# Proceedings of
# Umeå's 25$^{th}$ Student Conference in Computing Science

## USCCS 2022

L. Jiang, A. Jonsson, L. Vanhée (editors)

UMINF 22.01
ISSN-0348-0542

**Department of Computing Science**
Umeå University

# Preface

The Umeå Student Conference in Computing Science (USCCS) is organized annually as part of a course given by the Computing Science department at Umeå University. The objective of the course is to give the students a practical introduction to independent research, scientific writing, and oral presentation.

A student who participates in the course first selects a topic and a research question that they are interested in. If the topic is accepted, the student outlines a paper and composes an annotated bibliography to give a survey of the research topic. The main work consists of conducting the actual research that answers the question asked, and convincingly and clearly reporting the results in a scientific paper. Another major part of the course is multiple internal peer review meetings in which groups of students read each others' papers and give feedback to the author. This process gives valuable training in both giving and receiving criticism in a constructive manner. Altogether, the students learn to formulate and develop their own ideas in a scientific manner, in a process involving internal peer reviewing of each other's work and under supervision of the teachers, and incremental development and refinement of a scientific paper.

Each scientific paper is submitted to USCCS through an on-line submission system, and receives reviews written by members of the Computing Science department. Based on the review, the editors of the conference proceedings (the teachers of the course) issue a decision of preliminary acceptance of the paper to each author. If, after final revision, a paper is accepted, the student is given the opportunity to present the work at the conference. The review process and the conference format aims at mimicking realistic settings for publishing and participation at scientific conferences.

USCCS is the highlight of the course, and this year the conference received 10 submissions, which were carefully reviewed by the teachers of the course. As a result of the reviewing process, 6 submissions were accepted for presentation at the conference. We would like to thank and congratulate all authors for their hard work and excellent final results that are presented during the conference.

We wish all participants of USCCS interesting exchange of ideas and stimulating discussions during the conference, which this year is in hybrid mode.

Umeå, 5 January 2022

Lili Jiang
Anna Jonsson
Loïs Vanhée

# Organizing Committee

Lili Jiang
Anna Jonsson
Loïs Vanhée

# Table of Contents

# The User Experience of the Installation Process of Progressive Web Applications: A User Test

Lovisa Berggren

Department of Computing Science
Umeå University, Sweden
lobe0059@student.umu.se

**Abstract.** A Progressive Web Application (PWA) is a mobile application run in the browser, offering the user a similar experience to native applications. This concept was introduced in 2015 and has since been suggested as the possible ultimate cross-platform development method for mobile applications. However, little is known about the user experience, and previous studies show that there is an overall lack of research on PWAs. In this paper, we expand the existing literature by investigating the user experience of the installation process. We carry out a qualitative user test with 6 participants, who got to install both a native application and a PWA. The results show that the users prefer the native installation process, and that the PWA installation was more difficult and confusing. Furthermore, it was revealed that the participants were not aware of PWAs at all, and none of them understood that an application had been installed. The test shows clear differences in the approaches, proposing that there may be underlying issues with the installation process of PWAs. We propose future research on the subject of PWAs, especially concerning the end-user and their experience.

## 1 Introduction

In 2015, a new concept was introduced in mobile development: the Progressive Web Application [1]. This fairly new concept, also called PWA, is based on an approach set by the Google Web Fundamentals in an effort to present a viable alternative to the traditional native applications [1, 2].

A PWA is a web application that offers the user a similar experience as when using a native application. Unlike web applications, a PWA can be installed on the user's device, send notifications to the user and be used offline. This means a PWA is run in a browser but acts and looks like a native application. Because PWAs are run in a browser, they come with the added advantage that the app can be used without installation [1, 2].

Little is known about the end-user experience of PWAs. Since the introduction of the approach, many studies have analyzed and scrutinized the possibilities of PWAs, comparing them to other established approaches. It has been suggested that PWAs may be the approach that can replace native applications in the future, due to its advantages. Many of these studies focus on the technical side of

L. Jiang, A. Jonsson, L. Vanhée (Eds.): Umeå's 25$^{th}$ Student Conference in Computing Science USCCS 2022, pp. 1–14, January 2022.

PWAs, or if the concept is the ultimate cross-platform approach to mobile app development. However, there is a gap in research on how PWAs are perceived by the end user.

In principal, PWAs are superior to native applications since PWAs can be used with or without installation. However, PWAs are installed in the browser [1] instead of from an app store. One could argue that the installation process through an app store is more comfortable, since it is what most users are accustomed to. Is it a vision biased by habits or are there systematic issues with PWA user experience? To fully answer the question if PWAs can replace native applications, we have to understand the end-user. In this paper, we delve deeper into the user experience of PWAs, proposing a comparison of the installation process of PWAs with native applications.

Six test subjects participated in a user test followed by a short survey to collect information about the user experience. Half of the test subjects used Android, and the other half used iOS. A/B testing was used to allow the test subjects to compare the different installation processes. A/B testing is a method where each subject gets to test two different versions of the same variable. The methodology was chosen to gather qualitative data on the user experience, since the range of answers to the questions are relatively open at this point. The open coding technique was used to summarize the emotions of the participants, by labeling the reactions with words that give them meaning.

Using systematic user tests and interviews about the installation process of the different application types, we show the differences in the user experience. Furthermore, the test offered insights in other related issues with the user-awareness and knowledge of PWAs.

The structure of the paper is as follows. First, we define and compare native applications, PWAs, and other cross-platform development approaches. In section 3 we present earlier work and the current state of the research on PWAs. Then, in section 4 we explain the methods used to answer our research question. In section 5 we present our results from the user test. Lastly, we provide an analysis of the results of this paper, followed by recommended future work.

## 2 Mobile Application Development Approaches

Installable mobile applications can be categorized in five different ways; Native, Cross-platform with it's interpreter, cross-compilation and hybrid approach, and the Progressive Web Application.

### 2.1 Native Applications

A native application is a software which is used on a specific platform. The most common mobile platforms are Apple's iOS and Google's Android[1]. Because of

[1] Search Software Quality. *Native app.* `https://searchsoftwarequality.techtarget.com/definition/native-application-native-app`

this, an application has to be developed separately for each OS to work on multiple platforms. For example, different programming languages are needed for developing native applications for different platforms. Objective-C or Swift can be used for iOS, and Kotlin or Java can be used for Android [3]. Because native apps are developed for a specific platform, it can use hardware specific for the device, provide optimal performance and take advantage of the latest technologies[2]. Native applications can be distributed through app stores where users can discover and install the applications [4].

### 2.2 Cross-Platform Development Methods

The main characteristic of a cross-platform development framework is that one code base is created for one application, and it can be used on all platforms. How this is achieved however differs between cross-platform development methods [3]. These types of applications can be installed through the app stores and be used offline, like native applications [5].

There are three common cross-platform development methods:

- The interpreter approach.
- The cross-compilation approach.
- The hybrid approach.

With the **interpreter approach**, a JavaScript interpreter on the device will execute the code, and generate native components for each platform [2]. This is the method used by React Native [6].

The **cross-compilation approach** will compile the source code separately, creating runnable code for each separate platform. This approach results in native applications, and are described as truly native [1]. Xamarin uses this approach [2].

The **hybrid approach** is in some ways similar to web applications or websites. The application is developed using web technologies such as HTML or CSS, and then displayed to the user through a WebView [2]. The WebView is wrapped in a native container which displays the application, instead of a standard web browser [7]. A hybrid application can therefore be installed from an app store, like native applications [4]. The Ionic Framework and PhoneGap are both a part of the hybrid approach [1].

### 2.3 Progressive Web Applications

A Progressive Web Application is a Web app run in a browser that acts and feels like a native application. The application can be accessed through a browser with an URL and be used like a Web app, but they can also be installed and used offline though the home screen [2].

For an application to be classified as a PWA, it has to have the following features [2]:

---

[2] Search Software Quality. *Native app.* `https://searchsoftwarequality.techtarget.com/definition/native-application-native-app`

1. Progressive: accessibility for all users, independently of the browser.
2. Responsive: can be used with any platform or screen format.
3. Connectivity independent: correct behavior with or without internet connection.
4. Native app similarity.
5. Updates: uses *Service Workers* to continuously update content.
6. Safe: uses HTTPS.
7. Easy discovery: the app is easy for the user to find.
8. Engaging: uses features for higher user engagement, such as push notifications.
9. Installable.
10. Easy connection: allows the user to access the application without going through the installation process.

To install a PWA, the user saves it to the home screen like creating a short-cut to a regular website. Since the user can access the PWA directly in the browser through an URL, or installed on the home screen, they have the opportunity to test the application before installing [1]. After installation to the home screen, the necessary static files are downloaded to the device to be used offline [3].

Apple does not yet fully support the use of PWAs on iOS. PWAs can be used on iOS devices, but with some limitations since the specification for Service Workers are not implemented in Safari [3]. This however, does not affect the installation process of PWAs on iOS for the end-user. Even if PWAs on iOS are not as fully functional as they are on Android, the isolated process of installation can still be compared between the platforms.

### 2.4 Comparison of Native and PWA

The traditional way of developing native applications comes with an important disadvantage. Since one source code for each OS needs to be created using the native approach, the method makes it expensive to develop applications for multiple platforms. Thus, other cross-platform compatible approaches have emerged, such as the hybrid, interpreter and cross-compilation approaches [2, 3]. Seeing that PWAs are essentially Web Applications with some extra features, they can also be considered cross-platform compatible. However, since PWAs are not yet fully supported by Apple, it is stated that they cannot be labeled as a true cross-platform compatible approach [3].

The installation process is the feature that differentiates PWAs from native applications for the end-user. Native applications are, as previously stated, installed through an app store [4]. This also applies for the applications developed using the cross-platform approaches, since the end result is a native application. PWAs on the other hand, are installed through the browser by saving them to the home screen [1]. How this difference in installation affects the user experience will be explored and analyzed in this paper.

## 3 Earlier Work

Most studies found on the subject of PWAs only consider the development and/or technical side. One study notes that there is a "lack of academic involvement" concerning PWAs, and suggests that there is research potential [1]. Some studies also recommend more research on the impact of the *add to home screen*-button [1, 5]. The user experience of PWAs in general is also an area which researchers argue to be researched upon [2, 5].

The study *Analyzing User Experience in Mobile Web, Native and Progressive Web Applications: A User and HCI Specialist Perspectives* analyzes the user experience of PWAs, web applications and native applications. However, the study was fairly small with 8 test subjects, and focused on the experience of using the different applications. The experience of the installation process was not taken into account [8]. Here, we expand this literature by considering the installation process.

## 4 Method

In order to answer the research question a user test was conducted using A/B testing and the results were summarized with the open coding technique. Below, we explain in detail each method used and how the test was executed.

The test was conducted in a private setting on campus of Umeå University. The test subjects were allowed to perform the test sitting down at a table. One moderator was present on the opposite side of the table, instructing the test subjects, observing their behavior and gathering data through notes and audio recordings.

The test subjects used their own mobile phones to perform the tasks. The subjects were encouraged to "think out loud" while completing their tasks, explaining what they did and how they solved them. The tests were carried out with the consent of the test subjects.

### 4.1 Test Method

A/B testing was used to compare the installation process of PWAs and native applications. The test method allows us to compare two different variables on the same test subject[3], where the first variable A was the native application installation and B was the PWA installation. After each task, a short survey was done, and after the tasks a more in-depth interview was conducted, allowing the user to compare the different types.

This methodology was chosen to gather qualitative data on the user experience of installing a PWA and a native application. Since little to no research has been done on the user experience of PWAs, a large scale survey may be hard to conduct. A more quantitative method may result in uncertain or unclear data, since the range of answers is relatively open at this point. A qualitative, more controlled test is therefore preferable.

[3] Wikipedia. *A/B testing.* `https://en.wikipedia.org/wiki/A/B_testing`

### 4.2 Task Description

There were 6 test subjects who completed the user test[4]. Half of the test subjects used Android as OS and Google Chrome as browser, and the other half used iOS and Safari. All of the test subjects were engineering students at Umeå University, aged 19 to 30.

The test subjects were given two tasks. The first task was to install Twitter as a native app, and the second was to install the same app as a PWA.

The experimental protocol is as follows. First, the test subject was asked to download Twitter to their phone, and by observing the user we could get information about the experience of solving the task. After the task was completed, the following questions were asked:

- How did it go?
- Did you run into any problems?

In the second task, the test subject was asked to go to the website `mobile.twitter.com`. Then, the subject was asked to save the site to their home screen. By observing the user we could get information about the experience of solving the task. If the subject could not figure out how to solve the task, it was noted and the subject was guided in how to accomplish it. After the task was completed, the following questions were asked:

- How did it go?
- Did you run into any problems?
- What do you think that you have accomplished?
- In what way do you think you can use the saved application on your phone?
- Can you note any differences between this app and the previous one?

After the tasks were completed, some in-depth questions were asked about the experience.

- Which one of the methods did you prefer? Why?
- Did you know that you could save a website through "save to home screen"?
- Was it clear that the site you saved to your home screen is now downloaded to your phone?
- Do you know what a Progressive Web Application is?

### 4.3 Result Evaluation

After the user tests were conducted, the results were noted by combining the notes from the tests with a transcript of the audio recordings.

The observations of the participants emotions during the completion of the tasks, and while answering questions, were summarized with the open coding technique. The results are analyzed and tagged with words that give meaning to

[4] In a full study there should be more test subjects, but due to time constraints in this study 6 subjects were used.

| Tag | Description |
|---|---|
| Confidence | The subject expressed confidence. |
| Positive | The subjects seemed happy and/or positive. |
| Neutral | The subject acted indifferent or neither positively nor negatively. |
| Unsure | The subject expressed doubts, unsureness and/or carefulness. |
| Confused | The subject acted confused during a task or how to answer a question. |

**Table 1.** Explanations of tags used to summarize the observations of the subjects during the tasks and as they answered the question.

the data. This technique does not use pre-defined labels but labels that emerge from the data [9]. Table 1 contains the tags used in this study and a description for each tag. The tags represent the subject's verbal and body language. This is the same technique as the one used in the study *Analyzing User Experience in Mobile Web, Native and Progressive Web Applications: A User and HCI Specialist Perspectives* [8].

## 5 Results

The results of all 6 participants are considered in this section. The test subjects ages ranged from 19 to 30, with the average age of 22. All of the participants were students at Umeå University, studying computer science, engineering, energy engineering, interaction and design and/or tech. Three of the participants used Android and Chrome, and the other three used iOS and Android.

| Task | Application | Success rate | Observed emotions |
|---|---|---|---|
| 1 | Native | 6/6 | Confidence (6) |
| 2 | PWA | 5/6 | Unsure (4)<br>Confidence (2) |

**Table 2.** Success rate and observed emotions of task 1 and 2.

Table 2 shows the success rate of the tasks. Success implies that the subject accomplished the task, e.g. they installed the application on their own. In the first task, all subjects succeeded in installing the native application. In the second task, 5 out of the 6 subjects succeeded in installing the PWA. The subject who did not succeed used Android and Chrome.

The observed emotions during task 1 and 2 are also shown in table 2. All participants were confident in installing the native application. In the second task, installing a PWA, 2 participants expressed confidence and 4 participants acted unsure. It was also noted that the 4 subjects who acted unsure during the second task used trial and error as method.

### 5.1 Task 1

The first task was to install Twitter as a native application. After the task, the following questions were asked:

– Q1.1 : How did it go?
– Q1.2 : Did you run into any problems?

| Question | Answer | Expressed emotion |
|---|---|---|
| Q1.1 | Smooth (4)<br>Fast (2)<br>Good (2)<br>As expected (2)<br>Done before (2)<br>Easy (1) | Positive (6) |
| Q1.2 | No (6) | Confidence (6) |

**Table 3.** Answers and expressed emotions to the questions asked after the first task, where the subject was instructed to install the native application.

Table 3 shows that the answers to the questions of task 1 were positive. The most common answer to question 1.1 was that the installation went smoothly. None of the participants expressed that they ran into any problems during the task.

### 5.2 Task 2

The second task was to install Twitter as a PWA. After the task, the following questions were asked:

– Q2.1 : Can you find the page you saved?
– Q2.2 : How did it go?
– Q2.3 : Did you run into any problems?
– Q2.4 : In what way do you think you can use the saved application on your phone?
– Q2.5 : Can you note any differences between this app and the previous one?

Table 4 summarizes the answers to the questions followed by task 2, and observed emotions. The overall observed emotions were confusion, unsureness or neutral. In question 2.4, four of the subjects thought that the installed PWA could be used as a link or reference for the Twitter web page.

### 5.3 Concluding Interview

After the tasks were completed, the following questions were asked:

– Q3.1 : Which one of the methods did you prefer?
  • Q3.1.1 : Why?
– Q3.2 : Did you know that you could save a website through "save to home screen"?
– Q3.3 : Was it clear that the site you saved to your home screen is now downloaded to your phone?

| Question | Answer | Expressed emotion |
|---|---|---|
| Q2.1 | Yes (4)<br>No (1) | Confused (3)<br>Unsure (2) |
| Q2.2 | Good (2)<br>Difficult (2)<br>Not easy (1) | Neutral (1) |
| Q2.3 | Not easy (4)<br>No (1)<br>Yes (1) | |
| Q2.4 | As a link to a website (5)<br>As a link to a native application (1) | Unsure (4) |
| Q2.5 | No (3)<br>A little bit (3) | |

**Table 4.** Answers and expressed emotions to the questions asked after the second task, where the subject was instructed to install the PWA.

| Question | Answer | Expressed emotion |
|---|---|---|
| Q3.1 | App Store (5)<br>Don't know (1) | Unsure (1) |
| Q3.1.1 | Habit (4)<br>Smoother (3)<br>Easier (1)<br>Security (1) | Confident (5)<br>Unsure (1) |
| Q3.2 | Yes (3)<br>No (3) | Unsure (1) |
| Q3.3 | No (4)<br>Yes (1)<br>Maybe (1) | Unsure (1) |
| Q3.4 | No (6) | |

**Table 5.** Answers and expressed emotions to the concluding interview after the tasks were completed.

– Q3.4 : Do you know what a Progressive Web Application is?

Table 5 summarizes the answers to the concluding questions at the end of the test, with the observed and expressed emotions. Five of the participants preferred to use an app store over using the ”save to home screen” action. One of the participants were unsure. The most common answer to why they preferred to use an app store was because they were used to it. It was also mentioned that using an app store is smoother or easier. One of the participants mentioned security; that using an app store feels safer than downloading directly from the browser.

50% of the participants knew that they could save a website through the “save to home screen” option, and 50% did not. Four of the participants did not understand that the PWA was downloaded on their phone. None of the test subjects knew what a PWA was.

### 5.4 Android-iOS Comparison

The test subject who did not succeed in the second task used Android and Chrome. Two out of the three Android users could not find the PWA after installation in the second task, whereas all of the iOS users could. Otherwise, there were no differences in the answers to the questions or expressed emotions between the iOS and Android users.

## 6 Discussion

From the results of the user test in this paper, there is a clear difference in the user experience of installing a PWA and a native application. Looking at only the success rates, there seems to be no difference between installing a PWA or a native application. However, one cannot assume that this is always the case due to the small participant group. Also, the expressed emotions and the answers to the follow-up questions of each task shows a clear difference. All of the participants were confident in the first task; installing Twitter as a native application. In the second task, installing the PWA, there was more confusion and unsureness, and 4 out of the 6 participants solved the task due to trial and error. Five out of the six participants also stated that they preferred the native installation, mostly due to habit.

In comparing users with Android and iOS, there were not many differences. The subject who did not succeed in installing the PWA used Android, and 2 out of the 3 Android users did not find the saved PWA on their home screen. However, no conclusions can be drawn by these results since there were so few participants in the test group.

In the second task of the user test of this paper, the subjects were asked to save the site `mobile.twitter.com` to their home screen. This poses the question if any of the participants had succeeded in the second task if the task simply was to *install the PWA Twitter?*. The formulation of this question may have helped the participants in solving the task, since *saving to the home screen* was expressed by the test moderator. If this is the case, reformulating the task would have shown an even bigger difference in the installation processes, since none of the participants knew what a PWA is.

In addition to the fact that none of the participants were aware of PWAs, all of them thought that the saved icon on their home screen could be used as a link to the Twitter web site or the previously installed Twitter application. This shows that there is not only a difference in the user experience of the installation processes, there is also a knowledge gap of the user. The users are not aware of PWAs, and this is probably not their fault. The reason as to why the users thought they added a link to their phone, and not an application, may be because the PWA is viewed as a regular web application in the browser, and the process of installing a PWA is so similar to the process of adding a website to the home screen as a link, if not exactly the same. This may be the factors that led 4 out of the 6 participants to not understand that anything was downloaded in the process of the second task. This is misleading the user, which is clearly shown in one of the participant's statements in answering which of the methods they preferred; "I would have appreciated a link [instead of an application] on my previous phone because I did not have so much memory on it". Another participant mentioned that installing through an app store feels safer than downloading something from the internet, which is a valid statement. It may be an important security risk for the user if they do not know that anything is installed when saving a PWA. If the users were aware of PWAs and how they worked, the results of this study may have been different.

### 6.1 Target Group

In the user test of this paper, all of the participants were students in the ages 19 to 30, and all of them studied in a technical field. The target group can affect the result of a study, but in this case it was probably insignificant. Students of technical studies in a university has presumably a broad understanding of technology and applications overall. If the results would have been that there were no differences in the user experiences of installing a PWA and a native application, it may have been because of the user group. Then, a broader user group should have been used to examine the different processes. However, since the results did show a difference in the approaches with a user group of presumably high technical knowledge, the target group in this case did probably not affect the results. If the target group had been broader with people of many different areas of education, the results had probably only been more clear.

### 6.2 Test Design

In this paper, we used quite broad questions when asking the subjects about their experiences. This can result in answers that are hard to compare. However, since the field of PWAs is still new, it may be hard to ask more specific questions. We could have asked the users to rate their experiences on a scale of for example

1 to 5, to make the results easier to compare. However, since the target group of this study was small, with only 6 participants, it may have been hard to draw any conclusions from the scales anyways.

The way that the A/B testing was done in this study can be improved in the future, if the test was to be conducted again. In this study, A was task 1, to install a native application. B was task 2, to install a PWA. When implementing A/B testing, half of the group should do A before B, and the other half of the group should do B before A. In this study, all of the participants did A before B. How this affected the results is hard to say, but should be considered in the future.

Another element that could have been added to the test is time of task completion. Comparing the time it took for the participants to solve each task can give another indication to which task was easier, than just relying on expressed emotions. Emotions can be hard to analyze, and the way humans perceive emotions are clearly biased.

To give the test even higher quality, there should be more test moderators present during the test. In the case of this paper, there was only one test moderator asking the questions and observing the users. With two or more moderators, the results would be even more reliant. Since it can be hard to focus on many variables at the same time, it would be better if for example one moderator instructed the test subject, and another noted the user's behavior.

### 6.3 Suggested Future Work

In this paper, we have analyzed the differences in the installation process of PWAs and native applications, a comparison that has not been done before. The results of the user test that was conducted provided many new and interesting insights in the end-user's experience of PWAs. However, the study was fairly small, and there is a lot more to be explored in the future. Therefore, we suggest the following areas of future work:

- Does PWAs pose any security risks that native applications do not have? How does the user view PWAs in the light of security?
- Since this paper cannot show any differences between Android and iOS users in installing PWAs; how does PWA installation compare in iOS and Android users?
- In this paper, we did not explore the user experience of discovering PWAs. In the test, we gave the subjects a URL to the PWA. We suggest future research on the discovery process, since native apps can be browsed in an app store, and PWAs cannot.
- To validate the results of this paper; explore the user experience of installing PWAs on a larger target group.

In addition to this, we urge the research community to continue the scientific research of PWAs to fill the knowledge gap, especially with a focus on the end-user and UX. Qualitative and quantitative research on the user experience of PWAs can complement each other and improve the understanding of PWAs and their place in the world of mobile development.

### 6.4 Conclusions

This paper compared the installation processes of PWAs and native applications, from the end-user perspective. A qualitative user test was conducted which showed a clear difference in the user's experiences of the two approaches. Even though the majority of the participants succeeded in both installation tasks, native installation was preferred over PWA installation. The PWA installation process also seemed more confusing and difficult. The user test also showed that there is a knowledge gap about PWAs among the participants. None of them knew what a PWA was, and all of them misunderstood what they had actually accomplished with saving the PWA. The subjects thought the new icon on the home screen was a link to the web version of the application, or the previously installed native application.

These findings answer the research question of this paper, whether there is a difference in the user experience of installing PWAs and native applications, at least to some degree. Since the test group in this paper was fairly small, it may not be possible to draw any major conclusions from the results. However, the test showed clear differences in the approaches, proposing that there may be underlying issues with the installation process of PWAs. The revealed knowledge gap among the users is an important point for future research, and may affect the answer to the question if PWAs can replace native applications in the future.

## References

1. Bjørn-Hansen, A., Majchrzak, T.A., Grønli, T.M.: Progressive web apps: The possible web-native unifier for mobile development. In: Proceedings of the 13th International Conference on Web Information Systems and Technologies. (2017) 344–350
2. Carneiro, J.: Progressive web apps: Concepts and features. Kriativ-Tech **7**(7) (2018) 1–3
3. Majchrzak, T.A., Bjørn-Hansen, A., Grønli, T.M.: Progressive web apps: the definite approach to cross-platform development? In: Proceedings of the 51st Hawaii International Conference on System Sciences. (2018) 5736–5740
4. Tunali, V., Zafer, S.: Comparison of popular cross-platform mobile application development tools. Celal Bayar University: Faculty of Technology Department of Software Engineering, Maltepe University: Faculty of Engineering and Natural Sciences Department of Software Engineering (2015)
5. Rensema, D.J.: The current state of progressive web apps: A study on the performance, compatibility, consistency, security and privacy, and user and business impact of progressive web apps (2020)
6. Danielsson, W.: React native application development: A comparison between native android and react native. Master's thesis, Linköpings universitet (2016)
7. Chaudhary, P.: Ionic framework. International Research Journal of Engineering and Technology **5**(05) (2018) 3181
8. de Andrade Cardieri, G., Zaina, L.M.: Analyzing user experience in mobile web, native and progressive web applications: A user and hci specialist perspectives. In: Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems. (2018)

9. Corbin, J., Strauss, A.: Basics of qualitative research: Techniques and procedures for developing grounded theory. Sage publications (2014)

# Inter-tool Reliability of Three Automated Web Accessibility Evaluators

Maria Björkman

Department of Computing Science
Umeå University, Sweden
mabj0039@student.umu.se

**Abstract.** The aim of this study is to compare three popular automated web accessibility evaluation tools (AWAETs) and study the inter-tool reliability of the evaluators. The standard used is WCAG 2.1. The evaluators are also compared from a usability perspective. Nine websites are assessed by three different AWAETs. The evaluators are Mauve++, QualWeb and Wave. The statistical measure Fleiss' kappa is used to calculate inter-tool reliability. Based on the data the agreement between the evaluators is limited, and the inter-tool reliability is thereby poor. From a usability perspective Mauve++ is the most versatile of the three evaluators.

## 1 Introduction

Internet has become an unavoidable space since many everyday functions such as banking activities, paying taxes and booking transport are performed online. Accessibility, meaning that everyone despite disabilities should have the same access to the information provided on the web, is now a requirement for websites provided by the public sector. The disabilities can for example be in vision, physically or cognitive [1]. Even though web accessibility is by law regulated for the public sector and banking, so far there is no consensus on how to best evaluate or measure the level of accessibility for a website [2].

To help evaluate the accessibility of a website different standards have been developed. These standards are sets of guidelines that can be used to check the compliance of the accessibility of a website. Examples of standards are Web Content Accessibility Guidelines (WCAG) 2.1 and the United States Section 508 [2].

In order to measure accessibility there are four different approaches. These are automated evaluation, manual evaluation by experts, testing with end users or a hybrid of these. The advantage of automated testing is that more code can be analysed compared to manual testing. A disadvantage is that the guidelines that are of a more subjective nature are difficult to test automatically. Research has shown that the best way to measure accessibility is to involve users, but this is often costly and time consuming [2].

There are a variety of automated web accessibility evaluation tools (AWAETs) that can be used to check the accessibility of a site. Besides being used when

L. Jiang, A. Jonsson, L. Vanhée (Eds.): Umeå's 25$^{th}$ Student Conference in Computing Science
USCCS 2022, pp. 15–25, January 2022.

developing websites, the AWAETs are also used to ensure the quality of the accessibility [3]. Most of the tools do not provide information on how reliable the results from the assessment is, and research has shown that the coverage of the tools is narrow. It is therefore recommended to use many tools when assessing a website [3].

Reliability of an AWAET can be divided into intra-tool reliability and inter-tool reliability. Intra-tool reliability is whether the results can be obtained repeatedly by one evaluator under the same circumstances. The inter-tool reliability is whether different tools can obtain similar results when applied to the same data [4].

The focus of this study is to compare how reliable three different web accessibility evaluation tools are. The standard used is WCAG 2.1. The reliability will be compared by the inter-tool reliability. The chosen evaluators are popular, free of charge and can be used as an extension in the Firefox browser or by providing an URL of the web site. The evaluators are QualWeb, Mauve++ and Wave. Furthermore, an analysis of the evaluation results is made to see how the evaluators differ, and the usability of the evaluators is discussed.

## 2 Background

### 2.1 WCAG

The guidelines from WCAG are developed through the W3C process[1]. By using the guidelines the developer gets a better understanding of what is required in producing and rendering accessible web content. The latest guidelines are called WCAG 2.1 and has three levels of conformance: A (lowest), AA or AAA (highest). One objective with WCAG 2.1 compared to its predecessors has been to make the guidelines more testable. According to the W3C website there are today 77 AWAETs that are compliant with the WCAG 2.1 standard[2].

WCAG 2.1 is organised into 4 different principles. These principles are *perceivable*, *operable*, *understandable* and *robust*. For every principle there are a set of testable success criteria (SC). All in all there are 78 SC. An example of a SC is **1.4.3 Contrast** (that belongs to the principle perceivable and is of level AA), which says that large-scale text needs to have a contrast ratio of at least 3:1.

### 2.2 Inter-tool Reliability

Inter-tool reliability is related to the reproducibility of the results from an assessment. Vigo et al [3] says in their article that *"Ideally, metrics should be as independent from tools as possible, and show small variations when plugged to one or another tool."* This is not the case when different evaluators are conducting an assessment today. The inconsistency in results from one evaluator

[1] `https://www.w3.org/WAI/standards-guidelines/wcag/`

[2] `https://www.w3.org/WAI/ER/tools/?q=wcag-21-w3c-web-content-accessibility-guidelines-21`

to another is because the tests of the guidelines are implemented in different ways, and also how they crawl the website [3]. However, one could argue that the different AWAETs should not introduce artefacts in the results since they have been developed to test the same barriers.

One way to measure the reliability of agreement between raters is to use the statistical measure called Fleiss' kappa. This statistical measure is often used in medical and behaviour research, and in fields of content analysis [5]. This statistic can be used on nominal data and with any number of raters, when it is possible to assign categorical ratings to the data. A limitation is that it can not handle missing data. The Fleiss' kappa values range from -1 to +1, where Kappa = 1 means perfect agreement, and Kappa = 0 means the agreement is the same as would be expected by chance. The preferred value is 0.9[3].

### 2.3 Presentation of the evaluators

The inclusion criteria for the evaluators in this research are that they can be used free of charge and that they can be used in the Mozilla Firefox browser either as an extension or by providing an URL of the website in question. The evaluators should also have different engines and be able to evaluate according to WCAG 2.1 standard. Furthermore, the evaluators should be transparent with which SC they are evaluating, so that a comparison can be made between the three evaluators. See Table 1 for an overview of the three evaluators and Table 2 for more information of how the evaluators are used and what information they provide. The conformance level chosen is level AA of the WCAG 2.1 standard.

**Mauve++:** This evaluator is a project from the Institute of Information Science and Technologies (ISTI) and the National Research Council of Italy (CNR). The checkpoints are categorised into success, error or warning. A warning needs to be manually evaluated. After the evaluation the results are given either in a XML report, a report in EARL, a PDF report with the results or a web interface report of a single web page. Mauve can operate in three different ways: by evaluating a given URL, pasting HTML code, or by uploading files. Mauve is transparent in which guidelines that are being investigated and how many test that are run for each guideline. The service is found at `https://mauve.isti.cnr.it/`.

**QualWeb:** Can be used through the command line or by providing an URL on their website. Gives information of which techniques that are tested on the website that is evaluated. The outcome is classified as passed, failed, warning or not applicable. A warning needs to be manually inspected in order to determine if it is an error or not. The online service can be found at `http://qualweb.di.fc.ul.pt/evaluator/`.

---

[3] `https://support.minitab.com/en-us/minitab/18/help-and-how-to/quality-and-process-improvement/measurement-system-analysis/how-to/attribute-agreement-analysis/attribute-agreement-analysis/interpret-the-results/all-statistics-and-graphs/kappa-statistics/`

**Wave:** Provides a visual representation of the issues detected on a website. The tool assists by displaying the information of the issues within the evaluated webpage. Wave can be used as an extension in Firefox and Chrome. There is also an online service at `https://wave.webaim.org/`. The outcome is categorized as error, alert and features. Alerts are potential errors that needs manual inspection.

| Tool | Vendor | Version | Compliance | Supported formats | Type of feedback |
|---|---|---|---|---|---|
| Mauve++ | Hiis Lab ISTI - CNR | 2.0 | WCAG 2.1, WCAG 2.0 | WAI-ARIA, CSS, HTML, XHMTL, Images | Report of results, displaying information within web pages |
| QualWeb | Faculdade de Ciências da Universidade de Lisboa | 3.1 | WCAG 2.1, ACT-rules | WAI-ARIA, CSS, HTML, XHMTL, SVG | Report of results, displaying information within web pages |
| Wave extension | WebAIM.org | 3.0.3 | WCAG 2.1, WCAG 2.0, Section 508 | CSS, HTML, XHMTL, Images | Report of results, displaying information within web pages |

**Table 1.** Overview of the evaluators.

| Tool | How to use it | Severity classification | How to fix the issues |
|---|---|---|---|
| Mauve++ | Providing URL, files or code | Error, Warning, Success | No |
| QualWeb | Providing URL, command line | Passed, Failed, Warning, Not Applicable | No |
| Wave extension | Firefox, Chrome extension, online API service, providing URL, stand-alone API product | Error, Alerts, Features | Yes |

**Table 2.** Behaviour of the evaluators.

## 3 Related Work

The inter-tool reliability of three AWAETs were investigated by Molinero and Kohun in 2006 [4]. The AWAETs in the study were Watchfire Bobby, LiftNN for Dreamweaver and Ramp. Krippendorff's alpha was used to compute the inter-reliability. The Krippendorff's alpha used to compute the inter-tool reliability is a statistical method that can be used on nominal data. An advantage with this method is that it can handle missing data, which Fleiss' kappa is not able to. The guidelines from the United States 508 were used as a base for checking the accessibility. A conclusion made in this article was that the more subjective guidelines had a lower inter-reliability, and are problematic when performing automated assessments. Because of this, the authors recommend skill building for developers rather than using AWAETs when developing websites, at least until the AWAETs are more reliable [4]. The evaluators analysed in this study are not popular to use today.

An article from 2018 by Frãzao and Duarte [1] are comparing eight accessibility plug-ins. The inclusion criteria for the evaluators were that the chosen AWAETs can be used free of charge as an extension in the Chrome browser. The plug-ins are Lighthouse, Wave, aXe, Accessibility Insights, Arc Toolkit, Total Validator, ACCESS Assistant Community and Tenon Check. Ten websites were evaluated by the eight plug-ins and the evaluation results from the assessments were used as the test data for comparing the evaluators. The authors also looked into the usability aspect of the evaluators. A finding is that the evaluators have different ways of analysing the success criteria and are classifying the errors in their own ways. This article is supporting the results in other studies that shows that the coverage of a single AWAET is limited [2, 3, 6]. A limitation of the study is that it does not take into account true positives, false positives and false negatives. The authors have not calculated the inter-tool reliability of the AWAETs, only the coverage of SC for each evaluator. A conclusion is that if all eight tools are being used together the coverage of the SC rises from 10% to 40%.

A more recent article is focusing on the transparency of AWAETs [7]. The authors investigate Mauve++, Wave, AChecker and QualWeb. Transparency is for example how much information that is provided before and after an assessment of how well the guideline checkpoints are being evaluated. A finding is that AChecker is providing the least amount of information on how the tool operates. The authors argue that more transparency could help the users make better choices when performing accessibility assessments, such as understanding which evaluators are beneficial to use together [7].

Since previous research has not compared QualWeb, Mauve++ and Wave, this paper can add knowledge to the field of research of accessibility and automated testing by investigating the inter-tool reliability of these three evaluators.

## 4 Method

The goal of this research is to investigate the inter-tool reliability of the AWAETs, and based on the results analyse how the evaluators differ in their assessments. Another field of the investigation is to look into the evaluators from a usability perspective. The test data is based on the evaluation results from assessments from nine different websites, and the total number of errors to each SC for each evaluator is calculated. Based on this test data the inter-tool reliability is calculated by using the Fleiss' kappa measurement.

### 4.1 Selected Websites

The AWAETs will assess nine different websites. The websites are among the most popular websites in Sweden according to the Alexa rank[4] for October 2021. Websites that only consist of a login page or are very simple in other ways (such as google.se) are excluded. Among the chosen websites are websites from the

[4] `https://www.alexa.com/topsites/countries/SE`

commercial sector as well as the public sector. There is also a range of dynamic web pages to more static ones. See Table 3 for a description of the websites included in the study.

| No | Website | Description |
|---|---|---|
| 1 | https://www.reddit.com/ | An American social news and discussion website |
| 2 | https://en.wikipedia.org/ | The free online encyclopedia |
| 3 | https://www.aftonbladet.se/ | A daily newspaper based in Stockholm |
| 4 | https://www.avanza.se/ | Stockbroker and brokerage firm in Sweden |
| 5 | https://www.hemnet.se/ | Platform for home purchases |
| 6 | https://www.1177.se/ | The Swedish healthcare site |
| 7 | https://www.svt.se | The national public broadcasting company |
| 8 | https://www.tradera.com/ | Platform for selling items |
| 9 | https://www.amazon.se/ | E-commerce site |

**Table 3.** Description of the websites.

### 4.2 Testing Environment and Procedure

The test environment is the Mozilla Firefox Developer Edition browser. To ensure that the evaluators are assessing the same website, the evaluation of all the evaluators starts within a few seconds of each other. The number of occurrences for the reported errors for each website is saved. Warnings/Alerts and Passed SC are excluded in this study.

When all the nine websites have been evaluated by the three AWAETs, the analysis of the results starts. In order to calculate the inter-tool reliability, the number of total errors for each evaluator and each SC is summed. The SC that all three evaluators have found violations to is used to calculate the inter-tool reliability. Since Fleiss' kappa can not handle missing data, SC that only one or two of the evaluators found violations to, are excluded in the calculations. The number of occurrences is categorised into 25 different categories. The categories are 0 - 20, 20 - 40 and all the way up to 480 - 500. This categorisation is needed in order to easier calculate the kappa value.

There are online tools for calculating the Fleiss' kappa; in this study the statistics calculator is found online at `http://justusrandolph.net/kappa/`.

## 5 Results and Discussion

As can be seen in Table 4 there are big differences in the results from the evaluators. Wave found in total the least amount of errors (258), and Mauve++ the most amount (2704). Wave seems to be most reluctant to classify errors of the SC, and Mauve++ the most generous if looking at the total number of errors they reported. The website with the least amount of failures of SC is the online bank Avanza. The website with the most failures of SC is Reddit. Since Avanza is a bank it needs to be accessible, otherwise it is risking to get fined. Reddit is in the commercial sector and therefore it is not mandatory for Reddit to comply with any accessibility standard by law. As can be seen in Table 4, svt.se got a low rate of reported errors both from QualWeb and Wave, but since the assessment from Mauve++ generated 890 errors, svt.se got almost close to 1000 violations in total. Avanza is, apart from being the website with the least amount of reported errors, also the website closest to an agreement between the evaluators.

The big difference in the number of errors reported by each evaluator is pointing to that there is a vast difference in how they collect and classify the errors of the SC.

| **Website** | **Mauve++** | **QualWeb** | **Wave** | **Total** |
|---|---|---|---|---|
| 1 | 393 | 464 | 168 | 1025 |
| 2 | 347 | 167 | 3 | 517 |
| 3 | 242 | 194 | 9 | 445 |
| 4 | 14 | 1 | 7 | 22 |
| 5 | 113 | 3 | 52 | 168 |
| 6 | 234 | 0 | 9 | 243 |
| 7 | 890 | 1 | 1 | 892 |
| 8 | 185 | 27 | 5 | 217 |
| 9 | 286 | 250 | 4 | 540 |
| **Total** | 2704 | 1107 | 258 | 4069 |

**Table 4.** Total Number of Errors reported by each Evaluator for each Website

### 5.1 Measuring Fleiss' Kappa

In Table 5 the SC that all three evaluators found violations to is presented. These seven SC were used in the calculation for the inter-tool reliability by using the Fleiss' Kappa. Since the number of occurrences of the errors are in a range from 2 to 499, this data was divided into 25 categories to easier be calculated. The results from the calculations are that the overall agreement is 38.10% for these seven SC. The fixed marginal kappa value is 0.03, with 95% CI for fixed-marginal kappa [-0.37, 0.42]. The value of 0.03 implies poor agreement. To see all the SC that were flagged as errors, see Table 6 in Appendix A.

The SC with most reported errors of these seven SC is SC 1.4.3 which says the contrast between text and foreground is too low. SC 2.4.4 has the second most errors of these seven SC. Violation of SC 2.4.4 means that the purpose of a link or image button can not be determined from the link text alone or that the text is missing. The SC with the third most error is SC 1.3.1, which concerns how the semantic markup is used, that tables are correctly coded and that text labels are associated with form input elements.

During the evaluations there were instances where one evaluator classified an issue as an error and another evaluator classified the same issue as a warning that needs manual inspection. There were also instances where the same line

| SC | Mauve++ | QualWeb | Wave | Total |
|---|---|---|---|---|
| 1.1.1 | 48 | 56 | 24 | 128 |
| 1.3.1 | 285 | 36 | 15 | 336 |
| 1.4.3 | 7 | 436 | 166 | 609 |
| 2.4.1 | 3 | 8 | 2 | 13 |
| 2.4.4 | 8 | 499 | 24 | 531 |
| 3.3.2 | 16 | 3 | 11 | 30 |
| 4.1.2 | 36 | 11 | 2 | 49 |

**Table 5.** Number of Errors for each Success Criteria.

of code resulted in violations to different SC by the different evaluators. For instance all three evaluators flagged for the same line of code in the website for svt.se, but categorised the issues differently. The line of code is:

<a id="tab-focus-top" href="#" tabindex="-1" aria-hidden="true"></a>.

By QualWeb this was flagged as a failure of SC 2.4.1 (meaning the first focusable control on the Web page do not link to the top of the page). Wave flagged for the same line of code but as a failure of SC 2.4.4 (which means that there is a link with no text). Mauve++ classified the same line of code as a failure of SC 1.3.6 (ARIA landmarks are not used to identify regions of a page). This was an interesting finding and gives an example of how the evaluators differ in testing for accessibility barriers.

### 5.2 Usability analysis

The usability analysis is based upon experiences while collecting the data for this paper and not a formal procedure. While conducting the evaluations both Mauve++ and QualWeb stopped working for a few days. QualWeb could not analyse the YouTube website as intended and that website was thereby excluded from the data set. Wave can be confusing to use since the errors are displayed as icons on the inspected webpage. Wave seems to classify more issues as alerts than errors compared to the other two tools. The user interface of Mauve++ was the easiest to understand, and Mauve++ gives the option of getting the results in a pdf, to see them live on the website or in the code, which can be beneficial. A benefit with using Wave is that it gives an explanation of how to fix the error in question.

## 6 Conclusion and Future Work

### 6.1 Summary of Results

Nine websites were evaluated and the evaluation results were used as test data to compare the three AWAETs. The SC that all the evaluators found errors to were used in order to calculate the inter-tool reliability by using the statistical measurement called Fleiss' kappa. The standard used in the comparison was

WCAG 2.1. Based on this research the inter-tool reliability for the three AWAETs is poor. There is a vast difference in how many errors the evaluators find on a website, and there are discrepancies in how the errors are classified by the tools. From a usability perspective Mauve++ was most helpful and easiest to use, but had stability issues. Wave is most helpful in explaining how to fix the issues found on a website. This study shows the limitations of only relying on a single AWAET when doing an assessment. In web development this is important knowledge. It is crucial to have the skills to manually analyse websites from an accessibility perspective, since using automated evaluators has severe limitations. Even though WCAG 2.1 is made to be more testable, this research shows that the tools have difficulties in arriving at the same conclusion from the evaluations.

### 6.2 Limitations

This methodology for comparing the evaluators has several limitations. Firstly, even more websites could have been included to get a bigger sample. Secondly, the methodology of dividing the errors into 25 categories can possibly give artefacts and affect the agreement result. Different statistical measures could be used for calculating the kappa value. A limitation with the Fleiss' kappa is that it can not handle missing data, so the calculations are based on the seven SC where all the evaluators had an output. It would be interesting to use the Krippendorffs' alpha and see if this would yield a different result.

### 6.3 Future Work

Since this research has not looked into which evaluator classify the most errors correctly, this could be a topic for future studies. This could be done by having a human expert analysing the results from the evaluation. Another topic could be to look more into the alerts and passed SC and compare how they differ between the tools. It could also be good to use other statistical measures on this data such as the Krippendorff's alpha and see if that would yield any further insights. It could also be interesting to include in websites that are known to be poor from an accessibility perspective to the test data, and see how the evaluators would perform on these.

## A Appendix

In Table 6 is all the SC and number of errors found by each evaluator from the evaluations of the nine websites.

## References

1. Frazao, T., Duarte, C.: Comparing accessibility evaluation plug-ins. In: Proceedings of the 17th International Web for All Conference. (2020) 1–11

| **SC** | **Mauve++** | **QualWeb** | **Wave** | **Total** |
|---|---|---|---|---|
| 1.1.1 | 48 | 56 | 24 | 128 |
| 1.2.1 | - | 4 | - | 4 |
| 1.3.1 | 285 | 36 | 15 | 336 |
| 1.3.5 | 9 | - | - | 9 |
| 1.3.6 | 561 | - | - | 561 |
| 1.4.1 | 651 | - | - | 651 |
| 1.4.3 | 7 | 436 | 166 | 609 |
| 1.4.4 | 5 | 8 | - | 13 |
| 1.4.5 | 280 | 8 | - | 288 |
| 1.4.8 | 57 | 9 | - | 66 |
| 1.4.9 | 4 | 8 | - | 12 |
| 1.4.10 | 349 | - | - | 349 |
| 1.4.11 | 201 | - | - | 201 |
| 1.4.13 | 72 | - | - | 72 |
| 2.1.1 | - | 2 | - | 2 |
| 2.4.1 | 3 | 8 | 2 | 13 |
| 2.4.4 | 8 | 499 | 24 | 531 |
| 2.4.6 | - | - | 13 | 13 |
| 2.4.7 | 48 | - | - | 48 |
| 2.4.10 | - | 5 | - | 4 |
| 3.1.1 | 4 | - | 1 | 5 |
| 3.2.2 | 2 | 14 | - | 16 |
| 3.3.2 | 16 | 3 | 11 | 30 |
| 4.1.1 | 59 | - | - | 59 |
| 4.1.2 | 36 | 11 | 2 | 49 |

**Table 6.** Number of Errors for each Success Criteria.

2. Alsaeedi, A.: Comparing web accessibility evaluation tools and evaluating the accessibility of webpages: Proposed frameworks. Information **11**(1) (2020)
3. Vigo, M., Brown, J., Conway, V.: Benchmarking web accessibility evaluation tools: Measuring the harm of sole reliance on automated tests. In: Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility. W4A '13 (2013) 1–10
4. Molinero, A., Kohun, F.: Reliability in automated evaluation tools for web accessibility standards compliance. Issues in Information Systems **7** (01 2006)
5. Randolph, J.: Free-marginal multirater kappa (multirater Îºfree): An alternative to fleiss fixed-marginal multirater kappa. Volume 4. (01 2010)
6. Abduganiev, S.: Towards automated web accessibility evaluation: A comparative study. International Journal of Information Technology and Computer Science **9** (09 2017) 18–44
7. Parvin, P., Palumbo, V., Manca, M., Paternò, F.: The transparency of automatic accessibility evaluation tools. In: Proceedings of the 18th International Web for All Conference. W4A '21 (2021)

# Optimizing the accessibility of e-learning for older adults

Emmy Gulliksson

Department of Computing Science
Umeå University, Sweden
id18egn@cs.umu.se

**Abstract.** During COVID-19, the use of e-learning platforms increased substantially. Schools and businesses chose to educate online to reduce the risk of people being infected. Unfortunately, learning online and using e-learning platforms is a struggle for many. Especially people aged 60+ who are not as used to technology as the younger age groups who have grown up with the developing technology. This paper contains a literature review summarizing existing findings of research papers and articles about what attributes that should be considered to optimize an e-learning platform for older adults.

## 1 Introduction

This research paper contains a literature review summarizing previous research to identify the attributes that should be considered to optimize the accessibility of online learning for older adults. In this paper, older adults are considered people aged 60+.

Online learning is also referred to as e-learning and electronic learning. In an article named "*What is e-learning?*" [1], the author Tamm describes e-learning as: "*Learning that is enabled electronically*". It is typically implemented on the Internet hence the learning materials are accessible at any time, as long as the user has an Internet connection.

E-learning comes in many shapes and forms but is most commonly performed through online courses, online degrees, or online programs. One example is the Coursera platform[1] that for instance includes interactive certificate programs and language learning. Another one is the learning platform Udemy[2]. The platform is for businesses and contains online courses, webinars, e-Books and much more.

The possibility to retrieve knowledge online has resulted in a much more flexible way to learn and people who had geographical obstacles that can come with the traditional education now have the possibility to educate themselves.

[1] https://www.coursera.org/, Coursera homepage, an e-learning platform.
[2] https://business.udemy.com/, Udemy business homepage, an e-learning platform for businesses.

L. Jiang, A. Jonsson, L. Vanhée (Eds.): Umeå's 25th Student Conference in Computing Science USCCS 2022, pp. 27–38, January 2022.

Due to COVID-19, online education has increased substantially [2]. Likewise, has most businesses chose to use e-learning as their learning method to make it possible for the employees to educate themselves without having to meet people and risk being infected. It is important that all employees are comfortable with using online learning programs, but unfortunately, the age group of 60+ are overrepresented to be struggling with it [3]. This research paper should help designers create usable platforms by summarizing the research existing today.

The goal of this paper is to use the findings of existing work to summarize how to optimize the accessibility of e-learning for older adults. The research question for this paper is: *According to previous research, what attributes should be considered to optimize the accessibility of an e-learning platform for older adults?*

In an article written by MDN contributors accessibility is described as the practice of making websites usable for as many people as possible [4]. They write that it is traditionally thought of as being about people with disabilities, but the practice of making sites accessible also benefits other groups such as those using mobile devices, or those with slow network connections. If a website has good accessibility, everyone will have equal opportunities when using it, no matter what their ability or circumstance is. In this paper, that is how accessibility is thought of. When talking about the accessibility of an e-learning platform, it means how usable it is for older adults.

To optimize accessibility, the barriers that older adults have to face today must be addressed. Research shows that some of the barriers are that their physical and cognitive abilities needed when using a computer have slowed down, or that they show a lack of trust in the Internet, they are also faced with negative stereotypes, may have educational barriers, etcetera. This paper focuses on the kind of barriers that can be addressed by an organization or an individual, for example, the content and usability of the platform. The literature search was done to find possible solutions for those barriers. Literature that was considered to have high quality was chosen and its findings were read in order to find how to overcome the barriers.

The final research paper contains a brief explanation of what e-learning and accessibility are and further contains an analysis of what barriers are known in e-learning for older adults, for a better understanding of the problem. Finally, the paper contains a summary of the results from the literature chosen during the literature search. Those papers are existing publications taken from the different web pages with scientific articles and papers on the subject. That literature contains research about optimizing the accessibility of e-learning for adults by overcoming the existing barriers.

## 2 Background

Even though e-learning equals possibilities and flexibility for many, some are struggling to use it. Older adults aged 60+ belong to the age group that is known to be having the most problems with e-learning. Since accessibility means the practice of making websites usable, there is a need to identify the problems that exist in e-learning for older adults today. Many researches have been done to find the barriers in e-learning for older adults [5, 6, 7]. This section has an analysis of the current barriers in e-learning that older adults are facing today. The result is written in consideration of the detected barriers.

Notess and Lorenzen-Huber argue in their article that there are stereotypes that opine that older people are unable to learn how to use the Internet, or being negative and anxious about Internet use, to be a myth [6]. Despite that, barriers were found in the Internet use by seniors. They grouped the barriers into three categories of cohort differences, normal age-related changes, and stage of life. Notess and Lorenzen-Huber found that ageing implies that the physical and cognitive abilities are slowing down, which affects the capability to use a computer. When looking at the cohort differences, they found that the major reason why older adults do not use the Internet is because they lack access to a computer and/or broadband opportunity. There is also a barrier when it comes to trust. They write that "*while users of all ages need to evaluate websites for trustworthiness, inexperienced older users can be easy targets for Internet scams and schemes. Some of today's older adults may lack trust in the Internet in general and consequently distrust the content of online courses*". One last cohort-based barrier shown in the article was that many of the older adults who do not use the Internet believe to have no use for the Internet, thereby online learning is not be perceived as useful to them.

P.Githens studied e-learning programs for personal growth and social change, workplace learning and workforce development [5]. He also found that there are barriers to e-learning for older adults. Barriers that could hinder the older adults' full participation are issues such as negative stereotypes as well as issues caused by bad usability and interface design. P.Githens writes about a study that showed that negative stereotypes considerably reduced older adults' memory performance. However, it was shown that positive stereotypes had a positive effect on memory performance. Another barrier that was found due to negative stereotypes was the lack of training opportunities for older people in workplace settings. One problem lies in biases causing managers not no "waste money" on improving older workers. Another problem is that some older employees are unwilling to admit that they need or want to participate in learning events. This has also led to many elderly being anxious and lacking confidence when it comes to technical things like e-learning. Class and educational barriers were also found meaning that those who worked, or are working, in blue-collar jobs are less likely to join adult learning activities. Technical barriers can also be a recurring barrier and source of frustration for online students of all ages. That especially applies to people that have difficulties using computers. Another barrier that P.Githens mentions is the usability issues and problems in course design. The elderly can have difficulties hearing and seeing, which is something that the designers tend to forget. Undesirable features and inappropriate sequencing of courses can result in frustrated students, regardless of age. Lastly, the author mentions a problem with new technology, such as using games in e-learning programs. That often means that the user needs to proceed rapidly, with less time for thought and contemplation, which can be tough for older adults and create usability obstacles [8].

The most recent article written by A. Pappas et al. in 2019 argued that there still are barriers in e-learning for the elderly today [7]. The article is about a study whose purpose was to analyze the cognitive profile of older adults. This was made to identify their ways of learning, and also to analyze their attitudes, with the purpose to support the development of an e-learning platform adapted to the older adults' necessities. They mentioned similar barriers as the previously mentioned articles, that the barriers which hinder older adults to join online learning activities appear to be similar to those that hinder their usage of computers. Another finding of the study is that persons who have a minor or no use of technical devices are in danger of social exclusion and loss of social integration. Advanced age is often related to inequalities regarding access, intensity, and use of everyday technology. Finally, one last finding of the study was that a limited amount of education for older adults implies a risk for dementia. Contrariwise, older adults participating in education could give them protection against dementia. That shows the importance of including older adults when implementing an e-learning system.

Below is an assessment grid summarizing the detected barriers mentioned in this section.

| **Barrier** | **Description** |
|---|---|
| Physical and cognitive issues | It is slowing down, which affects the capability to use a computer. |
| Availability | They lack access to a computer and/or broadband availability. |
| Trust issues | Older adults may lack trust in the Internet in general and consequently distrust the content of online courses. |
| Negative stereotypes | Can significantly reduce memory performance in older adults. |
| Design issues | Such as usability and interface design, usability issues and problems in course design. The elderly can have difficulties hearing and seeing, which is something that the designers tend to forget. |
| Lack of training opportunities | In workplace settings, due to negative stereotypes. |
| Class and educational | those who worked, or are working, in blue-collar jobs are less likely to join adult learning activities. |
| Technical barriers | They lack basic technical knowledge. |
| Games in e-learning programs | That often means that the user needs to act fast, which means less time for thought and reflection. That can be tough for older adults and create usability obstacles. |
| Limited supply | There is a limited amount of education for older adults today. |

**Table 1.** The detected barriers in online learning for older adults.

## 3 Earlier work

Beyond the articles on barriers in e-learning exists more papers that explore the older adults relationship to learning and technology. These papers contribute by showing successful learning methods and findings on adults relationship to technology.

One of the papers is about adult learners' satisfaction in online education programs, written by Zher Ng and Sofia Baharom [9]. They performed tests on 200 adults to find out what items can be good predictors of the learner's satisfaction. A group at the University of Zaragoza has made a research paper providing a list of manuals and standards that could be considered to develop educational environments for older users [10]. They show the importance of, and how to, adapt services to meet the need of older people by conducting an extensive literature search to find e-learning educational environments adapted to older users.

In another article, L. Willis explores how technology usage can play a future and current role in the learning activities for older adults [11]. The primary interest lies in learning activities that are encountered in everyday life and that have to do with (1) the loss of computers and the Internet to acquire knowledge and skills, and (2) the knowledge and skills required to use computer and Internet technology.

Finally, an article could be found regarding e-learning through Augmented Reality for older adults [12]. It contains research about older adults' relationship with technology.

The findings of these articles can be used in this paper to gather solutions for the existing barriers in online learning for adults. This paper will also contribute to a wider image of what learning methods are successful.

## 4 Method

The purpose of this paper is to use existing literature to answer the research question: *According to previous research, what attributes should be considered to optimize the accessibility of an e-learning platform for older adults?* To answer the question, an analysis on how to overcome the barriers in e-learning for older adults was done. Based on that, the attributes needed to overcome the barriers are gathered and summarized in the result. This was done by a qualitative literature review, meaning that the literature chosen for the review is well-chosen, critically analyzed, and examined. The priority is to find valuable articles rather than finding many articles. The review should result in the advancement of knowledge about the research question, rather than a simple overview of the research area.

### 4.1 Literature search

Already existing articles or research papers on the subject are used to answer the research question. The material used for finding literature on the subject is different databases containing published scientific literature, such as Research Gate [3], Google Scholar[4] and Semantic Scholar [5].

Search strings are used for finding articles. Boolean operators, truncation and keywords are used in the search. The keywords are *e-learning, online learning, older adults, elderly, seniors, barriers, obstacles, accessibility, usability.* The Boolean operators "AND" and "OR" are used in the search, together with the truncation "*" that brings up different variations of a word.

The articles are selected by reading the titles and abstracts of each article to decide its relevancy. When the selection was done, there was time for further reading of the paper. The focus was on the findings of the paper, as that is what contributes to answering the research question. As an additional strategy to find relevant papers, references in the selected literature were being examined to identify the potential relevancy of those articles.

### 4.2 Choosing articles

The selected literature was analyzed to determine the quality of the papers. To check the quality, a checklist was used that examines these aspects:

**Value:** The literature needs to bring value to the research, meaning that it must describe necessary aspects of an e-learning platform for adults.

**Spelling, grammar, punctuation:** If the article contains a lot of these kinds of mistakes, it will be considered unprofessional and unreliable.

**Readability:** It should be well structured and have a logical sequence of the content.

**Accuracy:** The literature should not be older than 15 years. The findings of older research papers or articles will not be useful due to the fast-developing technology. That implies that people's relationships with technology are much different today compared to 15 years ago. Furthermore, the behaviour of the target group, older adults, will differ a lot in research made over 15 years ago compared to today. They are more familiar with the technical devices today than 15 years ago [13].

### 4.3 Compilation

The compilation began when 10 papers and articles that contributed to the research question were gathered. The different attributes were divided by the different kinds of barriers. Then a summary of the findings on how to overcome each barrier in e-learning for the elderly was written. Finally, a conclusion of the findings from the literature was created which discusses the main results, limitations, and possible future work.

[3] `https://www.researchgate.net/`, Research Gate homepage, an archive for publications and researches.

[4] `https://scholar.google.com/`, Google Scholar homepage, contains a huge number of scientific articles

[5] `https://www.semanticscholar.org/`, Semantic Scholar homepage, A research tool for scientific literature

## 5 Result

The accessibility issues are earlier identified as barriers in e-learning for older adults and are managed in this section. In this section, the attributes to consider when designing an e-learning platform for older adults are analyzed. They relate to the findings on the existing barriers. The results are divided into four different subsections containing solutions to the different barriers.

### 5.1 Overcoming barriers

There are barriers mentioned that can not be handled merely by improving the accessibility and can not be solved only by an organization or an individual. Such as the lack of access to a computer and/or broadband availability or the lack of training opportunities for older people in workplace settings due to negative stereotypes.

The problems that are addressed in this paper are the ones regarding accessibility. For example, barriers mentioned are that elderly can have difficulties hearing and seeing, which is something that the designers need to keep in mind. Also, many exercises on e-learning platforms mean that the user needs to proceed rapidly, with less time for thought and contemplation, which can be tough for older adults and create usability obstacles. These kinds of issues are managed in the chosen articles and summarized. The data gathered are divided by the different barriers that are found in the sections below.

**Negative stereotypes and bad confidence** were two of the barriers that older adults are facing. Macek writes in an article that the issue of bad confidence can be managed by meeting the student with a warm and inviting welcome announcement [14]. Macek wrote that they should be able to feel like they can communicate their concerns without reproach. One solution suggested was to ask the students to post to an "introduction" discussion forum. That allows students to exchange personal information and gain support from their fellow students.

Another study found that in order to make the older adults more confident they want a self-directed or personalized-learning approach [8]. The education activity should be performed in an informal learning environment, offering a flexible program that entails the possibility to execute the learning modules at their own pace and of their own choice. That can be accomplished by giving the student time and space to repeat and absorb the newly collected material.

**Consider their cognitive abilities.** A more self-directed and personalized-learning approach also favours the barrier considering that the senior's physical and cognitive abilities are slowing down [8]. Because of that, the material should be presented slow enough with a recap [15]. Further, they need good ways to remember the new material. In an article by Richard David Sheridan, he writes that the best way for seniors to learn is to link the material to as many different bits of intelligence as possible. Seven different cognitive approaches are recommended. The first one considers linguistic intelligence which can imply talking, reading, or writing about what is being learned, keeping a journal, interviewing an expert, or creating a mnemonic to help memorize the learning materials. The second intelligence to take into account is kinesthetic. This can mean finding an activity related to the topic being taught. The third intelligence mentioned is musical. It is favourably to play background music while learning the materials and using music as a way to remember facts or skills. The fourth one to consider is the intrapersonal aspect, relating the material to a personal feeling or inner experience. The fifth intelligence is logical-mathematical. It means conceptualizing, quantifying, or thinking critically about the new material. Interpersonal intelligence is the sixth cognitive approach to take into account. It implies inviting the student to determine how they feel about something and why. The last one is the interpersonal intelligence that exhorts the student to do a project with a partner [16].

To make the material easy to absorb, studies found that the e-learning modules should be short and extensive. Further should the educational material be divided into small modules [7, 17, 18, 8]. The learning material should neither have loads of text since the older learners find it hard to process large texts. The reason comes as a consequence of age, implying deteriorating cognitive skills and growing anxiety [7].

In another article, the authors argue that factors related to the learning process are motivation, metacognition, and self-regulated learning. By personalizing the learning content to the students' cognitive styles may facilitate the memorization of items and their memory [19].

**Usability and interface design.** Consequences of the deteriorating physical and cognitive abilities are difficulties hearing and seeing, which is something that the designers tend to forget. That motivates the need for a simple graphical user interface design. The interface should not use bright colours or excessive graphics. A straightforward navigation design is important to have for the online modules to reduce distractions [7]. Many other types of research have resulted in the same, that the older adults need a simple user interface [6, 20, 21, 7, 16].

An e-learning platform often consists of many different sites and if the interface design is not consistent, it can be difficult for the seniors to use. In Sheridan's research, he found that currently, the online courses are too complex for widespread adoption by older adults. He writes that for example, the user would have to find the site and negotiate through any login requirements. Then they had to navigate through the learning modules, case studies, and other components [16]. To ease the use of multiple sites, Notess and Lorenzen-Huber suggest having provisions on standards for concepts, terminology, layout, and navigation. By standardizing authentication, form fill-in and submitting, discussion browsing and participation the use of online learning platforms will be facilitated [7].

Other consequences of deteriorating physical and cognitive abilities were mentioned by Sheridan. He pays attention to mobility problems and mentions that there are many other restrictions in their ability to use an e-learning platform. In his research, he identified the problematic fact that web and print designers are often younger individuals who tend to develop their products for their younger peers. To help the web designers build sites that are easy to use for seniors, he made some recommendations. The first one is to use a large font size on the screen, 14 or larger. The sans-serif fonts, such as Arial, are easier to read. Some other tips are to maximize the contrast of characters to the background, minimize glare with a glare guard, and minimize clutter and irrelevant information. Another recommendation is to use multiple sensory modalities to communicate information. He also writes that the designer should highlight the important information on the screen and provide navigational aids as much as possible. Further, they should keep software procedures simple. Older adults may have difficulty remembering commands and complex procedures for performing tasks when using the software. Provide on-screen reminders of basic commands. One last recommendation from Sheridan was to consider that there will be a need for special assistive technologies such as screen readers, voice recognition, or optical readers [16].

**Finding suitable learning content** To ease the processing of the material of an e-learning platform, the content needs to be well organized and presented. Researchers pay attention to aspects of the layout of the learning process, the content of the modules, and the type of learning material.

Regarding the procedures of the process, the seniors want exercises after the end of an online course that clarifies their learning needs. They also desire clear learning goals presented before the course start. The e-learning environment should consequently provide an automated judging system for self-evaluation. One great resource mentioned for that is quizzes. Further are comprehension exercises, practical questions, assignments, etc. shown to be useful in research by A. Pappas et al. in 2018 [7]. They also argue that the learning modules are desired to be short in length. This is shown in several other types of research as well [17, 18, 22]. The modules should present the educational content divided into modules containing practice questions and examples. One arrangement shown to be desired by the older adults is to have assessment tests after the completion of each module. Further preferences found in that test were modules containing explanatory videos and special graphics [7]. Another study suggested that the short modules should use references to past events and build point goals that demonstrate the usefulness of each module for the elderly [17]. The importance of motivating the students was shown in research made by Notess et al. They found that older adults are not likely to participate in online learning activities that feel boring or lack motivating content. In their survey result, the interviewees demanded a better rating of online learners' achievement and satisfaction, clearer reward systems and better motivation for completing online learning. Another wish was to receive training in how to be an online learner [6]. Those features are also mentioned as important arguments to make a course usable according to Zaharias [22].

Another suggestion to make the modules more motivating for the student is to add an even more personalized touch to the teaching or let the elderly create content by themselves. However, that requires some basic technology skills. Another suggestion mentioned was to use well-designed, structured, and guided workshop activities in the modules along with the use of collaborative learning [15].

The type of learning material found to be most appreciated by the older adults was a more experience-based and active learning method rather than a passive one. Meaning that they felt positive towards the incorporation of practical exercises and valuation methods. Usage of practical-oriented learning material with examples and practise questions are desired. Therefore it is important to not isolate the learning content from its practical context [7].

Regarding the learning process, self-directed learning which is a combination of approaches suited to the individual learner's needs and capacity was argued to be a preferred approach to e-learning for the elderly [16].

## 6 Conclusion

The research question was answered with reference to 4 categories of barriers. Those barriers were the negative stereotypes and bad confidence, cognitive abilities, usability and interface design, and learning content. The most interesting data found in each category are summarized below together with the limitation of the findings and future work.

### 6.1 Summary of results

The research question was answered with different barriers in mind. The first one considered the bad confidence of the older adults and the negative aspects that they are facing. The e-learning platform needs to have a warm welcome where the student is introduced to the other students. Further, the platform should have a personalized-learning approach to make the student comfortable to overcome the barriers.

When considering the older adults' cognitive barriers, it should be faced by keeping in mind that their physical and cognitive abilities are slowing down. That can also imply a more personalized-learning approach as well as presenting the learning materials in smaller sections, slow enough with a recap.

When designing the interface, the deteriorating physical and cognitive abilities need to be addressed as well. Many types of research argue that a simple interface is crucial. The designer should use large font sizes and high contrasts and avoid using bright colours and excessive graphics.

Regarding the content of the platform, findings showed that the content needs to be well organized and presented. It should ease the processing of the learning material on the e-learning platform. Researches argued that the seniors appreciated an active and experience-based learning approach. Thereby the importance of not isolating the learning content from its practical context.

### 6.2 Limitations and future work

In the section about barriers that adults face today are more barriers mentioned than the ones addressed in the result section. The reason is that the solution to the barriers in the result is the ones that can be addressed by an organization or an individual. This paper does not focus on barriers that concern subjects such as the class and educational barriers and their lack of trust in the Internet in general.

One limitation to this literature review is that there are only 10 different literature reviews due to lack of time. A possible future work is to use more literature publications in the review. Another suggestion for future work is to focus merely on the designing of the interface of an e-learning platform. Then consider specifically an e-learning platform and how it is supposed to be designed. It was difficult to find literature on interface design solely focusing on e-learning platform sites when doing the search. One last suggestion is to find solutions to the barriers that are mentioned in this paper but not addressed in the result section.

## References

[1] Sander Tamm: What is e-learning? (2020)
[2] Cathy Li and Farah Lalani: The covid-19 pandemic has changed education forever. this is how. (2020)
[3] Alexander Seifert and Shelia R Cotten and Bo Xie: A double burden of exclusion? digital and social exclusion of older adults in times of covid-19. **76**
[4] MDN contributors: What is accessibility? (2021)
[5] Rod P.Githens: Older adults and e-learning: Opportunities and barriers. Quarterly Review of Distance Education **8(4)** (2007)
[6] Mark Notess and Lesa Lorenzen-Huber: Online learning for seniors, barriers and opportunities. **2007(5)** (2007)
[7] Marios A. Pappas and Eleftheria Demertzi and Yannis Papagerasimou and Lefteris Koukianakis and Nikitas Voukelatos and Athanasios Drigas: Cognitive-based e-learning design for older adults. **8(6)** (2019)
[8] Ralph Reimann and Yvonne Prinzellner and Claus Rainer Michalek and Ulrich Haas and Thomas Guggenberger and Mogens Andersen and Pauline Mieleszko and Milan Slavik and Veronika Pfaffenberger and Eqward Pierzgalski: E-learning: Didactical recommendations and quality assurance: An overview. (2012)
[9] Huey Zher Ng and Sakina Sofia Baharom: An analysis on adult learners' satisfaction in online education programmes. **12(7)** (2018)

[10] Inmaculada Plaza and Raul Igual and Carlos Medrano and Angel Sanchez and M Pilar Perez and Cristina Lopez: Education and training for older people: Preparing the future. **44003** (2007)
[11] Sherry L. Willis: Technology and learning in current and future generations of elders. (2006)
[12] Sergio Cicconi and Maurizio Marchese: Augmented learning: An e-learning environment in augmented reality for older adults. Proceedings of INTED2019 Conference (2019)
[13] Aaron Smith: Older adults and technology use. (2014)
[14] Jo Macek: 7 student myths of the online classroom. (2013)
[15] Minnamari Naumanen and Markku Tukiainen: Practises in old age ict-education. (2008)
[16] Richard David Sheridan: The language of digital learning: Developing an e-learning approach for the elderly. (2010)
[17] Leticia Rocha Machado and Patricia Alejandra Behar and Johannes Doll: Distance education for elderly: The construction of a pedagogic architecture. (2012)
[18] Simone Kimpeler and Peter Georgieff and Christoph Revermann: elearning for children and elderly people. (2007)
[19] Valeria de Palo and Pierpaolo Limone and Lucia Monacis and Flavio Ceglie and Maria Sinatra : Enhancing e-learning in old age. **58** (2018)
[20] Moulton, G.M.: Build it and they will come? (2007)
[21] National Telecommunications and Information Administration (NTIA) and Economics and Statistics Administration: A nation online: How americans are expanding their use of the internet. (2007)
[22] PANAGIOTIS ZAHARIAS: Usability and e-learning – the road towards integration. (2008)

# Mixed Connectivity in Series-Parallel Graphs

Markus Kauf

Department of Mathematics
University of Umeå, Sweden
maka0958@student.umu.se

**Abstract.** The concepts of edge- and vertex-connectivity are very well studied (see e.g. Menger's Theorem) and can be solved in polynomial time. Mixed Connectivity is a combination of both, first introduced by Heineke and Barary. Here one can remove edges and vertices at the same time to disconnect the graph. For an arbitrary given pair of integers $k$ and $l$ and an arbitrary graph the question whether disconnecting is possible by removing up to $k$ edges and up to $l$ vertices is NP-hard. In this paper we focus on the research question whether the problem is still NP-hard if we restrict it to a special subclass of graphs, the Series-Parallel Graphs. We show that then it indeed can be solved in polynomial time.

## 1 Introduction

Connectivity is a well-studied concept in graph theory, and it has various applications in, e.g., infrastructure planning. Generally and informally said the problem of Graph Connectivity is about the stability of networks. For example, given an electricity circuit that distributes power from one source to a number of customers, does this network stay connected even if some wires or switches break down? The electrical circuit is modelled as a graph in which the nodes represent switches and the edges represent wires. Edge- and vertex-connectivity is intuitively a measure on how many wires we can cut and switches we can turn off, respectively, without any customer losing power. Both problems are already very well studied. As early as 1927 Karl Menger gave a characterization by which a graph can be examined efficiently with respect to the question of edge- or vertex-connectivity. Here we consider the combination of both concepts, what we call mixed connectivity. This is of course also interesting in applications since both wires and switches can break down simultaneously. Mixed connectivity was first introduced by Beineke and Harary in 1967 [1] but there was not much known about it until recently when Johann, Krumke and Streicher [2] and Bonnet and Cabello [3] showed that mixed connectivity is much more complex and cannot be solved efficiently on general graphs. They distinguished between Global Mixed Connectivity where the question is essentially "what is the least number of vertices and edges that we have to remove from a graph in order to disconnect it?" and Rooted Mixed Connectivity where we are given two fixed vertices $s$ and $t$ in the graph and we ask "what is the least number of vertices and edges that we have to remove from the graph in order to disconnect $s$ and $t$?". Both have been

L. Jiang, A. Jonsson, L. Vanhée (Eds.): Umeå's 25$^{th}$ Student Conference in Computing Science USCCS 2022, pp. 39–55, January 2022.

shown to be NP-hard for general graphs. Though there is one problem with our questions above: what does it mean to have least number of vertices and edges? Since there is only a partial order on the pairs of integers $\mathbb{N}^2$ we ask for all pairs that are minimum with respect to this partial order. We come back to this later. For now we focus on the aspect of NP-hardness.

The NP-hardness of the mixed connectivity problems means that the time it takes for an application to solve them solutions to them can be exponential in the worst case. Ways to handle this can be computing approximations instead of exact solutions or for a special type of graph determining its structure and making use of that. Choosing the second option one often considers planar graphs, regular graphs, bipartite graphs or series-parallel graphs. For an early work of NP-hard problems efficiently solvable on series-parallel graphs see [4]. In this paper we work out how to solve Mixed Connectivity efficiently on the class of series-parallel graphs.

The main reason why we can solve Mixed Connectivity efficiently on the class of series-parallel graphs although it is NP-hard in general is their special structure. Especially for a series composition the question how to disconnect the resulting graph is quite easy to answer, just remove the vertex in the middle. For a parallel composition it is a little bit more difficult, there we have to remove two vertices in order to disconnect the graph, the source and the sink terminal. In the case of Rooted Mixed Connectivity we have one more restriction, of course we cannot remove one of our given vertices $s$ or $t$, that is not disconnecting. Thus, this case is where we have to spend the largest effort, what to do when we have a parallel composition and $s$ and $t$ are exactly the source and sink terminals. Intuitively the idea is that for each subgraph connecting them essentially there must have been either a series composition at some point in time and hence we can disconnect this subgraph by removing the middle vertex of this series composition. (The subgraph could also just consist of one single edge connecting $s$ and $t$ but that is easy to deal with.) Formalizing this becomes a bit technical unfortunately.

In this paper first, in Section 2, we recall edge- and vertex-connectivity to make it easier to understand the following, highlight that those problems can be solve efficiently for any graph and also introduce series-parallel graphs. The problem of Mixed Connectivity, which we introduce in Section 3, in contrary is NP-hard. That is why in Section 4 we concentrate on the class of series-parallel graphs and describe our algorithm for mixed connectivity at least on this class of graphs. Afterwards in Section 5 we conclude the paper.

## 2 Theoretical Foundation

For the beginning some notation: In the following we always consider undirected graphs $G = (V(G), E(G))$ where $V(G)$ is a finite set, the vertices of $G$, and $E(G) \subset 2^{V(G)}$ are the edges of $G$. For the sake of simplicity we assume that $G$ is connected and has no loops and no parallel edges. Moreover, for the graph $G = (V, E)$ and $W \subseteq V(G)$ a subset of the vertices $G[S] := (S, \{e \in E(G) :$

$e \subset S\}$ denotes the graph induced by $S$ and for $F \subseteq E(G)$ a subset of the edges $G - F := (V(G), E \setminus F)$ denotes the graph after removing the edges $F$. Furthermore the partial order on $\mathbb{Z}^2$ is defined as follows: $(a, b)$ is smaller or equal than $(c, d)$ with respect to the partial order if it is elementwise smaller, ie. $a \le c$ and $b \le d$.

### 2.1 Vertex- and Edge-Connectivity

A graph $G$ with at least $k$ vertices is called **$k$-vertex connected** if the graph remains connected after one removes up to $k-1$ arbitrary vertices, more formally if for every subset of the vertices $W \subseteq V(G)$ of size at most $k-1$ $G[V(G) \setminus W]$ is still connected. The **(global) vertex connectivity** is the largest number $k$ such that the graph is $k$-vertex connected, i.e. $k$ is the smallest number of vertices whose deletion disconnects $G$. See Figure 1 for an example. In contrast to this global view one can also consider disconnecting two given vertices: For two non-adjacent, distinct vertices $s$ and $t$, the **roots**, we ask whether the removal of up to $k-1$ vertices distinct from $s$ and $t$ leaves the roots $s$ and $t$ in the same connected component of $G$. This is called **rooted vertex connectivity** or $(s, t)$-vertex connectivity. In the above example in Figure 1 the $(2, 4)$-vertex connectivity is 3, since the removal of vertices $1, 3$ and 7 disconnects 2 and 4 while the removal of any two vertices does not disconnect the graph. For roots 1 and 7 the rooted vertex connectivity would be 2.

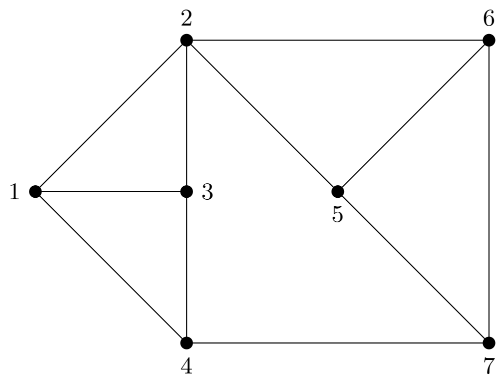



Fig. 1: Graph $G$. It is 1-edge connected (like every connected graph is), 2-edge connected, 3-edge connected, but not 4-edge connected (the removal of e.g. the edges $\{2, 6\}, \{2, 5\}, \{4, 7\}$ disconnects the graph). Therefore the global edge connectivity is 3.

Similarly to vertex connectivity we have **edge connectivity** where we remove edges instead of vertices: A graph $G$ with at least two vertices is called **(global) $k$-edge connected** if the graph remains connected after one removes up to $k-1$ arbitrary edges, that means if for every subset of the edges $F \subseteq E(G)$ of size at most $k-1$ $G - F$ is still connected. Again we also have the **rooted**

**edge connectivity** for two given, non-adjacent and distinct vertices $s$ and $t$. In the above example in Figure 1 $G$ is 1-vertex connected (again like every other connected graph), 2-vertex connected, but not 3-edge connected (the removal of e.g. vertices 2 and 4 disconnects the graph). Hence, the global vertex connectivity is 2. Moreover, for every choice of $s$ and $t$ the $(s,t)$-edge connectivity is 3.

For the rooted connectivity Menger gave a quite nice characterization using $s-t$-paths. We state some variants, for the proofs we refer to your favourite text book on graph theory or combinatorial optimization, e.g. [5]:

**Theorem 21 (Menger's Theorem, 1927)** *Let $G$ be an undirected graph and $s$ and $t$ two distinct and non-adjacent vertices, $k \in \mathbb{N}$.*

- *There are $k$ edge-disjoint $s-t$-paths if and only if the removal of $k-1$ edges cannot disconnect $s$ and $t$.*
- *There are $k$ internal vertex-disjoint $s-t$-paths if and only if the removal of $k-1$ vertices cannot disconnect $s$ and $t$.*

With that we get as a corollary for global connectivity:

**Corollary 22 (Whitney, 1932)** *Let $G$ be an undirected graph with at least two (for vertex connectivity at least $k$ respectively) vertices.*

- *$G$ is (global) $k$-edge connected if and only if for every pair of vertices $s$ and $t$ there are $k$ edge-disjoint $s-t$-paths.*
- *$G$ is (global) $k$-vertex connected if and only if for every pair of vertices $s$ and $t$ there are $k$ internally vertex-disjoint $s-t$-paths.*

Since for given $s$ and $t$ we can determine the number of edge-disjoint $s-t$-paths via $s-t$-flows in time $O(n^{2/3}m)$ with Dinic's algorithm, we can can solve global edge connectivity in $O(n^2 n^{2/3} m)$ by simply enumerating all choices of $s$ and $t$.

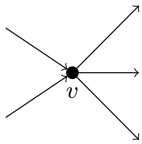
For global vertex connectivity we use the well-known construction of replacing a vertex $v \in V(G)$ by two vertices $v_{in}$ and $v_{out}$ connected by an edge $(v_{in}, v_{out})$ and yielding the directed graph $G'$ with $N_G(v) = N^-_{G'}(v_{in}) = N^+_{G'}(v_{out})$. See Figure 2 for a visualization. Hence, for every path in $G$ traversing a vertex $v$ the corresponding directed path in $G'$ must traverse the edge $(v_{in}, v_{out})$ and we can again use flows as above in the same asymptotic running time.

Thus, Menger's Theorem gives a pretty neat way to solve the problem of edge- and vertex-connectivity in polynomial time with respect to the size of the graph. But when we combine those concepts to Mixed Connectivity the problem becomes NP-hard.

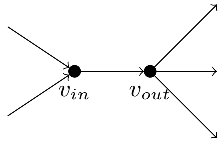
### 2.2 Series-Parallel Graphs

**Definition of Series-Parallel Graphs**

The class of series-parallel graphs has a simple, recursive structure and many NP-hard problems are solvable in polynomial time on them. We show that this also holds for Mixed Cut. Intuitively series-parallel graphs are very strong related




**a)** vertex $v$ in the given graph $G$




**b)** the corresponding construction in $G'$

Fig. 2: The construction of $G'$ in order to get vertex-disjoint paths via flows.

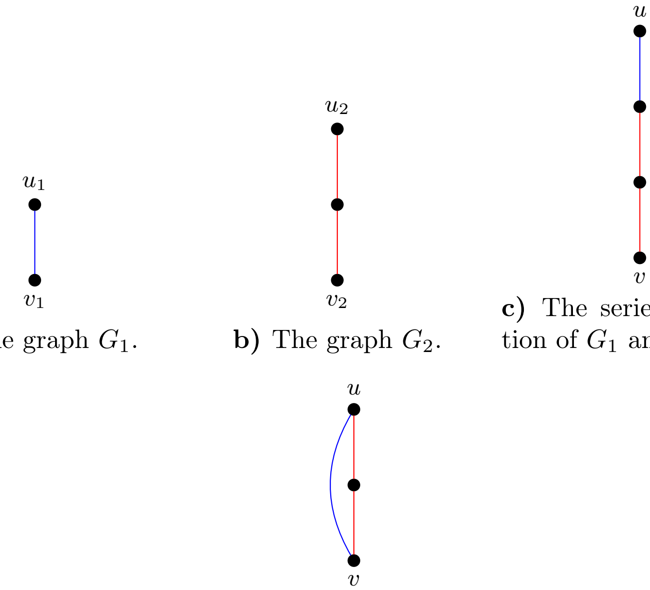
to the structure of series and parallel electricity circuits. The definition we give here is based on the definition in the book of Krumke and Noltemeier [6], first they were introduced by David Eppstein in 1992.

A series-parallel graph has two highlighted vertices, the source terminal and the sink terminal, and it is formed by recursive series- and parallel-compositions. Therefore the class of series-parallel graphs is also very strong related to the structure of series and parallel electricity circuits. To be more precisely a graph $G$ consisting only of two vertices $u, v$ and one edge connecting them is **series-parallel** with source terminal $u$ and sink terminal $v$ (see also graph $G_1$ in Figure 3a). Moreover, for two edge-disjoint series-parallel graphs $G_1, G_2$, each with source terminal $u_1$, $u_2$ respectively, and sink terminal $v_1$, $v_2$ respectively, we can connect them in series by merging the first sink terminal $v_1$ and the second source terminal $u_2$ to one new vertex. This is called series-composition of $G_1$ and $G_2$ and this graph is again series-parallel with source terminal $u_1$ and sink terminal $v_2$. For the parallel-composition we merge the source terminals and the sink terminals. The merged source terminal is now the source terminal of the new series-parallel graph and the merged sink terminal is the new sink terminal.

In Figure 3 there is a first example for that: Figure 3a shows the series-parallel graph $G_1$ with source terminal $u_1$ and sink terminal $v_1$ consisting of only a single edge, another series-parallel graph $G_2$ with source terminal $u_2$ and sink terminal $v_2$ that is the series-composition of two single edges is depicted in Figure 3b. Their series-composition with source terminal $u = u_1$ and sink terminal $v = v_2$ is shown in Figure 3c and their parallel-composition with source terminal $u = u_1 = u_2$ and sink terminal $v = v_1 = v_2$ in Figure 3d. A more advanced example can be found in Figures 4 and 5. For graph $G_2$ of Figure 4 we go a little bit more into detail. $G_2$ is a series-parallel graph with source terminal $u_2$ and sink terminal $v_2$. It consists of the parallel composition of $G[\{u_2, x_4, x_5, x_6, x_7, v_2\}]$ (on the left side) and a path of length 4 (on the right side). $G[\{u_2, x_4, x_5, x_6, x_7, v_2\}]$ then again is the series composition of the edge $\{u_2, x_4\}$ and $G[x_4, x_5, x_6, x_7, v_2]$. The structure of the last graph we omit here. To describe the structure in a shorter and more succinct way we next introduce the notion of a structure-tree.

### The Structure-Tree of a Series-Parallel Graph

We can also describe the recursive structure of a series-parallel graph, i.e. in




**a)** The graph $G_1$.

**b)** The graph $G_2$.

**c)** The series composition of $G_1$ and $G_2$.

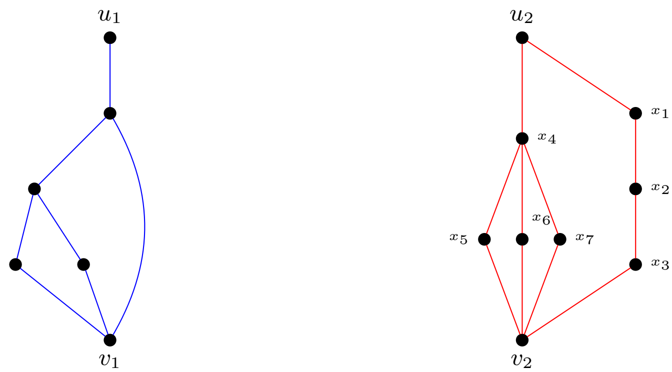
**d)** The parallel composition of $G_1$ and $G_2$.

Fig. 3: A simple example of series-parallel graphs.

how and in which order the series composition and the parallel composition are applied. This can be done by a binary rooted tree, which we call **structure-tree** of the graph. The leaves of the tree correspond to the base case of series-parallel graphs, the graph consisting of only one edge. This graphs is called $K_2$ (because it is the complete graph on two vertices). A node named "P" denotes the parallel composition of its children and analogously a node named "S" denotes the series composition of its children. Then the graph $G_2$ from Figure 4 corresponds to the structure-tree drawn in Figure 6. Note that it meets the written description given above.

Moreover, we define a mapping $\phi$ from the leafs of the structure-tree of $G$ to $\binom{V(G)}{2}$, the set of all subsets of $V(G)$ of size two. This $\phi$ assigns each of the basic components $K_2$ the name that these vertices of this $K_2$ get in the final graph $G$. For an example in Figure 6 the function value of $\phi$ applied to a leaf is written just below that leaf. We do not use the function $\phi$ in this paper anymore. Instead for better readability for $G_1$ a series-parallel subgraph of $G$ corresponding to a vertex $x \in T$ and all its descendants we just write $G_1$ contains a vertex $v \in V(G)$ if $G_1$ contains a vertex $v_{G_1}$ that in $G$ is named $v$. Formally said this means for $v_{G_1}$ contained in the basic component $K_2'$ that $\phi(K_2')$ has to contain $v \in V(G)$. This formal explanation is all we needed $\phi$ for.

Last but not least we state some properties of the structure-tree. Since the statements are quite intuitive we omit the proofs.




**a)** a series-parallel graph $G_1$ **b)** a series-parallel graph $G_2$

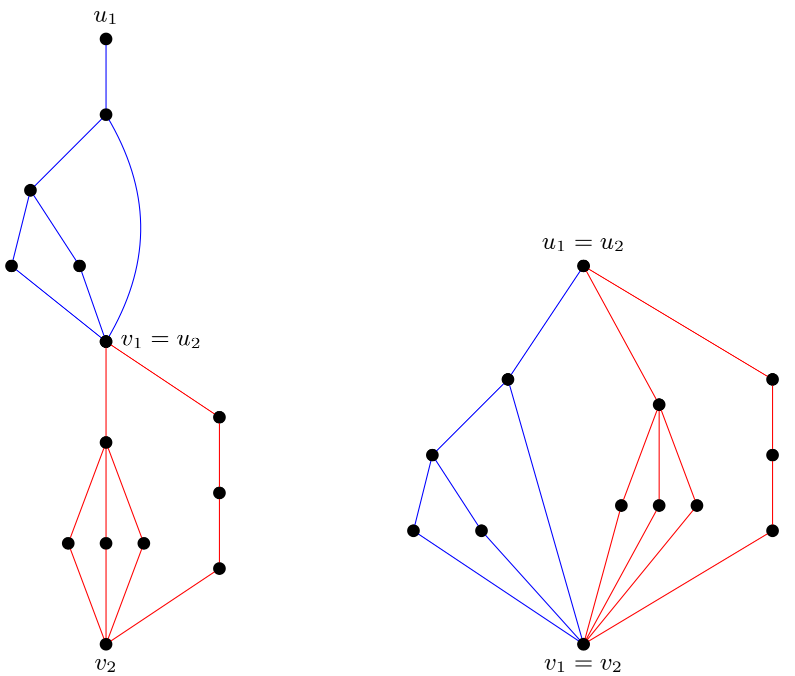
Fig. 4: An example of series-parallel graphs, part 1.

– For every series-parallel graph there exists a structure-tree (not unique).
– For every structure-tree such that all leafs are the basic components $K_2$ there exists a unique corresponding series-parallel graph.
– For every vertex $v$ in a series-parallel graph $G$ the number of leafs $l$ of the structure-tree such that $v$ is contained in $\phi(l)$ equals the degree of $v$ in $G$.
– If the structure-tree $T$ of a series-parallel graph $G$ remains connected after removing a vertex $x \in V(T)$ and all its descendants, then the remaining tree is again a structure-tree of some series-parallel graph $G'$. Moreover, $G'$ is isomorphic to some subgraph of $G$.

## 3 Mixed Connectivity

In the previous section, we had edge connectivity and vertex connectivity, and we asked for the minimum integer $k$ such that our graph gets disconnected after the removal of $k$ arbitrary edges or vertices, respectively. Now we combine these and remove edges and vertices at the same time. This means we ask for the “minimum" pair of integers $k$ and $l$ such that our graph gets disconnected after the removal of $k$ arbitrary edges and $l$ arbitrary vertices, ie. there exist subsets $W \subseteq V(G)$ and $F \subseteq E(G)$ with $|W| \leq k, |F| \leq l$ such that $G[V(G) \setminus W] - F$ is disconnected. Note that here the new graph $G[V(G) \setminus W] - F$ might have more than $l$ edges less than $G$ since by the removal of vertices we automatically also remove the incident edges. Because we only have a partial order for pairs of integers, ie. on $\mathbb{N}^2$, and no total order “minimum" here means minimum wrt. the partial order, which is component-wise minimum, and we can have several of those minimum pairs. Every of these is called a (global) connectivity pair. The pair consisting of the set of edges and the set of vertices whose removal disconnects $G$ is called a mixed cut.

To introduce these concepts more formally we first name the vertices and edges whose removal disconnects the graph. For a graph $G$, non-negative integers




**a)** the series composition of $G_1$ and $G_2$

**b)** the parallel composition of $G_1$ and $G_2$.

Fig. 5: An example of series-parallel graphs, part 2.

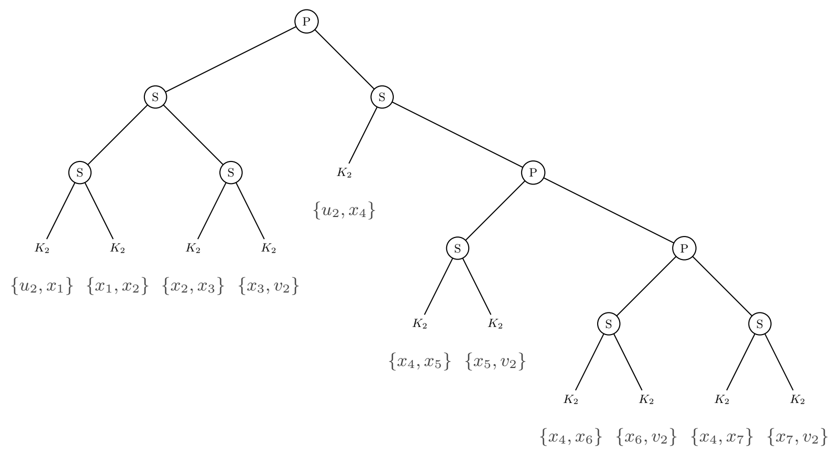



Fig. 6: the structure-tree of $G_2$ from Figure 4. In dark gray the value of $\phi$ applied to the leafs.

$k, l$ we call a subset $W \subseteq V(G)$ of the vertices and a subset $F \subseteq E(G)$ of the edges a **global** $(k, l)$**-mixed cut for** $G$ if and only if $|W| \leq k$, $|F| \leq l$ and the removal of $W$ and $F$ disconnects $G$, i.e. $(G - F) - W$ has two connected components. If for non-negative integers $k$ and $l$ there exists a global $(k, l)$-mixed cut but neither a global $(k-1, l)$-mixed cut for $G$ nor a global $(k, l-1)$-mixed cut, then we call $(k, l)$ a **global connectivity pair** in $G$. Hence, these are exactly the minimum pairs of non-negative integers $(k, l)$ for which a global $(k, l)$-mixed cut for $G$ exists. This matches our informal definition given above. Moreover, we this notation relates to edge- and vertex-connectivity. $k$-vertex-connectivity is equivalent to the lack of $(k - 1, 0)$-mixed cuts and the vertex connectivity of a graph equals $k$ if and only if $(k, 0)$ is a global connectivity pair in the graph. Analogously the $k$-edge-connectivity is equivalent to the lack of $(0, k - 1)$-mixed cuts and the edge connectivity of a graph equals $k$ if and only if $(0, k)$ is a global connectivity pair in the graph.

If we want to disconnect two certain vertices $s$ and $t$ then we call this rooted connectivity for $s$ and $t$ and the definitions just transfers. For clearness we state them fully again. Note the analogy to the definition above. Thus, for a given graph $G$, distinct vertices $s$ and $t \in V(G)$ and non-negative integers $k, l$, we call a subset $W \subseteq V(G) \backslash \{s, t\}$ of the vertices and a subset $F \subseteq E(G)$ a **rooted** $(k, l)$**-mixed cut for** $s$ **and** $t$ if and only if $|W| \leq k$, $|F| \leq l$ and the removal of $W$ and $F$ disconnects $s$ and $t$, i.e. $s$ and $t$ are in two distinct connected components of $(G-F)-W$. If for non-negative integers $k$ and $l$ there exists a rooted $(k, l)$-mixed cut but neither a rooted $(k - 1, l)$-mixed cut for $s$ and $t$ nor a rooted $(k, l - 1)$-mixed cut for $s$ and $t$, then we call $(k, l)$ a **rooted connectivity pair** for $s$ and $t$ in $G$ or sometimes $(s, t)$**-connectivity pair** for $G$. For any vertices $s, t \in V(G)$ a rooted $(k, l)$-mixed cut for $s$ and $t$ is always also a global $(k, l)$-mixed cut for $G$. The other direction is not necessarily true. Also a rooted connectivity pair for $s$ and $t$ in $G$ is not necessarily a global connectivity pair in $G$ since there indeed exists a corresponding global mixed cut but in contrast to the rooted mixed cut this global mixed cut may not be minimum anymore.

In the example in Figure 1, $(3, 0)$ and $(0, 2)$ are (global) connectivity pairs for $G$ (since both, the removal of 3 edges and the removal of 2 vertices, disconnects the graph $G$). Remember that the global edge connectivity of $G$ is 3 and the global vertex connectivity is 2. Also $(1, 1)$ is a (global) connectivity pair for $G$ since the removal of the edge $\{4, 7\}$ and of the vertex 2 together disconnects $G$. For the pair of non-negative integers $(2, 1)$ there exists also a (global) $(2, 1)$-mixed cut, but the pair $(1, 1)$ is smaller with respect to the partial order than $(2, 1)$ and that is why $(2, 1)$ is no connectivity pair. For the difference of rooted and global connectivity pairs consider the vertices 2 and 4, for them $(2, 1)$ is indeed a rooted mixed connectivity pair.

We summarize mixed connectivity into the following computational problems:

`GLOBAL-MIXED-CUT`

**Input:** An undirected graph $G$, and two positive integers $k$ and $l$.

**Question:** Can $G$ be disconnected by the removal of at most $k$ vertices in $V(G)$ and at most $l$ edges in $E(G)$?

ROOTED-MIXED-CUT

**Input:** An undirected graph $G$, two distinct vertices $s, t \in V(G)$, and two positive integers $k$ and $l$.

**Question:** Can the removal of at most $k$ vertices in $V(G)\backslash\{s,t\}$ and at most $l$ edges in $E(G)$ leave $s$ and $t$ in two distinct connected components?

For ROOTED-MIXED-CUT Johann, Krumke and Streicher [2] proved NP-completeness by reducing BIPARTITE PARTIAL VERTEX COVER to it. Bonnet and Cabello [3] use $k$-CLIQUE to show NP-completeness of GLOBAL-MIXED-CUT. Moreover, they observed that both problems become polynomial time solvable if we fix one of the two integers $k$ and $l$. Since we make use of this in the following we shortly repeat this here.

**Lemma 31** *Given an undirected graph $G$, two positive integers $k$ and $l$ (and two distinct vertices $s, t \in V(G)$) GLOBAL-MIXED-CUT and ROOTED-MIXED-CUT is polynomially solvable, if $k$ or $l$ is bounded by a constant.*

*Proof.* For better readability we only prove the statement for GLOBAL-MIXED-CUT and also only for $k$ bounded. For ROOTED-MIXED-CUT or $l$ bounded it works just analogously.

Let $k$ be bounded by the constant $c'$. The answer to GLOBAL-MIXED-CUT for our instance is "Yes" if and only if for some subset $S$ of the vertices with $|S| \leq c'$ we can disconnect $G[V(G) \setminus S]$ by removing at most $l$ edges. Since the latter can be checked in polynomial time with Menger's Theorem by enumerating all possible subsets $S$ we can solve GLOBAL-MIXED-CUT efficiently. The running time is $O(n^{c'} T_M)$ where $T_M$ is the time to check the edge-connectivity of $G - S$, that is $O(n^{2/3}m)$ for rooted edge connectivity and $O(n^2 n^{2/3} m)$ for global edge connectivity.

## 4 Mixed Connectivity in Series-Parallel Graphs

We show how to solve ROOTED-MIXED-CUT on series-parallel graphs in polynomial time. With that we can easily solve GLOBAL-MIXED-CUT on series-parallel graphs in polynomial time by just enumerating all possible $s$ and $t$.

For solving ROOTED-MIXED-CUT we essentially bound the number of vertices that we need to remove in order to disconnect two given, non-adjacent vertices $s$ and $t$, ie. we bound the $(s,t)$-vertex connectivity. Once for general series-parallel graphs we bounded this value by above we can use our previous Lemma 31 to solve ROOTED-MIXED-CUT efficiently. To be precise we cannot give a constant bound for an arbitrary series-parallel graph and an arbitrary choice of $s$ and $t$. Namely, in the case of a parallel-composition where $s$ and $t$ are the source and sink terminal the number of vertex-disjoint $s-t$-paths in $G$ cannot be bounded. (Remember: By Menger's Theorem we can look at vertex-disjoint $s - t$-paths instead of the $(s,t)$-vertex connectivity.) Therefore, we have to consider this case separately.

To show that for an arbitrary choice of graph $G$ and non-adjacent vertices $s,t \in V(G)$ the $(s,t)$-vertex connectivity of $G$ is indeed bounded by a constant, we consider the smallest subgraph of $G$ that contains $s$ and $t$ and examine the different structures that can occur. First we formally define what we mean with smallest subgraph that contains $s$ and $t$. For a series-parallel graph $G$ with structure-tree $T$ and vertices $s,t \in V(G)$ we choose a vertex $x \in V(T)$ such that the following three properties are full-filled:

1. both, $s$ and $t$, appear in some $\phi(l)$ for $l$ a leaf and descendent of $x$ (in other words the series-parallel graph corresponding to $x$ and its descendants contains $s$ and $t$),
2. The sibling of $x$ in $T$ does not contain both, $s$ and $t$.
3. $x$ is chosen such that the level of $x$ is minimum.

The series-parallel graph corresponding to the $x$ and its descendants in $T$ we call $\boldsymbol{G_{s,t}}$.

This definition is very technical but as we see later we need all these properties. Informally said we just look at the minimum series-parallel subgraph of $G$ that contains $s$ and $t$, where subgraph is stated with respect to the series-parallel structure of $G$. Moreover, if $s$ and $t$ are source and sink terminals of some parallel-composition, the graph $G_{s,t}$ includes all the parallel-compositions that have $s$ and $t$ as source and sink terminals. In other words $G_{s,t}$ includes all the subgraphs connected in parallel between $s$ and $t$. For an example see Figure 7. For $v_2$ and $x_4$ instead of $x_6$ and $x_7$ the graph $G_{x_4,v_2}$ would contain all the orange edges and vertices and also the vertex $x_5$ and its incident edges.

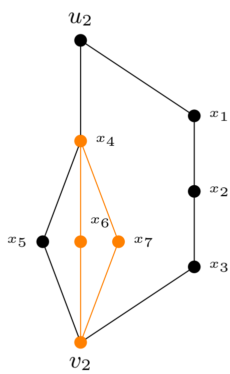



Fig. 7: The series-parallel graph $G_2$ and the (also series-parallel) graph $G_{x_6,x_7}$ in orange.

**Bounding the $(s,t)$-vertex connectivity for different cases**

Let a series-parallel graph $G$ and two non-adjacent vertices $s,t \in V(G)$ be given. In the following we consider the structure of $G_{s,t}$ and for most structures we can give a bound on the $(s,t)$-vertex connectivity of $G$. For a visualization of the different cases see Figure 8.

**Case 1** The last composition of $G_{s,t}$ is a parallel composition and none of $s$ and $t$ is source or sink terminal of $G_{s,t}$:
Removing the source and the sink terminal of $G_{s,t}$ disconnects $s$ and $t$ in $G_{s,t}$ and it also disconnects $s$ and $t$ in $G$. Hence, the $(s,t)$-vertex connectivity is at most 2.

**Case 2** The last composition of $G_{s,t}$ is a parallel composition and at least one of $s$ and $t$ is source or sink terminal of $G_{s,t}$:
First of all we show that both, $s$ and $t$, has to be source and sink terminal, respectively. Let $G_{s,t}$ be the parallel composition of $G_1$ and $G_2$. Wlog. $s$ is the source terminal of $G_1$ (and of $G_2$ then). If $t$ is not the sink terminal of $G_1$ and $G_2$ then $t$ is inside $G_1$ wlog. But then $G_{s,t}$ would be just $G_1$ by the minimality of $G_{s,t}$, ie. because of condition 3. in the definition of $G_{s,t}$. Hence, $t$ must be the sink terminal of $G_1$ and $G_2$, ie. of $G_{s,t}$. Case 2 is the case where we cannot bound the $(s,t)$-vertex connectivity. We come back to this case later.

**Case 3** The last composition of $G_{s,t}$ is a series composition and at most one of $s$ and $t$ is source or sink terminal of $G_{s,t}$:
Wlog. let $s$ be the source terminal of $G_{s,t}$. Moreover, let $G_{s,t}$ be the series composition of $G_1$ and $G_2$, ie. $s$ is the source terminal of $G_1$. Because of minimality of $G_{s,t}$ (=condition 3. in the definition) $t$ has to be in $G_2$ (otherwise $G_{s,t}$ would be $G_1$). Hence, removing the source and sink terminal of $G_2$ (called $v_{mid}$ and $v$ in the Figure), which are both not equal to $t$, disconnects $t$ from $s$ in $G_{s,t}$ and also in $G$. Therefore the $(s,t)$-vertex connectivity is at most 2.

**Case 4** The last composition of $G_{s,t}$ is a series composition and both, $s$ and $t$, are source and sink terminal, respectively, of $G_{s,t}$:
Let $G_{s,t}$ be the series composition of $G_1$ and $G_2$ and $s$ the source terminal of $G_1$ and $t$ the sink terminal of $G_2$. Removing the "middle vertex" $v_{mid} \in V(G)$ of $G_{s,t}$, ie. the source terminal of $G_2$, disconnects $s$ and $t$ in $G_{s,t}$ but not necessarily in $G$ (there could still be a path from $s$ to $t$ in the outside of $G_{s,t}$ if the remaining graph $G$ without $G_{s,t}$ is still connected). Therefore we have to argue a little bit more careful. Let $T$ be the structure tree of $G$ and $x \in V(T)$ the vertex of the definition of $G_{s,t}$. Consider the parent of $x$ in $T$, call it $y$. If there is no parent, ie. $x$ is the root of $T$, then $G_{s,t} = G$ and removing $v_{mid}$ also disconnects $s$ and $t$ in $G$. The label of $y$ must be a series composition by condition 2. of the definition of $G_{s,t}$: If $y$ would be the parallel composition of $G_{s,t}$ and $G_2$ then since $s$ and $t$ are the source and sink terminals of $G_{s,t}$ they must also be the source and sink terminals of $G_2$ and hence, $s,t \in V(G_2)$ and condition 2. of the definition is not full-filled. Hence, $y$ denotes the series composition of $G_{s,t}$ and some other graph, call it $G'$. Wlog. $s$ is the source terminal of the series composition, hence $t$ is the source terminal of $G'$. But now we are in a similar case as Case 3 and

removing $v_{mid}$ and the sink terminal of $G'$ disconnects $t$ from $s$. Therefore the $(s,t)$-vertex connectivity is at most 2.

We summarize our results in the following Lemma.4

**Lemma 41** *For any series-parallel graph $G$ and non-adjacent vertices $s,t \in V(G)$ either the graph $G_{s,t}$ has the structure of Case 2 above or the $(s,t)$-vertex connectivity in $G$ is at most 2.*

With this we can conclude:

**Theorem 42** *For any series-parallel graph $G$ and distinct vertices $s,t \in V(G)$ `ROOTED-MIXED-CUT` can be solved in polynomial time. The running time of our algorithm is $O(nn^{2/3}m)$.*

*Proof.* If $s,t$ are adjacent, ie. there is an edge $\{s,t\} \in E(G)$, replace the edge by a path of length 2 and apply the results from Lemma 42. At the end we replace the path by the edge $\{s,t\}$ again and if at some point we removed the middle vertex of this path we instead just remove the edge. Hence, we can assume that $s,t$ are non-adjacent.

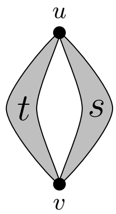
Let $k$ and $l$ be the two positive integers of the input of `ROOTED-MIXED-CUT`. The question is whether $s$ and $t$ can be separated by the removal of at most $k$ vertices in $V(G) \setminus \{s,t\}$ and at most $l$ edges in $E(G)$. By Lemma 42 we only have to consider two cases.

If the $(s,t)$-vertex connectivity in $G$ is at most 2, then for $k \geq 2$ the answer is "yes" and for $k = 0,1$ we apply Lemma 31 to decide if the removal of $l$ edges is enough. This can be done in time $O(n \cdot n^{2/3}m)$.

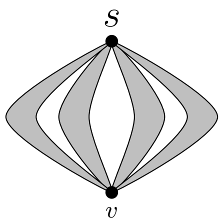
Now consider the case that the last composition of the series-parallel subgraph $G_{s,t}$ is a parallel composition and $s$ is the source and $t$ the sink terminal of $G_{s,t}$. Let $T$ be the structure tree of $G$ and $x \in V(T)$ the vertex of the definition of $G_{s,t}$. Denote with $y_1, y_2, \ldots, y_k$ all the descendants of $x$ in $T$ such that $y_i$ is not labeled with parallel-composition but the unique $y_i - x$-path in $T$ only contains vertices labeled with parallel-composition except for the vertex $y_i$ itself. Let $G_1, G_2, \ldots, G_k$ be the series-parallel subgraphs of $G$ such that $G_i$ corresponds to $y_i$ and all its descendants in $T$. Observe that $G_1, G_2, \ldots, G_k$ are connected in parallel and for each of them $s$ is the source terminal and $t$ the sink terminal. The other part of the graph, ie. $G$ without $G_{s,t}$ we call $G'$. For a visualization see Figure 9.

In the following we compute the $(s,t)$-connectivity pairs of $G_1, \ldots, G_k$ and of $G'$. Given these we can compute the $(s,t)$-connectivity pairs of $G$ and this is enough for solving `ROOTED-MIXED-CUT`: $s,t$ can be seperated by the removal of $k$ vertices and $l$ edges from $G$ if and only if there is a $(s,t)$-connectivity pair of $G$ that is smaller than $(k,l)$ with respect to the partial order.

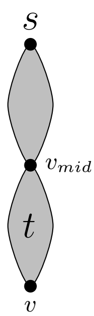
For every $i = 1, \ldots, k$ $G_i$ is either $K_2$ or a series-composition of let us say $G_i^1$ and $G_i^2$ (because the label of $y_i$ is not parallel-composition). Since we assumed $s$ and $t$ to be non-adjacent $G_i$ has to be a series-composition. Hence, removing the middle vertex $v_{mid}^i$ of each series-composition of $G_i$ disconnects $G_i$ and therefore



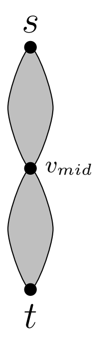

**a)** Case 1.




**b)** Case 2. Here the position of $t$ is not specified, but it has to be $v$ (see proof above).




**c)** Case 3.




**d)** Case 4.

Fig. 8: The different ways how $G_{s,t}$ could look like.

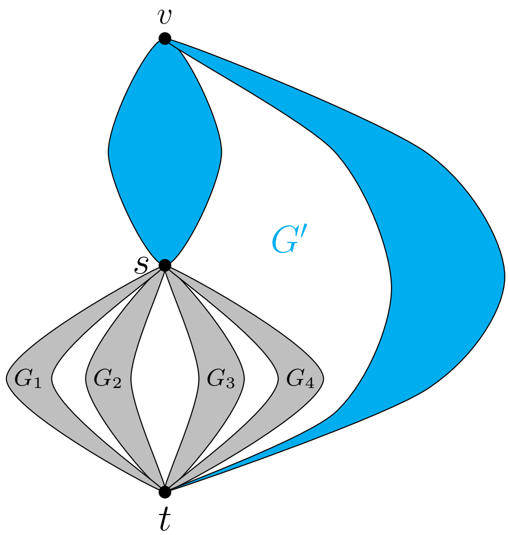



Fig. 9: $G_{s,t}$ and $G'$ for the Case 2 and $k = 4$. The cyan part is $G'$ in this example.

removing $v^1_{mid}, \dots, v^k_{mid}$ disconnects $s, t$ in $G_{s,t}$. This means that $(1, 0)$ is a $(s, t)$-connectivity pair for $G_i$. But this already means that there can be only one more $(s, t)$-connectivity pair, namely where we do not remove any vertices. This $(s, t)$-connectivity pair $(0, \cdot)$ can be computed by simply applying the algorithm for edge-connectivity in $G_i$. This can be done in total time $O(n \cdot n^{2/3}m)$.

But analogously to Case 4 above we also have to ensure that there is no $s - t$-path outside of $G_{s,t}$ and that is why we consider $G'$. Again we consider the parent of $x$ in $T$, call it $z$. If there is no parent of $x$ in $T$ then $G = G_{s,t}$ and the $(s, t)$-connectivity pair for $G'$ is $(0, 0)$. Therefore assume that $z$ exists. The label of $z$ must be a series-composition by condition 2 of the definition of $G_{s,t}$. Hence, removing the terminal vertex of the series-composition, that does not equal $s$ or $t$ (in Figure 9 called $v$), disconnects $s$ and $t$ in $G'$ and therefore $(1, 0)$ is a $(s, t)$-connectivity pair for $G'$. As before we compute the only remaining $(s, t)$-connectivity pair $(0, \cdot)$ with edge-connectivity in $G'$.

In order to get now the $(s, t)$-connectivity pairs of $G$ we merge all the $(s, t)$-connectivity pairs of $G_1, \dots, G_k$ and $G'$.

**Claim:** Every $(s, t)$-connectivity pair of $G$ is the sum of $(s, t)$-connectivity pairs of $G_1, \dots, G_k$ and $G'$. Vice versa for every sum of $(s, t)$-connectivity pairs of $G_1, \dots, G_k$ and $G'$ the result $(a, b)$ is a candidate for a $(s, t)$-connectivity pair of $G$ (ie. removing $a$ vertices and $k$ edges from $G$ disconnects $s, t$, but maybe $(a, b)$ is not minimal enough.).

*Proof.* $\longleftarrow$ Let $(a_i, b_i)$, $i = 1, \dots, k + 1$ be $(s, t)$-connectivity pairs of $G_i$ with $G_{k+1} = G'$. That means the removal of $a_i$ vertices and $b_i$ edges disconnects $s$ and $t$ in $G_i$. Then also the removal of all those vertices and edges in $G$ disconnects $s$ and $t$ in $G$.

$\longrightarrow$ Let $(a, b)$ be an $(s, t)$-connectivity pair of $G$. Consider $(W, F)$, a rooted $(a, b)$-mixed cut for $s$ and $t$ in $G$. Define for $i = 1, \ldots, k+1$ $W_i = W \cap V(G_i)$ the removed vertices in $G_i$ and $F_i = F \cap E(G_i)$ the removed edges in $G_i$ where $G_{k+q} = G'$. Since $(W, F)$ is a rooted $(a, b)$-mixed cut for $s$ and $t$ in $G$ for every $i$ $(W_i, F_i)$ is a rooted $(|W_i|, |F_i|)$-mixed cut for $s$ and $t$ in $G_i$ and with $\longleftarrow$ the pairs $(|W_i|, |F_i|)$ must also be minimal, ie. $(s, t)$-connectivity cuts. Moreover, clearly they sum up to $(a, b)$.

With this claim we can compute all the $(s, t)$-connectivity pairs of $G$ by choosing for each subgraph $G_1, \ldots, G_k$ or $G'$ a $(s, t)$-connectivity pair and add them up. Then we just check which of these pairs is minimal with respect to the partial order on $\mathbb{Z}^2$. But for choosing a $(s, t)$-connectivity pair for each sugraph we have two possibilities and therefore in total we have $2^{k+1}$ many possibilities, which is exponentially. Therefore we spend a little bit more effort to compute only the good candidates for the $(s, t)$-connectivity pairs: First of all note that there are at most $k+2$ $(s, t)$-connectivity pairs for $G$: The $(s, t)$-connectivity pairs are pairwise distinct. Since furthermore they are minimal for each integer there can be at most one $(s, t)$-connectivity pair with this integer at its first entry. But since $(k+1, 0)$ is a $(s, t)$-connectivity pair (for each subgraph $G_1, \ldots, G_k$ and $G'$ choose the $(s, t)$-connectivity pair $(1, 0)$) we have at most $k+2$ integers for the first component. Now order the $k+1$ subgraphs such that the second entry of their $(s, t)$-connectivity pair starting with 0 is non-increasing. Then for each $j = 0, 1, \ldots, k+1$ take for the first $j$ subgraphs of this order the $(s, t)$-connectivity pair starting with 1 and for the remaining ones the $(s, t)$-connectivity pair starting with 0 and sum them up. Then we get the (only) candidate for a $(s, t)$-connectivity pair starting with $j$. Finally, we just have to check those $k+2$ pairs for minimality. The total time for this summing up procedure is $O(n \log n)$ since the ordering dominates the running time. Hence, we got all the $(s, t)$-connectivity pairs for $G$ and with that we can solve our problem as stated above. The running time for this is $O(n \cdot n^{2/3} m)$. This completes the case that $G_{s,t}$ is a series-composition and $s$ and $t$ are the source and sink terminals.

Note that in fact we did not only solve the `ROOTED-MIXED-CUT` problem but while doing that we also computed all the $(s, t)$-connectivity pairs of $G$.

## 5 Conclusion

In this paper we discussed the problem of Mixed Connectivity, namely `GLOBAL-MIXED-CUT` and `ROOTED-MIXED-CUT`, a generalization of vertex- and edge-connectivity. Both problems were shown to be NP-hard by [2] and [3] therefore we focused on solving them for a certain class of graphs, the series-parallel graphs. Indeed we were able to solve our research question and show that for graphs with this special structure `ROOTED-MIXED-CUT` is solvable in polynomial time and with that also `GLOBAL-MIXED-CUT`. The first result holds essentially because we can disconnect a series composition pretty easily by just removing the vertex joining the two original graphs. More effort we have to spend on the case that there is no series composition between $s$ and $t$ but one (or possibly more) parallel compositions. But then there is series composition inside in each of the parallel compositions and being little bit carefully we can merge the single results in polynomial time. As a side effect we also compute the connectivity pairs for parallel compositions as well.

Now the natural question arises when does the problem become NP-hard. Can one for example also find polynomial time algorithms for the class of planar graphs or the class of graphs of bounded tree-width? Note that both classes are supersets of series-parallel graphs.

Regarding Mixed Connectivity on series-parallel graphs it might also be possible to find faster algorithms. The best way to approach this question, we think, is to first find faster algorithms for the vertex- and the edge-connectivity problem on series-parallel graphs. Then those algorithms immediately give a faster running time for solving Mixed Connectivity on series-parallel graphs.

## References

1. Beineke, L.W., Harary, F.: The connectivity function of a graph. Mathematika **14**(2) (1967) 197–202
2. Johann, S.S., Krumke, S.O., Streicher, M.: On the mixed connectivity conjecture of beineke and harary (2020)
3. Bonnet, E., Cabello, S.: The complexity of mixed-connectivity (2020)
4. Takamizawa, K., Nishizeki, T., Saito, N. In: Combinatorial problems on series-parallel graphs. Volume 3. (01 2006) 79–94
5. Korte, B., Vygen, J.: Combinatorial Optimization, Theory and Algorithms. Springer, Berlin, Heidelberg (2008)
6. Krumke, S.O., Noltemeier, H.: Graphentheoretische Konzepte und Algorithmen. B. G. Teubner (2005)

# Distributed drone swarm collision avoidance algorithms: a thorough comparison

Maurice Meedendorp

Department of Computing Science
Umeå University, Sweden
mrc20mmp@cs.umu.se

**Abstract.** Collision avoidance is an important topic in the robotics field. Decentralized approaches in particular are required to control large drone swarms where centralized control becomes prohibitive due to high computation and communication requirements. This paper presents a theoretical and empirical comparison between four distributed drone control algorithms: the well established ORCA and FMP algorithms, and the recent ASCA and BICARL algorithms. The theoretical comparison shows a clear advantage in the features of ASCA and FMP, and experimental results show that ASCA significantly outperforms the others in a simulated package delivery task and a scalibility test.

## 1 Introduction

Navigation, i.e. the problem of collision-free movement to a destination, is one of the core problems in the robotics field. The area of multi-agent collision avoidance, specifically, has attracted much recent interest in the research community [1], as the availability of quadcopters has grown, and the potential in various scenarios such as package delivery or search-and-rescue surveillance has become clear [2]. Since agents in these scenarios often operate close together, the problem of avoiding collisions in real-world environments has become more important.

An area of particular interest for multi-agent environments, where centralized planning might be unrealistic and computationally expensive, is distributed obstacle avoidance. In a distributed architecture, each agent plans its own path based only on observations of its direct environment. Because of their distributed nature, these algorithms can scale to a large numbers of agents with relative ease and are not reliant on communications to a centralized system.

Recent papers propose the Angular Swarm Collision Avoidance (ASCA) [3] and BICARL [2] algorithms for dynamic distributed obstacle avoidance in quadrotor drones. The ASCA algorithm is a geometric guidance algorithm (as defined in [1]), which uses geometric concepts to avoid collisions. BICARL, on the other hand, uses a neural network, which does not clearly fit any of the categories defined in [1]. Both algorithms are compared to well established algorithms such as FMP [4] and ORCA [5] over which they cite performance improvements. However, the ASCA and BICARL algorithms are not compared with each other.

L. Jiang, A. Jonsson, L. Vanhée (Eds.): Umeå's 25$^{th}$ Student Conference in Computing Science USCCS 2022, pp. 57–67, January 2022.

Comparing these algorithms with each other provides new insights into the advantages and limitations of each approach, and provides considerations to take into account when choosing one over the other in practice.

In this paper, a theoretical comparison of the four algorithms is made, and implementations of the ASCA and BICARL algorithms are compared in a package-delivery task.[1] The main goal is to compare the features, performance and reproducibility of these algorithms and their applicability in real-world scenarios.

The remainder of this paper is organized as follows: First, a brief overview of the different algorithms is given in Section 2 followed by theoretical comparisons between them in Section 3, describing some limitations and advantages of each approach. The simulation setup, tasks and performance metrics are described in Section 4, and the experimental results presented in Section 5. Lastly, the conclusions are drawn in Section 6, and the results are discussed in Section 7.

## 2 Background

### 2.1 ORCA

The optimal reciprocal collision avoidance (ORCA) algorithm [5] is a geometric guidance algorithm, like ASCA. Unlike the ASCA algorithm, ORCA is based on the idea of *velocity obstacles*, as it uses geometric principles in the velocity space. A velocity obstacle for agent A induced by agent B is the set of all relative velocities of A with respect to B that will result in a collision between the two. A geometric interpretation of this velocity obstacle can be seen in Figure 1.

In ORCA, each agent determines the velocity obstacles produced by all other agents in its environment, and adjusts the direction and magnitude of its velocity vector to avoid obstacles. All agents do this simultaneously and, under the assumption that each agent is running the ORCA algorithm, the choice of the velocity vector will avoid collisions between agents.

### 2.2 FMP

The Force-based Motion Planning (FMP) algorithm is a newer approach to distributed drone control that is based on potential fields [4]. The environment of the drones is modeled as a vector field of forces acting on the drone. The drone will then move in the direction that is the weighted sum of all forces it perceives.

In the FMP algorithm each agent is influenced by two types of forces: repulsion from other agents and attraction to the goal point. The repulsive force is generated only by agents or obstacles within a certain radius around the drone

[1] If this paper were to be published in a top conference, all four algorithms would be compared. Here, only two will be empirically compared. This allows full-fledged comparison while saving time on implementation and testing, which have little benefits with respect to what is taught in this course.

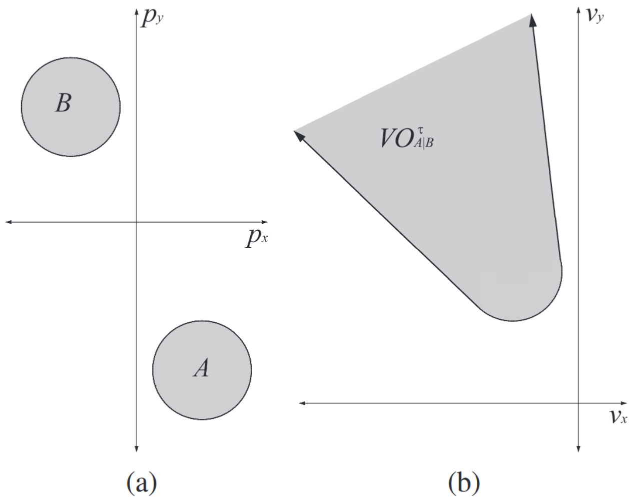



**Fig. 1.** An illustration of the ORCA algorithm. (a) The agents A and B. (b) the velocity space from the perspective of agent A. Selecting any velocity from the dark gray area will result in a collision with B. Agent A can avoid collision by selecting any velocity in $VO^{\tau}_{A|B}$.

and assumes the position of these obstacles can be measured perfectly using sensors or communicated in another way.

The authors of the paper provide a formal proof which guarantees that the algorithm is collision free, the separation distance between agents is larger than the minimum and that the algorithm always converges. This formal proof is for the ideal case in continuous time, which the algorithm – running in discrete time – tries to approximate.

The authors provide a comparison with the ORCA algorithm, which shows that the FMP algorithm greatly improves the transition and execution times of the algorithm. The transition time is the time an agent needs to travel from the starting position to the goal position, and the execution time is the time needed to calculate the next step in the algorithm.

### 2.3 BICARL

The BICARL algorithm is a reinforcement learning-based approach to decentralized drone control, consisting of a 2-layer neural network, each containing 64 nodes [2]. This relatively small network can be run on a drone micro-controller. The input to the network is the drone's own position, velocity, heading angle, heading angular velocity and goal position, as well as the distance, angle and velocity of it's nearest neighbour. The network then determines the desired angular velocity and desired velocity of the agent for the next time step. Before an implementation of the algorithm can be used, the neural network has to be trained in a simulation environment, complicating the implementation and reproducibility of the results.

The algorithm does not include a formal proof and since the results depend on the training of the neural network, it cannot be guaranteed to converge or be collision-free. Static obstacles are also not considered in BICARL.

### 2.4 ASCA

The Angular Swarm Collision Avoidance (ASCA) distributed control algorithm provides a geometric approach to distributed collision avoidance [3]. The algorithm calculates the direction of the velocity vector based on the position of the goal point and all neighbouring drones within its avoidance radius. When all motion is obstructed, the algorithm stops to avoid collisions. An illustration of the algorithm is shown in Figure 2.

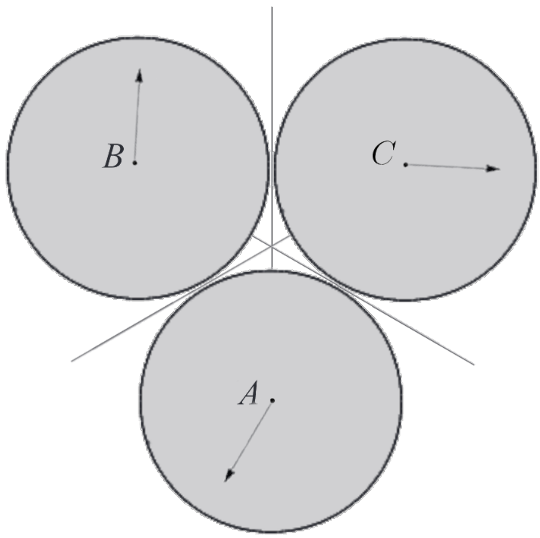



**Fig. 2.** An illustration of the ASCA algorithm. Agents 1, 2, and 3 each choose their velocities away from the others to avoid collisions.

## 3 Theoretical comparison

Comparisons between the algorithms outlined above can be made based on the theoretical description in the papers. This section compares various features of the algorithms that might be relevant in practical applications are compared, such as whether they generalize to 3 dimensional environments, how static obstacles are handled, whether a formal proof is supplied, and what the sensing requirements are.

### 3.1 3D environments

To be widely applicable in real-world quadcopters, algorithms must be able to avoid obstacles in three dimensions.

All four algorithms claim to offer support for 3D environments, though the details of a 3-dimensional variant of the algorithms are discussed in varying degrees. In the ORCA paper it is simply mentioned that a 3D extension is possible, while the FMP, BICARL and ASCA papers go into more detail. In the latter three papers, the equations for 3D motion planning are explained in detail and results are presented that verify the claims made.

### 3.2 Static obstacle avoidance

The algorithms show small, but important, differences when it comes to being able to avoid static obstacles. Static obstacles are a common problem in real-world scenarios, where agents are expected not to collide with non-agent obstacles.

In the case of ORCA, static obstacles are a special case of normal agents, modeled as a set of line segments. Both the ASCA and FMP algorithms also introduce a separate type of agent for representing static obstacles, but they are generalized as circles. The BICARL algorithm does not support static obstacles at all, and describes potential support for them as future work.

The algorithms that support static obstacle avoidance do not handle the case where an obstacle might be concave (in ORCA) or where two obstacles are placed such that a concave surface is formed between the circles (in all three). This might result in an agent getting stuck while trying to navigate towards a concave (pseudo-)obstacle, and not routing around it. The ORCA paper suggests using a global path planner to avoid this scenario, while the ASCA paper simply notes that obstacles should be placed such that an agent can fit between them, which might not be realistic in real world environments. These scenarios are not evaluated in the experiments conducted in this paper.

### 3.3 Formal proof

Collision-avoidance algorithms can offer a formal proof which mathematically guarantees that, under certain conditions, agents are collision-free. When these conditions can be satisfied in real-world scenarios the agents will never collide. Even in case the conditions cannot be met exactly, a formal proof provides a measure of certainty and might indicate in which situations collisions could occur.

Of the algorithms discussed in this paper, the ORCA, FMP and ASCA algorithms provide a formal proof of collision avoidance. BICARL, being a neural network-based algorithm cannot offer a formal proof. Since the network is trained by the user, it is essentially a black box where no guarantees about relations between inputs and outputs can be given.

### 3.4 Sensing requirements

The presented algorithms require different approaches to sensing the environment. Since the algorithms are distributed, agents do not communicate directly with each other, but are assumed to use sensors to observe other agents in their vicinity. More expansive sensing assumptions might require more processing power or more sensors, which are a serious trade-off on light-weight, battery powered agents.

In this group of four algorithms, ORCA and BICARL assume that agents are aware of the position and velocity of other agents, while the FMP and ASCA algorithms require only position information.

Furthermore, ORCA, FMP and ASCA assume knowledge about agents within a certain radius, while BICARL only uses information about the nearest neighbour.

### 3.5 Summary

In this section, the comparisons are summarized per algorithm. Table 1 is given to show the information about each algorithm. Each criterion is explained in the caption, and more detailed comparisons are made in the preceding sections.

| **Algorithm** | **Approach** | **3D** | **SOA** | **Proof** | **Sensing** |
|---|---|---|---|---|---|
| ORCA | Geometric guidance | ✗ | ✓ | ✓ | $p$, $v$ |
| FMP | Potential function | ✓ | ✓ | ✓ | $p$ |
| BICARL | Machine learning | ✓ | ✗ | ✗ | $p$, $v$ |
| ASCA | Geometric guidance | ✓ | ✓ | ✓ | $p$ |

**Table 1.** Summary of the theoretical comparison. Approach: one of the classes of collision avoidance strategies as defined in [1]. 3D: whether a 3D version of an algorithm is provided in a paper. SOA: whether static obstacle avoidance is supported. Proof: whether an algorithm can be proven to be correct. Sensing: information necessary about neighbouring agents ($p$: position, $v$: velocity).

## 4 Empirical comparison

To evaluate the real-world performance of these algorithms, their performance is tested in a package delivery task and a scalability test. The results are evaluated on multiple metrics, such as amount of packages delivered, trajectory length and minimum separation distance between the agents.

### 4.1 Task description

To empirically test the algorithms, a two-dimensional simulation was developed in Python, along with an implementation of both the ASCA and BICARL algorithms. For the machine-learning aspect of BICARL, stable-baselines [6] was used. Both algorithms were implemented using the pseudocode and hyper-parameters described in the respective papers. The simulation environment is shown in Figure 3

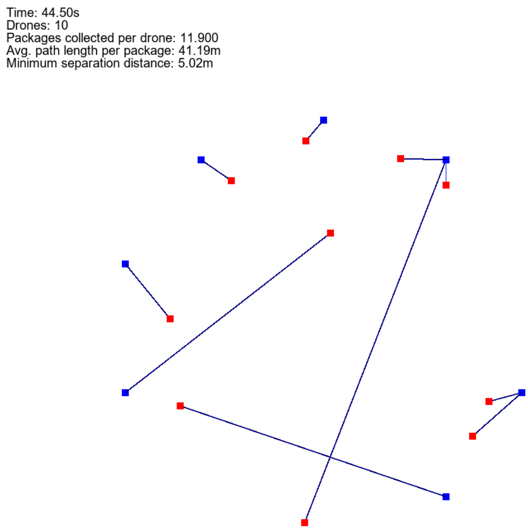



**Fig. 3.** A still image from the simulation with the ASCA algorithm. The drones are red squares, the goals are blue and the blue lines indicate which drone is moving to which goal position. The evaluation metrics and timestamp can be seen in the text at the top-left.

The simulated drones were assigned a package-delivery task. In a package delivery task, the drones are equally distributed on a fixed-radius circle, with randomly assigned goal positions among the set of initial positions. After reaching the goal position, another goal position is randomly assigned until the end of the simulation is reached. This scenario reflects potential real-world applications of these algorithms, where agents on unrelated tasks plan their path through an environment, avoiding other agents and obstacles.

The environment parameters relevant for the package delivery task are shown in Table 2. These parameters were chosen to be exactly the same as those used in the BICARL paper, which also satisfy the requirements of ASCA's collision-avoidance guarantee. To make a fair comparison between the two algorithms, it is assumed that the positions of other agents is known exactly.

To compare the algorithms, each was run 10 times in a package-delivery task with 10 agents. The goal positions for the drones over the 10 runs were the same for each algorithm.

| Parameter | Value |
|---|---|
| Simulation time | 100 $s$ |
| Time step size | 0.02 $s$ |
| Maximum drone velocity | 15 $m/s$ |
| Circle radius | 30 $m$ |
| Distance to sample new goal | 3.5 $m$ |
| Minimum separation distance | 5 $m$ |

**Table 2.** The simulation parameters used in all experiments

Additionally, the scalability of the algorithms was tested by running each algorithm once for 2 up to 100 agents. For the scalability test, the circle radius parameter was set to 160 meters, to allow 100 drones to fit on the circle. The computation times for the algorithms was not compared directly. While useful, the implementations differ too much to offer a clear comparison.

### 4.2 Evaluation metrics

The conducted experiments are evaluated along three metrics: (1) the average amount of packages delivered per drone within the simulation time, (2) the average travel distance per package delivered and (3) the closest separation distance between two agents during the simulation. These three metrics allow different aspects of the algorithms to be evaluated. The first two provide an indication of the overall performance of the algorithm, while the closest separation distance can be compared to the minimum separation distance specified in Table 2 to evaluate the robustness of the algorithm.

## 5 Results

In this section, the results from conducting the experiments described above are shown. The results for the package-delivery scenario with 10 agents are shown in Section 5.1. The scalability test results can be found in Section 5.2.

### 5.1 Direct comparison

The results for the direct comparison experiment, where the two algorithms were given the same 10 package delivery tasks, are presented in Table 3. The ASCA algorithm can pick up almost 5 times as many packages as BICARL in the same amount of time. The travel distances are approximately the same, and the closest separation distance between two drones with ASCA is within 1 centimeter of the specified minimum, while BICARL keeps the drones farther away from each other. Both algorithms respect the minimum separation distance parameter as claimed, which was set at 5 meters.

[2] Training the BICARL algorithm proved more difficult than expected, and a working version was not completed in time. Since in this course the process of writing a paper is more important than the result, numbers presented here are educated guesses based on the results reported in the BICARL paper.

| Algorithm | Packages | Travel distance (m) | Separation (m) |
|---|---|---|---|
| ASCA | 25.76 | 44.28 | 5.012 |
| BICARL[2] | 5.8 | 46.40 | 7.01 |

**Table 3.** The results for the direct comparison experiment with 10 drones on a 35m radius circle. *Packages* shows packages collected per drone, *Travel distance* is the average trajectory length per package collected in meters and *Separation* the closest distance between any two drones at any time during the simulation.

Figure 4 shows typical trajectories for both algorithms. With ASCA, drones meeting head-on always avoid each other in semi-circular paths as depicted. The drones keep moving at their maximum velocities, unless moving in any direction will result in a collision. Since the algorithm does not take velocities into account, sometimes one drone (A) will follow another drone (B) for a time, away from A's goal position, because B is obstructing the direct path to the goal for A at each time step. It is not immediately clear if this behaviour will transfer to a real world implementation of the algorithm.

Compared to ASCA, BICARL behaves a lot more defensively. Drones strongly avoid others and slow down when approaching each other. The trajectories are not necessarily straight to the goal, which is reflected in the numbers, as the travel distance for ASCA is shorter.

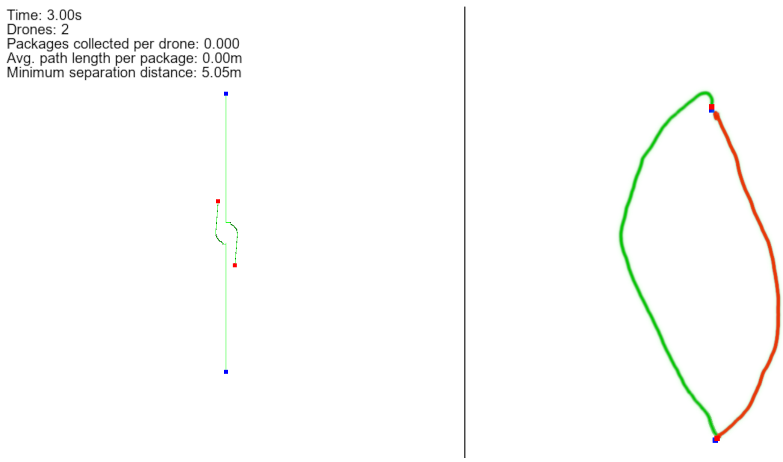



**Fig. 4.** Avoidance patterns typical for ASCA (left) and BICARL[3] (right). Coloured lines indicate drone trajectories. Blue and red squares are as in Figure 3.

## 5.2 Scalability

The results for the scalability test are shown in Table 4.

[3] The figure on the right is recreated from the results presented in the BICARL paper. See [2]

| Test setup | Packages collected | | Travel distance | |
|---|---|---|---|---|
| # of agents | ASCA | BICARL$^2$ | ASCA | BICARL$^2$ |
| 10 | 6.0 | 2.4 | 203 | 231 |
| 20 | 6.6 | 1.7 | 193 | 196 |
| 50 | 6.4 | 0.88 | 204 | 192 |
| 75 | 6.3 | 0.67 | 205 | 193 |
| 100 | 6.4 | 0.53 | 205 | 194 |

**Table 4.** A selection of the results for the scalability test experiment on a 160m radius circle. *Packages collected* shows packages collected per drone, *Travel distance* is the trajectory length per package collected

In all test setups, the ASCA algorithm collects more packages than BICARL. ASCA collects roughly the same amount of packages per agent as the number of agents increases, and the trajectory length per package delivered is also approximately constant.

In comparison, BICARL collects relatively little packages and the amount of packages seems inversely related to the number of agents in the simulation. Though less packages are collected in the same amount of time, the trajectory length is constant as more agents are added, which is caused by the fact that BICARL is more defensive in avoiding other agents and slows down to avoid collision. This leads to a shorter path being travelled in more time.

## 6 Conclusion

This paper presents an in-depth theoretical comparison between the ORCA, FMP, ASCA and BICARL distributed drone collision avoidance algorithms, along with an empirical comparison between ASCA and BICARL.

The theoretical comparison shows that that FMP and ASCA have the most desirable set of features. First, they require little information that has to be obtained through sensors or communications. Second, they offer a formal proof that guarantees avoidance of collisions. Third, they support 3D environments. Finally, they can easily be extended to avoid static obstacles. Both the ORCA and BICARL algorithm are missing at least two of these features.

The empirical comparison between ASCA and BICARL shows that the ASCA algorithm is much better suited to real-world distributed drone swarm control scenarios. ASCA significantly outperforms BICARL in both a package comparative delivery task and in a scalability test, demonstrating a superior ability to efficiently navigate through increasingly crowded environments. Both the ASCA and BICARL algorithms performed as specified in their respective papers.

While the empirical tests did not include the ORCA and FMP algorithms, similar comparison experiments were conducted in both the original papers presenting the BICARL and ASCA algorithms. Since both report an increase on multiple metrics in performance and scalability test, the tentative conclusion can be made that ASCA will outperform the other three in a full-fledged comparison.

## 7 Discussion

While the conclusions presented in this paper are clear, there are some improvements that could be made. A more conclusive comparison could be made by implementing the FMP and ORCA algorithms and adding them to the empirical comparison.

Furthermore, the scalability experiment only takes parameters from inside the simulation into account, while some external parameter such as execution time can also be a very important factor on computing-constrained agents. Harmonizing the algorithm implementations or testing on real world drones could lead to more insight into these parameters.

Lastly, the simulation environment assumes perfect knowledge of the states of the neighbouring agents, which is never the case in a real world scenario. To more accurately evaluate the real world performance of these algorithms in a simulation, an observation model as in [2] can be used.

## References

[1] Huang, S., Teo, R.S.H., Tan, K.K.: Collision avoidance of multi unmanned aerial vehicles: A review. Annual Reviews in Control **48** (2019) 147–164

[2] Ourari, R., Cui, K., Koeppl, H.: Nearest-neighbor-based collision avoidance for quadrotors via reinforcement learning. CoRR **abs/2104.14912** (2021)

[3] Qazavi, S., Semnani, S.H.: Distributed swarm collision avoidance based on angular calculations. CoRR **abs/2108.12934** (2021)

[4] Semnani, S.H., de Ruiter, A.H., Liu, H.H.: Force-based algorithm for motion planning of large agent teams. IEEE Transactions on Cybernetics (2020)

[5] Van Den Berg, J., Guy, S.J., Lin, M., Manocha, D.: Reciprocal n-body collision avoidance. In: Robotics research. Springer (2011) 3–19

[6] Dhariwal, P., Hesse, C., Klimov, O., Nichol, A., Plappert, M., Radford, A., Schulman, J., Sidor, S., Wu, Y., Zhokhov, P.: Openai baselines. `https://github.com/openai/baselines` (2017)

# An Evaluation of the Usage of Affective Computing in Healthcare

Ella Söderberg

Department of Computing Science
Umeå University, Sweden
elso0026@gapps.umu.se

**Abstract.** This paper is a literature review with a qualitative approach on the topic advantages and disadvantages of affective computing in healthcare. First, the paper gives an overview of what affective computing is. Later in the result part, there is a summary of different sources that explain the advantages and disadvantages of affective computing in healthcare. The paper also talks about some ethical issues regarding the data collecting of human emotions and the integrity issues that come with it. Following is a discussion about the found result. The paper concludes that it exists some disadvantages that could affect the users in such a way that they could receive the wrong type of diagnosis or treatment. The paper also concludes that the advantages of affective computing in healthcare are the diversity of its use and its ability to be applied to different areas.

## 1 Introduction

Affective computing is a field within Human-Computer Interaction (HCI) and is about the study and development of systems that can recognize, interpret, process and simulate human emotions [1]. Researchers have done different studies and papers with different approaches to affective computing. What is missing is a summary of the advantages and disadvantages of emotional technology when using it for designing future solutions. This paper will therefore aim to fill that hole and contribute to the field by gathering information and knowledge from existing papers and studies about emotional technology/affective computing and summarizing it. To pinpoint the search for answers, the paper focuses on emotional technology in the healthcare area. The reason for this is because there is a lot of software suppliers that have researched sensors and algorithms for affective computing to serve different kinds of fields in the healthcare industry [2]. This kind of research often aims to improve the life quality for those patients that suffer from various mental disorders such as stress, depression, or even autism.

An example of affective computing is a product named Feel. Feel is a product that is a wristband connected to a mobile application. The wristband can with its sensors monitor the user's psychological signals and recognize the user's emotional patterns. The application responds to these psychological changes and offers the user different advice, tools, exercises, and other resources for the user

L. Jiang, A. Jonsson, L. Vanhée (Eds.): Umeå's 25$^{th}$ Student Conference in Computing Science USCCS 2022, pp. 69–78, January 2022.

to be able to cope with their feelings. A response could for example be that the application helps the user with a breathing exercise [3]. However, there could be problems regarding affective computing such as the ability to interpret and recognize emotions. For example, if the system recognizes another emotion that the user is experiencing and responds in a faulty way that does not aid the user.

The goal of this paper will be to answer the following research question: *What are the pros and cons of affective computing when using it in healthcare?* This will be researched by the following structure: The first section consists of an introduction to the paper and a summary of affective computing. The second part of the paper, earlier work, describes an example of a paper with a similar research question but in a different field of science. This paper has been one of many to be inspired regarding the structure of the paper. The third section is a description of the method that was used for this research. Following the method, is a fourth section where the results of the research have been summarized. The fifth and last part is a conclusion of the result.

## 2 Earlier work

This section of the paper contains a brief overview of affective computing, what it is and how it is used. This section also provides an example of how to structure the paper when doing a literature review.

### 2.1 Affective computing

One of the pioneers regarding affective computing is Rosalind Picard. Her ideas in the 90s had an impact in the fields of artificial intelligence (AI) and human-computer interaction (HCI). The general idea she got regarding affective computing was that "it should be possible to create machines that relate to, arise from, or deliberately influence emotion or other affective phenomena" [4].

The goal for the field of affective computing is to achieve an interaction between humans and computer systems that can be identified as human-like or life-like. The system should also be able to interpret and adapt itself after the human's emotional state and aim to influence the human by various expressions and actions [4].

For computers to be able to interpret emotions and give reasonable feedback, many tools can be used to interpret emotions and give reasonable feedback that when designing solutions for affective computing. Tools for input could for example be facial recognition, voice recognition, body posture recognition, and other biological sensors. For expressing emotions when designing for affective computing there are other tools for output. For instance, emotions can be expressed as characters in the interface or as robot behaviors [4, 5]. The main goal for this branch of the field of HCI is for computing technology to be able to emulate empathy.

### 2.2 Literature review structure

An example of a paper that has a similar approach on the methodology and type of question is a paper named *"Web and Mobile Based HIV Prevention and Intervention Programs Pros and Cons – A Review"* [6]. The structure of the paper contains an introduction, followed by a method, result, and conclusion section [6]. Since the approach is similar to this paper, the method part is going to be inspired by it.

## 3 Method

This study aimed to review papers on the subject of affective computing in healthcare, to answer the research question *"what are the pros and cons of affective computing when using it in healthcare?"*.

This was done by a qualitative literature review, meaning that the sources chosen for this review are critically chosen, analyzed, and examined. The priority is to find a few qualitative articles rather than finding many articles that only scratch on the surface of the subject. The review should hopefully result in an advancement of knowledge about the research question, rather than a simple overview of the research area.

### 3.1 Research strategy

The research strategy for this paper is to find relevant articles that have been published on trusted platforms such as "Google Scholar". "Google Scholar" is a research platform that gather scholar literature from different libraries into one search engine were one can search for various topics, disciplines and sources [7].

The articles were selected by reading titles and abstracts to decide their relevance for the research question. To complement this strategy an organized search will be conducted where the following keywords will be searched for: Affective computing, emotional technology, healthcare, emotion recognition, affective, emotions, ethical computing.

### 3.2 The selection of articles

The selected literature is analyzed to determine the quality of the papers. To check the quality, a checklist was used that examined the following aspects:

**Value:** The chosen literature needs to bring value to the research, meaning that it must describe necessary aspects regarding the advantages and disadvantages of affective computing in healthcare.

**Readability:** The chosen literature should be well structured and have a logical sequence of content.

**Accuracy:** The chosen literature should not be older than 10 years. The reason for this is because findings of older literature will not be useful due to the fast-developing technology. That implies that the emotion technology that was used 10 years ago, may not have the same advantages and disadvantages as more recent technology has today. Therefore, the answer to the research question might differ depending on the age of the literature.

### 3.3 Presentation of research

The found research will be summarized and presented in the result part of the paper. The result part will mainly be divided into three sections, advantages of affective computing in healthcare, disadvantages of affective computing in healthcare, and ethical point of view of affective computing. In the end, the found result will be concluded in the paper to be able to answer the research question.

## 4 Result

This part of the paper presents the found literature on the subject of affective computing in healthcare. Here each finding is concluded and separated by different rubrics and sections.

### 4.1 Advantages of affective computing in healthcare

This section presents some of the positive outcomes of using affective computing in healthcare. Following are some examples of areas in healthcare where affective computing has been applied successfully. The examples are chosen based on the previous criteria mentioned in the method section and are supposed to represent some different areas in healthcare where affective computing has been applied. The purpose of the examples is also to emphasize the diversity of affective computing and its use.

**Emotion recognition using Google Glass:** In the field of healthcare, affective computing has for example been applied as a tool for helping people with autism to recognize emotions. To be able to do this, scientists used the Google Glasses [8] and added an application with advanced AI techniques. These techniques can, with the help of glasses, detect faces and facial expressions to interpret emotions. When an emotion is detected the application sends the emotion to the glasses and shows them as an emoji to the user. For the glasses to be able to recognize emotions, it first has to detect a face to be able to evaluate its facial expressions later. For the user to know when a face has been detected, a green light appears in the perceptual view and the user gets a confirmation that the face is currently being evaluated for emotion detection [9].

The use of this technology has helped people on the autism spectrum with their daily social interactions with other people. In the article "Upgraded Google

Glass Helps Autistic Kids 'See"', the author further writes about how this confirmation of a found face in the perceptual view has encouraged kids with autism to make more eye contact with people. When testing this technique on real kids with autism they got feedback from the families that further confirmed the improved eye contact these kids made after using the glasses [9, 10].

**Moving wheelchair with facial recognition:** Using emotion technology in healthcare does not necessarily mean using it for keeping track of the user's mental and physical health, it could also mean using it as a tool for example controlling a wheelchair. In an article in the USA TODAY TECH named "A smile can move this motorized wheelchair", it is described how a motorized wheelchair is moving by using facial expressions as commands [11]. Instead of using a joystick, the user could for example raise an eyebrow for making the wheelchair turn right.

In the prototype, the user uses an application for deciding what kind of facial expressions they want to be tied to the specific movements of the wheelchair. The wheelchair can be moved to the left, right, forward, or backward. For tracking facial expressions a combination of facial recognition software, sensors, robotics, and an Intel 3D RealSense Depth Camera (that is attached to the wheelchair) is used. The prototype then captures a 3D map of the user's face and uses AI algorithms for processing the data and then directs it to the wheelchair [11].

The algorithms in the prototype can also detect if a user, for example, is sneezing. The wheelchair can separate that action for an actual gesture that is used for moving the wheelchair, which prevents unintentional effects. This type of use of affective computing technology can help people with different kinds of disabilities to improve their everyday life.

**AI mental coach in smartphones:** Something that has been more and more common is AI therapy through your smartphone [12]. One of the existing applications is a solution called Woebot. Woebot is an AI chatbot that has been trained with cognitive-behavioral therapy (CBT). By talking to the user, Woebot builds an emotional model of the user to help them see their emotional pattern over time. By analyzing data, the system can offer the user tools for handling their emotions such as breathing exercises [13].

A clinical trial of Woabot was conducted where the test persons were divided into two groups. The first group used Woebot for therapy and the second group was directed to the National Institute of Mental Health ebook, "Depression in College Students". After two weeks of using the different tools of handling mental health, the test persons answer a questionnaire. One of the results that were discovered was that the group that had been using WoeBot significantly reduced their symptoms of depression over the trial period, while those in the control group that only had access to the Mental Health ebook had not. The clinical trial concluded that the conversation agent, Woebot, appeared to be a feasible, engaging and effective way to deliver CBT [14].

Another study that also was conducted with Woebot showed that the bond between the user and the AI agent was established after as quickly as 3–5 days and did not appear to diminish over time. The study also showed that the bond between the user and the agent did not interfere with the bond between the users and their human therapist [15]. The affective computing technology has in this example helped users to access CBT treatment. This kind of access makes it easier for the users to aid themselves whenever they need psychological support to help them improve their mental health.

### 4.2 Disadvantages of affective computing in healthcare

This section presents some of the negative outcomes of using affective computing in healthcare. Follows are descriptions of things that need to be paid attention to when designing for affective computing. If not, it can result in severe consequences.

**Miscommunication:** When using technology for emotion recognition there is a possibility that the emotions are misinterpreted. This is because human emotions are complex, meaning that a smile does not necessarily mean that the human is happy, the human can at the same time feel sadness or anger. Multiple emotions can exist at the same time in one person, which is a problem since today's facial recognition systems only can detect one emotion at a time. These emotions are detected by using facial expressions, for example, joy is detected by the corners of the mouth being raised [16].

If the emotions are misinterpreted it could lead to a negative impact on a user's life. An example that was given in a paper named "Pitfalls of Affective Computing" was that if a person is expecting an interaction to be awkward based on the emotion recognition systems prediction the person might behave differently than usual. This could lead others involved in the interaction to be off-balance, since the behavior of the user is off-putting, and further lead to awkwardness and making it seem like the system is correct in its prediction [17].

By giving these kinds of predictions, the system affects the outcome of the interactions and may interfere with the interaction itself which creates a dissonance between the emotions the user feels and the emotions the system detects [17]. If that is the reality, using affective computing technology be beneficial for the user but instead affect them negatively.

**Security risks:** Some technology in healthcare might need to process, store and generate different kinds of data to be able to provide correct diagnosis or treatment for the patients. For example, if a device is designed for detecting anxiety the system of the device needs to evaluate and process data to be able to learn the patterns of that specific patient. This kind of system uses machine learning which requires loads of data for detecting these abnormalities and providing help for the patient. By doing so the system creates self-learning algorithms so that machines can learn from themselves [18].

However, the gathered data can be misused if accessed by third parties. Since the data probably is stored on an external cloud service it may be possible for others to access it. It becomes a question of cybersecurity where the risk of exposed patient data may lead to serious consequences.

For example, if a patient's data somehow has been altered or even deleted, it might lead to them getting the wrong diagnosis or treatment. This could further lead to severe consequences regarding the patient's health, since the patient and the healthcare workers may rely on the system to give them the correct treatment plan [19].

There is also a risk of third parties accessing medical devices with the intent of altering their functionality. Therefore such devices need to have correct security protocols for preventing the device from malfunctioning [19].

### 4.3 Ethics of affective computing

The following section provides information about the ethics of affective computing and how it is applied and handled in the field.

As mentioned in the introduction, the goal for the field of affective computing is to achieve an interaction between humans and computer systems that can be identified as human-like or life-like. To be able to do this, data about the user's vital information needs to be collected for the system to process and evaluate for the system to be able to act out accordingly. However, some ethical issues need to be considered when handling such sensitive data.

An ethical issue to be aware of is the privacy of the user's emotions, since emotions are very personal. By then collecting information about users' emotions could result in integrity issues. For example, if a computer detects that a human is angry, is it okay for the computer to manipulate that feeling so the human will not be angry anymore? And how does that situation differ when comparing it to a human-human interaction? If the interaction is between human-human and one of the humans detects that the other one is angry there are two approaches to take. Either one can try to manipulate the angry person by trying to make them feel another emotion, or, depending on the situation, let the anger be if the angry human wants the other person to recognize his emotion [20].

Humans are constantly detecting emotions by recognizing facial expressions, body movements, gestures, and voice [21]. By picking up information about those things, the human can recognize the emotion and respond in a way that can be considered highly ethical. For example, a response could be to cheer someone up with a joke if a person senses that another is feeling sad [20]. The issue that arises is whether the emotion-sensing computer is free to respond to any human emotion and manipulate the emotion in a way that fits the computer protocol.

## 5 Discussion

This paper has reviewed some of the work in the field of affective computing. The purpose of the review was to find out the advantages and disadvantages of affective computing in healthcare and thereby answer the research question.

What has been found is that affective computing can be applied to multiple different areas within healthcare and has a lot of diversity in its use. In the result section, the paper presents three different areas where affective computing has been successfully applied to. For example the use of emotion recognition to help people with autism to determine the feelings of the people they interact with [10] and the use of facial gesture recognition for helping disabled people control their wheelchair [11].

The diversity of the use of affective computing is a big advantage since it opens doors where only the imagination is the limit. However, when designing technology for affective computing some things need to be considered or else it could lead to negative consequences.

As mentioned in the result part, one of the concerns when using affective computing is if the system misinterprets the data and returns the wrong input. If this kind of miscommunication happens when using the technology within the healthcare industry, it could lead to such consequences that could mean that patients receive wrong diagnosis or treatment. Another thing that could lead to the same consequences as if the data is misused by third parties. If the security of the system is bad and others can access such sensitive and personal data it will be problematic.

Further on the paper also talks about the ethics of affective computing. An issue that was described was about the privacy of the user's emotions and if it is okay or not for the computer to detect and manipulate those feelings. An example of this scenario would be if the system makes the user go from sad to happy. This type of issue could both be an advantage and disadvantage since it depends on users and how they want their emotions to be managed by the system. However, it is important to think about this when designing for future solutions and be aware that users sometimes, even if they use the system frequently, do not want or need their feelings to be managed.

## 6 Conclusion

Based on the result and discussion of the research question, *"what are the pros and cons of affective computing when using it in healthcare?"*, can be answered and concluded. The found advantage of affective computing in healthcare was found to be the diversity of the use of the technology. The technology can be used in multiple different areas within the healthcare industry and help users to improve their everyday life. The technology can also be a tool for hospital workers for helping them to diagnose and treat patients.

The found disadvantages were the risks of miscommunication and security. If the collected data from the user is misinterpreted, it could lead to consequences related to wrong output from the system which could lead to wrong treatment or diagnosis. Also if the security of the system is low, it could make it easy for none authorized parties to access personal information about the users and misuse it.

Suggested future work for this topic would be to conduct a more comprehensive investigation on the topic. Since this paper was written by one author as a school assignment on a deadline, the time limit and resources could have affected the result of the research.

## References

[1] A, B.: What is affective computing? (2016) https://www.bbvaopenmind.com/en/technology/digital-world/what-is-affective-computing/, accessed 2021-12-06.
[2] Silicon, A.: Computer vision with affective intelligence benefits healthcare (2020) https://www.activesilicon.com/news-media/news/affective-computing-in-healthcare/, accessed 2021-12-06.
[3] Feel: Feel emotion sensor (2021) https://www.myfeel.co/feel-program accessed 2021-10-18.
[4] K, H.: Affective computing (2021) https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/affective-computing, accessed 2021-10-18.
[5] Moshfeghi, Bedasso, B.: Emotion recognition from eeg signals using machine learning. https://www.diva-portal.org/smash/get/diva2:831471/FULLTEXT01.pdf, accessed 2021-10-18 (2013)
[6] S, N., E, M., T, N., E, G.: Web and mobile based hiv prevention and intervention programs pros and cons – a review. https://pubmed.ncbi.nlm.nih.gov/28508813/, accessed 2021-11-05 (2017)
[7] Scholar, G.: About google scholar (2022) https://scholar.google.com/intl/sv/scholar/about.html accessed 2022-01-02.
[8] Wikipedia: Google glass (2021) https://sv.wikipedia.org/wiki/Google_Glass, accessed 2021-10-22.
[9] N, H., C, V., D, W.: Upgraded google glass helps autistic kids "see" (2020) https://spectrum.ieee.org/upgraded-google-glass-helps-autistic-kids-see-emotions/particle-8, accessed 2021-10-22.
[10] J, D., J.N, S., C, V.: Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism (2021) https://doi.org/10.1038/s41746-018-0035-3, accessed 2021-10-22.
[11] E, B.: A smile can move this motorized wheelchair (2018) https://eu.usatoday.com/story/tech/talkingtech/2018/12/03/ai-wheelchair-controlled-smile-other-facial-expressions/2183372002/ accessed 2021-11-11.
[12] A., A.: Five tech innovations that changed mental health in 2020 (2020) https://www.forbes.com/sites/forbestechcouncil/2020/11/25/five-tech-innovations-that-changed-mental-health-in-2020/?sh=4e1427331e9c accessed 2021-11-23.
[13] Woebot: For users (2021) https://woebothealth.com/for-users/ accessed 2021-11-23.

[14] Fitzpatrick K, Darcy A, V.M.: Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial. https://pubmed.ncbi.nlm.nih.gov/28588005/, accessed 2021-11-23 (2017)

[15] Darcy A, Daniels J, S.D.W.P.R.A.: Established with a digital conversational agent: Cross-sectional, retrospective observational study. https://formative.jmir.org/2021/5/e27868/PDF, accessed 2021-11-23 (2021)

[16] Recfaces: Emotion recognition: Introduction to emotion reading technology (2021) https://recfaces.com/articles/emotion-recognition accessed 2021-11-11.

[17] Cooney M, Pashami S, P.S.A.F.Y.N.S.: Pitfalls of affective computing: How can the automatic visual communication of emotions lead to harm, and what can be done to mitigate such risks? https://people.ict.usc.edu/~gratch/CSCI534/Readings/OHAC-14-Ethics.pdf, accessed 2021-11-11 (2018)

[18] iTech: The importance of data processing in machine learning (2020) https://itechindia.co/blog/the-importance-of-data-processing-in-machine-learning/ accessed 2021-12-06.

[19] Data, G.: The disadvantages of technology in healthcare (2021) https://www.galendata.com/disadvantages-of-technology-in-healthcare/ accessed 2021-11-17.

[20] R, P.: Affective computing: Challenges. https://affect.media.mit.edu/pdfs/03.picard.pdf, accessed 2021-11-23 (2003)

[21] Wikipedia: Emotion recognition (2021) https://en.wikipedia.org/wiki/Emotion_recognition accessed 2021-11-23.

# Author Index