Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Bachelorarbeit Nr. 333

Game-driven Investigation and Optimization of Graph Visualizations

Dennis Heller

Course of Study: Softwaretechnik

Examiner: Prof. Dr. Albrecht Schmidt

Supervisor: Dipl.-Ing. Valentin Schwind,
Jakob Karolus, M.Sc.

Commenced: May 23, 2016

Completed: November 22, 2016

CR-Classification: H.1.2, H.5.2, I.3.6

Abstract

In the era of "Big Data", information visualisation becomes more and more important to draw conclusions and implications from the massive amount of data. One research area of the information visualisation studies the visualisation of graphs in the form of node-link-diagrams. Under this topic, we are interested in the effects of graphs' aesthetic aspects on human task performance. There has already been research about some of the aesthetic criteria of graph visualisation, but - to the best of our knowledge - not about the ratio of node size and edge width. In this bachelor thesis, we investigated the effect of node size, edge width and the ratio of both on human task performance and on subjective aesthetic perception. This field study was wrapped in three online games. We tested 27 combinations of node size and edge width together with three levels of difficulty. Our post-game questionnaire polling the participants subjective perception of node/edge-ratio surfaced interesting results: there are clear optima for the node size and edge width and their ratio. We also found out that a too low ratio of node size and edge width is significant less appealing than a too high one. However, there was no significant correlation between the graph parameters and the human task performance measured by a self-derived score system.

Kurzfassung

Im Zeitalter von „Big Data“ wird die Visualisierung von Informationen zunehmend wichtiger, um aus großen Datenmengen Schlussfolgerungen ziehen zu können. Ein Forschungsbereich der Informationsvisualisierung beschäftigt sich mit der Visualisierung von Graphen in Form von Knoten-Kanten-Diagrammen. Innerhalb dieses Themas interessiert uns die Auswirkungen der ästhetischen Eigenschaften von Graphen auf das menschliche Verständnis. Einige dieser ästhetischen Eigenschaften wurden schon erforscht, aber darunter war - soweit wir wissen - nicht das Verhältnis von Knotengröße zu Kantendicke. In dieser Bachelorarbeit untersuchten wir die Auswirkungen von Knotengröße, Kantendicke und das Verhältnis dieser beiden auf die menschliche Leistung beim Lösen von Problemen und auf die subjektive Wahrnehmung. Diese Feldstudie wurde durch drei Online-Spiele mit Highscore-Tabellen umgesetzt. Wir haben 27 Kombinationen aus Knotengröße und Kantendicke zusammen mit drei Schwierigkeitsgraden getestet. Unser Fragebogen am Ende jedes Spiels, der die subjektive Meinung der Teilnehmer über Knotengröße, Kantenbreite und deren Verhältnis erfasste, brachte interessante Ergebnisse: Es gibt ein klares Optimum für Knotengröße, Kantendicke und deren Verhältnis. Außerdem fanden wir heraus, dass ein zu kleines Verhältnis deutlich negativer wahrgenommen wird als ein zu großes Verhältnis. Zwischen obigen Graph-Parametern und der menschlichen Leistung, die wir durch ein selbst erstelltes Punktesystem maßen, besteht jedoch kein signifikanter Zusammenhang.

Contents

1 Introduction

> The goal is to turn data into information, and information into insight.
>
> *(Carly Fiorina [Fio04])*

Information visualisation played an important role in human history for a long time, especially in the coming era of "Big Data" [CZ14]. From the first maps of countries and shipping routes to financial statements of the last quarter, data can only be used efficiently if it is visualised for people, who can draw conclusions about it.

In this work, we focused on a subfield of information visualisation: the visualisation of graphs. "Graphs are abstractions of node-link diagrams studied by mathematicians, and the discipline of graph drawing has developed to investigate methods for laying out graphs according to a set of aesthetic principles that are assumed to improve readability" [BETT99]. Graphs have a large bandwidth of applications, for example in street maps and navigation, sociograms, machine learning, games and networks. Because of this, there has been much research and interest in the topic of graph theory and graph visualisation. The problem of finding good visualisations for graphs may be trivial for small or sparse graphs but gets harder for large and dense ones. Possible visualisations of graphs are node-link-diagrams, matrices and hybrids. Evaluating aesthetic criteria of node-link-diagrams can be challenging as there are many influencing factors. Some of them are the absolute node and edge size as well as their ratio. We believe that the aesthetics of graphs influences human understanding and task performance. Thus, we investigated the effect of these three variables on human understanding and task performance through an empirical field study. This study is wrapped in three online games, measuring the players' success through the resulting score.

Structure

The thesis is structured as followed:

Chapter 2 – Related Work contains the summaries of related work.
Chapter 3 – Method describes the composition and structure of the experiment and the online games.
Chapter 4 – Results reports the data of the experiment.
Chapter 5 – Discussion evaluates the data and discusses implications.
Chapter 6 – Conclusion summarises the results of this work, gives guidelines and contains an outlook on future work.

2 Related Work

Graphs are a well-known kind of visualisation of relational data. In our context, a graph means the same as a node-link-diagram, although the computer science differentiates between these two terms, where a node-link-diagram is just one possible visualisation of a graph, which is an abstract set of nodes and edges.

“The visualisation produced by a graph drawing subsystem should illuminate application data. That is, should help the user to understand and remember the information being visualised. [...] Designers of graph drawing algorithms and systems claim to illuminate application data by producing layouts that optimise measurable aesthetic qualities, on the assumption that considering these aesthetics makes the drawing ’nice’ and easier to read” [PCJ97].

2.1 Aesthetic Criteria

There has been much research about the different aspects of visualisation of graphs in the past 20 years. The aesthetic is an important aspect in visualisation because it “affects our receptiveness to viewing something further or making the effort to understand it” [BRSG07]. To quantify the aesthetic of graphs, there are some aesthetic criteria:

- minimising the number of edge bends and edge crossings
- maximising the angles between the edges leaving the nodes
- symmetry
- orthogonality
- optimise the area of the graph and the length of edges
- clustering
- shape, size and colour of nodes and edges

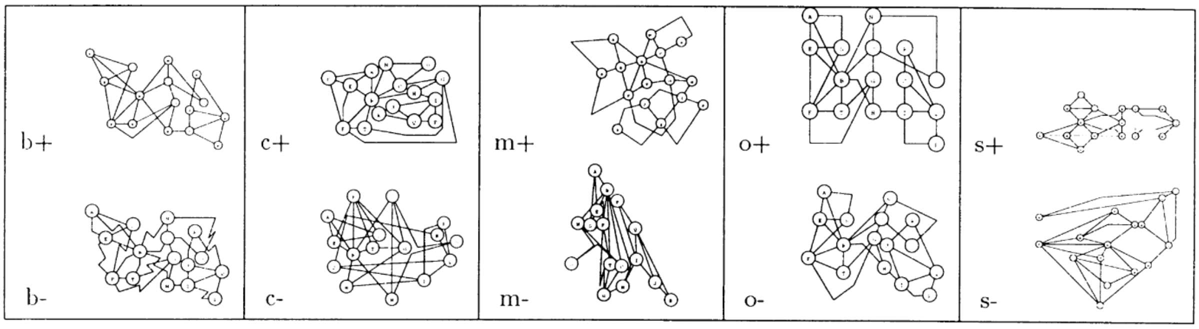

Figure 2.1: Different Graphs used for the experiment in the work of Purchase [Pur00]. Each two graphs focus on one aestetic criterium, from which one is very good (+) and the other graph is very bad (-) according to that criterium. The criteria are minimising bends (**b**), minimising edge crossings (**c**), maximising angles (**m**), orthogonality (**o**) and symmetry (**s**). Permission to use granted by the author.

Some of these criteria are contradicting, e.g. the number of edge crossings and the number of edge bends. The better the one gets, the worse the other one gets. This is the reason why there has been research on the relevance of these criteria, that is on their effect on human understanding.

There are studies which had the aim to examine the effect of these criteria on understandability and user task performance. Usually, the participants had to answer questions about some graphs shown to them, where each graph focusses on exactly one aesthetic criterion while trying to be neutral from the perspective of all other aesthetic criteria. "The performance of a person in answering these questions about a graph drawing was considered an appropriate indicator of their understanding of the drawing" [PCJ97].

Helen Purchase has published many studies of that kind. She considered the five criteria edge bends, edge crossings, angles between edges, symmetry and orthogonality. Participants had to answer questions like "How long is the shortest path between two given nodes?" [Pur00] or "What is the minimum number of nodes that must be removed in order to disconnect two given nodes?" [Pur00] on the graphs shown in Figure 2.1. She detected that the number of bends and the number of edge crossings have a significant impact on the error rate and, together with the symmetry, have a significant impact on reaction time, whereas the angles and orthogonality had no impact. In the discussion, she points out that "There is no doubt that the evidence is overwhelmingly in favour of crosses as being the aesthetic that affects human relational graph reading the most". [Pur00; Pur97]

Purchase et al. [PCJ95; PCJ97], examined the effect of symmetry, the number of bends and the number of crossings on the error rate on answering questions about the graphs.

Similar to before, they found out that the number of bends and crossings has a significant effect on human understanding, while it is inconclusive whether symmetry has an impact on it.

Huang, Hong, and Eades [HHE08] observed a very special aesthetic criteria: the angles of edge crossings. They came to the conclusion "that crossing angles had significant impact on response time in tracing a path" [HHE08]. Furthermore, "drawing graphs with large-angle crossings can be equally effective as drawing graphs without crossings, in terms of response time for path tracing tasks" [HHE08]. It is not clear, however, in which degree these results can be transferred to graph-related tasks other than path tracing.

Another study on path tracing tasks was performed by Ware et al. [WPCM02], which proved that the aesthetic criteria of "path continuity is an important factor in perceiving shortest path" [WPCM02], where continuity was defined as the sum of all angular deviations at the nodes of the path. The smaller the continuity value (or the straighter the path), the easier it was to find for the subjects. Comparing path continuity with the number of edge crossings, they estimated that "the cognitive cost of a single crossing is approximately the same as 38 degrees of continuity" [WPCM02].

As outlined, there has been much research on the layout aesthetic criteria like the number of bends, the number of edge crossings, the angles between edges, symmetry and orthogonality. To the best of our knowledge, the effect of node and edge sizes on human understanding has not been investigated yet. Thus we expect those aesthetic criteria to have an effect on task performance during our studies, too. Following methodology from previous work, we keep other criteria neutral while manipulating node and edge size and measure the effect.

2.2 Syntax and Semantics of Graph Understanding

According to Purchase, Cohen, and James [PCJ97], there are two different kinds of graph understanding: syntactic and semantic. Syntactic graph understanding is application-independent and more generic than semantic understanding, it enables to answer questions like "what is the shortest path between A and B?" [PCJ97]. In contrast, semantic graph understanding is application-specific and interpretative, for example, the understanding of a UML diagram. An example question for semantic understanding would be "what object classes would be affected by changing the external interface to class X?" [PCJ97].

Purchase, Carrington, and Allder [PCA02] have shown that it is important to consider this when choosing a graph layout algorithm because there is "a clear signal that algorithms that are designed for abstract graph structures, with no consideration of their ultimate use, will not necessarily produce useful visualisations of semantic information" [PCA02]. In this case, she found out that orthogonality is an important aesthetic criterion for UML diagrams, whereas it is unimportant for generic graphs.

Our study will focus on syntactic understanding because we will use graphs without a specific application context. However, the origin of our second game, Scotland Graph, includes some kind of semantic, where the application context is an underlying public transport network.

2.3 Gamification in Empirical Studies

Wrapping empirical studies in games is a common approach, that has some drawbacks but some benefits, too. The most important factor for using games in studies is that they are self-motivating. "People naturally want to win games they play and winning is a reward and a source of motivation" [BBR04]. This is dependent on the quality of the game, e.g. moderate difficulty, variety and fancy graphics. Another advantage of using games are the different expectations of game players: "Game players do not, on the whole, mind having the results of their games recorded since this would, anyway, be necessary for features such as level progression and high-score tables" [BR07] and they "are tolerant of having arbitrary looking changes made to the way games are presented – it would be seen as part of the game" [BR07]. Especially the latter point is important for us since we need to change the node and edge sizes of the graphs.

To increase the number of subjects, such games should be accessible via Internet. This leads us to some drawbacks when using games for field studies: For multiplayer games, we need to implement an artificial intelligence (AI), because we can not let subjects wait for another subject logging in to play against him, which would decrease the number of played games. Luckily, the AI "does not need to play perfectly, just well enough to provide a challenge for a novice human. The most important thing is that the human subjects should not be able to adopt very trivial strategies in order to win. For example, a user who is trying to block a path between two nodes should not be able to do so by simply picking off all the neighbours of one end point" [BBR04]. Especially for *online* games, there is another drawback: "We have no control over who is playing the games or how. [...] maybe a group gathered around a computer" [BR07] and we do not know "whether the subjects concentrate and give their best or whether they, for example, check their e-mails at the same time" [BR07]. For this, we have to and do trust in the good will of our subjects. Also, we expect the variance to be minimised by a large number of participants.

We can use games for scientific studies, because “we can assume that playing a game on a graph typically involves a similar cognitive understanding of the graph to that needed for other kinds of graph applications” [BR07].

The challenge is to design games that

- ... “can only be played successfully by understanding the structure of the displayed graph” [BBR04].
- ... are “sufficiently enjoyable and rewarding that enough people will play enough games to provide data” [BR07].
- ... “let us vary the experimental parameters independently” [BR07] (in our case node and edge size).
- ... are “playable by people who do not know anything about graph theory and graph terminology” [BR07].
- ... can be played by exactly one human player (we need an AI for multiplayer games).

Bovey and Rodgers [BR07] implemented a Treasure Hunt Game (Figure 2.2) and the Shannon Switching Game (Figure 2.3) to investigate the effect of moving nodes on the success of the human players, e.g. their win rate. They found out that the subjects had a significantly higher win rate in the Shannon Blocking Game (which we will use in our study, too), when the nodes were moving, “despite the extra cognitive and motor load required to find and click on moving objects” [BR07].

2.3.1 Questionnaires Inside Games

Frommel et al. [FRB+15] investigated the effect of questionnaires inside games on the players’ performance. They found out that interrupting the game decreases the experienced “presence” of the game, where “presence” was defined as “[experiencing] the virtual world as a natural environment” [FRB+15]. So they concluded that such questionnaires should not interrupt the game but instead “fit into that world as naturally as possible” [FRB+15].

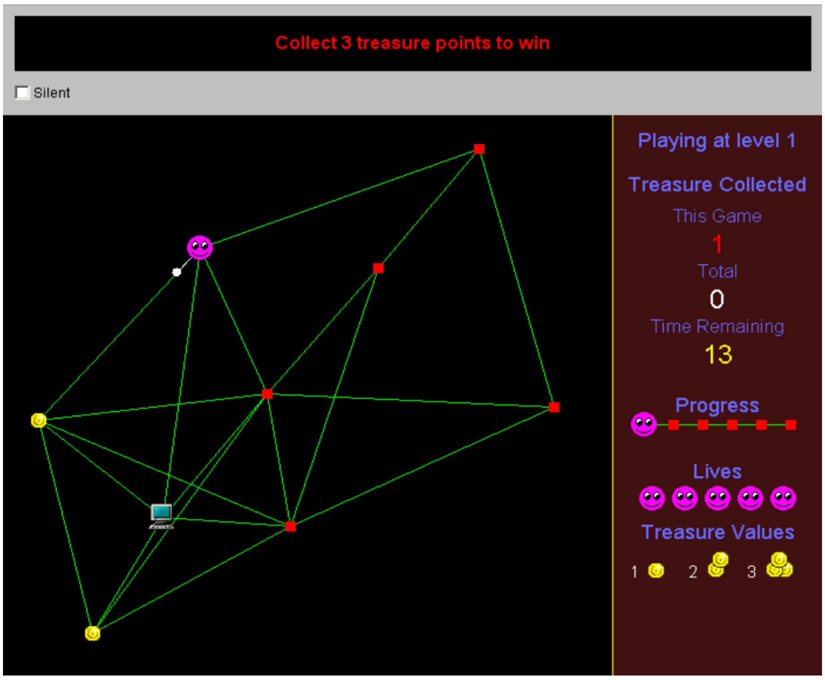

Figure 2.2: The Treasure Hunt Game in [BR07], used to investigate the effect of moving nodes on the proportion of won games. Permission to use granted by the authors.

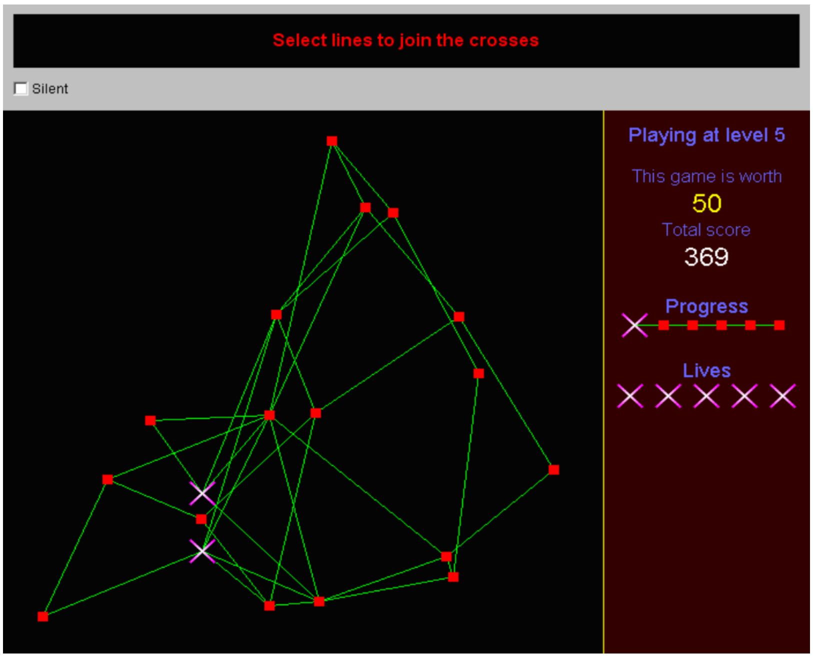

Figure 2.3: the Shannon Switching Game in [BR07], used to investigate the effect of moving nodes on the proportion of won games. Permission to use granted by the authors.

2.4 Cognitive Aspects of Graph Visualisation

The science of graph visualisation is not far apart from psychology. Another way to investigate human understanding is by looking at the human cognition process instead of only modifying the visualisation, which treats the human mind as a black box.

Huang and Eades [HE05] did so by performing an eye-tracking study. Their subjects had to find shortest paths in some graphs, while their eye focus was monitored. After examining the eye-tracking-videos, they concluded that most people

- ... start searching at the top-left corner
- ... avoid dense areas of the graph
- ... do not test all possible paths but use a heuristic process

Considering the graph layout, they propose a radial layout being best for human understanding. Their subjects, too, “claimed this kind of layout preferable” [WPCM02]. Also, they proved the result of Ware et al. [WPCM02] that path continuity is an important factor for the performance of path-finding tasks.

Huang, Eades, and Hong [HEH09] interrogated the traditional methods of performance measuring by considering only response time and accuracy. They introduced a new formula for calculating “visualisation efficiency” not only out of the former two variables but also out of the mental effort. “When visualizing large and complex data, the efficient use of limited human cognitive capacity stands out and becomes an important factor determining the success of a particular visualization technique” [HEH09].

3 Method

3.1 Hypotheses

As mentioned in Section 2.1 we will investigate the effect of node size and edge width in graph visualisation on human understanding. The hypotheses for our experiments are:

- The ratio of node size and edge width has an effect on the human task performance (such as task completion time) of graph-related tasks.
- There is an optimum for the ratio of node size and edge width in terms of human task performance.
- There is an optimum for the ratio of node size and edge width in terms of subjective aesthetic appeal.

We will also have a look at the two variables node size and edge width on their own and whether they have an effect on human task performance and on aesthetic appeal. In contrast to Related Work, we restrict our study to round black nodes and solid black edges.

3.2 Pilot Study

On the one hand, we want to test a large bandwidth of possible node sizes and edge widths to be sure to capture the assumed optimum of node-edge size ratio. On the other hand, we want to test at a fine granularity to get a more exact result in the end. But this would lead to many node-edge size combinations and therefore many conditions. Together with three difficulty levels, there would be too many conditions to handle. So we need to pre-select some node-edge size combinations, that are feasible and discard outlier combinations. In order to limit the number of possible combinations and to determine the potential cluster centre of the most appealing node-edge size ratio, we conducted a pilot study.

edge size $[mm]$ / node size $[mm]$	0.5	1	1.5	2	2.5	3	4
6			1	2			
7		3	4	5	6		
8	7	8	9	10	11	12	
9	13	14	15	16	17	18	19
10		20	21	22	23	24	
12			25	26	27		

Table 3.1: The IDs of the 27 ordered combinations of node size and edge width.

We suspected the optimal node-edge size ratio to be somewhere between 1 and 30. So we created Figure 3.1, a sheet with 74 little graphs with node sizes from $3\,mm$ to $15\,mm$ and edge sizes from $0.5\,mm$ to $5\,mm$. We then printed this sheet in DIN A3 format and posted it on a dashboard.

The participants were asked to stand in front of the sheet and had to point out exactly one small graph which is, in their opinion, the most aesthetically appealing.

17 people from the University of Stuttgart, the University of Konstanz and the Max-Planck-Institute Tübingen participated in the pilot study.

The results of the pilot study are shown in Figure 3.2. This data led us to choose the node-edge size combinations in the area with the light red background, to find a compromise between covering all voted combinations and keeping the number of conditions small. These 27 combinations are ordered as shown in Table 3.1.

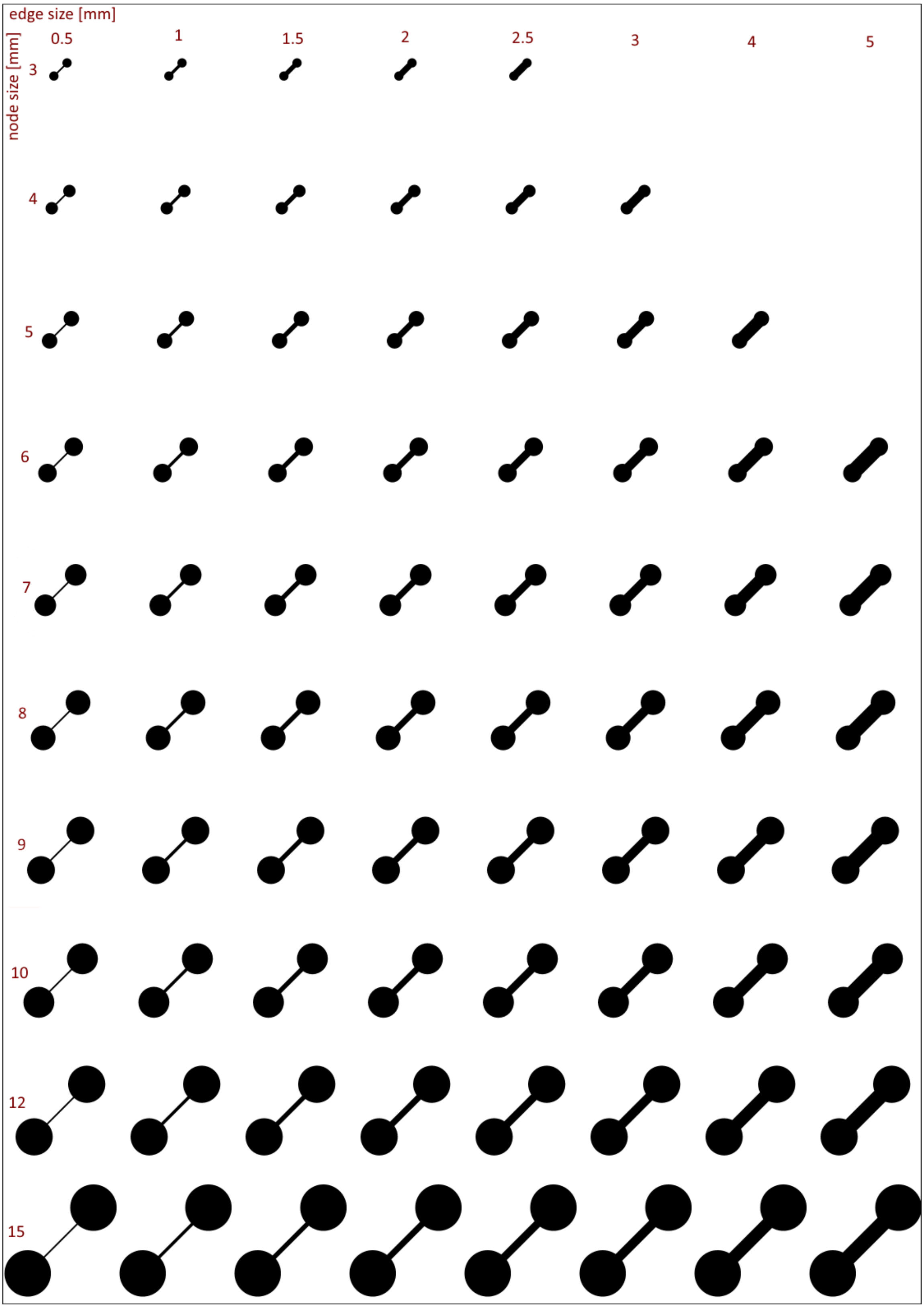

Figure 3.1: Sheet handed out during the pilot study. The red words and numbers were added later for better overview.

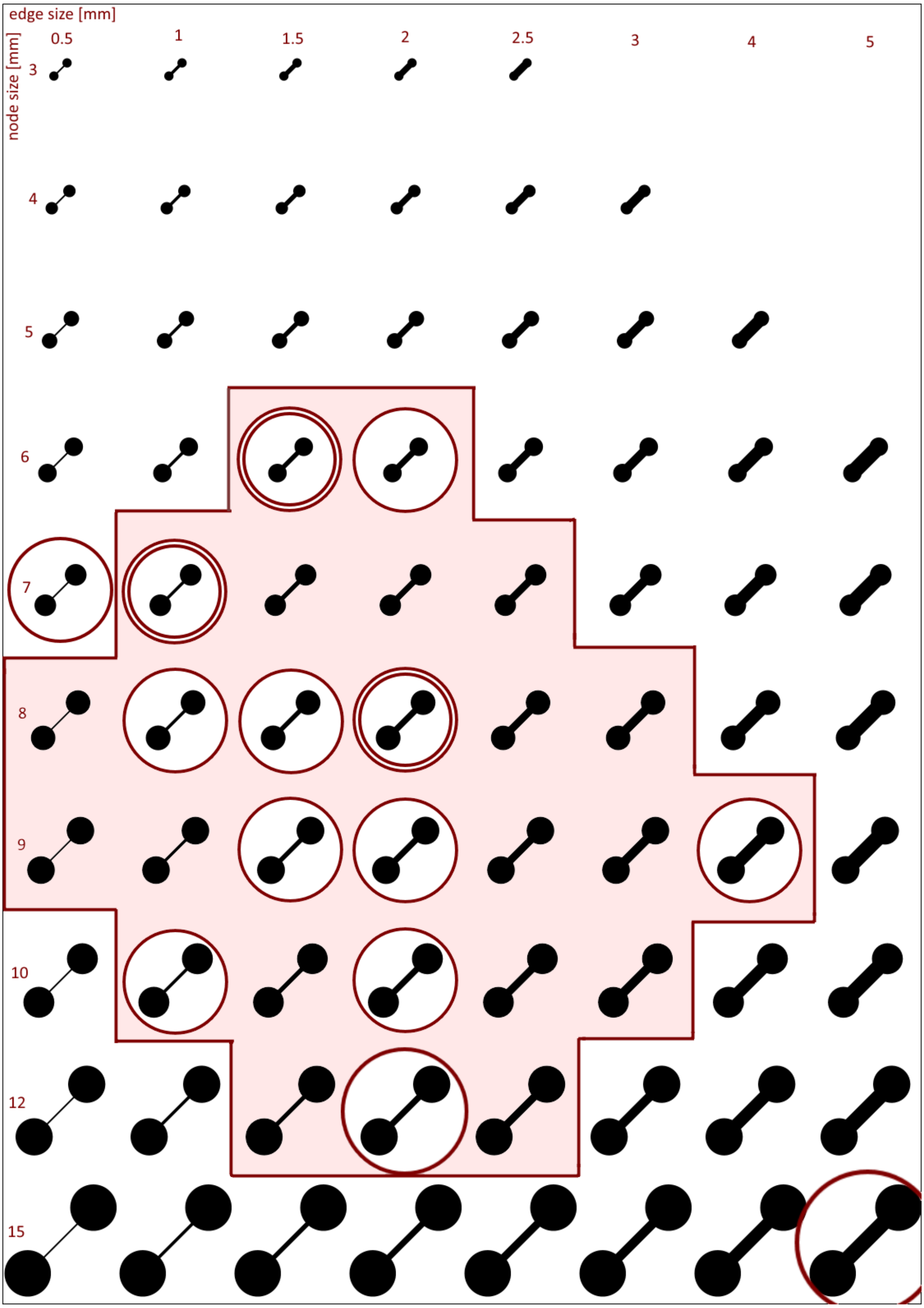

Figure 3.2: Results of the Pilot Study: Each circle around a little graph means one vote for this node/edge size combination.

3.3 Study Design

This section describes the scientific structure of the experiment, i.e. the operationalization, such as conditions and measured task performance.

3.3.1 Gamification

We decided to wrap the study in an online game due to the reasons mentioned in Section 2.3. Especially with such a large number of conditions, we need many participants. This is only possible if the participants can play the game anytime and at home.

3.3.2 Variables

We have four independent variables: the difficulty and the node-edge size combination. The latter one technically consists of the two variables node size and edge width, which are not independent on their own, because we did not test all possible combinations of them. The independent variable (IV) “ratio of node size and edge width” is added for clarity, it is actually derived from the node size and the edge width, but we will investigate its effect on the dependent variables, too. Table 3.2 shows a summary of all variables.

name	**type**	**scale**	**values**
difficulty	IV	ordinal	$\{1, 2, 3\}$
node size	IV	ratio	see Table 3.1
edge width	IV	ratio	
ratio of node size and edge width	IV	ratio	$[2.25, 18]$
game result	DV	nominal	$\{win, loss\}$
needed time	DV	ratio	$\mathbb{R}^+$
needed moves	DV	ratio	$\mathbb{N}$
score	DV	ratio	$\mathbb{N}$
in Planarity only			
proportion of useful moves[1]	DV	ratio	$\mathbb{N}$

Table 3.2: The independent (IV) and dependent (DV) variables.

[1]proportion of moves in which the number of edge crossings decreased

3.3.3 Latin Square

From 27 combinations of node size and edge size and three difficulty levels, we have 81 conditions which need to be evaluated on their own afterwards.

One common problem in experimental design is the varying performance of the participants over the time. “This would not be too much of a problem if this were merely random fluctuation in performance, because then it should cancel out across conditions. However, systematic variations in performance pose a much more serious problem. Participants become fatigued, bored, better practised at doing the set tasks, and so on. These systematic or ’confounding’ effects may interact with our manipulations of the independent variable of interest, rendering our results uninterpretable.” [FH02]. An approach to counterbalance this “order effect” or “learning effect” is to split the participants into groups of equal size and let the different groups perform the experimental conditions in different orders. This is possible for a few conditions, but for 81 conditions there are $81! = 5.8 * 10^{120}$ permutations and we would have to create as many groups.

“However, by using a Latin Squares design, we can cut down on the number of groups we need to run, and yet still avoid order effects” [FH02]. A Latin Square is a square of numbers in which each number occurs exactly once in each row and each column. This 81 out of $5.8 * 10^{120}$ permutations are a representable subset and therefore reduce the learning effect if each condition A occurs after condition B just as often as B occurs before A. Like this, a potential learning effect of condition A on condition B or vice versa is counterbalanced over all participants.

In our specific case, when a new user logs in on the website, a row of the Latin Square is associated with him. To determine this row, the whole database is searched on which condition currently has the fewest entries. The row that starts with this condition’s ID is then chosen. The subsequent levels played by the user are based on the subsequent numbers of “his/her” row of the Latin Square.

Given a condition ID n, we calculate the difficulty and the node/edge-combination ID as described in (3.1) and Equation (3.2). From the node/edge-combination ID, we can determine the node and edge size by looking up in Table 3.1.

$$\text{difficulty} = \left\lfloor \frac{n}{27} \right\rfloor + 1 \tag{3.1}$$

$$\text{node/edge-combination ID} = n \mod 27 \tag{3.2}$$

Figure 3.3: Our 81x81 Latin Square

Our 81×81 Latin Square, shown in Figure 3.3, was generated with the so-called checkerboard algorithm[2]. It let us create this Latin Square out of nine 9x9-Latin Squares. Afterwards, we randomly permutated the rows and columns.

3.4 Procedure

The study is carried out through three online games: Planarity, Shannon Switching Game and Scotland Graph. This section describes the procedure of playing the two games from the participant's point of view.

3.4.1 Registration

Each participant needed to register first by entering his/her age, sex and the display size (diagonal in inch) of the currently used screen. To find out the display size, the

[2] http://www.ciphersbyritter.com/ARTS/PRACTLAT.HTM

participant could also enter the vendor or model name of the display. If it was found in our list of known displays, the display size was automatically inserted. Our list contains a total of 758 screens from common vendors such as Acer, Apple, ASUS, BenQ, Dell, HP, LG, Samsung, Lenovo. The aim of this list was to make the process as easy as possible to not lose possible participants already in this early step, who would be overwhelmed by finding out the size of their display.

We asked for the age and sex of the participants because perception may be dependent on these factors, and therefore we could evaluate the data concerning this fact.

If the participant visits the website again later, he has the possibility to continue as the previous user or create a new one.

3.4.2 The Games

The player generally has the choice between the two games. Before each level, he sees the instructions for the game, including the mouse controls and the factors on which the score depends on. As soon as he clicks on “Start Game”, the stopwatch starts.

On the right side, there is always a sidebar, on which the player can see the current level number, the current score (which is automatically updated every half second) and the factors on which the score depends on, e.g. the number of moves and the time. Additionally, in the right bottom corner, there is a help button which opens the game instructions in an overlay. After finishing each level, the player sees the score of the previous level.

Also after each level, the player has to answer a questionnaire about his opinion on the previous graph. Like this, we have both an objective (the score) and a subjective result and can compare the two. The questionnaire popup is laid over the graph of the previous level so that the player still can see the graph. We asked the users about their opinion on the node size, the edge width and the ratio of the former two. Possible answers were items from a five-level Likert scale, as shown in Figure 3.4. The player had to answer all questions in order to continue.

3.4.3 Planarity Shannon Game

This game is a combination of two well-known games in the graph theory: *Planarity* and the *Shannon Switching Game*. The result of Planarity is a planar graph, which is then used to play the Shannon Switching Game on.

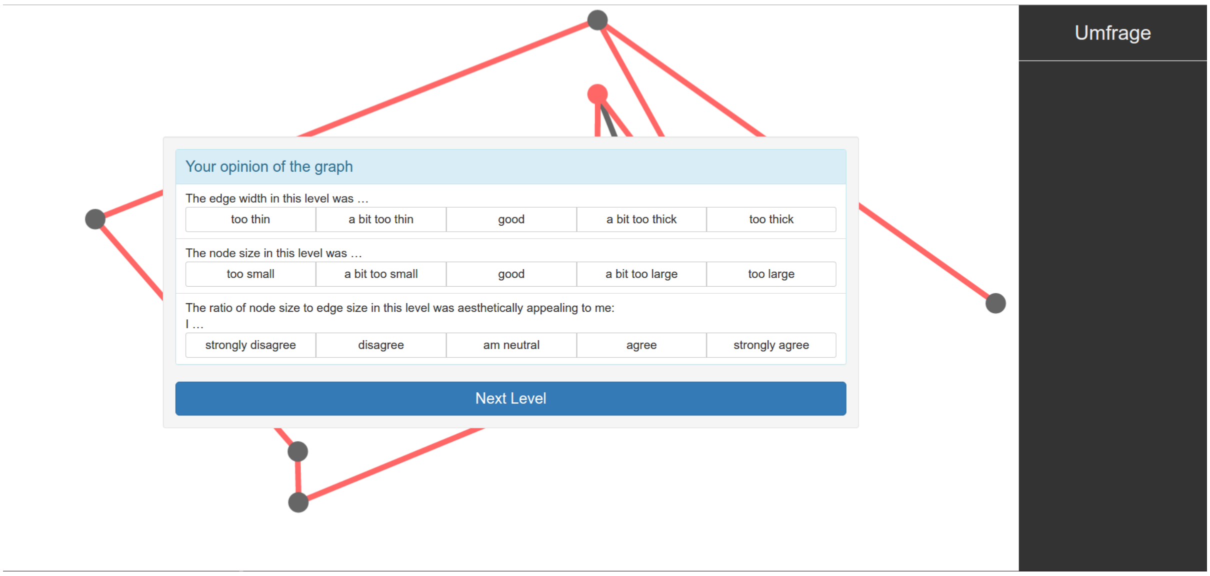

Figure 3.4: Screenshot of the questionnaire at the end of each level

Planarity

Planarity is a graph game originally invented by Tantalo [Tan05]. The player sees a muddled graph. He then has to reposition the nodes so that the graph becomes planar, e.g. it has no more edge crossings. Nodes can be moved with the mouse by drag-and-drop. Figure 3.5 show the instructions page and Figure 3.6 shows an example game. The implementation is described in Section 3.6.3.

Shannon Switching Game

The Shannon Switching Game is a well-known game in computer science, especially in the Graph Theory, invented by Claude Shannon. It has been part of some studies about combinatorial games and is therefore sufficiently investigated on the aspects of winning strategies. A combinatorial game is “a game with two players each having separate turns and making different moves to achieve a winning positive” [Woo12].

The rules are very simple: Given an undirected connected graph, two of the vertices are highlighted (“distinguished”). In our case, they are painted in red contrary to the other black vertices. One player (“short”) marks an edge in each of his turns, which is then painted in red. This edge can be any edge, it does not need to be connected to another marked edge. *Short* wins as soon as there is a marked path connecting the two highlighted nodes. The other player (“cut”) deletes an edge in each of his turns, that

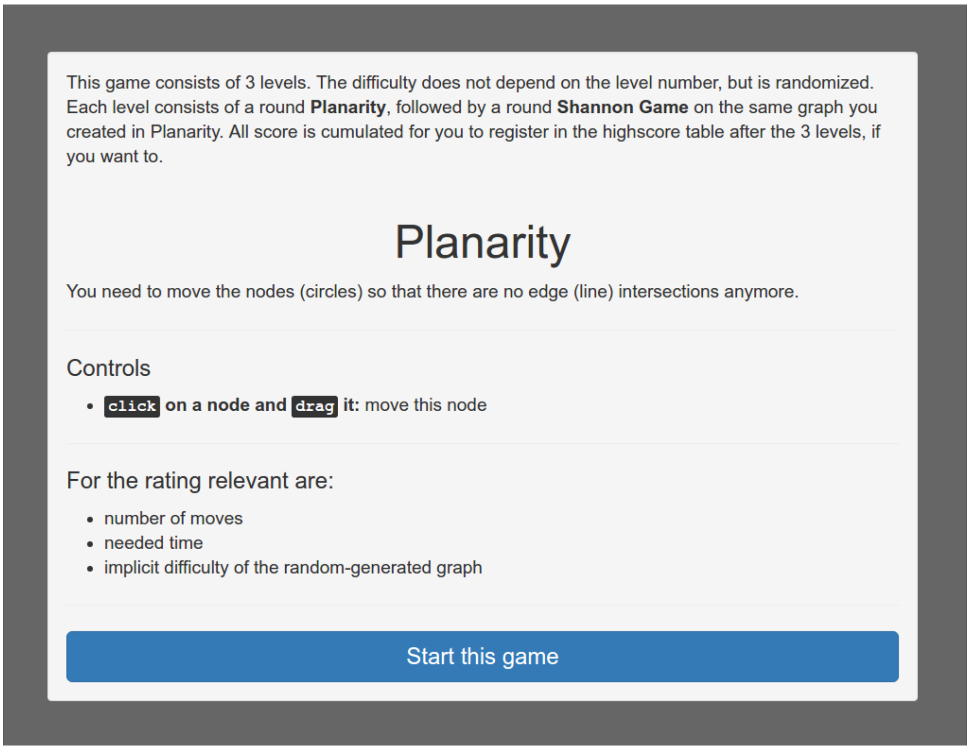

Figure 3.5: Instructions for the Planarity Game, including controls and score-composing factors.

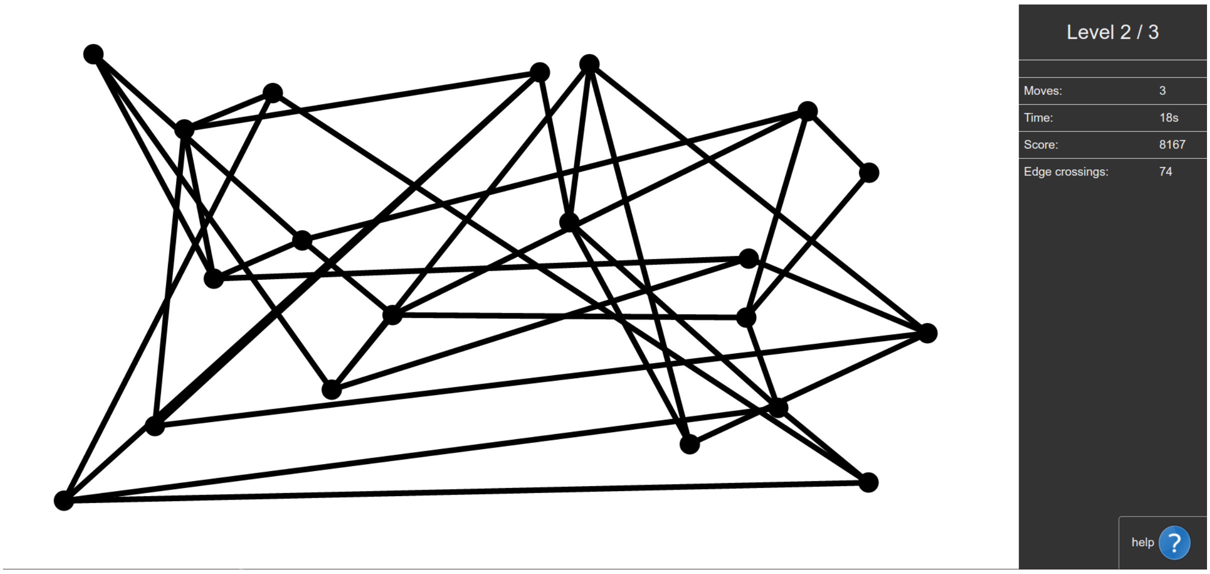

Figure 3.6: Screenshot of a Planarity game

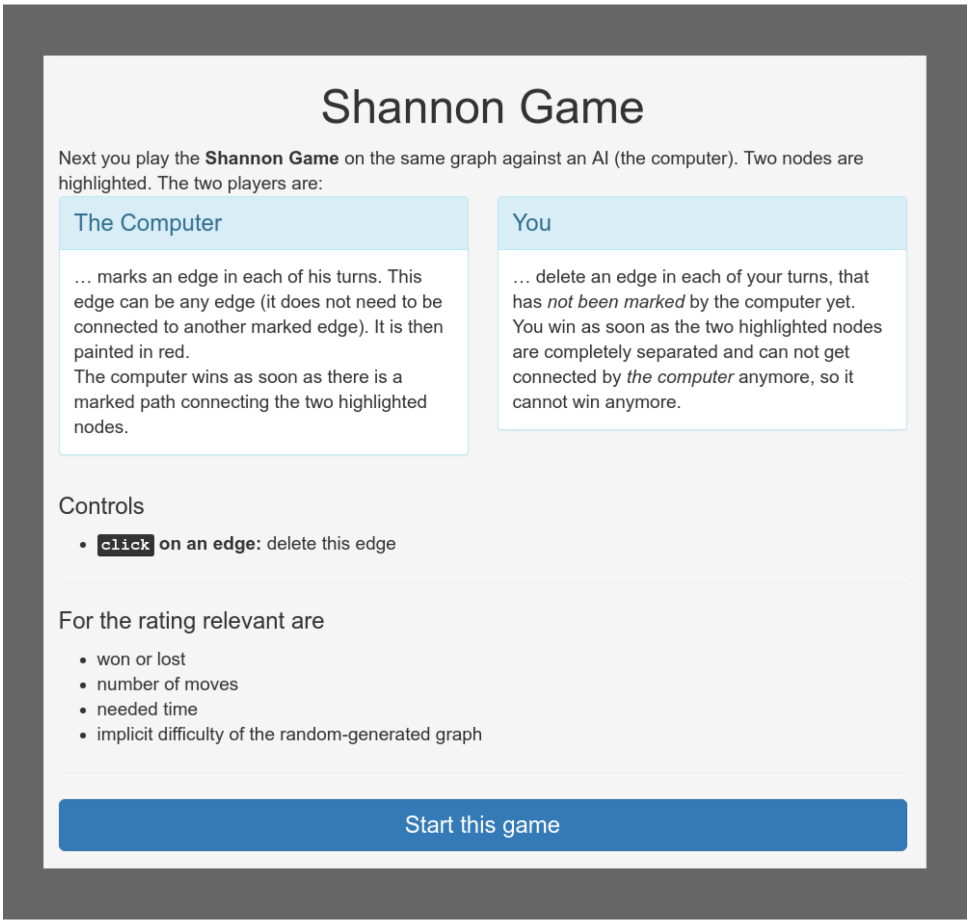

Figure 3.7: Instructions for the Shannon Switching Game, including controls and score-composing factors.

has not been marked by *short* yet. *Cut* wins as soon as the two highlighted nodes are not connected anymore, so *short* cannot win anymore.

The participants always play *cut*. They can remove edges by clicking on them. Figure 3.7 shows the instructions page and Figure 3.8 shows an example game after six moves have been passed. The implementation is described in Section 3.6.4.

3.4.4 Scotland Graph

This game is inspired by “Scotland Yard” from Ravensburger from 1983 [SGI+83]. In the original game, some *detectives* had to capture *Mister X* on the public transport network of London. At the beginning, each player stands on a node, which was randomly chosen from a set of starting nodes. They can then move to adjacent nodes via the edges.

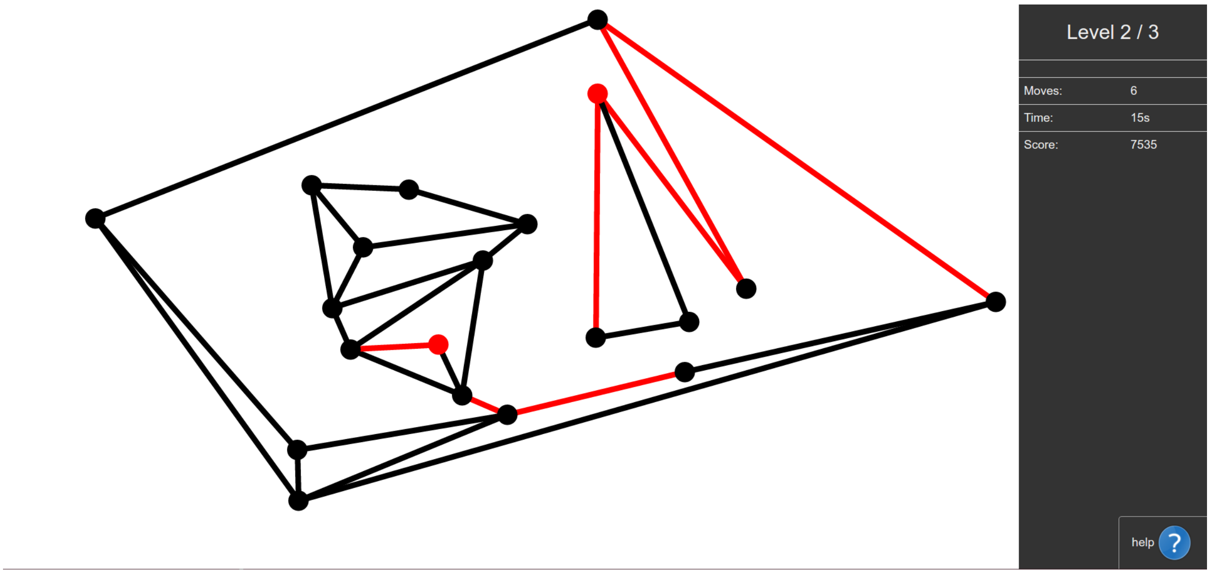

Figure 3.8: Screenshot of a Shannon Switching game

The game is round-based. Players must do exactly one movement each round if it is possible for them to do so. In contrast to the original game, we have no taxi, bus and underground tickets, but endurance. Every move to an adjacent node costs endurance, in fact, the endurance costs are proportional to the length of the edge. The detectives win as soon as one of them stands on the same node as Mister X. Mister X wins if he survives until the detectives have run out of endurance. The detectives start with 400 endurance. Mister X start with 350 endurance, but it recovers by two points each round. In order to give Mister X any chance at all, he is invisible to the detectives most of the time. His current location is only revealed to the detectives every four rounds. If someone's endurance has been depleted, he cannot perform moves anymore.

The player can see the costs of all edges at any time by hovering the edges with the mouse. To move to an adjacent node, he needs to click on that node. Mister X and the detectives cannot move to a node on which another detective is standing on.

The participant always plays *Mister X*. Figure 3.9 shows an example game. The implementation is described in Section 3.6.5.

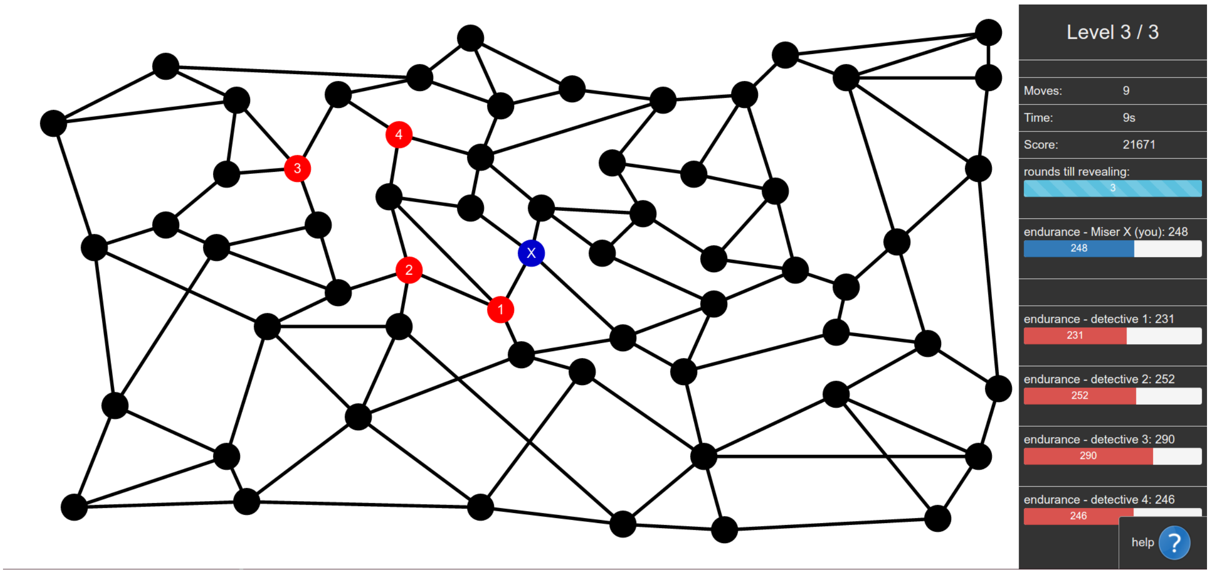

Figure 3.9: Screenshot of a Scotland Graph game

3.5 Participants

In total, there are 72 registered participants who finished at least one game and are differing in sex, age or user agent string. We cannot exclude the possibility that some of them registered multiple times with different information.

The ages reach from 17 to 53 years, with an average of 24.7 years. 79% are male, 20% are female and 1% have chosen "other" as sex. The participants' screen sizes reach from $1024px \times 768px$ to $2880px \times 1800px$, with an average of $1795px \times 1048px$.

3.6 Implementation

In this chapter, we describe technical details, mainly the used algorithms. These algorithms are useful for other purposes that are related to graph-based visualisations or games, too.

The game logic of both games is based on graphs. One game consists of three subsequent levels and is meant as completed only if all three levels are completed. Unlike in most games, the level number is not an indicator of difficulty.

The node size, edge width and difficulty differ at each played level and each player. The exact procedure is explained in Section 3.3.3.

We base our games on undirected graphs without circuit edges and without multi-edges.

3.6.1 Platform Architecture

The three online games are hosted on http://bachelor.dennis-heller.de/. The frontend is written in HTML, CSS and JavaScript, the backend is written in PHP. The database is a MySQL-database on the same host.

When the user starts a game, he receives a “game ID”. The game is then played on the client side and afterwards, the collected data is sent back with the game ID to the server via AJAX.

3.6.2 Calculation of the screen-independent dimensions

Unfortunately, HTML and CSS do not give the developer a possibility to pass size values in metrical or similar units, but only pixel-based units. The units `cm`, `mm`, `in`, `pt` and `pc` that are specified in CSS deceive, they assume that the display has a pixels-per-inch (PPI) value of 96, so it is always $1\,in = 2.54\,cm = 25.4\,mm = 72\,pt = 6\,pc = 96\,px$. While the PPI value of most PC screens is around 96, this is not precise enough for our claims.

To guarantee that the same node sizes and edge sizes will be drawn with the same size on all screens, each user has to enter the display diagonal of the screen he is going to play on. Furthermore, he is requested not to change the device while he is playing.

The application then calculates the PPI value of the user's screen according to this formula:

$$PPI = \frac{screenResolutionHeight\,[px]}{displayDiagonal\,[in]} * \sqrt{\left(\frac{screenResolutionWidth\,[px]}{screenResolutionHeight\,[px]}\right)^2 + 1} \quad (3.3)$$

Each time an object (node or edge) needs to be drawn with a given size in millimetres, the size of this object in pixels is calculated:

$$size\,[px] = size\,[mm] * PPI/25.4 \quad (3.4)$$

This value differs naturally on different screens, but the object's size in millimetres will be the same on all screens.

It is important to notice that a “CSS-pixel” must not necessarily be equal to a “device pixel”. This is only the case if the browser’s zoom level is 100%. A little JavaScript-Function detects the zoom level and asks the user to adjust the value if necessary. The user is not allowed to continue playing until he has done so.

3.6.3 Planarity Shannon Game

The three levels of difficulty are the number of nodes and edges:

(1) 10 nodes with 16 edges
(2) 15 nodes with 26 edges
(3) 20 nodes with 36 edges

The number of nodes was chosen according to subjective feelings of how many nodes are “easy”, “intermediate” and “hard”. The number of edges was chosen according to the human player having a chance to win against the AI (see Section 3.6.4):

Generating a Planar Graph

Not every graph is planar. In Planarity, the computer has to generate a planar graph and then shuffle the nodes, so it is not planar anymore, but the player can remake the graph planar by repositioning the nodes. Algorithm 3.1 creates such a shuffled planar graph, whose nodes are well-distributed over the pitch. It uses a Delaunay Triangulation[3] on a set of randomly distributed vertices to create a planar graph. This graph has more edges than it is supposed to have, so we gradually delete an edge between two nodes with high degree. Afterwards, we distribute the vertices evenly over the pitch. For this purpose we use numbers of the Halton Sequence[4] with a random start index. The problem in using random positions instead would be the danger of two vertices being placed overlapping so that the player cannot clearly distinguish them.

Calculating the Number of Edge Crossings

Each time a node was moved by the player, the computer needs to calculate the number of edge crossings of the graph to find out whether the player has finished the task or not.

[3] https://en.wikipedia.org/wiki/Delaunay_triangulation
[4] https://en.wikipedia.org/wiki/Halton_sequence

Algorithmus 3.1 Generating a planar graph with n nodes and m edges over an area of size $width * height$

Require: n vertices
  procedure GENERATEGRAPH(n,m,$width$,$height$)
    $V \leftarrow$ distribute n vertices randomly over the pitch
    $E \leftarrow$ the *Delaunay triangulation* of V
    $G \leftarrow (E, V)$
    while $|E| > m$ **do** // delete some edges
      $v \leftarrow$ one of the vertices with the highest degree
      $w \leftarrow$ one of the vertices with the highest degree, that is adjacent to v
      delete the edge (v, w)
    end while
    $i \leftarrow 1$
    $hx \leftarrow$ a random number $\in [0, 10]$
    $hy \leftarrow$ a random number $\in [0, 10]$
    for all $a \in V$ **do** // distribute the nodes evenly over the pitch
      $a.x \leftarrow width*$ the $(hx + i)$-th entry of the Halton Sequence with prime base 2
      $a.y \leftarrow height*$ the $(hy + i)$-th entry of the Halton Sequence with prime base 3
      $i \leftarrow i + 1$
    end for
  end procedure

Figure 3.10 shows the mathematical problem of finding an edge crossing. Given two edges in parametric forms, we are searching a solution for t and u, so that Equation (3.5) is true.

$$\vec{p} + t\vec{r} = \vec{q} + u\vec{s} \tag{3.5}$$

By applying the cross product with $\vec{s}$ on both sides

$$(\vec{p} + t\vec{r}) \times \vec{s} = (\vec{q} + u\vec{s}) \times \vec{s}$$

and taking into account that the cross product of a vector with itself is zero:

$$\vec{p} \times \vec{s} + t(\vec{r} \times \vec{s}) = \vec{q} \times \vec{s}$$

Solving for t leads to

$$t = (\vec{q} - \vec{p}) \times \vec{s}/(\vec{r} \times \vec{s})$$

Similar by applying the cross product with $\vec{r}$ we find u:

$$u = (\vec{q} - \vec{p}) \times \vec{r}/(\vec{r} \times \vec{s})$$

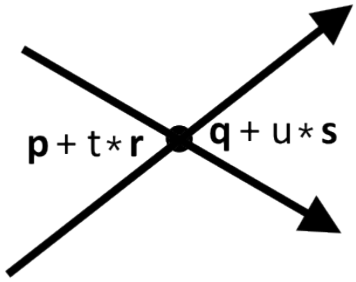

Figure 3.10: Mathematical background in the calculation of a crossing of two edges.

The two given edges are crossing if and only if t and u are defined (which is the case if $\vec{r} \times \vec{s}$ is not zero) and if t and u are between 0 and 1.

This procedure needs to be executed for each pairwise combination of edges. It would be sufficient to iterate over these combinations and stop if a crossing was found. To show the remaining number of crossing to the participants, we additionally calculate all further edge crossings. Furthermore, we tracked the number of moves in which the number of edge crossings got lower, that is the number of "successful" moves.

The Score

Given the needed time t (in minutes), the number of moves m and the number of edge crossings in the initial graph k, the score of the player is calculated as shown in Equation (3.6).

$$score(t, m, k) = 10000 * 0.5^{t+m/k^2} \tag{3.6}$$

We used exponential decrease and a high starting value of 10000 to avoid draws in the top five of the highscore table and not to demotivate players, who are unsuccessful with a too low score. We wanted to honour the used time independent of the other parameters, that is why we take the sum of t and the other factors. Between the number of moves and the number of edge crossings at the beginning, we approximated a quadratic dependency, e.g. the minimum number of moves is approximately quadratic to the number of edge crossings at the beginning. Actually, the problem of finding the minimum number of moves is NP-hard and hard to approximate for a given graph [GKO+09].

3.6.4 Shannon Switching Game

The three levels of difficulty are the same as in the Planarity game because this game is played on the resulting graph of the preceded Planarity game.

Difficulty and the Number of Edges: an Excursion into Game Theory

According to [Leh64], for Shannon Switching Games exactly one of the following properties is true:

- The cut player plays first and the short player can win against all possible strategies of the cut player. (“positive game”)
- The short player plays first and the cut player can win against all possible strategies of the short player. (“negative game”)
- The player who plays first, regardless of identity, can win against all possible strategies of the other player. (“neutral game”)

In our case, the human player always is the cut player and the computer (the short player) always begins. So we need “negative” games. To decide whether a game is negative, we first need some definitions:

Definition 3.6.1 (Spanning Tree)
[Woo12] A spanning tree in a graph G is a selection of edges in G where there are no cycles and the edges are all connected. We say two spanning trees are disjoint when they share no common edges.

Definition 3.6.2 (Near Spanning Tree)
[Tan09] If one edge from a spanning tree on the path from the one to the other distinguished vertex is removed, the result is a near spanning tree.

Tantalo [Tan09] established the following theorem:

- A game is positive if the edge set of its graph can be partitioned into two spanning trees.
- A game is negative if the edge set of its graph can be partitioned into two near spanning trees.
- A game is neutral if the edge set of its graph can be partitioned into one spanning tree and one near spanning tree.

Assuming $|E|$ the number of edges and $|V|$ the number of vertices in a game, the preceding rules can be simplified to: [Tan09]

- A game is positive, if $|E| > 2 * |V| - 3$
- A game is negative, if $|E| < 2 * |V| - 3$
- A game is neutral, if $|E| = 2 * |V| - 3$

As our graphs are not very difficult, we chose them all to have exactly $2 * |V| - 4$ edges. If the AI would be perfect, the participants could win only if they played perfectly, too. But since the AI is not perfect, the participants are allowed to do some wrong moves.

The AI of the Shannon Game opponent

We designed and implemented an algorithm for the AI of the *short* player of the Shannon Switching Game. Algorithm 3.2 shows the simplified algorithm in pseudo-code. It iterates over all paths between the two distinguished vertices and counts the occurrences of each edge in the shortest paths. Here, we consider only paths with minimum length or one above. This is because we found out that otherwise, especially in large graphs, the algorithm would prefer edges in dense areas of the graph (where many paths go through) instead of crucial ones. Furthermore, we weight each occurrence, so paths who are shorter in total have a slightly higher impact on the selection of the chosen edge.

The Score

Given the needed time t (in minutes), the number of moves m and the number of edges in the initial graph k, the score of the player is calculated as shown in Equation (3.7).

$$score(t, m, k) = \begin{cases} 10000 * 0.5^{t+m/k} & \text{if player has won} \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

We used exponential decrease with the same reasoning like in Planarity. Here we approximated that the minimum number of moves is proportional to the number of edges, so we used their ratio in the formula.

Algorithmus 3.2 Determining the edge to choose as *short* player in the Shannon Switching Game

Require: Graph G(E, V), two distinguished vertices a and b
  procedure SHANNONAI(G, a, b)
    $P \leftarrow$ list of all paths from a to b without circuits
    for all $p \in P$ **do**
      $p.distance \leftarrow$ length (number of edges) of that path, where marked edges count as zero
    end for
    min = minimum of all path distances
    $P' \leftarrow \{p \in P \mid p$ has distance min or $min + 1\}$
    for all $p \in P'$ **do**
      for all edge $e \in p$ **do**
        if e is not marked **then**
          edgeWeight(e) $\leftarrow$ edgeWeight(e) + $1/(p.distance)^2$
        end if
      end for
    end for
    return the edge $e \in E$ with the heighest edgeWeight(e)
  end procedure

3.6.5 Scotland Graph

The three levels of difficulty are the number of detectives: three, four or five.

The AI

The AI for the detectives of the Scotland Graph game is very complex because there are more factors to consider: The many possible locations of Mister X, the preferences of Mister X to choose nodes with specific properties (e.g. far away from the detectives and with many connected edges) and to avoid multiple detectives moving to the same node, which is illegal. The ideas from Nijssen and Winands [NW12] inspired us to write the following algorithm:

Algorithmus 3.3 Performing the moves for the detectives in the Scotland Graph Game.

Require: Graph $G(E, V)$, possible positions of Mister X before hist last turn $P \subseteq V$
  procedure SCOTLANDGRAPHAI(G, P)
    if Mister X has just revealed his position **then**
      $P' = \{$ current position of Mister X $\}$

```
else
    P' ← ∅
    for all p ∈ P do
        P_local ← ∅
        for all adjacent nodes q of p do
            if no detective stands on q then
                if cost(p, q) ≤ endurance of Mister X before his last turn then
                    P' ← P' ∪ {q}
                end if
            end if
        end for
        if P_local = ∅ then
            // Mr. X could not have moved his last turn, if he would have been at p
            P' ← P' ∪ {p}
        else
            P' ← P' ∪ P_local
        end if
    end for
end if
P ← P'
for all p ∈ P do
    degree ← degree of node p
                            // distances are measured as 1 + cost(edge)/100 per edge
    minDistance ← minimum distance from p to the next detective
    avgDistance ← average distance from p to all detectives
    nodeRating(p) ← degree * minDistance^3 * avgDistance
                                            // attractiveness of that node for Mister X
end for
for all detectives D do
    for all adjacent nodes q of D's current position d do
        if (cost(d, q) ≤ endurance of D) then
            for all p ∈ P do
                weightedDist2MX(q, p) ← (distance from q to p)/nodeRating(p)
            end for
            targetsRating(D, q) ← average of weightedDist2MX(q, x) for all x
            possibleMovesCount(D) ← possibleMovesCount(D) + 1
        end if
    end for
end for
for all detectives D in ascending order of their possibleMovesCount(D) do
    repeat
```

$t \leftarrow$ targetNode with the highest $targetsRating(D, t)$
delete $targetsRating(D, t)$
until no other detective is standing on t
$moveDetective(D, t)$
end for
end procedure

First, it calculates all possible positions of Mister X. Then, it rates them based on node degree and distance to the detectives. The higher this rating, the more likely Mister X is standing on that node. Afterwards, each detective looks for all nodes on which he could move and calculates the average weighted distance to all possible locations of Mister X. Finally, each detective moves to his adjacent node with the lowest weighted distance to Mister X, if another detective has not already moved there.

A drawback of this algorithm is that the detectives are likely to clump together because they all prefer the same node to move to. But the graph is very small anyway, so it is hard enough for the human players.

The Score

Given the needed time t (in minutes), the remaining endurance at the end e and the number of detectives d, the score of the player is calculated as shown in Equation (3.8).

$$score(t, d, e) = \begin{cases} 10000 * \frac{d}{3} * 0.5^{t/2} * (1 + \frac{e}{350}) & \text{if player has won} \\ 1000 * \frac{d}{3} * 0.5^{t/2} & \text{otherwise} \end{cases} \quad (3.8)$$

We used exponential decrease with the same reasoning like in Planarity and Shannon Switching Game. The number of detectives seemed to be an important indicator of the difficulty, so we weighted this factor with a starting value between 10000 at three detectives and 16667 at five detectives. In case of a won game, we also wanted to honour players who played very economical by multiplying the resulting score with a factor between 1 and 2 (the proportion of the beginning endurance of 350). In practice, this had no major impact since it was very hard to have more than 50 endurance at the end.

4 Results

Dependent variables (DVs) and IVs are written *cursive*. Whenever it is written *ratio* without any further specification, we mean the *ratio of node size and edge width*.

As mentioned in Section 3.6, one played game consists of three levels and we have 81 different conditions. Table 4.1 shows an overall statistic about all played games. The largest number of finished games by a single user was 152.

Before analysing the data, we deleted 21 data records because they were obviously hacked data in malicious purpose. It would have been too much effort to implement constant server-side validation. Instead, the server script only did some rudimentary tests on the submitted data, e.g. whether the number of nodes or detectives fits the difficulty level. Yet, our recorded data is sufficient to discern computer-generated games from actual played participants. Furthermore, we removed four data records because they contained extreme outliers.

We tested the data for correlations between *node size*, *edge size* (and their *ratio*) and *difficulty* and the dependent variables from Section 3.3.2 using an repeated-measures Analysis of Variance (ANOVA). For variables with an (ordinal) Likert-scale such as the questionnaire values, it is controversial whether an ANOVA is the right choice or other tests like Kruskal-Wallis are more suitable. Simon [Sim08] says that the difference is negligible in practice.

	Finished games	Played levels	Avg. data records per condition
Planarity Shannon Game	338	1014	12.5
Scotland Graph	254	762	9.4
Both Games	592	1776	21.9

Table 4.1: Overall statistic about all played games.

4.1 Questionnaires

As described in Section 3.4.2, there were five possible answers from a Likert Scale, which we mapped on whole numbers from -2 to 2. In the scales about the players' opinion on *node size* and *edge width*, the values reached from -2 (too small) over 0 (nice) to 2 (too big), therefore we searched a minimum of the absolute values. In the scale about the opinion on the *ratio of node size and edge width*, the values reached from -2 (very bad) to 2 (very nice), therefore we searched a maximum of the values.

Figure 4.1 shows a colored table containing the means of the DV *node size questionnaire*. As expected, there is a gradient of colour from left to right, switching between red over green to red. This means that the aesthetic optimum of the *node size* indeed is somewhere in the middle of our range. There is no gradient from top to bottom, which means that the *edge width* did not have a major impact on this questionnaire variable, which is to be expected because we asked for the node size.

Figure 4.2 shows a colored table containing the means of the DV *edge width questionnaire*. As expected, there is a gradient of colour from top to bottom, switching between red over green to red. This means that the aesthetic optimum of the *edge width* indeed is somewhere in the middle of our range. There is no gradient from left to right, which means that the *node size* did not have a major impact on this questionnaire variable, which is to be expected because we asked for the edge width.

Figure 4.3 shows a colored table containing the means of the DV *ratio questionnaire*. The equivalence line of the *ratio of node size and edge width* goes nearly diagonal from the top left corner to the bottom right corner. As expected, there is a blurred green spot in the centre and the bottom left and the top right corner tend to be red. This means that the aesthetic optimum of the *ratio* indeed is somewhere in the middle of our range.

Difficulty: all

	6	7	8	9	10	12
0.5			0	0.2		
1		-0.2	0.1	0.1	0.3	
1.5	-0.6	-0.4	-0.1	0.1	0.3	0.7
2	-0.8	-0.6	-0.4	0	0.1	0.7
2.5		-0.7	-0.3	-0.2	0.1	0.5
3			-0.4	-0.3	0.1	
4				-0.5		

Difficulty: 1

	6	7	8	9	10	12
0.5			0.1	-0.1		
1		-0.1	0.2	0	0.5	
1.5	-0.5	-0.4	-0.1	0.1	0.1	0.7
2	-0.7	-0.8	-0.5	-0.1	0.1	0.6
2.5		-0.6	-0.2	-0.1	0.1	0.6
3			-0.6	-0.3	0.3	
4				-0.5		

Difficulty: 2

	6	7	8	9	10	12
0.5			0.1	0.1		
1		-0.4	0	0.1	0.3	
1.5	-0.6	-0.4	0	0.1	0.3	0.7
2	-0.8	-0.2	-0.4	-0.2	0.1	0.8
2.5		-0.7	-0.2	-0.4	0.1	0.5
3			-0.3	-0.3	0	
4				-0.5		

Difficulty: 3

	6	7	8	9	10	12
0.5			-0.1	0.4		
1		-0.4	0	0.1	0.2	
1.5	-0.6	-0.4	-0.1	0	0.4	0.8
2	-0.9	-0.6	-0.3	0.2	0	0.8
2.5		-0.7	-0.4	-0.1	0.1	0.5
3			-0.2	-0.3	-0.1	
4				-0.4		

Figure 4.1: Colored tables of the means of the *node size questionnaire* (DV), categorized by difficulty (IV). Rows mean *edge sizes* (IV) and columns mean *node sizes* (IV). The color scale is linear from red (too small or too large) to green (good).

Difficulty: all

	6	7	8	9	10	12
0.5			-0.7	-1		
1		-0.4	-0.3	-0.5	-0.6	
1.5	-0.1	-0.1	0	-0.1	0	-0.4
2	0.2	0.3	0.2	0.2	0.1	0.1
2.5		0.6	0.7	0.5	0.6	0.6
3			0.9	0.8	0.8	
4				1.4		

Difficulty: 1

	6	7	8	9	10	12
0.5			-0.7	-1.1		
1		-0.4	-0.3	-0.4	-0.6	
1.5	-0.5	0	0	0	0	-0.5
2	0.1	0.4	0.3	0	0.1	0.2
2.5		0.7	0.4	0.2	0.4	0.7
3			0.8	1	0.6	
4				1.3		

Difficulty: 2

	6	7	8	9	10	12
0.5			-0.6	-0.7		
1		-0.4	-0.2	-0.6	-0.5	
1.5	0.1	-0.2	-0.1	0	0	-0.3
2	0.2	0.3	0.1	0.3	0.3	0.1
2.5		0.5	0.8	0.6	0.5	0.5
3			0.9	0.8	0.8	
4				1.4		

Difficulty: 3

	6	7	8	9	10	12
0.5			-0.7	-1.1		
1		-0.5	-0.4	-0.4	-0.6	
1.5	-0.1	0	0	-0.2	-0.1	-0.5
2	0.3	0.2	0.1	0.3	0.1	0.1
2.5		0.6	0.9	0.7	0.8	0.7
3			1	0.7	0.9	
4				1.4		

Figure 4.2: Colored tables of the means of the *edge width questionnaire* (DV), categorized by difficulty (IV). Rows mean *edge sizes* (IV) and columns mean *node sizes* (IV). The color scale is linear from red (too thin or too thick) to green (good).

Difficulty: all

	6	7	8	9	10	12
0.5			0.1	0.1		
1		0.4	0.5	0.4	0.3	
1.5	0.1	0.4	0.7	0.8	0.6	-0.1
2	-0.1	-0.1	0.2	0.4	0.5	0.3
2.5		-0.2	0	0.2	0.4	0
3			-0.4	-0.1	0	
4				-1		

Difficulty: 1

	6	7	8	9	10	12
0.5			0.3	0		
1		0.4	0.6	0.4	0.2	
1.5	0.2	0.5	1	0.8	0.6	0
2	0	-0.3	0.1	0.7	0.7	0.5
2.5		-0.2	0.4	0.6	0.5	0.1
3			-0.4	-0.4	0.1	
4				-1.1		

Difficulty: 2

	6	7	8	9	10	12
0.5			0.3	0.5		
1		0.3	0.5	0.4	0.5	
1.5	0	0.7	0.6	0.9	0.4	-0.1
2	-0.2	0.4	0.2	0.5	0.2	0.4
2.5		-0.2	-0.1	0.1	0.2	0.4
3			0	-0.2	-0.2	
4				-0.8		

Difficulty: 3

	6	7	8	9	10	12
0.5			-0.2	-0.2		
1		0.5	0.5	0.3	0	
1.5	0.2	0.1	0.5	0.6	0.8	0
2	-0.3	-0.4	0.2	0	0.6	0
2.5		-0.2	-0.2	-0.1	0.3	-0.3
3			-0.7	0.3	0.1	
4				-1		

Figure 4.3: Colored tables of the means of the *ratio questionnaire* (DV), categorized by difficulty (IV). Rows mean *edge sizes* (IV) and columns mean *node sizes* (IV). The color scale is linear from red (aesthetically disappealing) to green (aesthetically appealing).

4.1.1 ANOVAs

Table 4.2 shows the results of the ANOVAs of the questionnaire data, based on all four independent variables (node size, edge width, ratio of node size and edge width, difficulty). Each DV was significantly influenced by the IVs. The *difficulty* on its own, however, had no significant effect on the questionnaire DVs, except on the *ratio questionnaire* in the Planarity Game ($p < 0.02$).

Planarity & Shannon Game			
Questionnare (DV)	F	p	R^2
Node size	$F(4, 1005) = 77.145$	< 0.001	0.235
Edge width	$F(4, 1005) = 180.844$	< 0.001	0.419
Ratio	$F(4, 1005) = 22.021$	< 0.001	0.081
Scotland Graph			
Questionnare (DV)	F	p	R^2
Node size	$F(4, 757) = 99.800$	< 0.001	0.345
Edge width	$F(4, 757) = 125.893$	< 0.001	0.399
Ratio	$F(4, 757) = 26.700$	< 0.001	0.124

Table 4.2: Results of the ANOVAs of the questionnaire data. The values are related on all independent variables.

4.1.2 Regression

Figure 4.4 shows a regression of the mean absolute *node size questionnaire* values over the *node size*. We used a quadratic regression function because we suggested an optimum (minimum) to exist somewhere in the centre of our range of node sizes, which was confirmed by the plotted data points. The regression resulted in an optimum at $9.07\,mm$ for the *node size*.

Figure 4.5 shows a regression of the mean absolute *edge width questionnaire* values over the *edge width*. We used a quadratic regression function because we suggested an optimum (minimum) to exist somewhere in the centre of our range of node sizes, which was confirmed by the plotted data points. The regression resulted in an optimum at $1.69\,mm$ for the *edge width*.

Figure 4.6 shows a regression of the mean absolute *ratio questionnaire* values over the *ratio of node size and edge width*. By an ANOVA, we found a significant effect from the latter *ratio* on this variable ($p < 0.006$). The regression resulted in an optimum at 5.97 for the *ratio*.

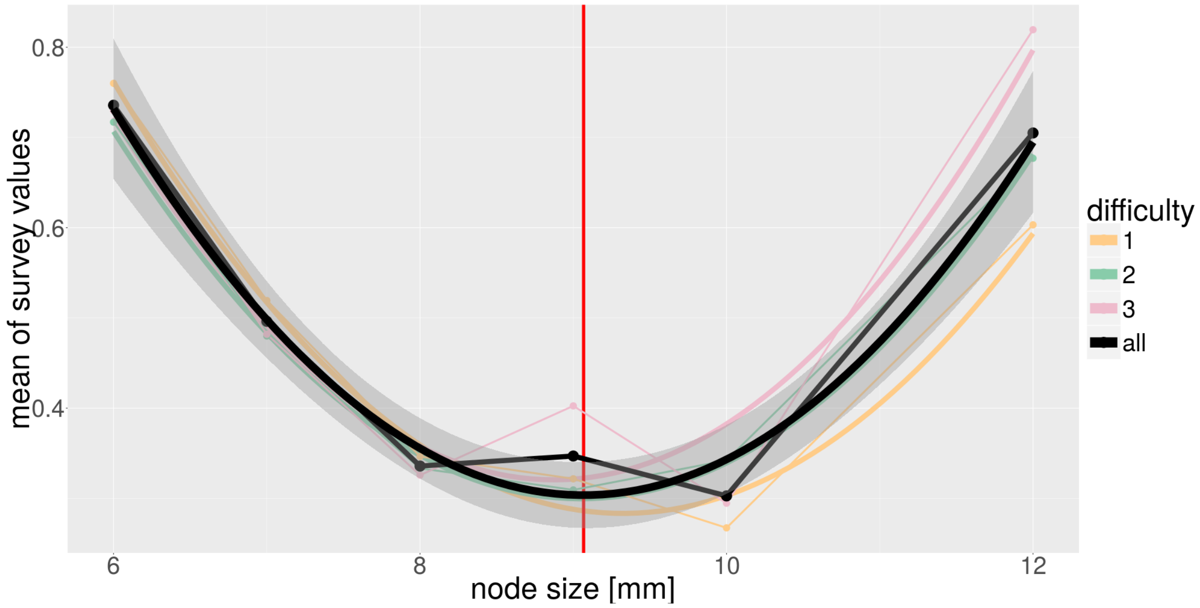

Figure 4.4: Regression of the means of the absolute values of the *node size questionnaire* (DV) over the *node size* (IV). The blue curve is a quadratic regression function ($R^2 = 0.9789$). The grey shadow shows its 95% confidence band. The red vertical line marks its optimum at *node size*$= 9.07\,mm$.

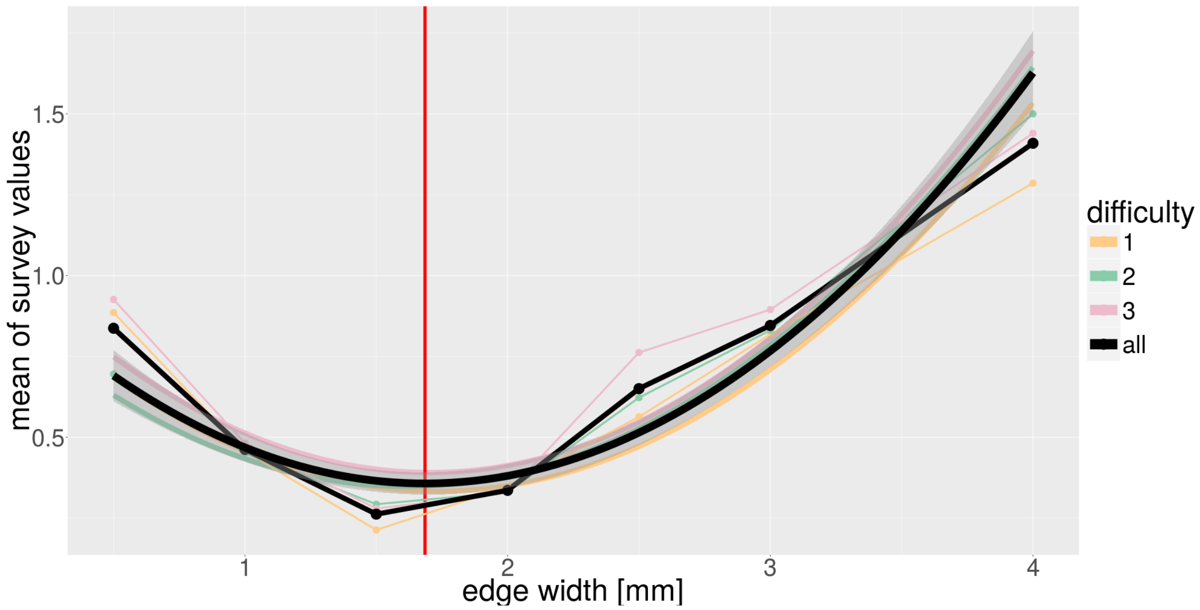

Figure 4.5: Regression of the means of the absolute values of the *edge width questionnaire* (DV) over the *edge width* (IV). The blue curve is a quadratic regression function ($R^2 = 0.9161$). The grey shadow shows its 95% confidence band. The red vertical line marks its optimum at *edge width size*$= 1.69\,mm$.

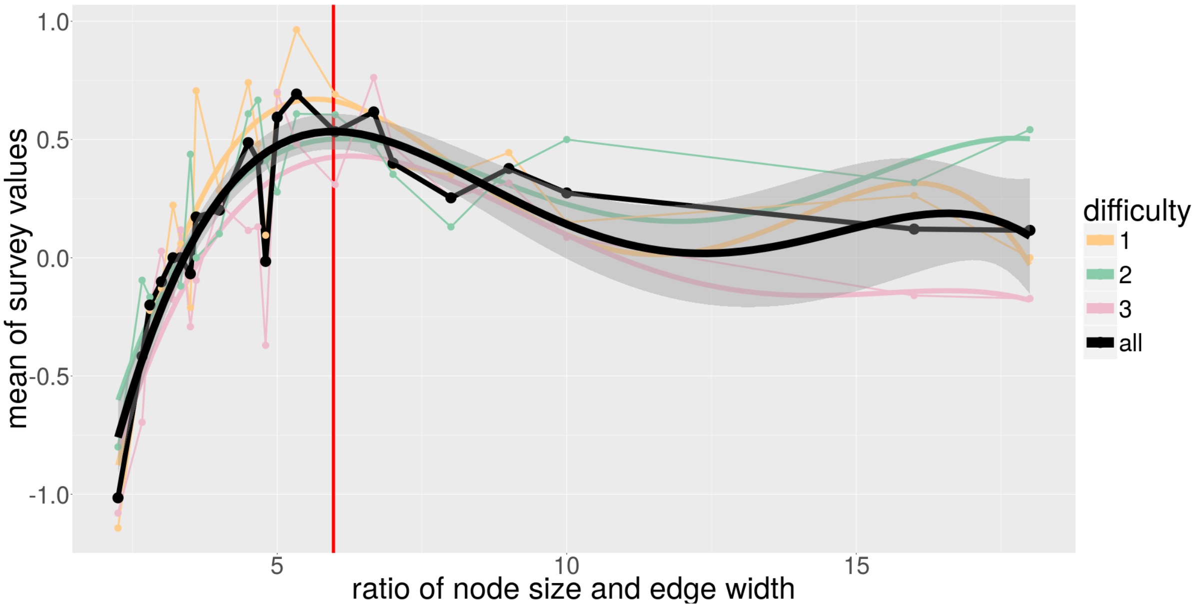

Figure 4.6: Regression of the means of the absolute values of the *ratio questionnaire* (DV) over the *ratio of node size and edge width* (IV). The blue curve is a polynomial regression function of degree 4 ($R^2 = 0.8532$). The grey shadow shows its 95% confidence band. The red vertical line marks its optimum at *ratio*$= 5.97$.

Besides the comparisons of each questionnaire with its corresponding variable, there were significant effects in some other combinations, too. Table 4.3 shows the results all regressions.

IV [mm] / DV [mm]	node size	edge size	ratio of node size and edge width
node size questionnaire	$9.07\,mm$	—	(6.40)
edge width questionnaire	—	$1.69\,mm$	(5.72)
ratio questionnaire	$(9.41\,mm)$	$(1.32\,mm)$	5.97

Table 4.3: The optima of the regressions on the data of the questionnaires.

4.2 Planarity

From our IVs, only the *difficulty* had a significant impact on the DVs, in fact on all of them. Each other IV showed no significant effect on the DVs. Solely, the *node size* had a small impact on the *proportion of useful moves* but this correlation was not significant ($p < 0.071$).

4.2.1 ANOVAs

A multiple linear regression was calculated to predict the *score* based on *nodesize*, *edge size*, *ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 277.320$, $p < 0.001$), with an R^2 of 0.525. The *difficulty* was the only significant predictor of the *score*. Figure 4.7 shows a boxplot of the correlation between the *difficulty* and the *score*.

A multiple linear regression was calculated to predict the *needed time* based on *nodesize*, *edge size*, *ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 121.176$, $p < 0.001$), with an R^2 of 0.325. The *difficulty* was the only significant predictor of the *score*.

A multiple linear regression was calculated to predict the *number of moves* based on *nodesize*, *edge size*, *ratio of node size and edge size* and *difficulty*. A significant regression

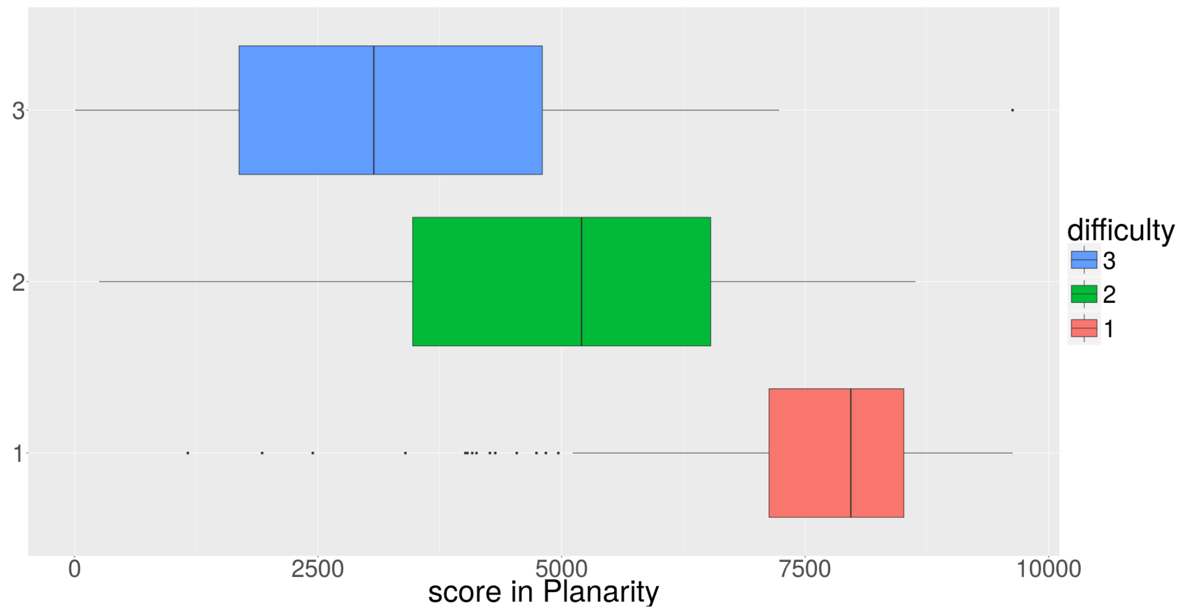

Figure 4.7: Boxplot of the *score* (DV) over the *difficulty* (IV) from the Planarity Game.

equation was found ($F(4, 1005) = 170.629$, $p < 0.001$), with an R^2 of 0.404. The *difficulty* was the only significant predictor of the *number of moves*.

A multiple linear regression was calculated to predict the *proportion of useful moves* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 46.923$, $p < 0.001$), with an R^2 of 0.157. The *difficulty* was the only significant predictor of the *proportion of useful moves*.

The *node size* and the *edge size* had no significant effect on all DVs.

4.3 Shannon Switching Game

Like in Planarity, the *difficulty* had a significant effect on all DVs. But in contrast to Planarity, the *ratio of node size and edge width* correlated with two DVs.

4.3.1 ANOVAs

A multiple linear regression was calculated to predict *score* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 12.061$, $p < 0.001$), with an R^2 of 0.046. The *difficulty* and the *ratio of node size and edge size* were significant predictors of the *shannon score*.

A multiple linear regression was calculated to predict *game result* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 10.233$, $p < 0.001$), with an R^2 of 0.039. The *difficulty* and the *ratio of node size and edge size* were significant predictors of the *game result*. Figure 4.8 shows a stacked bar chart of the correlation between the *difficulty* and the *game result*.

A multiple linear regression was calculated to predict *needed time* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 33.041$, $p < 0.001$), with an R^2 of 0.116 Only the *difficulty* was a significant predictor of the *needed time*. Figure 4.9 shows a boxplot of the correlation between the *difficulty* and the *needed time*.

A multiple linear regression was calculated to predict *number of moves* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 1005) = 90.916$, $p < 0.001$), with an R^2 of 0.266 Only the *difficulty* was a significant predictor of the *number of moves*.

The *node size* and the *edge size* had no significant effect on all DVs.

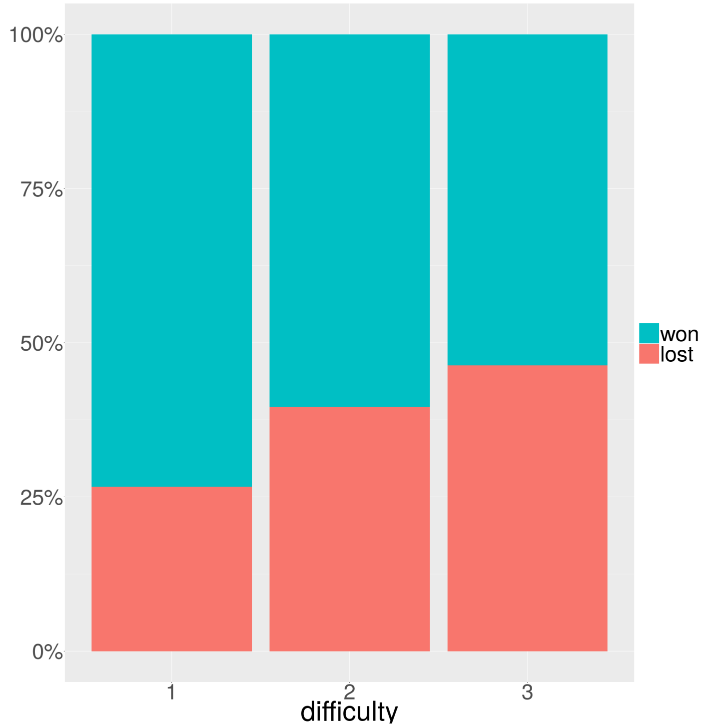

Figure 4.8: Stacked bar chart of the *game result* (DV) over the *difficulty* (IV) from the Shannon Switching Game.

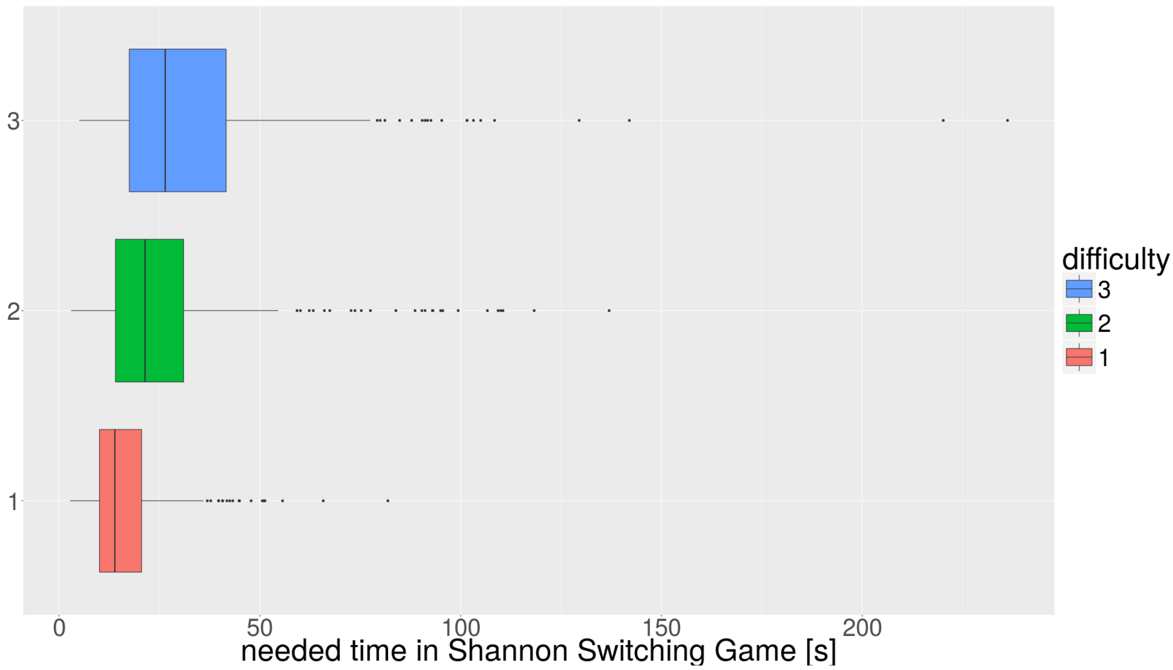

Figure 4.9: Boxplot of the *needed time* (DV) over the *difficulty* (IV) from the Shannon Switching Game.

4.3.2 Regressions

We performed regressions for the effects of the *ratio of node size and edge width* on the *game result* and on the *score*. Figure 4.10 and Figure 4.11 show the regression results of the means of these variables over the *ratio of node size and edge width*. The regression resulted in an optimum for the *ratio* at 5.91 from the *game result* and at 4.22 from the *score*. The global maxima of both data sets, however, are at a *ratio* of 18 at the right end of the scale.

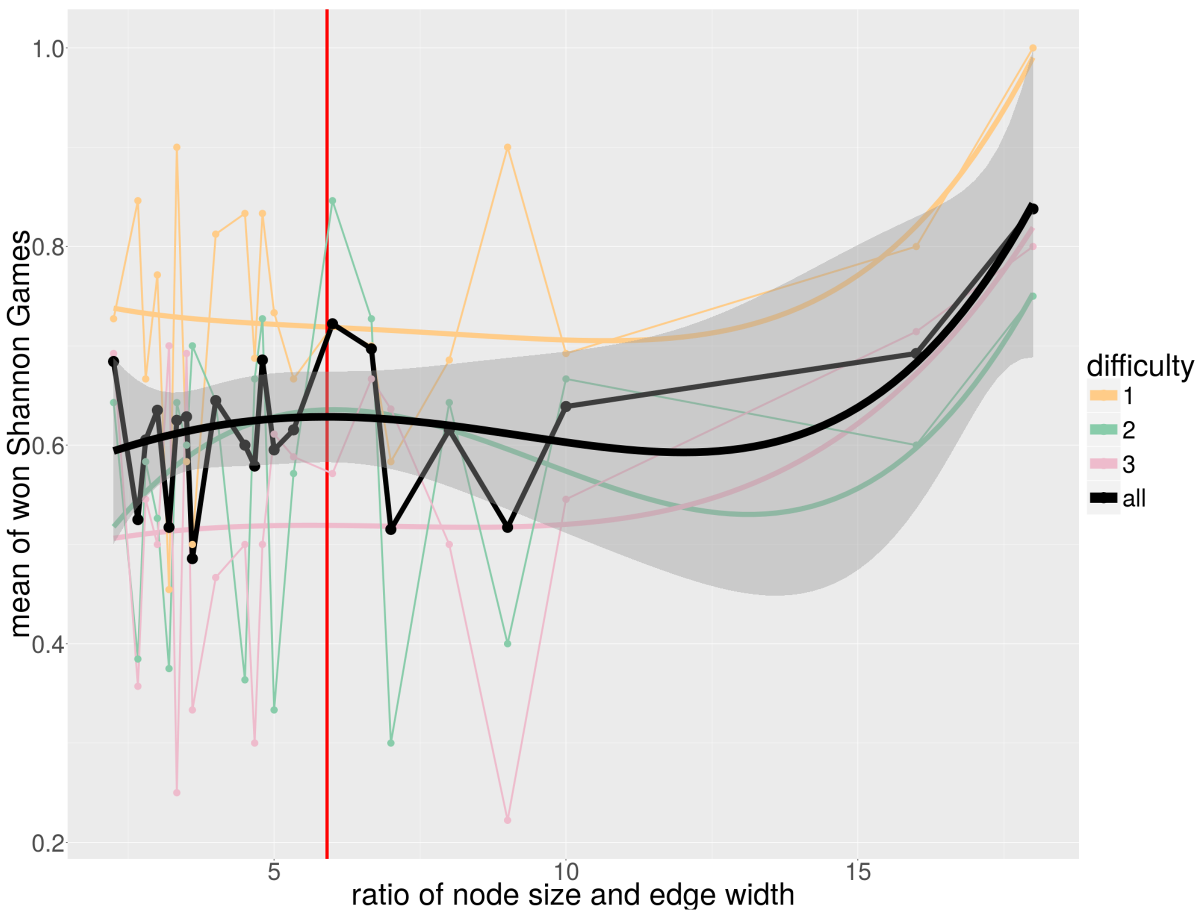

Figure 4.10: Regression of the *game result* (DV) over the *ratio of node size and edge width* (IV) from the Shannon Switching Game. A won game was counted as 1, a lost one was counted as 0. The blue curve is a polynomial regression function of degree 4 ($R^2 = 0.4116$). The grey shadow shows its 95% confidence band. The red vertical line marks its local optimum at *ratio*$= 5.91$.

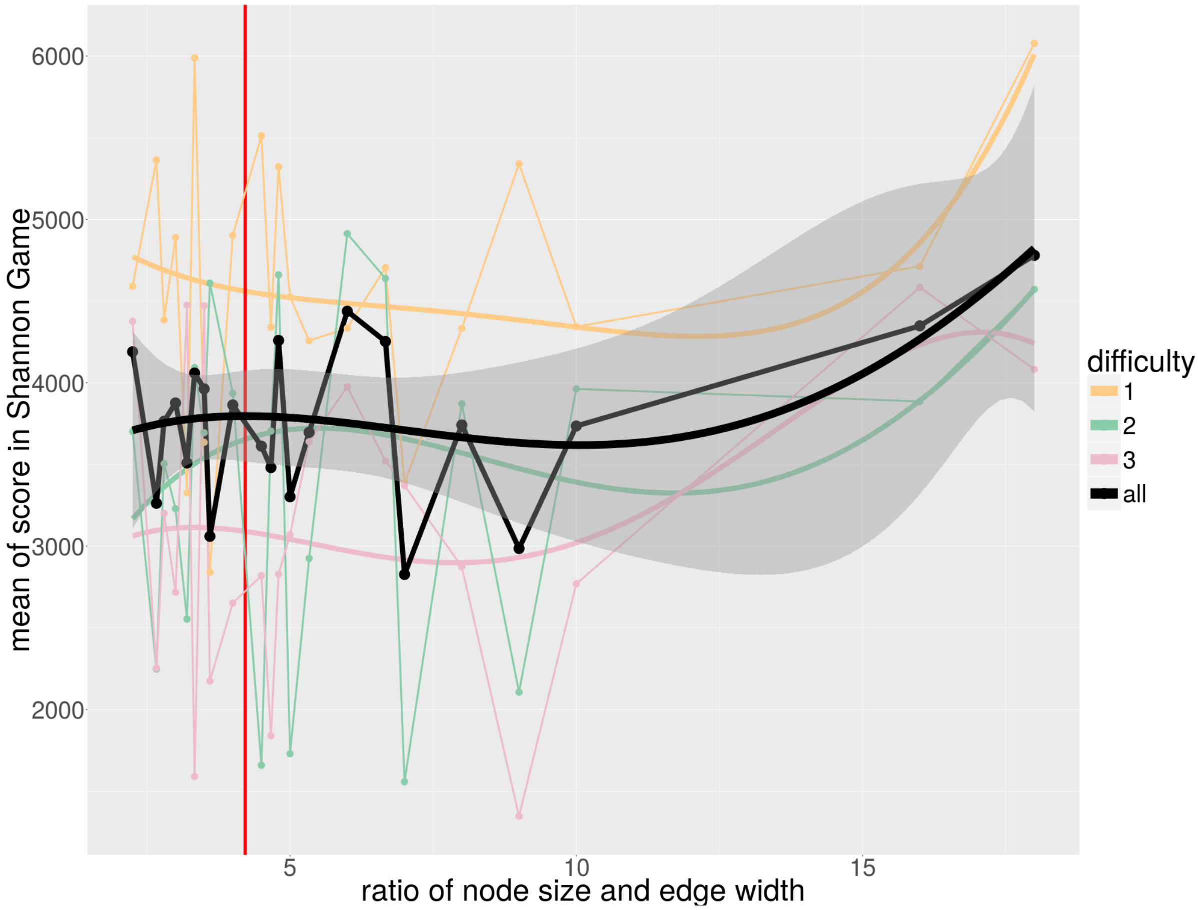

Figure 4.11: Regression of the *score* (DV) over the *ratio of node size and edge width* (IV) from the Shannon Switching Game. The blue curve is a polynomial regression function of degree 4 ($R^2 = 0.3017$). The grey shadow shows its 95% confidence band. The red vertical line marks its local optimum at *ratio*$= 4.22$.

4.4 Scotland Graph

Like in Planarity, only the *difficulty* had a significant effect on all dependent variables, except, in this case, on the *score*.

4.4.1 ANOVAs

A multiple linear regression was calculated to predict the *score* based on *nodesize*, *edge size*, *ratio of node size and edge size* and *difficulty*. No significant regression equation was found ($p < 0.773$), none of the IVs was a significant predictor of the *score*.

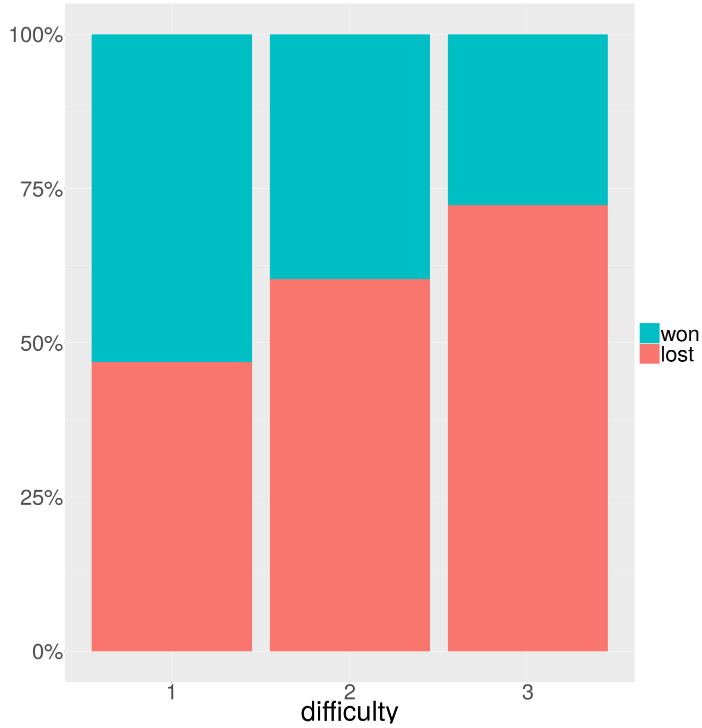

Figure 4.12: Stacked bar chart of the DV *game result* over the IV *difficulty* from the Scotland Graph Game.

A multiple linear regression was calculated to predict the *game result* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 757) = 9.386$, $p < 0.001$), with an R^2 of 0.047. Only the *difficulty* was a significant predictor of the *game result*. Figure 4.12 shows a stacked bar chart of the correlation between the *difficulty* and the *game result*.

A multiple linear regression was calculated to predict the *needed time* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 757) = 4.828$, $p < 0.001$), with an R^2 of 0.025. Only the *difficulty* was a significant predictor of the *needed time*. Figure 4.13 shows a boxplot of the correlation between the *difficulty* and the *needed time*.

A multiple linear regression was calculated to predict the *number of moves* based on *nodesize, edge size, ratio of node size and edge size* and *difficulty*. A significant regression equation was found ($F(4, 757) = 7.579$, $p < 0.001$), with an R^2 of 0.039. Only the *difficulty* was a significant predictor of the *number of moves*.

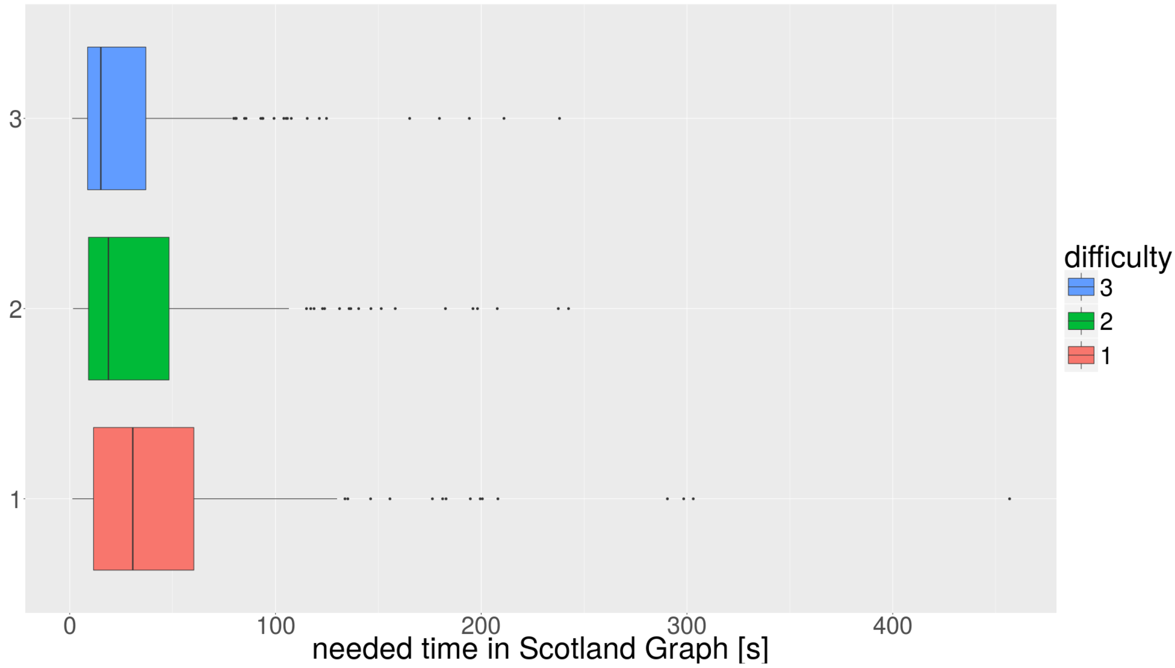

Figure 4.13: Boxplot of the DV *needed time* over the IV *difficulty* from the Scotland Graph Game.

5 Discussion

5.1 Questionnaires

The questionnaire data showed clear correlations between *node size*, *edge width*, their *ratio* and the subjective aesthetic appeal. Our quadratic regressions for the *node size questionnaire* and *edge width questionnaire* fitted very well (see their R^2 values in the plots' captions).

Through polynomial regressions, we could estimate the optimal values of the former three independent variables. Based on the regression results in Table 4.3, we suggest the following values for graph drawing algorithms that aim to produce aesthetically appealing visualisations:

- node size: $\sim 9.1\,mm$
- edge width: $\sim 1.7\,mm$
- ratio of node size and edge width: ~ 6

These results fit well with the result of our Pilot Study, where the centre of the red area in Figure 3.2 is nearly at *node size*=8.5 and *edge width*=2. Figure 5.1 shows the four combinations of *node size* and *edge width* that are located around the optima. If we take the ratio of the above node size and edge width, which is $\frac{9.1\,mm}{1.7\,mm} \approx 5.35$, we notice a slight deviation of 11% compared to the optimal ratio (~ 6). It seems that the optima for *node size* and *edge width* do not fit perfectly with the optimum for the *ratio of node size and edge width*. This is possible because the optima for *node size* and *edge width* are absolute values, which means that their optimum can vary for different use cases or other factors, whereas the *ratio* is a relative value. In Section 5.6, we will further discuss this problem.

For the questionnaire about the *ratio of node size and edge width*, it is worth mentioning that the participants voted a very low ratio worse than a very high ratio. This can be seen from the very steep beginning of the curve at low values for the *ratio* in the corresponding diagram (Figure 4.6) in contrast to the flat ending for high *ratio*-values.It

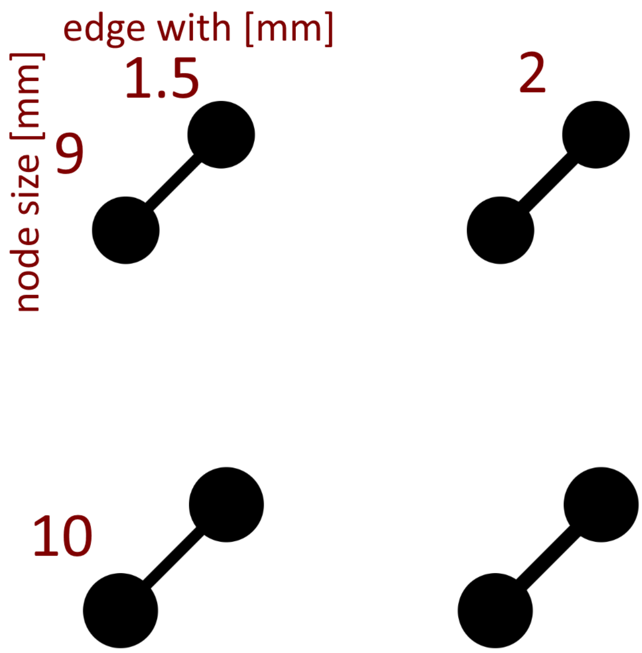

Figure 5.1: Ranges of the most aesthetically appealing node sizes and edge widths according to the questionnaire data.

seems that ratios below our suggested value are repulsive, whereas higher ratios are not so bad at all but still worse than the optimum.

5.2 Planarity

It is clearly visible that players got a lower score at increasing *difficulty*, although we counterbalanced the difficulty in the score formula. Remember, the *difficulty* was expressed as the number of nodes and edges. The larger the graph, the higher was the expected value of edge crossings at the beginning of the game and therefore the more work it was for the player to remove all these edge crossings. This explains the impact of the *difficulty* on the *needed time*. Because the *score* was mainly dependent on the *needed time*, the impact of the *difficulty* on the *score* implicated from the precedent explanation.

Contrary to our expectations, the *node size*, *edge width* and their *ratio* have no significant effect on all DVs. So these graph parameters seem to have no effect on human task performance of graph-related tasks. The players did not even need more time to hit smaller nodes with the mouse cursor than bigger ones.

5.3 Shannon Switching Game

The impact of the *difficulty* on the *game result* of the Shannon Switching Games does not come unexpectedly. The larger the graph the more difficult it is for the player to keep an overview of the possible paths the AI could take. So it is more likely in large graphs to make wrong turns and, theoretically, one wrong turn is sufficient to lose the game. Since the *score* depends on the game result (a lost game results in a score of zero), this has also an impact on the *score*.

The *needed time* increased with the *difficulty*, too, because it takes longer to connect or disconnect the two distinguished nodes when there are more edges.

Like in Planarity, the *node size* and the *edge width* had no significant effect on all DVs. However, there was an effect of the *ratio of node size and edge width* on the *game result* and the *score*. It was surprising that the highest mean of won games and score was at a *node/edge-size ratio* of 18. The players who won games at this condition have in average already played the same number of games like winners at other conditions and we have nearly the same number of data records for all conditions, so this cannot be the reason. Nevertheless, we focus on the local optima at 5.91 from the *game result* and at 4.22 from the *score* because there are only three data points in the right half of the diagram (near the *ratio* of 18), but 19 data points in the left half, which decreases the significance of the above stated outliers in the area of high *ratios*. Also, the qualities of these regressions are lower than those of the questionnaire data regressions ($R^2 = 0.4116$ and $R^2 = 0.3017$ in comparison to an average of $R^2 = 0.9161$ at the three questionnaire data regressions).

5.4 Scotland Graph

Like in the other two games, the difficulty was again the main predictor for the DVs. In Scotland Graph, the difficulty consisted of the number of detectives. With a higher number of detectives, there was a significantly lower chance of winning on such a small graph because there are fewer escape possibilities (and there is no "double-move" like in the original Scotland Yard game). The average of the *number of moves* and the *needed time* decreased with an increasing *difficulty* because the players got caught by the detectives earlier at higher difficulties. The *score*, however, kept unaffected of these changes, which shows that we weighted the *difficulty* well in our score formula to counterbalance the *difficulty*-caused fluctuations of the other dependent variables.

Again, the *node size*, the *edge width* and their *ratio* had no significant effect on all DVs.

5.5 Summary

There was no significant effect of the *node size*, *edge width* and their *ratio* on the dependent variables of the three games, apart from the effect of the *ratio of node size and edge width* on the *win rate* and *score* in the Shannon Switching Game. The *difficulty*, however, had a significant effect on all dependent variables, except the *score* and the *number of clicks* in Scotland Graph.

This is interesting since our original aim was to mitigate the effect of the *difficulty* on the *scores* by including it in the score formulas. It seems we were successful in it in the formula of Scotland Graph, but not in the formulas of Planarity and the Shannon Switching Game.

Summarising, we need to discard the first two hypotheses: there was no clearly measurable effect of *ratio of node size and edge width* on human task performance (first hypothesis), hence there is no optimum (second hypothesis). However, we can accept the third hypothesis: the *ratio of node size and edge width* has an effect on the subjective aesthetic appeal of the graphs with given optima. Furthermore, there were optima for the *node size* and *edge width* on their own.

Usually, when asking for an optimal aesthetical ratio, a common answer is the golden ratio $\phi \approx 1.61803399$. This differs greatly from our result for the optimal ratio of node size and edge width, namely ~ 6. The golden ratio indeed is an extraordinary number and it creates aesthetically appealing geometrical objects, but not for each aesthetic criterion and only when it is applied accordingly. Boselie [Bos84] says to this: "Only as far as the realization of a golden section entails equivalences between parts of a pattern,

(a) ϕ = ratio of node diameter and edge width **(b)** ϕ = ratio of node radius and edge width

Figure 5.2: Little graphs visualising the golden ratio ϕ.

will the aesthetic appeal of a pattern be enhanced" [Bos84]. Figure 5.2 shows a little graph for comparison whose ratio of node size and edge width is the golden ratio. If we ever considered the golden ratio as relevant for the aesthetic criterion of this ratio, we needed to use it as the ratio for node radius, insted of node diameter, and edge width (right subfigure). But this is controversial because the node radius is not a dimension that the viewer sees directly. So the golden ratio does not apply to the case of node diameter and edge width because these two are not parts of *one* pattern.

Although we needed to separate objective and subjective results because of their different outcome, one should not underestimate their commonship. It is evident that humans are more productive on things that they like or find appealing, so they are more productive on aesthetically appealing graphs. Like this, we can bridge the gap between the subjective and the objective and offer our suggestions not only to use cases, where the graph needs to look nice but also to use cases, where human performance is important.

5.6 Limitations

Our measure may not have been the best to quantify human task performance. We did a great effort to draw all nodes and edges at the same size on all displays, but this effort may have been corrupted by the fact that the participants' distance from the display varied. This is a general limitation of online studies. So although the optimum for *ratio of node size and edge width* has generic significance in terms of application contexts, this is not the case for the optima of *node size* and *edge width*.

Another limitation was the strong effect of the *difficulty* on the score in Planarity and the Shannon Switching Game. It was not only unfair to the players (it was rather a matter of luck by getting a low difficulty than skill) but made it harder for us to investigate the effect of the other independent variables on human performance.

Another problem was the large number of experimental conditions and the consequent low amount of data records per condition. Except for the questionnaire data, we would have needed a larger number of participants to find possible correlations with a statistical significance. On the other hand, we tested “only” 27 possible combinations of node size and edge width, hence there could possibly be other optima outside of our tested ranges, but this is unlikely because we tested a large interval of both variables which was also supported by our pilot study.

As a conclusion, we should have omitted the levels of difficulty to reduce the number of conditions or should have improved the score formulas incrementally by better testing to make the games fairer. One might consider examining a similar experiment like ours using only a single difficult level.

Since we restricted our study to solid black egdes and round black nodes on white background, there might be different optima at other edge types, node shapes or colours.

6 Conclusion

In this bachelor thesis, we investigated the effect of node size, edge width and their ratio in node-link diagrams on human task performance and aesthetic appeal. We did so by performing an online study wrapped in three online games and measuring the players' performance in this games as well as asking them for their opinion on the aesthetic appeal of the graphs by a questionnaire. The node size and edge width differed at different games, where we had 27 combinations of node size and edge width. Additionally, there were three levels of difficulty in all games.

After analysing the game data, we did not find a significant correlation between the independent variables (node size, edge width and their ratio) and the players' success at the games. However, there were significant optima for node size and edge width and their ratio in the questionnaire data. Our data suggests an optimal ratio of around 6 and optima for the node size at around $9.1\,mm$ and for the edge width at around $1.7\,mm$. The latter two are restricted to use cases in which the viewer's distance to the graph drawing is the average distance between screen and eye.

These optima did not differ between our games, hence they are general optima for graph-related applications and are therefore useful for graph drawing algorithms that aim to create aesthetically appealing graphs. Furthermore, we discovered that too high ratios of node size and edge width are not as aesthetically repulsive as too low ratios.

Although our study had 81 conditions, our online game approach has proven feasible for collecting enough data points. Yet, a controlled user study in the lab has to be executed with fewer conditions. Since the additional independent variable *difficulty* made it harder for us to detect correlations by increasing the number of conditions and having a significant effect on all dependent variables, one could repeat our experiment without difficulty levels.

Another option were to adapt our experiment to other graph-related games to detect if the detected general optimum of node/edge-ratio still holds. Examples of such games are graph-colouring games[1], Treasure Hunt games (like the one in Figure 2.2) and path-finding games (shortest paths or an Eulerian path).

[1] a list of some graph-colouring games can be found at http://www.fernuni-hagen.de/MATHEMATIK/DMO/graphcolor.html

Bibliography

[BBR04] J. Bovey, F. Benoy, P. Rodgers. “Using Games to Investigate Movement for Graph Comprehension.” In: *Proceedings of the working conference on Advanced visual interfaces*. ACM. 2004, pp. 71–79 (cit. on pp. 12, 13).

[BETT99] D. BATTISTA G, P. Eades, I. G. Tollis, R. Tamassia. *Graph drawing: algorithms for the visualization of graphs*. 1999 (cit. on p. 7).

[Bos84] F. Boselie. “The aesthetic attractivity of the golden section.” In: *Psychological Research* 45.4 (1984), pp. 367–375 (cit. on pp. 56, 57).

[BR07] J. Bovey, P. Rodgers. “A Method for Testing Graph Visualizations Using Games.” In: *Electronic Imaging 2007*. International Society for Optics and Photonics. 2007, 64950S–64950S (cit. on pp. 12–14).

[BRSG07] C. Bennett, J. Ryall, L. Spalteholz, A. Gooch. “The Aesthetics of Graph Visualization.” In: *Computational Aesthetics* 2007 (2007), pp. 57–64 (cit. on p. 9).

[CZ14] C. P. Chen, C.-Y. Zhang. “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data.” In: *Information Sciences* 275 (2014), pp. 314–347 (cit. on p. 7).

[FH02] A. Field, G. Hole. *How to design and report experiments*. Sage, 2002 (cit. on p. 22).

[Fio04] C. Fiorina. *Information: the currency of the digital age*. http://www.hp.com/hpinfo/execteam/speeches/fiorina/04openworld.html. 2004 (cit. on p. 7).

[FRB+15] J. Frommel, K. Rogers, J. Brich, D. Besserer, L. Bradatsch, I. Ortinau, R. Schabenberger, V. Riemer, C. Schrader, M. Weber. “Integrated Questionnaires: Maintaining Presence in Game Environments for Self-Reported Data Acquisition.” In: *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM. 2015, pp. 359–368 (cit. on p. 13).

[GKO+09] X. Goaoc, J. Kratochvil, Y. Okamoto, C.-S. Shin, A. Spillner, A. Wolff. “Untangling a planar graph.” In: *Discrete & Computational Geometry* 42.4 (2009), pp. 542–569 (cit. on p. 33).

[HE05] W. Huang, P. Eades. “How People Read Graphs.” In: *proceedings of the 2005 Asia-Pacific symposium on Information visualisation-Volume 45*. Australian Computer Society, Inc. 2005, pp. 51–58 (cit. on p. 15).

[HEH09] W. Huang, P. Eades, S.-H. Hong. “Measuring Effectiveness of Graph Visualizations: A Cognitive Load Perspective.” In: *Information Visualization* 8.3 (2009), pp. 139–152 (cit. on p. 15).

[HHE08] W. Huang, S.-H. Hong, P. Eades. “Effects of Crossing Angles.” In: *2008 IEEE Pacific Visualization Symposium*. IEEE. 2008, pp. 41–46 (cit. on p. 11).

[Leh64] A. Lehman. “A solution of the Shannon switching game.” In: *Journal of the Society for Industrial and Applied Mathematics* 12.4 (1964), pp. 687–725 (cit. on p. 34).

[NW12] P. Nijssen, M. H. Winands. “Monte carlo tree search for the hide-and-seek game Scotland Yard.” In: *IEEE Transactions on Computational Intelligence and AI in Games* 4.4 (2012), pp. 282–294 (cit. on p. 36).

[PCA02] H. C. Purchase, D. Carrington, J.-A. Allder. “Empirical Evaluation of Aesthetics-based Graph Layout.” In: *Empirical Software Engineering* 7.3 (2002), pp. 233–255 (cit. on pp. 11, 12).

[PCJ95] H. C. Purchase, R. F. Cohen, M. James. “Validating graph drawing aesthetics.” In: *International Symposium on Graph Drawing*. Springer Berlin Heidelberg. 1995, pp. 435–446 (cit. on p. 10).

[PCJ97] H. C. Purchase, R. F. Cohen, M. I. James. “An experimental study of the basis for graph drawing algorithms.” In: *Journal of Experimental Algorithmics (JEA)* 2 (1997), p. 4 (cit. on pp. 9–11).

[Pur00] H. C. Purchase. “Effective information visualisation: a study of graph drawing aesthetics and algorithms.” In: *Interacting with computers* 13.2 (2000), pp. 147–162 (cit. on p. 10).

[Pur97] H. Purchase. “Which aesthetic has the gratest effect on human understanding?” In: *International Symposium on Graph Drawing*. Springer. 1997, pp. 248–261 (cit. on p. 10).

[SGI+83] W. Schlegel, D. Garrels, F. Ifland, M. Burggraf, W. Scheerer, W. Hoermann. *Scotland Yard*. http://luding.org/Skripte/GameData.py/DEgameid/169. 1983 (cit. on p. 27).

[Sim08] S. Simon. *Using ANOVA for a sum of Likert scaled variables*. http://www.pmean.com/08/LikertSum.html. 2008 (cit. on p. 39).

[Tan05] J. Tantalo. *Planarity.net*. http://planarity.net/. 2005 (cit. on p. 25).

[Tan09] P. Tantalo. *Some Remarks on Strategy for Shannon's Switching Game*. 2009. URL: https://classes.soe.ucsc.edu/cmpe177/Summer09/remarks.pdf (cit. on pp. 34, 35).

[Woo12] K. Wood. "Switching the Shannon Switching Game." In: (2012) (cit. on pp. 25, 34).

[WPCM02] C. Ware, H. Purchase, L. Colpoys, M. McGill. "Cognitive Measurements of Graph Aesthetics." In: *Information Visualization* 1.2 (2002), pp. 103–110 (cit. on pp. 11, 15).

All links were last followed on October 16, 2016.

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature