approach was consistent and rational.

Another interesting insight from this graph is, the majority of the gold medalists did not choose "shoes" as a first task. Our first graph in Fig. 1. shows that during the first 30 minutes, there have been almost 200 submissions on "shoes". By analyzing the current graph, we can say that future "gold" medalists have been a very small portion of these submissions. As seen from the graph, "shoes" were not the second target of the "golds", either. We may infer two possible hypotheses from here: either the "golds" have started to work on the tasks in the order they got them from the envelope ("rect", "split" and "shoes" – there is no fact or record on the actual order of tasks) or there is a common strategy of successful contestants – start with the hardest and then do the easy tasks.

On the contrary, on day 2, they start with the "lines" problem, which was mentioned as "easy" in the previous analysis. This fact leads to the first option, which states the stronger contestants probably approach the tasks by their order in the envelope. Even if this is the strategy, by the similar pattern of the colors, it is discernible that they have a steady pace on problem-solving. As mentioned earlier, the selection of the first task is not the unique characteristic of the contestant – there is a particular pattern of switching (or not doing so) from one task to the previous tasks.

The knowledge cannot be gained only on the top positive and limited set of information, so to see the impact of the task-switching on the results, we analyzed the distribution of the task switches for all contestants, grouping them by the medals (and also not getting one). By the medal we mean the final reward of the contestant – it
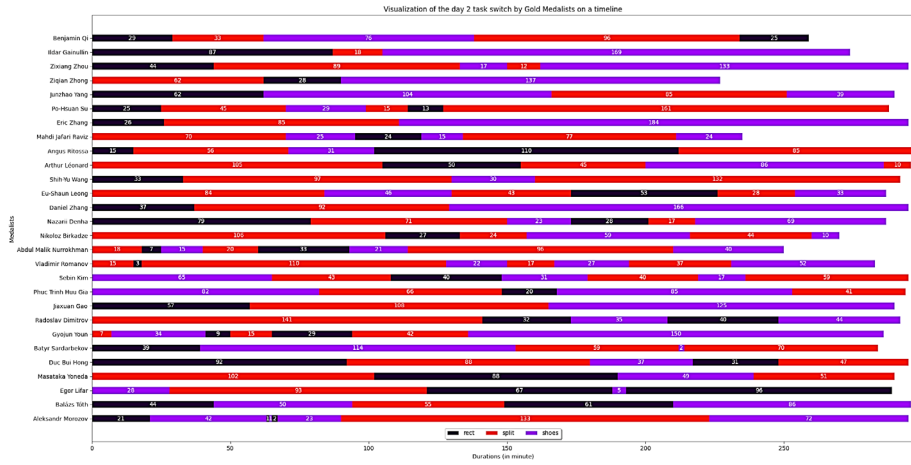
Day 1
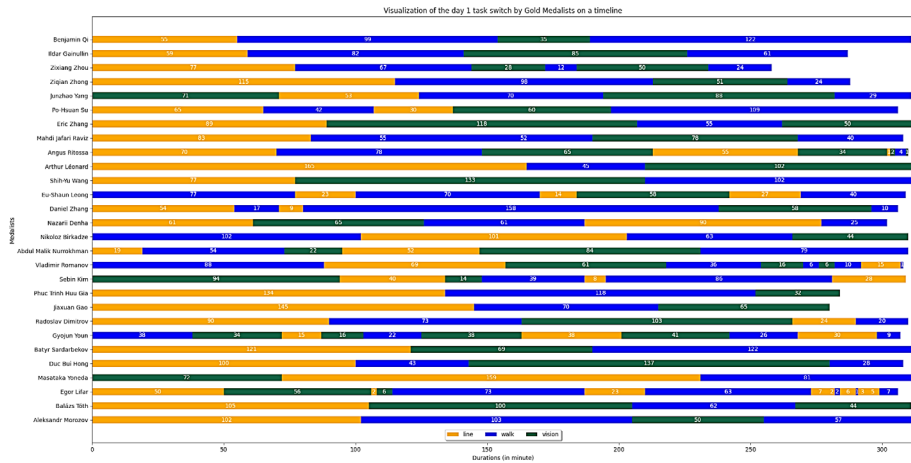
Day 2

Fig. 3. Gold medalists (sorted by standing) working on the tasks.

shall not be confused with the current standing. Fig. 4 shows the distribution of the amount of time spent on each task for each contestant category. The outliers with the minimum values (falling below Q1-1.5*IQR) have been removed from the data set. These are mainly the erroneous or random behavior (like working on a task for a few minutes, submitting and switching back to another task), that is not a subject of our analysis. On the contrary, the outliers with the maximum values are definitely valuable for the analysis.

As seen from Fig. 4, on day 1, almost 50% of the gold medalists spent from 30 to 70 minutes of time per task. The average time spent on a task in this category is around
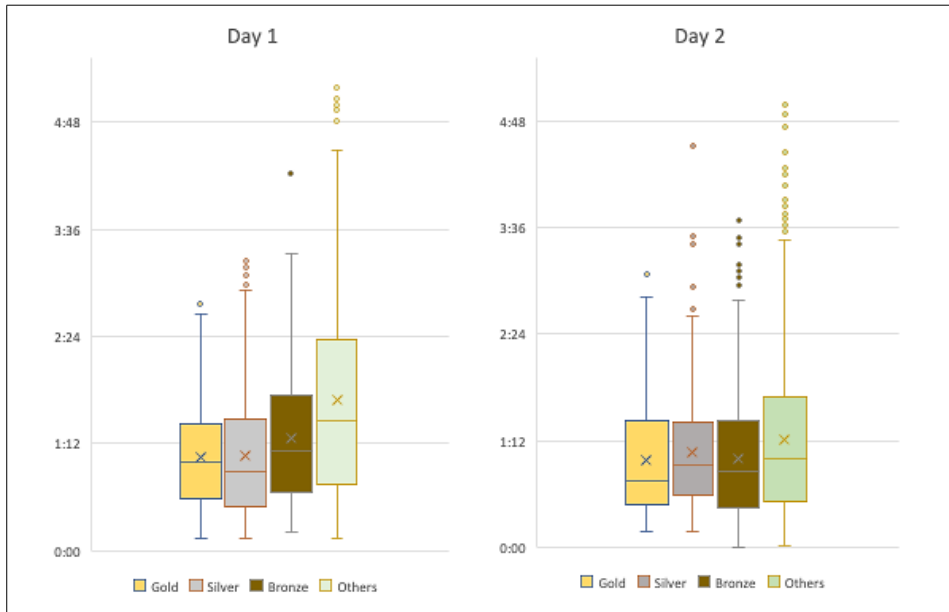
Fig. 4. Time spent on a task.

1 hour. It might be considered as a good time frame to keep the focus on a task. The silver medalists have the same average but with a wider variance and more outliers, that fall behind 2.5 hours bar. Despite silver medalists, the distribution of the bronze medalists does not share the common average with the "golds" and is shifted up. The average time spent on a task by the non-medalist category stands beyond 75% of the gold and silver medalists. 25% of non-medalists have spent more than half of the contest time on a single task and some contestants (outliers) have totally spent the contest time solving one task.

It seems the lessons learned (analyzed by the coaches or contestants themselves) applied to day 2:

- Gold medalists still had the same variance but some long attempts have been replaced with the small iterations.
- Silver medalists are still organized and follow the same plan. Some more outliers might be observed – contestants know their standings after day1 and work on the tasks that they are sure to score at.
- Bronze medalists show better performance – almost repeat the distribution of "golds". To secure their place for a medal, they need to be rational – not to spend time on a task, if they cannot succeed.

In Fig. 5, we show another distribution that is connected to the previous one – the number of repetitive switches made by each contestant. The number of repetitive switches is counted as the total number of switches from one task to another (including the first one) minus 3 (the number of tasks per day) – working on the tasks themselves for the
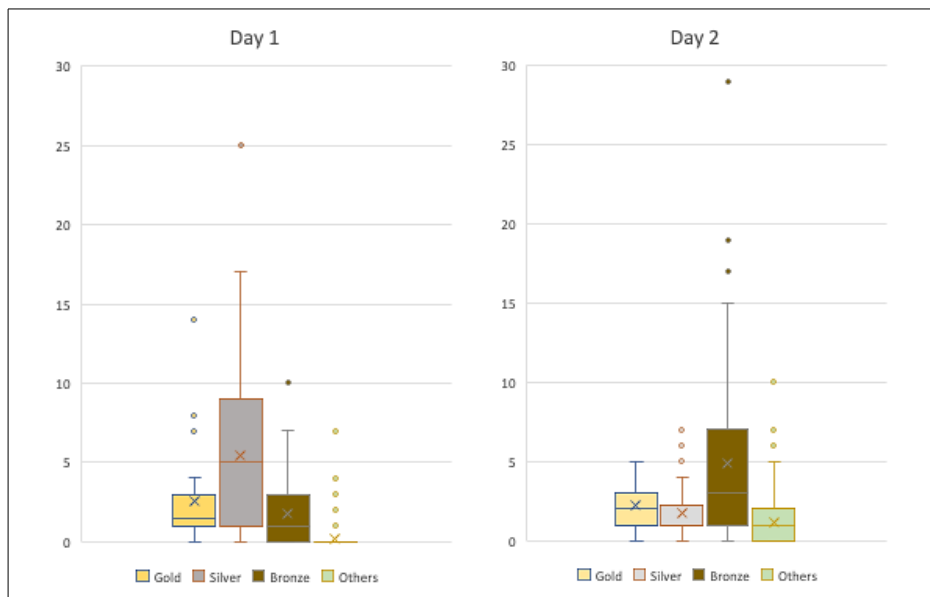
Fig. 5. The number of task switches.

first time is not counted as a switch. So in Fig. 3 day 2, Benjamin and Ildar both have made 1 repetitive switch.

Fig. 5. shows the number of switches which kind of is an addition to the previous graphs. As seen from day 1 graphs, despite the similar distribution of spent average time, "silvers" had way more task switches than "golds". On both days, gold medalists make no more than 5 switches, except a few cases shown as outliers for the day 1 report. On the contrary, "bronzes" had fewer task switches, and contestants who did not get any medal almost had no switch on day 1.

The lessons learned from day 1 had decreased the task switch count for "silvers". As explained for the task duration case, the "bronzes" started to score as much as possible by leaving the tasks, where they had no progress. These dynamics have touched the other contestants who did not receive any medal (as a result) – they stopped getting stuck at one or two tasks and started to try to have an iterative way.

According to the insights, the problem-solving process can be set as the following algorithm:

A1. Spend time to understand task X (say, the next one you get from the envelope). If the solution is obvious, move to A2. Otherwise, move to step A1.

A2. Start working on a task. Do you get positive/promising scoring?
- If it has been more than Y minutes, then move to A1, Otherwise, keep working.
- else, move to A1.

As mentioned earlier, the "Y minutes" value should be identified and set individually for each contestant by the coach.

2.4. *A Team Performance*

Teams that are associated with contestants' countries are usually introduced once in the Opening Ceremony of the event and occasionally mentioned in the cultural night events. Although IOI considers only individual contestants in the ranking and medal list, the medalists usually demonstrate their link to the particular country during the award ceremony. There is something more than just patriotic feelings in team recognition – a victory of the contestant is the victory of the coach and the system (educational, Olympiad movement in that country, the materials written and taught, trials, selection, and many others) that are not mentioned or noted as a part of the event. Regular country reports may include brief statistics about the country statistics, but it does not go deeper into the team performance analysis. One of the recent works provided an interesting analysis of the countries' performance in Science Olympiads, focused mainly on the particular regions (Jovanov *et al.*, 2018).

Taking this opportunity, we decided to show another report that shows the success of the teams according to their performance in IOI 2019 (Fig. 6). The graph shows the list of teams that owned at least one medal and a range of absolute scores that were made by their members. Teams with a single medalist have a single short vertical line. The teams are sorted from left to right by the decreasing order of the maximum team score. There are also three horizontal lines correspondingly showing the border for the "gold", "silver" and "bronze" medals scores.

As seen from Fig. 6, this graph resembles that ranking order, where team "USA" leads with Benjamin's highest absolute score 547.09. This report might help infer some more information, but the order of teams does not say anything about the performance of teams: although team "USA" got the highest score, the adjacent team "RUS" has better
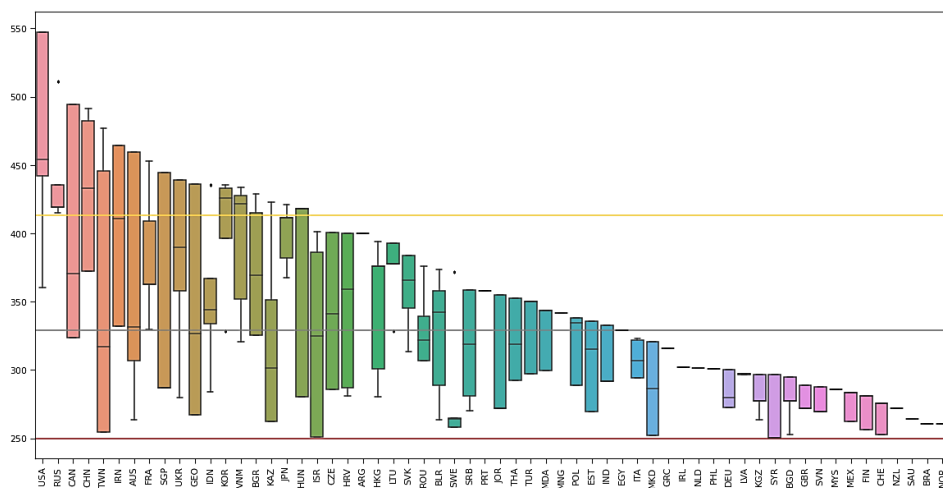

Fig. 6. Teams with medals and their score ranges (sorted by maximum value).

results – all 4 contestants got "gold" medals. Or team "TWN" in total scored more than team "JPN" (477.14 vs 420.89) but the range of scores are way much wider ([254.77, 477.14] vs [367.33, 420.89]). As mentioned earlier, having less variance might be a good criterion to evaluate the preparation of the team. One of the options to evaluate the team success could be the ordering of the team standing by the highest values of their mean, median or minimum scores.

3. Summary

As stated at the beginning of the paper, this short report covers the initial phase of the "IOI data analysis" project. The current phase covers the analysis of the CMS database records for the IOI 2019 event and provides insights into the activities that happened in that year. More detailed insights will be covered in the next phase, which is going to cover the year on year analysis, a high-level analysis of the code that was submitted (how the code is being evolved, what is being added/changed on each successful attempts, and so on) and also have a software that shows the introduced graphs in real-time. This software will be visible to the Technical and Scientific teams to better understand what is happening behind the scenes and prevent the possible problems before they might happen.

Summarizing the current report, the following outcomes might be useful for event organizers, committees, and coaches:

- The teams are mentioned but never evaluated or rewarded. Team members that keep almost the same standing or steadily increase their overall position from year to year are indicators of the systematic approach (having fair and decent Nationals, trials, selection processes, preparation camps, participation in other contests, and so on).
- The organization processes shall be refined as much as possible. The data of the previous events and the previous contest day shall be analyzed to offer an efficient contestant seating plan, contestant desk design, the content of snacks, and other details. Using graphical icons and tags in contestant requests (like in IOI 2019 – #water, #banana, #question, etc) may help better categorize the requests.
- The dynamics of the task submissions in the previous years might be analyzed for efficient task type selection and the creation of the test cases.
- Coaches may analyze the behavior of the successful contestants and teams, to find patterns that lead to success and adjust their strategy accordingly.

The sources of the reports and results of the work can be found in a public GitHub repository of the project (GitHub, 2021b).

Acknowledgements

References

Alnahhas, A., Mourtada, N. (2020). Predicting the performance of contestants in competitive programming using machine learning techniques. *Olympiads in Informatics*, 14, 3–20. DOI: 10.15388/ioi.2020.01.

CMS (n. d.). *Contest Management System*. https://cms-dev.github.io

Fantozzi, P., Laura, L. (2020). Recommending tasks in online judges using autoencoder neural networks. *Olympiads in Informatics*. 14, 61–76. DOI: 10.15388/ioi.2020.05.

GitHub (2021a). *ADA-SITE-JML / IOI-Grant: Structure of the CMS DB*.
https://github.com/ADA-SITE-JML/ioi-grant/tree/master/database

GitHub (2021b). *ADA-SITE-JML / IOI-Grant: Project Page*.
https://github.com/ADA-SITE-JML/ioi-grant

IOI Statistics (2020a). *International Olympiad in Informatics Statistics*.
https://stats.ioinformatics.org/olympiads/

IOI Statistics (2020b). Rezults The ranking of IOI 2020 contestants.
https://stats.ioinformatics.obrg/results/2020

Jovanov, M., Mihova, M., Kostadinov, B., Stankov, E. (2018). New approach for comparison of countries' achievements in science olympiads. *Olympiads in Informatics*, 12, 53–68. DOI: 10.15388/ioi.2018.05.

Kostadinov, B., Jovanov, M., Stankov, E. (2018). Platform for analysing and encouraging student activity on contest and e-learning systems. *Olympiads in Informatics*, 12, 85–98. DOI: 10.15388/ioi.2018.07.

Lodi, M. (2020). Informatical thinking. *Olympiads in Informatics*, 4, 113–132. DOI: 10.15388/ioi.2020.09.

Pankov, P.S., Kenzhaliyev, A.A. (2020). Pattern recognition and related topics of olympiad tasks. *Olympiads in Informatics*, 14, 143–150. DOI: 10.15388/ioi.2020.11.

Tsvetkova, M.S., Kiryukhin, V.M. (2020). Top 10 key skills in Olympiad in Informatics. *Olympiads in Informatics*, 14, 151–167. DOI: 10.15388/ioi.2020.12.

Vegt, W. van der, Schrijvers, E. (2019). Analyzing task difficulty in a Bebras contest using cuttle. *Olympiads in Informatics*, 13, 145–156. DOI: 10.15388/ioi.2019.09.

J. Hasanov is an Assistant Professor of Computer and Information Sciences in the School of IT and Engineering at ADA University. Dr. Hasanov is mainly focused on image processing and machine learning problems covering text and digital object recognition domains. Additional to the research field, Dr. Hasanov teaches the management aspects of the IT in production and operation environments. Dr. Hasanov has been an ITC member for the period of 2017–2020 and led HTC during IOI 2019.

H. Gadirli is a student in a dual-degree master program of "Master of Science in Computer Science and Data Analytics" at ADA University and George Washington University. Habil received his B.Sc. in computer science from ADA University in 2019.

A. Bagiyev is a first-year student in the dual-degree master's program of "Master of Science in Computer Science and Data Analytics" offered by ADA University and George Washington University. Aydin has been a member of the HTC team during IOI 2019 and contributed to the automation of the contestant environment management and image distribution processes. He also works as a full-time computer science instructor at ADA School, to contribute to the enhancement of the youths in the country.