UNIVERSITY OF SOUTHAMPTON

Consent Mechanisms in Privacy Engineering

by

Dorota Filipczuk

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Monday 18th October, 2021

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Dorota Filipczuk

As the number of online services powered by personal data is growing, the technology behind those services raises unprecedented concerns with regard to users' privacy. Although there are significant privacy engineering efforts made to provide users with an acceptable level of privacy, often users lack mechanisms to understand, decide and control how their personal data is collected, processed and used. On one hand, this affects users' trust towards the service provider; on the other, under some regulatory frameworks the service provider is legally required to obtain user's consent to collection, use and processing of personal data. Therefore, in this thesis, we focus on privacy engineering mechanisms for consent. As opposed to the simple act of clicking 'I agree', we view consent as a process, which involves the formation of user's privacy preferences, the agreement between the user and the service provider and the implementation of that agreement in the service provider's system.

Firstly, we focus on understanding the user's consent decision-making. Specifically, we explore the role of privacy knowledge in data sharing. To that end, we conduct an experiment, where we inform participants how they stop allowing the collection of their online activity data. We compare the behaviour of two groups with an increased knowledge of data collection: one provided only with actionable information on privacy protection, and one additionally informed about the details of how and by whom the collection is conducted. In our experiment, we observe no significant difference between the two groups. Our results suggest that procedural privacy knowledge on how users can control their privacy has impact on their consent decisions. However, we also found that the provision of factual privacy knowledge in addition to procedural knowledge does not effect users' prevention intent or behaviour. These outcomes suggest that the information about privacy protection itself may act a stimulus for users to refuse consenting to data collection.

Secondly, we investigate the idea of agent-based privacy negotiations between a user and a service provider. To that end, we propose a novel framework for the implementation of semi-automated, multi-issue negotiation. Our findings suggest that such a framework is more suitable for negotiation in the privacy domain that the 'take-it-or-leave-it' approach or setting privacy preferences manually, because it allows for a collaborative search for mutually beneficial agreements: users consent to data use more often, consent is more consistent with users' data-sharing sensitivity and it requires less users' effort. Moreover, in order for an agent to accurately represent the user, the agent needs to learn the user's privacy preferences. To address this problem, we compare two approaches to privacy preference elicitation through a user study: one where the preferences are personalised for each user based on their previous consent and one where the user classified into one of the three privacy profiles and later re-classified if their consent decisions reflect a change. We find that the latter approach can represent the user more accurately in the initial negotiation rounds than those of the former.

Finally, we look at the implementation of consent on the service provider's side after the agreement regarding data use has been made. In more detail, we consider a scenario where a user can deny consent to process certain data for certain purposes. To that end, the existing approaches do not allow service providers to satisfy the user's consent in the optimal way. Therefore, we propose a novel graph-theoretic model for the service provider to store consent, which indicates the kinds of data processing that can be performed under the privacy agreement. Then, we formalise the consent problem as a constraint satisfaction problem on graphs. We provide several algorithms to solve the problem and compare them in terms of their trade off between execution time and quality of the solution. Our algorithms can provide a nearly optimal solution in the face of tens of constraints and graphs of thousands of nodes in a few seconds.

The research presented in this thesis contributes to understanding users' consent decision-making and addresses an emerging need for technologies that can help service providers manage users' consent. We propose ideas for potentially fruitful lines of exploration within this area.

Contents

List of Figures **ix**

List of Tables **xi**

List of Algorithms **xiii**

List of Accompanying Material **xv**

Declaration of Authorship **xvii**

Acknowledgements **xix**

1 Introduction **1**
1.1 Research Challenges . . . 5
1.2 Research Requirements . . . 7
1.3 Research Contributions . . . 9
1.4 Thesis Outline . . . 10

2 Background **13**
2.1 Origins of Consent Mechanisms . . . 14
2.1.1 Consent in Early Policy Debates . . . 14
2.1.2 Consent as a Legal Requirement . . . 16
2.1.3 Criticism of the Consent Requirement . . . 19
2.2 Consent Decision-Making . . . 20
2.2.1 Privacy Concerns . . . 21
2.2.2 Privacy Paradox . . . 22
2.2.3 Privacy Knowledge . . . 23
2.3 Automated Consent . . . 25
2.3.1 Privacy Self-Management Tools . . . 25
2.3.2 Privacy Negotiations . . . 27
2.3.3 Privacy Preference Elicitation . . . 29
2.4 Consent Infrastructure . . . 31
2.4.1 Privacy Agreement Representations . . . 31
2.4.2 Enforcement of Privacy Agreements . . . 34
2.4.3 Consent Propagation . . . 36
2.5 Summary . . . 37

3 Informed Consent **39**
3.1 Preliminaries . . . 40

3.1.1 Online Tracking Technologies . . . 40
3.1.2 Online Tracking Countermeasures . . . 42
3.1.3 Web Mirror . . . 43
3.2 Motivation . . . 44
3.3 Experiment . . . 47
3.3.1 Apparatus . . . 48
3.3.2 Methodology . . . 49
3.3.3 Participants . . . 51
3.3.4 Results . . . 53
3.4 Discussion . . . 54
3.5 Limitations . . . 56
3.6 Summary . . . 57

4 Negotiable Consent **59**
4.1 Preliminaries . . . 60
4.1.1 Automated Negotiation . . . 60
4.1.2 User Preference Elicitation . . . 61
4.1.3 Automated Negotiation Agent for Permission Management . . . 62
4.2 Motivation . . . 63
4.3 Negotiation Framework . . . 65
4.3.1 Negotiation Setting . . . 65
4.3.2 Negotiation Protocol . . . 66
4.3.3 Utility . . . 66
4.4 Negotiation Agent . . . 68
4.4.1 Models of Uncertainty . . . 69
4.4.2 Privacy Preference Elicitation . . . 70
4.4.3 Negotiation Strategy . . . 71
4.5 Negotiation of Privacy . . . 73
4.5.1 Negotiation Domain . . . 73
4.5.2 Overview of the Negotiation . . . 74
4.5.3 Agent 1: Individual Preferences . . . 75
4.5.4 Agent 2: Type-Based Preferences . . . 75
4.6 Experimental Evaluation . . . 76
4.6.1 Offer Completion . . . 76
4.6.2 Apparatus . . . 76
4.6.3 Methodology . . . 79
4.6.4 Participants . . . 81
4.6.5 Results . . . 82
4.7 Discussion . . . 86
4.8 Limitations . . . 88
4.9 Summary . . . 88

5 Implementable Consent **91**
5.1 Preliminaries . . . 92
5.1.1 Computational Complexity Theory . . . 92
5.1.2 Graph Theory . . . 93
5.1.3 Graph-cutting Problems . . . 94

5.2 Motivation . . . 96
5.3 Consented Data Workflow Problem . . . 98
5.3.1 Model . . . 99
5.3.2 Problem Formulation . . . 100
5.4 Complexity Analysis . . . 101
5.5 Additive Model . . . 104
5.6 Algorithms . . . 106
5.7 Optimality of Solutions . . . 109
5.8 Experimental Evaluation . . . 113
5.8.1 Methodology . . . 113
5.8.2 Results . . . 116
5.9 Discussion . . . 124
5.10 Limitations . . . 125
5.11 Summary . . . 126

6 Conclusions **129**
6.1 Future Work . . . 132

A Questionnaire **135**

Bibliography **139**

List of Figures

2.1 An example of a P3P privacy policy. 32
2.2 An example of an APPEL rule set. 33

3.1 The interface of the Web Mirror. 44

4.1 Sequence diagram of a negotiation that ended after n rounds. 67
4.2 The interaction model between the user, the agent and the opponent. . . 69
4.3 The interface of the experimental tool. 78
4.4 Number of times the participants granted permissions to each of the resources. 83
4.5 Means of self-reported data sensitivity scores on a 7-point Likert scale. . . 84
4.6 The accuracy of Agent 1 (A1), Agent 2 (A2) and the manual negotiation (MN) in each scenario. 85

5.1 Data workflow in the system described in the motivating scenario (U – user vertex, A – algorithm vertex, P – purpose vertex). 97
5.2 Illustration of the data workflow (U – user vertex, A – algorithm vertex, P – purpose vertex). 99
5.3 Example importance and utility values in the special instance of the data processing system (U – user vertex, A – algorithm vertex, P – purpose vertex). 105
5.4 A data processing model where $V^U = \{v_1\}$, $V^A = \{v_2\}$, $V^P = \{v_3, v_4\}$ and $E = \{(v_1, v_2), (v_2, v_3), (v_2, v_4)\}$. 110
5.5 A data processing model where $V^U = \{s_1, s_2\}$, $V^A = \{v_1\}$, $V^P = \{t_1, t_2\}$ and $E = \{(s_1, v_1), (s_2, v_1), (v_1, t_1), (v_1, t_2)\}$. 111
5.6 The number of constraints vs. the runtime of the algorithms in graphs (dataset 1a). 116
5.7 The number of constraints vs. the runtime of the algorithms in graphs (dataset 1b). 117
5.8 The number of constraints vs. the runtime of the algorithms in graphs (dataset 1c). 117
5.9 The number of constraints vs. graph utility after applying the algorithms on graphs (dataset 1a). 118
5.10 The number of constraints vs. graph utility after applying the algorithms on graphs (dataset 1b). 119
5.11 The number of constraints vs. graph utility after applying the algorithms on graphs (dataset 1c). 119
5.12 Number of paths vs. runtime (dataset 1c). 121
5.13 Number of paths vs. utility (dataset 1c). 122

5.14 Path length vs. time in sparse graphs (dataset 2). 122
5.15 Graph size vs. runtime (dataset 3). 123
5.16 Graph size vs. utility (dataset 3). 123

List of Tables

3.1 Demographics of the study participants. 52

4.1 A summary of the treatments used in the experimental evaluation. 79
4.2 A summary of the user study procedure as experienced by a participant. . 80
4.3 The number of participants of each privacy type per treatment. 81
4.4 Mean, median and standard deviation of the perceived effort of negotiation in each treatment on a scale from 0 to 20. 86

5.1 Parameter configurations for datasets 1, 2 and 3. 115
5.2 Comparison of the graph's utility after applying RemoveMinMC and BruteForce. 120

List of Algorithms

1 RemoveRandomEdge 106
2 RemoveFirstEdge 107
3 RemoveMinCuts 108
4 RemoveMinMC 108
5 BruteForce 109

List of Accompanying Material

The material accompanying this thesis is available electronically at: `https://doi.org/10.5258/SOTON/D2003`.

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signed: Date:

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Dr Enrico H. Gerding and Dr George Konstantinidis. Thank you for guiding me throughout this research, extensive conversations, constructive feedback, continuous support and for believing in me.

I would also like to thank my examiners, Dr David Millard and Dr Nadin Kökciyan for their time and effort invested in reviewing this thesis, an interesting discussion and valuable comments on my research.

This work would not have been possible without generous funding from the EPSRC. I am honored to have been awarded the EPSRC Enhanced Studentship. I would also like to acknowledge the use of the IRIDIS High Performance Computing Facility and associated support services at the University of Southampton in the completion of this work. I would like to thank each and every one working on these services for allowing me to run my computationally intensive experiments.

Moreover, I would like to thank my collaborators, Dr Richard C. Gomer and Dr Tim Baarslag, who reviewed several versions of the work that ended up in Chapters 3 and 4 accordingly. I am grateful that I had the opportunity to learn from you and build upon your work. Thank you for all your feedback, discussions and support.

Furthermore, I would like to thank my colleagues from the Electronics and Computer Science department. In particular, I would like to thank Dr Paolo Pareti and Dr Luis-Daniel Ibáñez Gonzalez for interesting conversations that inspired some of my ideas in Chapter 5 of this thesis. I would also like to thank Dr Taha Doğan Güneş for immediate answers to my questions when I was getting started with IRIDIS, and Dr Adriana Wilde and Dr Stephen Snow for all the mentorship on this PhD journey.

Last but not least, I would like to thank Edoardo – for everything, really. Resilience is one thing but courage is another. Thanks for showing me every day what true Italianness looks like. I love you too.

Chapter 1

Introduction

Over the past half century, issues involving privacy have become more persistent and concerning than ever before. With the proliferation of mass data collection systems, surveillance technologies, numerous high-profile scandals[1] and data breaches[2], data privacy received a great deal of attention from scholars, policy makers, governments and businesses. At the same time, no universal definition of privacy has yet been established. In fact, the concept of the 'right to privacy' was first proposed by Warren and Brandeis (1890) who defined it as 'the right to be let alone'. However, some argue that 'one aspect of privacy is the withholding or concealment of information' (Posner, 1978) and that it refers to 'the control we have over information about ourselves' (Fried, 1968). Alternatively, Westin (1968) described privacy as:

> *'the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others'* (Westin, 1968).

Westin's desire for individuals to retain ultimate control over their personal information comes in light of privacy threats posed by early technological advancements. This includes how much of it and to whom it should be disclosed, how it should be maintained and how disseminated (Westin, 1968). Since then, both the scale of data collection, processing and sharing, and the amount of personal information involved have been constantly growing. Fundamentally, both the number of online users and the number of connected devices is now measured in billions (Nordrum, 2016; Roser et al., 2015), and the volume of digital records worldwide – in zettabytes (Seagate, 2012). As data processing is becoming embedded in everyday objects and linked with the people's daily lives, an overwhelming majority of online users express concerns regarding its potential harms (Paine et al., 2007).

[1] E.g. the Cambridge Analytica scandal (Berghel, 2018; Isaak and Hanna, 2018); the 'Cookiegate' scandal (Ring, 2015).

[2] E.g. the Equifax data breach (Zou et al., 2018).

Thus, as a result of public debates on data privacy, many countries have imposed strict rules for the collection, processing and handling of personal information. As early as in 1973, a report by the Department of Health, Education and Welfare of the United States (US) listed the Fair Information Practice Principles to address concerns about the increasing digitization of data (US Department of Health, Education, and Welfare, 1973). These principles were embodied selectively in various statutes in the US and helped shaping the OECD Privacy Guidelines of 1980 (Organisation for Economic Co-operation and Development, 1980), the European Data Protection Directive 95/46/EC of 1995 (European Parliament and the Council of the European Union, 1995) and, recently, the European Union's General Data Protection Regulation (GDPR) (European Parliament and the Council of the European Union, 2016). The ultimate objective of these regulations is to provide online users with the ability to control the flow of their information (Bygrave, 2004; Guarda and Zannone, 2009).

Consequently, online service providers are incentivised to enable such privacy control. For example, a failure to comply with the GDPR may result in penalties of up to 4% of the service provider's annual global revenue of the preceding fiscal year or €20 million, whichever is higher (European Parliament and the Council of the European Union, 2016). Since the GDPR came into force in 2018, several enforcement actions have been taken against businesses that violated the regulation, including large online service providers such as Google and Facebook (Houser and Voss, 2018). However, apart from legal compliance, there is another and, perhaps, more important incentive for service providers to empower users with privacy controls: trust. That is, there is significant evidence suggesting that privacy concerns affect users' trust in online services (Milne and Boza, 1999; Wu et al., 2012). In particular, Martin (2016) finds that 'violating informal privacy norms negatively impacts trust in the website even when the information exchange conforms to or is not mentioned in the privacy notice'. She argues that users rely on privacy norms, because they are 'vulnerable to information asymmetries and uncertainty', and that 'respecting privacy norms is key to trust online'.

Nonetheless, implementing data privacy is a non-trivial problem, as it 'requires the translation of complex social, legal and ethical concerns into systems requirements' (Anthonysamy et al., 2017). This is because the process of engineering systems with privacy in mind, also known as *privacy engineering* (Spiekermann and Cranor, 2008), 'requires integrating privacy requirements into the typical systems engineering activities' (Gürses et al., 2011). In the realm of increasing data exchanges between billions of users and various service providers, 'this effectively amounts to privacy management on an ultra-large-scale' (Anthonysamy et al., 2017). Therefore, to address this problem, the concept of *privacy by design* has been proposed (Cavoukian, 2009; Schaar, 2010). Gürses et al. (2011) define this concept as follows:

> '*"Privacy by design" consists of a number of principles that can be applied*

> *from the onset of systems development to mitigate privacy concerns and achieve data protection compliance*' (Gürses et al., 2011).

However, like for the definition of privacy itself, there is some disagreement on what the consequences of these principles are. On the one hand, some researchers believe that the fundamental step in engineering systems in line with the privacy by design principles is data minimisation. According to this approach to privacy engineering, only the data that is absolutely necessary to fulfill the functionality of a system should be analysed (Gürses et al., 2011). In fact, based on this idea, the so-called privacy-enhancing technologies (PETs), have been developed (Goldberg et al., 1997). In more detail, van Blarkom et al. (2003) define PETs as a system of information and communication technology (ICT) measures:

> '*Privacy-Enhancing Technologies is a system of ICT measures protecting informational privacy by eliminating or minimising personal data thereby preventing unnecessary or unwanted processing of personal data, without the loss of the functionality of the information system*' (van Blarkom et al., 2003).

As such, research on PETs focuses on developing cryptographic privacy protections and systems with provable privacy guarantees. However, Spiekermann and Cranor (2008) criticise this approach because of its sacrifice of usability and point out that PET researchers 'tend to favor systems that prevent access to individuals and their information at all cost'. While they agree that the advances in PETs 'may lead to deployable solutions with strong privacy guarantees', they also highlight that for PET researchers:

> '*The goal is to make access to the individual tamper-proof and to build a technological infrastructure based on nonidentifiability of users even vis-à-vis governments. Often, unfortunately, achieving this ambitious goal undermines system usability and drives system cost to a point where marketability and adoption of the solution becomes difficult*' (Spiekermann and Cranor, 2008).

On the other hand, the second approach to privacy engineering focuses on the idea that 'information may be collected for useful purposes such as personalized services' (Spiekermann and Cranor, 2008). As opposed to seeing the online service provider as a 'privacy attacker', Spiekermann and Cranor (2008) describe the researchers of this viewpoint as consent-oriented:

> '*For them, the threat model is what is commercially feasible to do and not what is theoretically doable. This group's goal is to give people control through informed consent to personal data use*' (Spiekermann and Cranor, 2008).

In this thesis, we follow the latter approach to privacy engineering. More specifically, our goal is to give people control of the flow of their information through consent to personal data use (collection, processing, sharing, etc.). At the same time, we do acknowledge the fact that preventing unnecessary or unwanted processing of personal data is a crucial aspect of a well privacy-engineered software system, as research on PETs has shown. However, we take the stand that the decision of which functionality of the system is important should be a result of an agreement between the user and the specific service provider. We refer to such agreement as the *privacy agreement*.

In fact, this is in line with Article 4(11) of the GDPR, which states that consent 'signifies agreement to the processing of personal data' (European Parliament and the Council of the European Union, 2016). Yet, in the scientific literature, many researchers refer to the 'consent process' which goes beyond a single action of simply signifying the privacy agreement (e.g. Luger and Rodden (2013), Bashir et al. (2015)). For instance, Luger and Rodden (2013) say:

> '*The act of consent is significantly more than a box-ticking exercise*' (Luger and Rodden, 2013).

In fact, Van Der Geest et al. (2005) brings attention to the fact that in the health care sector consent on the use and application of personal data is defined more extensively:

> '*Informed consent is the process by which a fully informed user participates in decisions about his or her personal data. It originates from the legal and ethical right the user has to direct what happens to his or her information, and from the ethical duty of organisations using personal data to involve the user in the control, use and maintenance of these data*' (Van Der Geest et al., 2005).

What is more, Coles-Kemp and Kani-Zabihi (2010) argue that since users have privacy concerns but are prepared to trade their privacy for a reward, consent should be a 'micro dialogue':

> '*In order to support service users in making situated decisions about the deployment of privacy controls and exercising of privacy practices, there needs to be a dialogue between the service user and service provider which enables the service user to understand the implications of their privacy practices* (Coles-Kemp and Kani-Zabihi, 2010).

If we take a closer look at this dialogue, we can see that the consent process starts way before the user signifies the privacy agreement. To this end, we look at consent as a process of reaching and honouring the agreement, which involves:

- the user forming their privacy preferences, based on any knowledge and concerns they have, which shape their consent decision-making behaviour;
- the user and the service provider agreeing on certain terms of the agreement which the user then consents to;
- the service provider complying with the agreement in all aspects of data use (collection, processing, sharing, etc.), while ensuring the best possible quality of the service.

In order to enable the such a consent process, adequate privacy engineering mechanisms need to be in place. Such mechanisms are what we call *consent mechanisms*.

1.1 Research Challenges

Existing consent mechanisms have received a lot of criticism from the research community. One popular mechanism that much research attention has focused on so far are privacy notices (Bannihatti Kumar et al., 2020; Sadeh et al., 2013; Schaub et al., 2015, 2017; Wilson et al., 2016). In particular, critics highlight the fact that privacy notices are ill-equipped to educate users of corporate data practices (Luger et al., 2013; Waldman, 2020). Although there have been proposals of other mechanisms such as 'nutrition labels' (Kelley et al., 2009, 2010), educational videos (Stein et al., 2020) and comics (Anaraky et al., 2019; Knijnenburg and Cherry, 2016), information about corporate privacy practices is limited. As a result, studies suggest that users in general tend to have a low level of privacy knowledge (Bartsch and Dienlin, 2016; Bashir et al., 2015; Dommeyer and Gross, 2003; Nowak and Phelps, 1992; Park, 2013; Turow, 2003; Turow et al., 2005), which may lead to difficulties making a reasonable evaluation of the risks and benefits of decisions to consent (Sloan and Warner, 2014; Solove, 2013). For example, Bashir et al. (2015) conducted a survey that assessed users' knowledge and opinions of online privacy issues. Their results expose several key knowledge gaps, demonstrating a problem of information asymmetry between users and internet services providers, and strong dissatisfaction with the current system. These findings demonstrate that there is insufficient comprehension and voluntariness in the consent process for users to give informed consent to the collection and management of their personal information. We therefore argue that the first challenge in engineering consent mechanisms is about making consent actually *informed*. To that end, consent mechanisms should provide users with adequate privacy knowledge to ensure that there is comprehension and voluntariness in the process.

Furthermore, an increasing number of consent requests users have to respond to brings another challenge to engineering consent mechanisms. When there are too many service providers asking a user for consent to data use, at some point the user starts simply

granting consent without understanding the consequences (Choi et al., 2018). As a result, this leads the user to the feeling of a loss of control and a sense of 'weariness' towards privacy issues (Choi et al., 2018). To that end, critics of the current status quo of consent mechanisms conclude that they struggle to scale (Solove, 2013). As a solution to this problem, Luger and Rodden (2013) suggest automating this process. In fact, users could be represented by personal data managers that negotiate with other agents the appropriate level of exposure and control:

> '*The issue then is how to architect future ubiquitous computing environments to embed what might be termed 'consent by design'. Current approaches to design operate on the presumption of both availability of information and that the system has permission to process the information without engaging users. However, we might envisage a future environment where consent is given the same primacy as security and that we develop models where the user is aware of the processing that takes place on their personal data. This might take the form of a personal data manager that negotiates with agents in the embedded environments, to discover the appropriate level of exposure and control. For example, as your enter a store, the environment might request an appropriate level of data collection permission with a personal manager, based on previously encoded preferences.*' (Luger and Rodden, 2013).

Similarly, Krol and Preibusch (2015) envisioned 'effortless privacy negotiations'. As opposed to the prevailing take-it-or-leave-it approach where the user either accepts the service provider's conditions or stops using the service (Bender, 2011; Polykalas, 2017), they argue that privacy negotiations have a potential to support reaching a compromised privacy agreement, which is beneficial for both parties. Therefore, if consent is truly to involve a bidirectional dialogue between the user and the service provider, then the second challenge in engineering consent mechanisms is about ensuring that consent is by design *negotiable*.

Last but not least, service providers struggle to implement consent. For instance, studies show that the EU websites, which are legally required to implement consent of the user, tend to collect personally identifiable data without consent (Matte et al., 2020; Sanchez-Rola et al., 2019). In the UK, researchers estimate that only 11.8% of the websites meet the minimum requirements based on the EU regulations (Nouwens et al., 2020). To address this problem, researchers proposed database-level solutions to support consent implementation. Inspired by the privacy tenet of the Hippocratic Oath, Agrawal et al. (2002) envisioned databases that include privacy as a central concern and enunciated the key principles for such databases. However, some of the proposed implementations of such databases are based on the assumption that data is manually accessed by an employee, e.g. solutions by Byun et al. (2005); Byun and Li (2008); Massacci et al. (2006); Ni et al. (2010); Petković et al. (2011), whereas in modern large-scale

systems, when user's personal information enters a data processing workflow, it is processed automatically. Other solutions are designed for relatively small data processing scenarios, e.g. those by Agrawal et al. (2002); Ashley et al. (2002a,b, 2003); LeFevre et al. (2004); Karjoth et al. (2002). Currently, the existing solutions do not support automatic implementation of consent in large-scale systems, where data is processed for different purposes and by several service providers. Therefore, the third challenge in engineering consent mechanisms is about ensuring that consent, which signifies the privacy agreement, is *implementable*.

1.2 Research Requirements

As the problem of mechanisms that support the consent process is quite broad, in this thesis we focus specifically on the challenges identified in Section 1.1 – that is, ensuring that consent empowered by adequate privacy mechanisms is *informed*, *negotiable* and *implementable*. Therefore, the research requirements that guide the exploration presented in this thesis are formulated around these three concepts.

Firstly, in order to support the user in controlling the flow of their personal data, the user must be *informed* about the extent and conditions of their data flow, and how they can exercise control over that flow. While in certain regulatory regimes, such as the GDPR, provision of that information is legally required when asking users for consent (European Parliament and the Council of the European Union, 2016), studies show that users tend to approve consent requests without fully understanding their meaning and consequences (Felt et al., 2011, 2012; Kelley et al., 2012). In contrast, relevant literature provides some evidence that users generally have a relatively low level of privacy knowledge (Bashir et al., 2015; Bartsch and Dienlin, 2016; Dommeyer and Gross, 2003; Nowak and Phelps, 1992; Park, 2013; Turow, 2003; Turow et al., 2005) and that the low privacy knowledge correlates with lack of efforts to protect privacy (Acquisti and Gross, 2006; Turow, 2003). In fact, there are calls within the research community for an investigation of the impact of privacy knowledge on users' disclosure behaviour (Bashir et al., 2015; Brough and Martin, 2020; Park, 2013; Smith et al., 2011). Since disclosure of information is part of the consent process, in this thesis, we aim to investigate whether efforts to increase users' privacy knowledge could result in more informed consent decision-making. Therefore, we pose the following research question:

> **RQ1:** *How does the increase of privacy knowledge impact users' disclosure behaviour?*

We investigate this topic in Chapter 3.

Secondly, if consent signifies an agreement between the user and the service provider, consent should be *negotiable*. As such, consent mechanisms should allow both parties to express their preferences in detail in order for them to reach a compromise that both can be comfortable with. With regard to that, related work offers some ways of clustering users with potentially similar preferences into small numbers of privacy preference profiles and inferring preferences of those groups of users (Agarwal and Hall, 2013; Baarslag et al., 2017; Lin et al., 2012; Liu et al., 2014; Nakamura et al., 2016; Mugan et al., 2011; Ravichandran et al., 2009). For example, Baarslag et al. (2017) proposed assigning users into three profile categories (*fundamentalists*, *pragmatists* and *unconcerned*) according to their general level of privacy concern measured using a three-question survey instrument called Privacy Segmentation Index (Kumaraguru and Cranor, 2005). However, Jensen et al. (2005) observed that while those classified as fundamentalists have consistent privacy concerns, they do not appear to form a cohesive group with respect to decision-making. In fact, the investigation by Woodruff et al. (2014) found no correlation between the categories assigned by Privacy Segmentation Index and behavioral intent, as well as a lack of correlation between these categories and peoples' reactions to specific consequences of their privacy decisions. To that end, our second research question involves addressing the differences between individual users in negotiable consent mechanisms:

> **RQ2:** *How can privacy negotiations be conducted such that the user's privacy preferences are automatically taken into account?*

We explore this research question in Chapter 4.

Finally, when a privacy agreement is reached, the agreement must be *implemented* in all aspects of data processing within the service provider's infrastructure. In particular, under legal frameworks such as the GDPR, data processing can only be considered lawful if the consent relates to one or more specific purposes of data processing (European Parliament and the Council of the European Union, 2016). In fact, the user may refuse to consent to some of the data processing, which may significantly affect the utility of the service provider. If the data processing system is large, the service provider may be able to decide *how* the user's consent is implemented. In a large-scale data processing workflow, there may be several ways of satisfying the user's constraints. To that end, the existing solutions do not support service providers in finding the most optimal ways to implement privacy agreements. Thus, our third research question is formulated as follows:

> **RQ3:** *How can service providers satisfy the user's privacy constraints in an optimal way?*

We propose a solution to this problem in Chapter 5.

1.3 Research Contributions

Against this background, we propose several contributions to the development of consent mechanisms that can make consent *informed*, *negotiable* and *implementable*. More specifically, we satisfy the research requirements outlined in Section 1.2 as follows.

Firstly, we show that making consent *informed* has impact on users' privacy behaviour, and therefore, their privacy preferences. In more detail, we explore the relationship between the provision of information and users' consent behaviour in a specific situational context of online tracking. Our results indicate that the provision of actionable information about privacy protection can motivate 22% of users to change their consent decision. At the same time, we find no significant effect of the provision of factual information about the extent of data sharing on the consent decision. Importantly, we present results of the first user study reporting on the adoption of anti-tracking protection techniques that measures participants' actual behaviour.

Secondly, we propose a novel framework for making consent *negotiable*. As part of the framework, we introduce an alternating-offers, multi-issue negotiation protocol for automated negotiation of privacy agreements. Using this protocol, an agent representing the user can autonomously negotiate on the user's behalf by specifying some terms of the privacy agreement, and the service provider's agent completes those terms. Findings from our user study suggest that such a framework is more suitable for negotiation in the privacy domain that the take-it-or-leave-it approach or setting privacy preferences manually, because it allows for a collaborative search for mutually beneficial agreements: users consent to data use significantly more often, consent is more consistent with users' data-sharing sensitivity and it requires significantly less users' effort. In fact, this negotiation framework generalises to other domains where the relationship between the negotiating parties is asymmetric in terms of power, as it is between the individual user and the service provider.

Thirdly, making consent *negotiable* means allowing a user agent to learn the user's privacy preferences to accurately represent them. We compare a new approach, where the preferences are personalised for each user based on their previous consent (granted or refused), to the approach of Baarslag et al. (2017), where the user is classified into one of the three privacy profiles and later re-classified if their consent decisions reflect a change. The results of our user study show that offers proposed by the latter approach is more accurate in the initial negotiation rounds than those of the former. However, we can observe a rising trend of the accuracy of the new agent, which in the end, exceeds the accuracy of the agent of Baarslag et al. (2017).

Finally, we propose a novel approach to making consent *implementable*. That is, we model data processing as a graph and use graph-cutting algorithms to translate consent constraints into data processing policies which allow service providers to know

what kind of data processing of specific pieces of the user's personal data is consent-compliant. Specifically, given a set of consent constraints from the user, our algorithm determines the stages of data processing that cannot be performed on the particular piece of user data. We present our theoretical results which prove the complexity of the problem. Furthermore, we compare five algorithms that can solve the problem in terms of accuracy and performance. We argue that our data processing model, apart from the regulatory compliance, may benefit large service providers through a higher degree of transparency of algorithmic data processing, as well as explainability. In addition, our theoretical results in this area and the proposed algorithms generalise to other graph-cutting problems with a similar additive-weight model.

1.4 Thesis Outline

The remainder of this thesis is structured as follows. In Chapter 2, we expand the discussion on consent mechanisms. In particular, we provide a broader context by presenting a review of the relevant literature and highlight the gaps in the literature that motivate our research.

In Chapter 3, we focus on the *informed* aspect of consent. That is, we explore the relationship between the provision of information and users' consent behaviour in a specific situational context of online tracking. To that end, we first introduce the context of online tracking and protection techniques against it, as well as visualisation tools that can help to educate users about data practices in this context. Then, we explain our motivation and problem formulation in this initial study. After that, we report on the methodology, participants' profile, results and limitations of our experiment. Lastly, we discuss the broader implications of our findings in the light of privacy knowledge and consent behaviour.

In Chapter 4, we discuss a novel approach to *negotiable* consent. Specifically, we illustrate how agent-based negotiation can support the process of reaching a privacy agreement between the user and the service provider, and how a user can be represented by an agent that learns their privacy preferences. Therefore, we first introduce the related theory on agent-based negotiation. Then, we formally propose our framework for such negotiation, with a protocol for bilateral multi-issue negotiation. After that, we demonstrate how the general framework can be implemented to develop an agent that represents the an individual. Next, we apply the framework and its proposed implementation to the privacy permission negotiation domain. Furthermore, we describe the apparatus, methodology, results and limitations of our experimental evaluation. Ultimately, we discuss the implications of our results for future design choices for privacy management and automated negotiation.

In Chapter 5, we propose a novel approach to *implementable* consent. In particular, when the privacy agreement is reached and as part of it, the user refuses to consent to some forms of data processing, that refusal must be implemented by the service provider to ensure consented information flow in the data processing system. We propose a novel data processing model and formulate the corresponding graph-theoretic satisfaction problem. To that end, we first provide some background about concepts from complexity and graph theory, as well as the relevant graph-cutting problems. Then, we formally propose our framework and problem definition. After that, we prove that the problem in general is $\mathcal{NP}$-hard and perform experimental evaluation of algorithms for a simplified instance of the problem. Ultimately, we discuss the implications of our results for future large-scale data processing infrastructures.

In Chapter 6, we draw overall conclusions from the results presented in this thesis and outline possible directions for future work. Specifically, we reflect on potential paradigm shifts in areas such as the presentation of consent options to the user (from binary take-it-or-leave-it to an informed negotiation), everyday privacy management (from being handled manually by the user to automation) and consent management systems (from data-centered to privacy-centered).

Chapter 2

Background

In this chapter, we expand on the previous discussion by examining the existing literature on consent mechanisms in privacy engineering. Specifically, we focus on three phases of the process:

- understanding the user's preferences which are based on their privacy attitudes, values, intentions, perceived risks and benefits,
- reaching a data sharing agreement between the user and the service provider, and
- acknowledgement of that agreement by all affected elements of the service provider's data processing infrastructure.

To that end, the goal of this chapter is twofold. First, we provide a broader context for the discussion by looking at the series of events that led to the need for consent mechanisms in privacy engineering. Since the current state of the art is based on the prevailing data protection legislation, we examine the origins of the current consent mechanisms in this area. Secondly, we identify the gaps in the existing literature that serve as motivation for the rest of the thesis. Specifically, we explain how the legal requirements, the unknowns about users' consent decision-making and the existing privacy engineering techniques motivate our results in Chapters 3, 4 and 5.

The chapter is divided into four main sections. In Section 2.1, focus on how consent mechanisms became a requirement in personal data processing. There, we briefly summarise the early policy debates on privacy threats in data processing, the important data protection laws that those debates led to being enacted, and the criticism of the data protection approach taken then. In Section 2.2, we look at how users make decisions when they grant consent to data processing. In particular, we examine the existing literature on users' related beliefs and attitudes, as well as on the dichotomy between their attitudes and decisions. In Section 2.3, we move on observing how consent decisions can

be automated. To this end, we show how users' privacy preferences can be learnt, as well as how a software agent can negotiate and make consent decisions on behalf of the user. In Section 2.4, we focus on honouring the privacy agreements. Specifically, we review the literature on how consent is represented and implemented, as well as how it propagates to all aspects of data processing. Finally, we give a brief summary in Section 2.5.

2.1 Origins of Consent Mechanisms

We begin with a brief introduction to the origins of consent mechanisms. In more detail, various social, economic and political factors contributed to the currently prevailing consent requirement in global data protection legislation. In this section, we first study the concept of consent in the context of early policy debates that took place globally in response to the advancements in computerised personal data processing. Then, we compare the consent-related requirements in policy outputs in different countries. Last but not least, we report on the scholarly criticism of those requirements.

Notably, in our review, we focus on the specific events that led to the requirement of consent mechanisms in today's software systems. While the full history of privacy principles and data protection legislation provides a valuable background for those interested in the topic, we select the presented material with consent mechanisms in mind. For details on the history of privacy principles or privacy law, see e.g. Bennett (1992); Gellman (2019); van Alsenoy (2019).

Before we continue, it must be clarified that the field of law and policy, which much of the literature cited in this section originates in, has been increasingly adopting a nomenclature that avoids explicit reference to privacy. That is, the term *data protection*, derived from the German term *Datenschutz*, has gained broad popularity in Europe and, to a lesser extent, elsewhere. Whilst Bygrave (2004) argues that supplementing it by the term *data privacy* is more appropriate as 'it better communicates the central interest(s) at stake and provides a bridge for synthesising North American and European policy discussions' (Bygrave, 2004), for simplicity, in this thesis the terms *privacy* and *data protection* are used interchangeably.

2.1.1 Consent in Early Policy Debates

Although legal experts have advocated for a right to privacy since 1890 (Warren and Brandeis, 1890), it was not until the 1960s when the growing use of automated data systems sparked a serious discussion on the topic (Bygrave, 2004). As society continued its transition from an industrial to a post-industrial economy, the advancements in computing technology made it much easier to collect, store, assemble, correlate and use information about individuals than ever before (Bennett, 1992). Simultaneously, the

requirement on citizens provide governments with sensitive information on their financial, employment, health, and educational histories has led to policy debates around the world focused on the asymmetry of power between an individual and a state (Bennett, 1992).

Particularly in the US, where privacy is viewed as a civil liberty (Regan, 2000), terms such as 'privacy', 'surveillance', 'personal freedom' and 'trust' started being used to evaluate the implications of computerised processing of personal data (see Westin (1968)). Notably, pioneering works of Alan Westin exercised considerable influence on debates about privacy-related threats in the US and other countries.

Among those countries was the United Kingdom (UK), where in 1972, a Committee on Privacy chaired by Kenneth Younger recommended a set of principles that should apply to the handling of personal information by computers (Gellman, 2019). While consent was not explicitly required for data collection, the first one of those principles stated:

> '*Information should be regarded as held for a specific purpose and not to be used, without appropriate authorization, for other purposes*' (Gellman, 2019).

Around the same time, in 1973, the US Department of Health, Education and Welfare's Advisory Committee on Automated Personal Data Systems issued a similar report (US Department of Health, Education, and Welfare, 1973) where they also proposed a set of principles for protecting the privacy of personal data in record-keeping systems. Similarly, the original formulation of the so-called Fair Information Practices (FIPs) included the following principle:

> '*There must be a way for an individual to prevent information about him obtained for one purpose from being used or made available for other purposes without his consent*' (Gellman, 2019).

Even though consent was present in both reports as a purpose limitation safeguard, it is impossible to judge how one committee may have influenced the other (Bennett, 1992). Subsequently, there were analogous reports published and other related activities reported from several other countries, including Sweden, Germany, France, Canada and Australia (Bygrave, 2004; Gellman, 2019). Soon after that, the public debates on privacy threats posed by complex record-keeping systems prompted global movements towards data protection legislation.

2.1.2 Consent as a Legal Requirement

In 1948, the United Nations adopted the Universal Declaration of Human Rights (United Nations General Assembly, 1948), including Article 12, i.e. the Right to Privacy. The Article stated:

> '*No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks*' (United Nations General Assembly, 1948).

The Declaration laid the foundation for the human rights protections. It is widely recognized as having inspired, and paved the way for, the adoption of other human rights treaties at global and regional levels.

From the beginning, consent has been an important mechanism in data protection legislation. As early as in 1970, the first data protection act in the world was passed in the State of Hesse of the Federal Republic of Germany. In accordance with the Hesse Data Protection Act of 1970, those charged with the handling of data were only allowed to share such data with others when this was authorized by law or when there was consent of 'those entitled to exercise control' over the data (van Alsenoy, 2019). However, the decision-making power over the disclosure of the data did not lie with the individual the data was related to, since as noted by van Alsenoy (2019), that the Act did not specify who exactly was authorized to grant such consent or under what conditions.

In that regard, the French Law no. 78–17 concerning Informatics, Files and Liberties (LIFL) of 1978 was the first data protection act that, unless exceptions applied, required consent of the individual concerned for the processing of sensitive personal information. Specifically, under LIFL, data revealing racial origin, religious, philosophical or political opinions, or union membership could only be processed with the express consent of the individual (van Alsenoy, 2019). Notwithstanding, exceptions to this rule were provided for religious, philosophical and political organisations, as well as for data processing performed in public interest (van Alsenoy, 2019).

As data protection laws started being enacted in European countries, international institutions engaged in regulating the international implications of data processing. Among them was the Organisation for Economic Cooperation and Development (OECD) which in 1980 proposed the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (Organisation for Economic Cooperation and Development, 1980). Under the OECD Guidelines, consent was safeguarding the use of personal data, unless required by the authority of law. To this end, the Use Limitation Principle in the Guidelines was similar to the requirement in LIFL:

> '*Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with [the Purpose Specification Principle] except: a) with the consent of the data subject; or b) by the authority of law*' (Organisation for Economic Cooperation and Development, 1980).

At the same time, when it came to data collection, consent of an individual was equivalent to ensuring their knowledge about data collection taking place. In more detail, the Collection Limitation Principle set out in the Guidelines stated:

> '*There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject. (...) The knowledge or consent of the data subject is as a rule essential, knowledge being the minimum requirement. On the other hand, consent cannot always be imposed, for practical reasons.*' (Organisation for Economic Cooperation and Development, 1980).

Nonetheless, countries such as Canada, where the OECD Guidelines have been especially influential (Organisation for Economic Cooperation and Development, 2013), did impose the consent requirement in addition to the knowledge requirement. In general, the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) readily maps to the basic principles of the OECD Guidelines. However, the PIPEDA covers consent as a separate Consent Principle with the requirement of obtaining consent for the collection of personal information and its subsequent use or disclosure. Additionally, it is highlighted that both knowledge and consent are required under the PIPEDA (Department of Justice, Canada, 2000).

As pressure grew in Europe for more uniformity in data protection legislation, countries of the European Union (EU) decided on a common standard of protection with regard to the processing of personal data. In 1995, the European Union (EU) adopted Directive 95/46/EC which in 2018 was replaced by the General Data Protection Regulation (GDPR). Notably, both under the Directive and under the GDPR consent was one of the six legal bases under which data processing could be considered lawful (European Parliament and the Council of the European Union, 1995; European Parliament and the Council of the European Union, 2016). Specifically, Article 4(11) of the GDPR defines consent as follows:

> *'Consent of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'* (European Parliament and the Council of the European Union, 2016).

Moreover, the provision of relevant information is required when asking people for consent. That is, processing of data is fair only if it is transparent and effectively communicated to users, including in the use of information notices. In particular, Article 5(1)(b) of the GDPR predicates the Purpose Limitation Principle, which mandates personal data to be 'collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes' (European Parliament and the Council of the European Union, 2016). To this end, GDPR Article 6 emphasises that data processing can only be lawful if the consent relates to one or more specific purposes. Further, Article 7 sets out additional conditions for valid consent, including keeping records to demonstrate consent, as well as prominence and clarity of consent requests.

Soon after the GDPR came into force, it became 'one of the most demanding and comprehensive privacy regulations of all time' (Linden et al., 2020) and a global reference point for data protection. Importantly, many other regulatory frameworks around the world started introducing similar requirements. For example, the Personal Information Security Specification, introduced in 2018 to accompany China's Cybersecurity Law, also requires telecommunication operators and Internet service providers to inform users of the purposes, methods and scope of the users' personal data processing (Liu, 2014). Although it appears to be less strict than the GDPR, explicit consent is similarly necessary for collection of personal sensitive information (Pernot-Leplay, 2020). In addition, the California Consumer Privacy Act (CCPA), which came into force in 2020, requires websites to collect the consent of minors and to allow users to opt-out of the sale of their personal data (Hils et al., 2020).

Currently, countries without data privacy laws are heading towards a minority (Greenleaf, 2014). Yet, the basic approach to protecting privacy has remained largely unchanged since the 1970s. Manifested in various legal frameworks around the world, data protection law provides people with a set of rights to enable them to make decisions about how to manage their data. These rights consist primarily of rights to notice, access and consent regarding data collection, use and disclosure. The goal of this bundle of rights is to provide people with control over their personal data. Through this control, people can decide for themselves how to weigh the costs and benefits of the collection, use or disclosure of their information (Solove, 2013). This approach to privacy was described as *privacy self-management* (Solove, 2013).

Interestingly, even some countries that have no single principal data protection legislation for the private sector likewise endorse the privacy self-management approach. This is the case in the United States[1], where the Federal Trade Commission and the White House

[1]There are, however, industry-specific regulations such as the 1996 Health Insurance Portability and Accountability Act which gives patients control over the use and disclosure of their medical data, and the 1999 Gramm-Leach-Bliley Financial Services Modernisation Act which requires financial institutions to disclose their privacy policies and allows consumers to opt out of data sharing with non-affiliated third parties.

issued a framework concerning consumer privacy in 2012, with the objective to 'make information collection and use practices transparent' and provide consumers with the 'ability to make decisions about their data at a relevant time and context' (Federal Trade Commission, 2012). The so-called *notice-and-consent* (or *notice-and-choice*) framework requires that users are notified and grant their permission before information about them is collected and used (Federal Trade Commission, 2012).

2.1.3 Criticism of the Consent Requirement

The implementations of the privacy self-management concepts, including the notice-and-consent framework, have come under sustained scholarly criticism (Cate, 2006; Nissenbaum, 2011; Pascalev, 2017; Radin, 2012; Schermer et al., 2014; Schwartz, 1999; Sloan and Warner, 2014; Solove, 2013; Susser, 2019; Waldman, 2020). One of the highlighted implementation issues is the fact that privacy notices are ill-equipped to inform users of corporate data use practices (Waldman, 2020). Designed to help users understand where and under what conditions their personal data may flow, and how to exercise control over that flow (Cranor, 2012), privacy notices are generally so long (Milne et al., 2006), abstruse and legalistic (Cate, 2006; Nissenbaum, 2011; Schermer et al., 2014) that even experts find them misleading (Reidenberg et al., 2015). Estimates conclude that it would take a user an average of 244 hours per year to read the privacy notice of every website they visit (Cranor, 2012). Because of this, the consent granted by a user often fails to meet the requirement of an informed choice.

Concerning is also the binary nature of the choice: in practice, users can either accept service providers' unlimited use of their personal data, or they can give up using the service (Nissenbaum, 2011; Schermer et al., 2014; Susser, 2019). Nowadays, even just using a website, an app, a wearable device, or a smart home appliance is often interpreted as consent to the data practices of the provider (Schaub et al., 2015). This failure to offer the user any real choices has been described as the *take-it-or-leave-it* approach to privacy (Bender, 2011; Polykalas, 2017). In a sense, such consent cannot be considered an indication of a user's wishes, as users almost always grant consent when it is required by the service provider (Cate, 2010).

Another issue relates to the *consent transaction overload*, i.e. a situation when there are too many consent requests for an individual to consider (Cate, 2006; Schermer et al., 2014). Consequently, the increasing difficulty on the user's side in managing their personal data leads to them feeling a loss of control and a sense of 'weariness' towards privacy issues, in which users believe that there is no effective means of managing their personal data (Choi et al., 2018). As a result, excessive consent requests lead to *consent fatigue* (Schermer et al., 2014), which reflects users' tendency to simply accept a privacy notice without reading it (Choi et al., 2018). Thus, although in theory privacy

self-management centered around consent may be regarded as a utopia, critics conclude that it 'does not scale well' (Solove, 2013).

Moreover, there is currently no established form of a consent request (Santos et al., 2020). That is, consent is often implemented as an to equivalent to *not opting-out* of data collection. In particular, continuing to use a service is interpreted as consent to the data practices of the service provider – this is termed as *implicit consent*. While this is legal under some regulatory frameworks, implicit consent cannot be classified as a statement or a clear affirmative action, as required by the GDPR. Despite this, research shows that even European websites, where the GDPR applies, tend to collect visitor's data without consent, register consent even when the user has explicitly opted out or nudge users towards granting consent by pre-selecting options (Matte et al., 2020; Sanchez-Rola et al., 2019). Actually, implied consent seems to be ubiquitous – research suggests that only 11.8% of UK websites meet the minimum requirements based on European law (Nouwens et al., 2020). In essence, Santos et al. (2020) highlight that not complying with the legal requirements for a valid consent renders the consent invalid.

On top of this, users generally lack sufficient knowledge to make a reasonable evaluation of the risks and benefits of their decision to consent (Sloan and Warner, 2014; Solove, 2013), and those who do have the knowledge often lack enough expertise to adequately assess the consequences of such decisions (Solove, 2013). In particular, users greatly struggle to factor in the potential harm caused by individual consent decisions in the future, when non-sensitive data is combined and analysed to reveal sensitive facts (Solove, 2013). One reason for this is the fact that privacy notices tend to be vague about future uses of data (Solove, 2013). Simultaneously, the consequences of giving up data are quite complex if explained in sufficient detail to be meaningful and may seem abstract at the time of decision-making (Solove, 2013). As such, we focus on the users' decision-making in the next section.

2.2 Consent Decision-Making

Understanding how users arrive at consent decisions is important for the potential design of consent mechanisms. Thus, in this section, we review the existing literature on consent decision-making and highlight the gap in the research approaches up to date. Firstly, we focus on privacy concerns, which influence users' intentions to grant consent to data sharing. Secondly, we discuss the relationship between those intentions and users' actual consenting behaviour. Finally, we look at the potential impact of privacy knowledge on the consenting behaviour and motivate our exploration of this topic.

2.2.1 Privacy Concerns

Central to the discussion on users' consent decision-making is the issue of privacy concern. In this context, privacy concerns refer to an individual's subjective views of fairness within the privacy context (Campbell, 1997). The reason why they are important for consent decision-making is that privacy concerns – both dispositional and situational – have been found to affect users' intention to share personal information (Joinson et al., 2010; Smith et al., 1996; Stewart and Segars, 2002).

Consistently, studies have concluded that the overwhelming majority of people are 'concerned' or 'very concerned' about threats to their online privacy while online (Paine et al., 2007). For example, Paine et al. (2007) mention an early study from 1998 which reported that 87% of Internet users are 'concerned' about threats to their privacy while online, with 56% of them being 'very concerned'. More recent public opinion polls show that this feeling remains unchanged: 79% of American adults assert that they are 'very concerned' or 'somewhat concerned' about how companies are using the data they collect about them (Auxier et al., 2019).

Unsurprisingly, privacy concerns are a widely studied topic. In particular, a privacy research pioneer, Alan Westin, conducted over 30 surveys examining the general level of privacy concern between 1978 and 2004, and classified the public into three categories: *fundamentalists*, *pragmatists* and *unconcerned* (Kumaraguru and Cranor, 2005). The three-question survey instrument he developed (Privacy Segmentation Index), has been hugely influential in the debate over privacy attitudes (Woodruff et al., 2014) and deployed in other research studies (e.g. Consolvo et al. (2005); Malheiros et al. (2013)). However, Jensen et al. (2005) observed that while people classified as fundamentalists have consistent privacy concerns, they do not appear to form a cohesive group with respect to decision-making. In fact, the investigation by Woodruff et al. (2014) indicates a lack of correlation between Westin's categories and behavioral intent, as well as a lack of correlation between Westin's categories and individuals' reactions to specific consequences of their decisions.

Subsequently, variations of Westin's segmentation model categorised users into additional groups along a continuum of privacy concern (Elueze and Quan-Haase, 2018; Sheehan, 2002), or according to their personal dispositions (Bansal et al., 2010) or attitudes towards privacy boundaries (Milne and Bahl, 2010). As the measurement of users' privacy concerns has evolved (Malhotra et al., 2004; Smith et al., 1996; Preibusch, 2013; Sheehan and Hoy, 2000), variation in privacy concerns has been attributed to differences in three key dimensions: chronic privacy attitudes, information sensitivity and context (Brough and Martin, 2020).

Within these dimensions, several factors have been identified that shape users' privacy concerns. They include personal experiences of Internet use (Fogel and Nehmad, 2009;

Jensen et al., 2005; Miyazaki and Fernandez, 2001; Paine et al., 2007; Park et al., 2012; Yao et al., 2007; Youn, 2009), socio-demographic factors (Fogel and Nehmad, 2009; Jensen et al., 2005; O'Neil, 2001; Yao et al., 2007; Youn, 2009), trust in institutions and in other people (Chellappa and Sin, 2005; Okazaki et al., 2009; Park et al., 2012), political orientation and ideology (Acquisti, 2004; Yao et al., 2007).

2.2.2 Privacy Paradox

While the majority of users express privacy concerns and consider privacy to be important (Brandimarte and Acquisti, 2012; Jensen et al., 2005), research has shown that users simultaneously tend to disclose their personal information most of the time (Adjerid et al., 2013). This dichotomy between people's information disclosure intentions and their actual disclosure practices, referred to as the *privacy paradox* (Norberg et al., 2007), has been confirmed in many studies (e.g. Acquisti and Grossklags (2005a); Spiekermann et al. (2001); Tsai et al. (2011); see Barth and De Jong (2017); Gerber et al. (2018); Kokolakis (2017) for detailed reviews).

Different theories have been proposed to explain the reasons behind this phenomenon. One of them considers the decision to disclose information to be caused by factors that make users unable or unwilling to consider the risks. Specifically, such factors include bounded rationality (Acquisti and Grossklags, 2005a), incomplete information (Acquisti and Grossklags, 2005a), psychological distortions (Acquisti, 2004), desire for immediate gratification (Acquisti, 2004), lack of knowledge of privacy-protective behaviours (Hargittai and Marwick, 2016) or lack of understanding of the privacy risk involved (Hargittai and Marwick, 2016).

Another theory frames the privacy paradox around the concept of *privacy calculus* (Culnan and Armstrong, 1999). According to this point of view, the decision to disclose information is a result of the user's rational assessment of some economic or social benefits gained in exchange for it, and subject to a belief that the information will be used fairly and that they will not experience negative consequences (Milne and Gordon, 1993). However, some researchers criticised this explanation, saying that the attitude-behaviour dichotomy is also affected by misperceptions of those benefits and costs, social norms, emotions, and heuristics (Acquisti et al., 2015).

Other explanations combine the previous two theories. For example, Zafeiropoulou et al. (2013) view the privacy paradox as part of a process of structuration. In that sense, users' attitudes and values are tempered by external structures such as the situations they are in and the context of disclosure. In other words, the consent decision is seen as a rational trade-off influenced by a set of structures that make users deviate from their beliefs. Similarly, Li et al. (2011) suggest that situation-specific reactions of users influence their decision-making and possibly override the effect of general privacy concerns on users'

behaviours. In their study, privacy concerns had a significant influence on participants' privacy risk belief, but did not significantly affect the formation of participants' privacy protection belief.

Although a significant volume of research has aimed to explain this the privacy paradox (Gerber et al., 2018), it remains a wide open issue (Kokolakis, 2017). Despite the fact that this dichotomy is a complex topic, conducting research to understand this phenomenon is also challenging. According to Norberg et al. (2007), there are three main challenges in investigating the privacy paradox. Firstly, privacy perceptions of users vary widely – not just within populations or even different segments of a population, but also depending on the personal data type considered. Secondly, previous researchers used different research methods to examine the topic, causing confusion regarding the implications that can be drawn from the existing literature. For example, the diverse measurements include attitudes toward privacy, concern for privacy, privacy-related behavioral intentions. Finally, research into users' actual data sharing behaviours has been far more limited than that measuring concerns, attitudes or intentions.

On this point, Smith et al. (2011) point out that in the previous research there is a common assumption that participants' actual behaviours will match their stated intentions, inferred through references to the Theory of Reasoned Action (Fishbein and Ajzen, 1975). While researchers tend to measure the stated intentions instead of the actual behaviours, it has been highlighted that to the extent the privacy paradox holds, such assumption might be misguided in privacy research (Norberg et al., 2007; Smith et al., 2011). Thus, Kokolakis (2017) suggests that future studies should use evidence of actual behaviour rather than self-reported behaviour.

2.2.3 Privacy Knowledge

While much attention has been focused on privacy concerns, researchers have been calling for an investigation into the impact of users' privacy knowledge on their disclosure behaviour (Brough and Martin, 2020; Park, 2013; Smith et al., 2011). Although privacy concerns affect users' motivation to protect their personal data, Brough and Martin (2020) argue that even among those who are highly motivated the privacy behaviour may vary greatly depending on their privacy knowledge. Since none of the factors that constitute users' privacy concerns – namely, chronic privacy attitudes, information sensitivity and context – explicitly accounts for differences in users' privacy knowledge, this aspect may be the key to understanding the reasons behind the privacy paradox (Brough and Martin, 2020).

In that context, privacy knowledge is strongly associated with the digital literacy in the control of personal information online (Park, 2013). As highlighted by Park (2013), privacy knowledge refers to users' "understanding of data flow and its implicit rules

for users to be able to act". Furthermore, Brough and Martin (2020) identified three dimensions of privacy knowledge:

- *factual knowledge* which refers to the awareness of any privacy-related risks, users' rights, corporate practices and law;
- *procedural knowledge* which refers to the understanding of how to use privacy-enhancing strategies, tools, and skills to protect personal information;
- *experiential knowledge* which refers to the general familiarity with online technology and any first-hand experience with privacy violations.

To that end, privacy knowledge also includes the extent to which users realise when they share any personal data (Brough and Martin, 2020).

In contrast with privacy concerns, several studies have concluded that users generally have a relatively low level of privacy knowledge (Bartsch and Dienlin, 2016; Dommeyer and Gross, 2003; Nowak and Phelps, 1992; Park, 2013; Turow, 2003; Turow et al., 2005). This is particularly worrying in the light of evidence that they also tend to approve consent requests without fully understanding their meaning and consequences (Felt et al., 2011, 2012; Kelley et al., 2012).

In fact, there is some evidence suggesting that the lack of privacy knowledge correlates with the lack of efforts to protect privacy. For example, Turow (2003) observed that American adults who use the Internet at home seem not to understand the flow of their data online and tend not to take steps to learn about ways to control their information online. Despite strong concerns about privacy online, 64% of the study participants reported that they have never searched for information about how to protect their data online; 40% admitted that they know 'almost nothing' about stopping sites from collecting information about them and 26% said that they know just 'a little'.

Additionally, misunderstandings and ignorance of the common data-collection practices are also very common Acquisti and Gross (2006); Turow (2003); Turow et al. (2005), and users tend to have inaccurate perceptions of their own privacy knowledge (Jensen et al., 2005). In another study, Acquisti and Gross (2006) observed that most of their participants were unaware of Facebook's data-collection practices. Specifically, 67% of their participants mistakenly believed that Facebook was not collecting information about them from other sources; 70% believed that Facebook was not combining information about them collected from other sources; 56% believed that Facebook was not sharing personal information with third parties. At the same time, they were happy to reveal their personally identifiable information, regardless of their level of privacy concerns. Even among those who expressed the highest concern, more than 48% provided at least their sexual orientation and almost 47% provided at least their political orientation.

Surprisingly, few studies have examined the impact of users' privacy knowledge on their privacy protection behavior (Brough and Martin, 2020). Not only is it an important step towards unpacking the privacy paradox, but it may also provide evidence to focus public efforts on the crucial initiatives that help users manage their personal data in ways that align with their privacy preferences. In particular, a suggestion to bring research attention to the need for such educational programs was made by Norberg et al. (2007) who first explored and coined the term 'privacy paradox'. They highlighted this by saying:

> '*There must be the realization that, unless consumers make the effort to truly understand what they are granting permission to, and to whom they are giving their personal information, their sense of personal privacy will continue to deteriorate. Especially, as people expand their usage of data-rich transaction channels such as the Internet, the need to comprehend where the data go increases dramatically*' Norberg et al. (2007).

In this thesis, we focus on this open topic in Chapter 3. Specifically, we investigate whether expanding users' privacy knowledge, particularly the factual and procedural knowledge, influences their opt-out consent decisions.

2.3 Automated Consent

In this section, we explore ways of automating user's consent decision-making. Specifically, we focus on aligning the automated consent with user's intentions which can be based on the user's privacy preferences. As such, we first look at the previous research on privacy-self management that support users in handling the growing number of consent requests. Second, we discuss the gap in the literature on privacy self-management tools that allow users to negotiate their consent to data use in return for a service, and motivate our research on this topic. Finally, we review the related work on how users' privacy preferences can be derived to reflect their personal values, and perceptions or risks and benefits.

2.3.1 Privacy Self-Management Tools

With the number of data-sharing transactions growing vastly in the recent years, there has been an emergence of initiatives and tools whose aim is to aid users in handling the consent transaction overload and controlling their personal data. Early efforts[2] include the Platform for Privacy Preferences (P3P) project developed by the World Wide Web

[2]For details on the history of early privacy self-management tools, see: Hochheiser (2002).

Consortium (W3C), which aimed to enable machine-readable privacy policies (Cranor, 2002). Such privacy policies could be automatically retrieved by Web browsers and other tools that can display symbols, prompt users, or take other appropriate actions. Some of the so-called *user agents* were also able to compare each policy against the user's privacy preferences and assist the user in deciding when to exchange data with websites (Cranor, 2002).

However, P3P has been widely criticised (Reay et al., 2007). One of the criticisms was that P3P did not provide any mechanisms for imposing limits on the collection and use of users' personal data, as proposed in privacy models such as the OECD Guidelines (Hochheiser, 2002). Moreover, P3P was limited to privacy of personal data collected from Web browsing, leaving e-mail and other online activities is beyond its scope (Hochheiser, 2002). In addition, there was no enforcement of the P3P mechanisms and, therefore, no way to enforce adherence to stated privacy policies (Hochheiser, 2002). Last but not least, P3P has also suffered from semantic inconsistencies (Li et al., 2006), and from an unwillingness or inability of users to make privacy-preserving decisions (Acquisti and Grossklags, 2005a; Spiekermann et al., 2001).

Later, several tools have been created to support users in privacy self-management. Among the academic efforts, the main approaches focused on raising users' awareness of data collection, included providing privacy nudges (Acquisti, 2009; Balebako et al., 2011; Liu et al., 2016; Zhang and Xu, 2016), visualizing privacy information (Gates et al., 2014; Harbach et al., 2014; Kelley et al., 2009, 2013; Van Kleek et al., 2017) and detecting possible data leaks (Balebako et al., 2013; Egele et al., 2011; Enck et al., 2014). While such approaches do help users control their personal information, they do not address the fundamental criticism regarding the binary nature of the choice that users are confronted with.

Whereas, the key challenge for today's privacy self-management tools is moving away from the prevailing take-it-or-leave-it approach and instead, providing fine-grained solutions for users to communicate their privacy preferences to the service provider. For instance, a study by Preibusch et al. (2013) has found that users distinct clearly between mandatory and optional data, and they selectively decide what information they want to disclose, if possible. Conversely, where opting out of disclosure is impossible, researchers proposed solutions that allow users to explore the trade-off between functionality and privacy by instead generating fake or 'mock' personal data (Beresford et al., 2011; Zhou et al., 2011). In fact, results of a 2015 survey by Symantec in Europe show that one in three respondents admitted to falsifying personal data online in order to protect their privacy (Thomson et al., 2015). Such behaviour may have consequences for the quality of users' data the service providers obtain (Krol and Preibusch, 2015) and, thus, the quality of the services they can offer.

On the other hand, giving a user the opportunity to choose whether they want to provide certain information about themselves may be beneficial to service providers. Specifically, Preibusch et al. (2013) provide evidence that optional disclosure delivers a 'good data return' for service providers. In contrast, they also find that increasing the amount of required data jeopardises voluntary disclosure for the remaining information. This result suggests that as the society becomes more privacy-knowledgeable, fine-grained consent options may be preferable by users and beneficial to service providers.

While nowadays some operating systems such as iOS and Android[3] already allow service providers to explicitly ask for consent via a dialog box when their app requires access to the user's protected resources, it is important to highlight that the overwhelming number of consent requests is not solely a problem of mobile app permissions. In fact, a large number of other service providers have started requesting access to personal data through Web browsers, desktop apps, Smart TVs, other the Internet of Things (IoT) devices and even authentication mechanisms such as Facebook Login.

For example, the uncertainty around sanctions for non-compliance led many websites to embed a Consent Management Provider (CMP), which defines legal terms and conditions, presents these to users via embedded consent dialogues, stores the resulting consent signals and shares them with third-parties (Hils et al., 2020). However, Hils et al. (2020) have found that the consent dialogues offered by CMPs impose a significant time cost on privacy-aware users who attempt to opt out of data collection. As the number of services asking for permissions grows, there is an increasing risk that more and more users may start experiencing the consent transaction overload and, as a result, stop being able to manually control their personal data flow at all. Thus, inevitably, there is a need for automation in the area of privacy self-management – not just for websites and mobile apps, but on all data collection platforms.

2.3.2 Privacy Negotiations

While some of the data that service providers collect is necessary in order to be able to provide the user with the expected functionality, (e.g. a food delivery platform may require the user's home address, so that the order can be delivered to that address), in 2011, Felt et al. (2011) showed that a significant number of services were over-privileged: they were requesting access to sensitive data that were not necessary for the their core functionality. As demonstrated by Sophus Lai and Flensburg (2020), even after options for users to manually grant access to certain resources were introduced in mobile operating systems, this situation remains unchanged. Relying on users' unawareness, such practices do not result in fair agreements between the user and service provider. For

[3]A new dynamic permissions model was first introduced in Apple's iOS 6 in 2012 and in Google's Android 6.0 Marshmallow in 2015.

instance, user cannot choose to pay for a free app instead of receiving targeted advertising, nor can they, as suggested by Preibusch et al. (2013), opt to share extra information in return for an incentive (e.g. a discount).

In fact, there is evidence suggesting that, for various reasons, many users prefer to exercise fine-grained control and selectively choose the personal information they are happy to disclose rather than refuse to disclose any information at all. Specifically, in a study by Krol and Preibusch (2016), participants were able to choose the option of erasing their personal data from all optional form fields, which would save them a lot of time. Regardless, the majority of the participants preferred to go from question to question and decide whether to share each data item, even though it required more effort on their side. While the reasons behind these choices varied (e.g. some participants simply enjoyed disclosing the information), some participants did so hoping for some benefit in return for the optional information.

Therefore, the fact that more data is being collected than necessary opens an opportunity for *negotiating* consent to collection and use of non-essential personal data for an incentive. As such, Krol and Preibusch (2015) envision consent negotiations working as follows:

> '*Service providers can offer incentives such as discounts or work with smart defaults to guide their customers. In return, users navigate effortlessly through a series of simple choices, often either providing or withholding information at will or deciding on permissible uses for a data item. Customers are free to decide whether they want the free version of an app that shows advertisements or to pay for an ad-free experience, sustained through the monetization of their personal information*' (Krol and Preibusch, 2015).

In the article, they pose the question of how such 'effortless privacy decisions' could be achieved, highlighting that making such negotiations effortless is challenging and 'might not be achievable' at all (Krol and Preibusch, 2015). Not only may fine-grained options result in even more consent fatigue, but they also cite Korff and Böhme (2014) who have shown that more privacy options to choose from results in users experiencing more negative emotions, more regret and less satisfaction with their consent decisions. In this thesis, Chapter 4 offers a response to this question. Specifically, we argue that such negotiation can be automated to *reduce* the user's effort.

Indeed, a number of previous studies have explored the idea of automated negotiations in the online privacy context. Particularly in the context of horizontal privacy, automated negotiations were used in resolving privacy conflicts among users in social networks (Kekulluoglu et al., 2018; Kökciyan et al., 2017; Kökciyan and Yolum, 2016). For example, a negotiation protocol was proposed as a conflict resolution method to find

adequate compromises among multiple users (Such and Criado, 2016). Similarly, a negotiation framework and protocol was developed to help users manage their agreements with others (Mester et al., 2015).

In the context of vertical privacy, several studies based their approach on P3P. Although P3P itself lacks a negotiation mechanism, researchers utilised P3P enhancements to enable privacy policy negotiations between the user and the service provider (Bennicke et al., 2003; Cheng et al., 2007; Maaser and Langendoerfer, 2005; Maaser et al., 2006; Kalyani and Adams, 2006; Preibusch, 2006). For example Cheng et al. (2007) proposed a model for automatic privacy policy conflict detection and resolution; Preibusch (2006) extended P3P with a negotiation process modelled as a Bayesian game where the service provider faces different user privacy types.

Closest to our work in Chapter 4, Yassine and Shirmohammadi (2009) proposed an intelligent agent-based system to quantify and measure privacy payoff through private data valuation and privacy risks quantification. Their system includes five different agents, including one responsible for negotiation with the service provider. However, the agent works on behalf of a list of users whose privacy preferences need to be entered by the users themselves prior to the negotiation. Therefore, such approach is not suitable for our requirement of utilising automated negotiation to reduce the users' effort.

Instead, we extend the work of Baarslag et al. (2017) who proposed an automated negotiation agent to represent the user in negotiations with the service provider. The agent classifies the user into a category as specified by Westin's Privacy Segmentation Index and derives the user's privacy preferences based on historic preferences of other users in the same group. This allows the agent to successfully reduce the burden on the user's side. To that end, we develop a generalised theoretical framework for negotiations between users and service providers. Additionally, since Westin's categories do not correlate with users' actual intentions (Woodruff et al., 2014), we compare the agent of Baarslag et al. (2017) to a new variant which derives the user's privacy preferences regardless of their category.

2.3.3 Privacy Preference Elicitation

In order for an automated agent to represent the user accurately in a negotiation, the agent needs to learn the user's preferences. Early approaches to agent-based automation of consent were based on the assumption that users would communicate their privacy preferences manually. For example, P3P came with a standard language allowing users to express their privacy preferences to the agent, called *A P3P Preference Exchange Language* (APPEL) (Cranor, 2002). Due to several shortcomings of APPEL, researchers at IBM later developed a new preference language for P3P called XPref (Agrawal et al., 2003b, 2005).

However, recent research has proposed ways of deriving the user's preferences automatically. The so-called *preference elicitation* is defined as 'the process of extracting and determining preferences of a human user over a set of possible outcomes in a decision problem' (Qin et al., 2008). Such process may be performed though direct querying, observing activity or choices of the user, and inferring or predicting the preferences based on any evidence obtained (Qin et al., 2008).

In particular, preference elicitation is important for privacy-enhancing technologies, because users themselves are often unable to articulate their privacy preferences (Sadeh et al., 2009). Moreover, those who are able, do not act in accordance with their stated privacy preferences (Berendt et al., 2005) and even the experienced ones tend to set them incorrectly (Madejski et al., 2011). To that end, Berendt et al. (2005) proposed a requirement that privacy-enhancing technologies should learn users' preferences by observation and change settings dynamically, on a per-service level.

In accordance with that requirement, a number of previous studies have looked at learning users' privacy preferences. For example, a study by Sadeh et al. (2009) has shown that machine learning techniques have the potential not only to reduce the users' effort of specifying privacy preferences, but also to learn privacy preferences more accurately than any preference rules defined by the users themselves. However, they highlight that such techniques are usually 'configured as black boxes', significantly restricting the ability of users to understand the preference rules that have been learned, let alone modify them. They suggest that it is more effective to deploy machine learning-based privacy preferences in the form of suggestions that users can either accept or reject.

To that end, Kelley et al. (2008) has proposed a *user-controllable policy learning approach* based on incremental improvements that allows the user and the system to work hand in hand on refining common privacy preference model. Such approach is consistent with a requirement specified by Berendt et al. (2005) that privacy-enhancing technologies should always be under the full control of the user. In our work on privacy negotiations, we adhere to this requirement: the user's agent learns the user's preferences in order to negotiate the most suitable agreement, but it is the user who decides whether or not to grant consent.

Soon after that, a number of studies provided evidence that with a limited amount of user input, it is possible to build an accurate machine learning model that concisely describes a user's privacy preferences and use this model to configure the privacy settings automatically. While the work of Kelley et al. (2008) relied on simulations, Cranshaw et al. (2011) have shown that a combination of machine learning and user-controllable incremental improvements can quickly converge to be just as accurate as standard non-incremental learning, but the user-controllable models require far fewer changes. Similarly, Fang and LeFevre (2010) and Fang et al. (2010) have presented a 'wizard' for privacy settings on Facebook based on uncertainty sampling, i.e. the wizard initially

asks a user to specify settings for selected friends and then uses this input to construct a classifier, which can in turn be used to automatically assign privileges to the rest of the user's friends.

Furthermore, another effective approach to privacy preference elicitation involves using machine learning techniques to identify clusters of similar users. To that end, several studies explored different methods for generating small numbers of user-understandable privacy preference profiles based on clustering users with similar preferences (Agarwal and Hall, 2013; Lin et al., 2012, 2014; Liu et al., 2014; Nakamura et al., 2016; Mugan et al., 2011; Ravichandran et al., 2009). In general, such profiles have been found to influence users decisions to share significantly more data without a substantial difference in comfort (Wilson et al., 2013).

However, Wilson et al. (2013) also observed that what brought their participants to satisfaction were subsequent edits the participants made to their privacy preference settings. This observation may suggest that although profiling may be useful for proposing to users some options for possible default settings, relying solely on privacy preference profiles fails to take any differences between individual users into account. Therefore, in Chapter 4, we compare two approaches to privacy preference elicitation: the approach of Baarslag et al. (2017) where the outcomes are personalised solely according to the user's privacy profile, and one where the privacy preference predictions are relying on the user's privacy profile only initially and later based on the behaviour of the individual user. Then, we illustrate how an agent negotiating on the user's behalf can utilise these preferences to reach a agreement with the service provider.

2.4 Consent Infrastructure

After the user and service provider reach an agreement and consent is obtained, the service provider must register that agreement and honour it in all aspects of data processing. Therefore, in this section, we focus on honouring the privacy agreements service providers make with users. Firstly, we explore the ways of representing such agreements in the service provider's data processing infrastructure. Secondly, take a detailed look at how the agreements are currently enforced. Finally, we investigate how the state-of-the-art data processing systems comply with the consent constraints placed by the users and motivate our work on honouring consent in algorithmic data processing.

2.4.1 Privacy Agreement Representations

Once user's consent is obtained by the service provider, whether as a result of any agent-based negotiation or options manually selected by the user, the consent must be modelled, stored and honoured in all aspects of data processing. In other words,

```
<POLICY>
  ... ...
  <STATEMENT>
    <PURPOSE>
      <contact required="opt-in"/>
    </PURPOSE>
    <RECIPIENT><ours/></RECIPIENT>
    <RETENTION><business-practices></RETENTION>
    <DATA-GROUP>
      <DATA ref="#user.home-info.online.email"/>
      <DATA ref="dynamic.miscdata">
        <CATEGORIES><purchase/></CATEGORIES>
      </DATA>
    </DATA-GROUP>
  </STATEMENT>
</POLICY>
```

FIGURE 2.1: An example of a P3P privacy policy.

the *privacy agreement* (Oberholzer and Olivier, 2006), made by the user and the service provider, that the consent is associated with must be recorded before the service provider can use the data.

To this end, several ways of representing privacy agreements have been proposed. In particular, in P3P described in Section 2.3.1, consent is represented by a set of preference rules regarding the service provider's privacy policy. Specifically, P3P provides service providers with a way of encoding their data-collection and data-use practices in a machine-readable XML format, creating what is known as the P3P policy (Cranor, 2002). Such policy includes statements which list the purposes for which information is collected, the intended recipients of the information, the retention period for which the information is going to be kept and a list of individual data items that are collected for the purposes stated in the statement. An example of such a statement adapted from a policy presented by Agrawal et al. (2005) is presented in Figure 2.1. Then, languages such as APPEL allow users to communicate their privacy preferences to service providers, which are expressed as a list of rules (Cranor, 2002) such as the ones from Agrawal et al. (2005) in Figure 2.2.

More recently, the Consent Receipt standard has been developed by the Kantara Initiative (Lizar and Turner, 2018), which provides a way of representing the consent granted by the user in JSON[4]. As opposed to rules in P3P, the Consent Receipt is meant to serve the user as *a posteriori* documentation of the consent they granted. For this reason, in addition to a description of the data collected, the purposes for that collection and details on how that information will be used or disclosed, the standard includes requirements for links to privacy notices and policies. Although such solutions can help users understand how their personal data is used by the service provider (i.e. provide *transparency*) through the provision of a human-readable receipt, enterprise-level representations of

[4]JavaScript Object Notation, `https://www.json.org/`.

```
<appel:RULESET>
  <appel:RULE behavior="block">
    <POLICY>
      <STATEMENT>
        <PURPOSE appel:connective="or">
          <contact/>
          <telemarketing/>
        </PURPOSE>
      </STATEMENT>
    </POLICY>
  </appel:RULE>

  <appel:RULE behavior="request"/>
    <appel:OTHERWISE/>
  </appel:RULE>
</appel:RULESET>
```

FIGURE 2.2: An example of an APPEL rule set.

the privacy agreements are needed to support compliance with these agreements on the service provider's end.

Thus, a number of languages have been developed to support representation of privacy agreements on the enterprise level (Kumaraguru et al., 2007). One example of this is the Enterprise Privacy Authorization Language (EPAL) developed by researchers at IBM, which defines lists of hierarchies of data categories, data users and purposes for data use, as well as sets of any privacy actions, obligations and conditions that describe the authorisation rules for the data use (Ashley et al., 2003). Other examples include privacy-focused extensions to the eXtensible Access Control Markup Language (XACML) (OASIS, 2013), which on its own does not allow specifying attributes that correspond to privacy preferences (Kolter et al., 2007; Kounga et al., 2010). While the privacy preferences of the user are communicated as rules that put certain constraints on the privacy policy of the service provider, these languages are intended for modelling authorisation rights. With regard to that, this approach is similar to the approach we propose in Chapter 5, where constraints expressed by the user are translated into the authorisation rights.

Furthermore, ontologies have been commonly used to model privacy agreements (Fatema et al., 2017; Jang et al., 2008; Kirrane et al., 2018; Kökciyan et al., 2017; Squicciarini et al., 2006; Pandit et al., 2019a). Based on open and interoperable standards such as the Resource Description Framework (RDF)[5] for information representation and the Web Ontology Language (OWL)[6] for representation of modeling, ontologies are by design extendable, which makes them suitable for application across use cases. For example, Kirrane et al. (2018) utilised RDF vocabularies to define data usage policies as well as any data processing and sharing events. Similarly, Pandit et al. (2019a) developed

[5]Resource Description Framework, `https://www.w3.org/RDF/`.
[6]Web Ontology Language, `https://www.w3.org/OWL/`.

an OWL ontology to express information associated with consent and its associated information such as provenance, specifically to address the requirements of the GDPR.

Then, such ontology-based representations of consent can be used to verify that data processing or sharing on the service provider's end complies with the relevant privacy agreements. For instance, Kirrane et al. (2018) used OWL reasoning to verify that the service provider's data processing and sharing events comply with the corresponding usage policies specified by users. Likewise, Pandit et al. (2019b) prototyped a test-driven solution, which generates and checks adherence to the GDPR requirements and persists the results towards compliance documentation. While these systems are useful for demonstrating the ongoing compliance, they do not prevent user data from being used without consent.

2.4.2 Enforcement of Privacy Agreements

At the heart of consent mechanisms is the enforcement of privacy agreements. That is, consent mechanisms should ensure that a user has consented to data processing activities such as collection, use, sharing and disclosure before these activities are performed. However, some large service providers may not even know what types of personal data they are collecting, where it is stored, what kind of consent the user has granted nor what the legal regulations are that apply to personal data of the specific user (Karjoth et al., 2002). Furthermore, other enterprises that process or store data collected by the service provider are also unable to enforce these privacy agreements (Karjoth et al., 2002). To that end, Ashley et al. (2002b) wrote:

> *The missing piece is enterprise privacy management technology. This technology must be the focal point for defining and enforcing an enterprise wide privacy policy. It must enable monitoring, enforcement and auditing of the the policy across the whole IT infrastructure of the organization. It must also allow for management and enforcement of individual privacy preferences* (Ashley et al., 2002b).

Thus, to complement P3P on the enterprise level, Karjoth et al. (2002) developed the Platform for Enterprise Privacy Practices (E-P3P) with its own policy language and (Ashley et al., 2002a) extended their work with an authentication mechanism for E-P3P. Importantly, E-P3P was designed to manage consent on a per-person and a per-record basis, even if the data was disclosed to another enterprise (Karjoth et al., 2002). However, the enforcement of those privacy practices still relied on the privacy officer and the security administrator.

In order to develop consent mechanisms that can support automatic compliance with the privacy agreements, researchers proposed to enforce those agreements directly where

personal data is being collected: on the database level (Agrawal et al., 2002). At the time, research on privacy-preserving database systems was largely focused on problems such as providing statistical information about individuals without compromising sensitive information. Within that area, the techniques are broadly classified into query restriction and data perturbation (Adam and Worthmann, 1989). Specific examples of concepts that later developed in that area include k-anonimity (Sweeney, 2002), l-diversity (Machanavajjhala et al., 2007), t-closeness (Li et al., 2007) and differential privacy (Dwork, 2008). Although these methods successfully prevent disclosure of personal data by masking or altering it, they do not aim to take into account an individual's consent.

Another well-researched area in privacy-preserving databases was access control, concerned with preventing unauthorised access. There, the work could largely be grouped into discretionary and mandatory (Ramakrishnan and Gehrke, 2000). While the mandatory access control model involved a single set of rules governing access to the entire database system, discretionary access control allowed an administrator to grant and revoke access privileges. In particular, specific refinement of discretionary access control was role-based access control which allowed this type of privileges to be granted not just to an individual user, but to a user group or role (Sandhu et al., 1996). However, while access control models and consent mechanisms share the goal of preventing disclosure of private information, access control models do so by limiting access to data, as opposed to enforcing compliance with privacy agreements.

Therefore, Agrawal et al. (2002) proposed the vision of Hippocratic databases, i.e. 'database systems that take responsibility for the privacy of data they manage'. Inspired by the ethics tenet of the Hippocratic Oath, they defined ten principles of such systems, rooted in the privacy regulation and guidelines. Among them, there is the Purpose Specification principle, which requires personal information stored in the database to be associated with the purposes it was collected for. With regard to those purposes, the Hippocratic databases' Consent principle states:

> *The purposes associated with personal information shall have consent of the donor of the personal information* (Agrawal et al., 2002).

Other principles of the Hippocratic databases refer to the limitation of data collection, use, disclosure and retention, as well as the accuracy and safety of the data, the right of access (openness), and the ability to demonstrate compliance with these principles.

Moreover, the authors suggested that Hippocratic databases could add the enforcement dimension to the P3P initiative (Agrawal et al., 2002). Consequently, Agrawal et al. (2003a) implemented a system for validating the P3P privacy policies against users' privacy preferences within the database systems, which relies on a mapping of the P3P

policy schema into a relational database schema. Furthermore, LeFevre et al. (2004) developed an architecture for incorporating privacy policy enforcement into existing applications and database environments.

At the same time, another line of research suggested that existing access control mechanisms could be used as a starting point for enforcing privacy agreements (Powers et al., 2002). To achieve that, existing access control mechanisms have been augmented with the notion of purpose (Byun et al., 2005; Byun and Li, 2008; Massacci et al., 2006; Ni et al., 2010; Petković et al., 2011). As part of the so-called *purpose control*, the intended purposes of data processing are typically treated as a label attached to the data; upon receiving an access request, the purpose given as part of the request is verified against the purposes attached to the data (Byun et al., 2005; Massacci et al., 2006). While the advantage of this approach is that it unifies privacy policy enforcement efforts and access control policy enforcement into one access control model, it is based on the assumption that the data is manually accessed by individuals, which is not the case in our work presented in Chapter 5.

2.4.3 Consent Propagation

In order to hold service providers accountable for processing personal data for the intended purpose after consent has been granted, any flows and usages of the data need to be identified. Specifically, Karjoth et al. (2002) proposed that the following pre-requisites have to be established in order to use the E-P3P system:

- a *business-process model* for the collection and use of user data, which specifies the other enterprises that use collected data, the data they use, as well as how and what purposes they use the data for,
- *informal privacy policies* that govern the use of personal data in the business processes.

Then, when the E-P3P authorization engine receives an access request, it outputs a decision whether the access to the data is allowed or denied. In more detail, a request consists (or is decomposed into a set of requests) of a single data user, a single data category, a single purpose, a single action and any context data as defined in the policy, and is processed based on the set of authorization rules specified in the agreed privacy policy (Ashley et al., 2002a).

However, all of the approaches to consent enforcement described in Section 2.4.2 are based on scenarios where data processing is a relatively limited. For example, Agrawal et al. (2002) considered a scenario where an online bookseller collects information about the customer such as their name, shipping address, credit card number, e-mail address

and purchase history for the purposes of providing the service (book sales and delivery), as well as offering book recommendations and publishing information about books popular in a certain region. Then, Massacci et al. (2006) extended that scenario with additional purposes such as credit assessment and notification, where some of the data processing is performed by other parties than the bookseller.

In contrast to the early computing systems, where personal data was collected and passed on to a small number of service providers for few specific purposes, modern computing has created a proliferation of large volumes of user data across thousands of companies. That is, in such data processing systems, user data enters the system where it is automatically processed, possibly by several entities which create new information that is used to finally fulfill a certain purpose. This new information is the output of advanced analytical algorithms performing predictions and inferences about individuals and, thus, it constitutes additional personal data relating to those individuals (Wachter and Mittelstadt, 2019).

Therefore, it is important to enforce compliance with the privacy agreements in large-scale data processing systems. To that end, some solutions have been proposed to the problem of honouring user's consent. They include methods to ensure that datasets used by the data-processing algorithms are policy-compliant (Debruyne et al., 2019, 2020), as well as a privacy-enabling user modeling framework for deriving inferences, which takes into account the consent constraints of individual users regarding their personal data and data processing techniques applied on it e.g. clustering techniques, rule-based reasoning, incremental machine learning (Wang and Kobsa, 2007). In Chapter 5, we contribute to this goal by proposing a novel method for service providers to establish how consent constraints users to place on the business-process model propagate within the business model so that any data processing performed complies with the relevant privacy agreement.

2.5 Summary

In this chapter, we review the related literature on consent mechanisms. In Section 2.1, we begin with an investigation of the origins of the consent requirement. In particular, we highlight that there was a long series of privacy concerns in the society and regulatory interventions that led to consent being a legal basis for data processing in the most influential data protection regimes today. Although consent is not the only legal basis for data processing, it is one that is often used in the kind of situations we look at in this thesis.

Then, we focus on three main aspects of consent: the user's decision-making, the agreement between the user and the service provider and the enforcement of the agreement on the service provider's side. To that end, in Section 2.2, we review the existing literature

on users' privacy concerns, knowledge and disclosure behaviour. More specifically, we discuss the dichotomy between users' intentions and behaviour, as well as any factors that may influence that behaviour. Furthermore, in Section 2.3, we explore the ways of automating user's consent decisions. In particular, we report on the current privacy self-management tools and opportunities that automation opens. This brings us to a discussion on how the user's preferences can be learnt, so that automated consent aligns with them. Last but not least, in Section 2.4, we look at the existing consent infrastructure that can support compliance with privacy agreements. Specifically, we review how privacy agreements can be represented and enforced.

Along the way, we show that this corpus of literature is fragmented and spans across different areas of computer science. Firstly, we find that while much attention has been focused on privacy concerns, there is some evidence that user's privacy knowledge may have a significant impact on the disclosure behaviour. Secondly, we learn about the potential for negotiating consent, which, although may be beneficial to both users and service providers, is currently not practised and not well-researched. Finally, we observe that the existing infrastructure-level solutions do not allow service providers to translate the consent constraints into data processing policies.

The contributions we present in this thesis address the gaps in the literature and provide a structured study of consent mechanisms in privacy engineering. More specifically, in Chapter 3, we study the impact of privacy knowledge on disclosure behaviour, which allows us to learn about users' consent decision-making. Then, in Chapter 4, we explore how consent could be negotiated and how different methods of learning user's privacy preferences affect the negotiation. After that, in Chapter 5, we propose a novel method that determines what kind of data processing can be performed on the user's data, given their consent constraints.

Chapter 3

Informed Consent

While much attention has been focused on privacy concerns, some researchers argue that there may be another factor in play that affects users' motivation to protect personal data: the lack of privacy knowledge (Brough and Martin, 2020). In fact, none of the factors that constitute users' privacy concerns explicitly accounts for differences in their privacy knowledge. As we show in Chapter 2, little studies have examined the impact of users' privacy knowledge on their decisions to consent. Undoubtedly, it is hard to call consent *informed* if it is granted without relevant privacy knowledge.

In this chapter, we take a first step towards addressing this gap. Specifically, we explore the impact of factual and procedural privacy knowledge on implicit consent in a specific context of online tracking. To that end, we conduct a between-subject experiment with 50 participants who previously were not taking steps to adopt protection measures against online tracking. The participants were divided into two groups: one that was provided with procedural privacy information and one that was presented both procedural and factual privacy information personalised specifically to their own tracking profile. As part of the study, we observe the behaviour of the participants after they are informed about tracking and protection against it.

The work presented in this chapter suggests that procedural privacy knowledge has impact on privacy behaviour. Specifically, our results show that the provision of procedural information about anti-tracking protection motivated 22% of the participants (11 out of 50) to refuse implicit consent by adopting anti-tracking protection. At the same time, we observe no significant effect on the behaviour of the provision of factual information about the extent of tracking. These findings not only confirm the general trends in the related work on privacy knowledge and disclosure behaviour, but they also suggest that making consent *informed* requires the provision of relevant actionable information. Importantly, we present results of the first study on the adoption of anti-tracking protection where participants' actual behaviour is observed, as opposed to self-reported intentions.

The remainder of this chapter is structured as follows. In Section 3.1, we first introduce the concept of online behavioural tracking and protection against it, as well as visualisation tools that can help to educate users about data practices in this context. Then, in Section 3.2, we discuss the motivation of this study and present the related conceptual framework. After that, in Section 3.3, we report on the methodology, participants' profile, results and limitations of our experiment. Next, in Section 3.4, we discuss the broader implications of our findings in the light of privacy knowledge and consent behaviour. In Section 3.5, we list the limitations of our study. Lastly, in Section 3.6 we provide a brief summary of the work presented this chapter.

3.1 Preliminaries

Before we continue to the conceptual framework, we introduce the topic of online behavioural tracking. Specifically, in this section, we first explain on how tracking technologies collect user data. Second, we discuss the strategies that users can use to prevent tracking from taking place. Finally, we provide details on the tracking visualisation tool that we later use in our experiment.

3.1.1 Online Tracking Technologies

When browsing the Web became a daily activity for millions of users, advertisers seized the opportunity to use online data about those users to personalize and target advertisements at them. This phenomenon is called *online behavioral advertising* (OBA). In a simple example, Boerman et al. (2017) explains how a company that serves OBA on thousands of websites, called an *advertising network*, tracks a user's website visits:

> '*If a consumer visits several websites about cars, the network assumes the consumer is interested in cars. The network can then display ads for cars only to people (presumed to be) interested in cars. Consequently, when two people visit the same website at the same time, one may see car ads while the other (who had visited websites about furniture) may see furniture ads.*' (Boerman et al., 2017)

However, website visits are not the only data that the advertising network can use to personalise OBA. While the collected data can include articles read, videos watched and everything searched for with a search engine (Boerman et al., 2017), it is the information that can be predicted based on this data that raises privacy concerns. For example, in 2010 the Wall Street Journal found that based on the data collected to serve OBA Microsoft's popular Web portal, MSN.com, was able to predict a sensitive personal

information about a user such as their age, ZIP code, gender, income, marital status, presence of children and home ownership (Angwin, 2010). In fact, according to the Wall Street Journal, collecting data about the activities of online users is one of the fastest growing businesses on the Web (Angwin, 2010). We refer to this kind of data collection as *online tracking*.

Although the detail of information that can be gathered through online tracking may seem concerning, the technology behind online tracking was initially intended to enhance the user's interaction with the Web. That is, the so-called 'Magic Cookies' (also known as HTTP cookies) were first introduced by a Web browser (client) developer, Netscape, in Navigator 1.1 to enable a user to return to a site and resume interaction where it was left off on the previous visit (Randall, 1997). In more detail, Felten and Schneider (2000) explain how the HTTP cookies work:

> '*The cookie itself is a relatively small amount of data, 'written' by the server and stored by the client as part of a normal HTTP request. Clients volunteer the contents of the cookie back to the same server on subsequent HTTP accesses as part of the HTTP request. Clients can store cookies across multiple browsing sessions; a cookie may have an expiry time associated with it beyond which it will be discarded by the client*' (Felten and Schneider, 2000).

Furthermore, trackers can use cookies to collect information about users' activities across multiple websites. In such case, the website is referred to as the *first party* and the tracker – the *third party* (Mayer and Mitchell, 2012). In fact, a small number of third parties have increasingly served advertising on a larger number of first-party websites, enabling these few trackers to track a user's activities across large portions of the Internet (Krishnamurthy and Wills, 2009).

Nonetheless, online tracking through HTTP cookies can be controlled. That is, the rejection of the HTTP cookies can be selected in Web browser settings, which allows users to opt out of the placement of HTTP cookies on their hard drive (Sipior et al., 2011). Additionally, the HTTP cookies already placed on the drive can be deleted through the browser or erased by anti-spyware software (Sipior et al., 2011). Consequently, in the advertising business, frequent deletion of HTTP cookies by users started leading to significant overestimation of the number of true unique visitors to websites and overpayment for advertising impressions (Soltani et al., 2009).

Thus, in an attempt to increase the reliability of online tracking technology, an online advertising company, United Virtualities, developed the so-called persistent identification element (PIE). In more detail, PIE was using the 'local shared objects', also called the Flash cookies – a feature of Adobe's Flash Player plug-in which at the time was installed on over 98% of computers (Soltani et al., 2009). By default, Flash cookies could

be stored or retrieved whenever a user accessed a page containing a Flash application (Sipior et al., 2011). While HTTP cookies expire at the end of a session, Flash cookies by default had no expiration date. Moreover, when a user deleted HTTP cookies, a Flash cookie was able to recreate those cookies (Sipior et al., 2011). However, Flash cookies were stored in a different location than HTTP cookies and browser privacy controls, such as erasing HTTP cookies, clearing history, or erasing the cache, were ineffective on them (Soltani et al., 2009). In fact, most users were neither aware of Flash cookies nor knew how to delete them (Soltani et al., 2009).

Furthermore, other online tracking technologies are even harder to control, less detectable and more resilient to blocking or removing, including Web beacons (that can track typed entries on a page and mouse movements) (Angwin, 2010; Sipior et al., 2011), browser fingerprinting, canvas fingerprinting, evercookies and 'cookie syncing' (Acar et al., 2014). In some cases, HTML5 client-side storage APIs, such as Web Storage, Web SQL Database and Indexed Database API were adopted as a way to enhance the capabilities of online tracking technologies (Belloro and Mylonas, 2018).

3.1.2 Online Tracking Countermeasures

At the same time, studies consistently show that users are concerned about online tracking. For example, Turow et al. (2009) found that 87% of their survey participants would not allow advertisers to track them online if given a choice, even if their activities were to remain anonymous. In another study, Wills and Zeljkovic (2011) built a website that provided users with personalised information about the third-party sites that were tracking their online activity and reported that 63% of the participants expressed concerns about third-party tracking. While users are not completely against targeted ads, they describe online tracking as 'scary', 'creepy' and 'invasive' (McDonald and Cranor, 2010; Ur et al., 2012).

Therefore, to address these concerns, research and commercial work has been carried out to enable protection against online tracking. Following the related literature (Felten and Schneider, 2000; Shirazi and Volkamer, 2014), we refer to these activities, browser functionalities and tools as online tracking *countermeasures*. In particular, results of several studies indicate that the most well-known countermeasure is cookie deletion (Leon et al., 2012; McDonald and Cranor, 2010; Shirazi and Volkamer, 2014; Ur et al., 2012). Other countermeasures offered by popular Web browsers include blocking cookies completely or browsing in incognito mode which prevents the browsing history from being stored.

Moreover, users can protect themselves against tracking by using browser-based blocking extensions which take different approaches to blocking trackers from loading and executing content. Specifically, Mathur et al. (2018) classifies these extensions into three

types: ad blockers, tracker blockers and content blockers. Firstly, since ad blockers block advertisements, they also block the advertising networks that are responsible for tracking users. For example, AdBlock[1] and AdBlock Plus[2] work by blocking from loading any browser requests that match patterns corresponding to known advertising networks. However, unless specifically augmented to do so, they both fail to block several other non-advertising third-party trackers. Secondly, tracker blockers block third-party trackers more generally, as opposed to just blocking advertising networks. They include tools such as Disconnect[3], Ghostery[4] and PrivacyBadger[5]. Finally, content blockers aim to function as general-purpose blockers. In addition to blocking both advertisements and other third-party trackers, they also block malware domains by default, enable blocking of pop-ups and contain other content-filtering features. Examples of content blockers include uBlock[6] and uBlock Origin[7].

3.1.3 Web Mirror

As part of our probe to understand how to provide users with privacy knowledge specific to online tracking, we considered different tools that personalise information to an individual user. Many of the available tools only provide information about trackers on the particular website, e.g. Ghostery, whereas we intended to educate users about tracking more generally. Moreover, tools that provide such an aggregated view of trackers, e.g. Lightbeam[8], need to be in use for an extended period of time before users can see their tracking activity, which was impractical for our study. Therefore, we decided to use Web Mirror developed by Gomer (2018) specifically with the objective of increasing users' understanding of online tracking. Below, we provide a brief summary of the functionalities of the tool based on its description from Gomer (2018).

Unlike Ghostery or Lightbeam, Web Mirror is not implemented as a browser add-on and does not need to be used for a period of time before results about tracking patterns become available. Instead, it is a Web application, designed to provide a sophisticated and immediate aggregate view of tracking activity based on an analysis of a sample of the user's browsing history. Once the browsing history is provided, an instrumented Web browser visits each link, recording any trackers connected to the website and the page content itself. This results in a personalised graph of the online tracking network, including both first- and third-party trackers. Additionally, during the crawl, key topics are extracted from each website to build a profile of topics for each visited website.

[1] AdBlock, `https://getadblock.com/`.
[2] AdBlock Plus, `https://adblockplus.org/`.
[3] Disconnect, `https://disconnect.me/`.
[4] Ghostery, `https://www.ghostery.com/`.
[5] PrivacyBadger, `https://privacybadger.org/`.
[6] uBlock, `https://ublock.org/`.
[7] uBlock Origin, `https://ublockorigin.com/`.
[8] Lightbeam 3.0, `https://addons.mozilla.org/en-GB/firefox/addon/lightbeam-3-0/`.

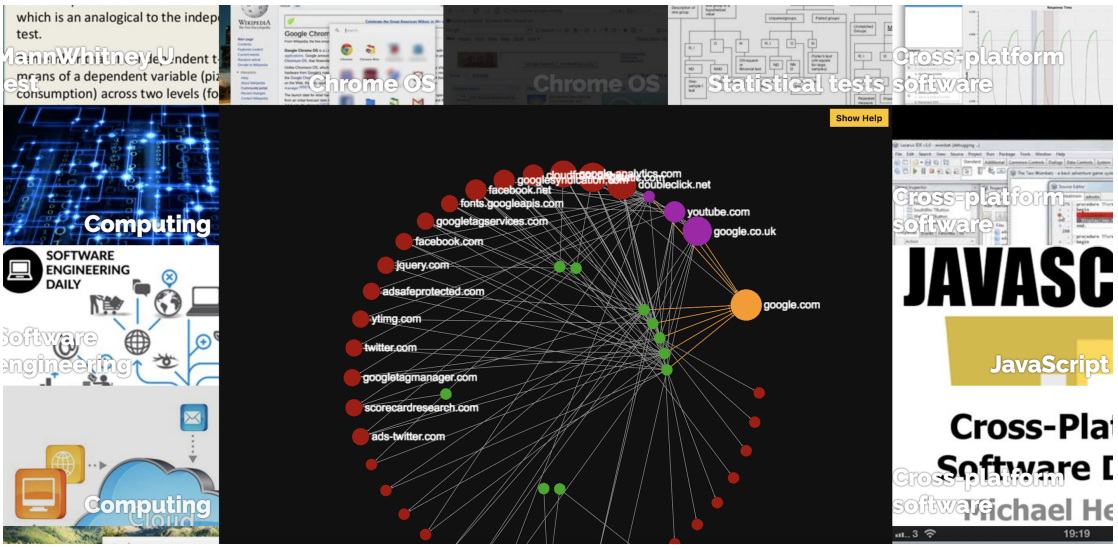


FIGURE 3.1: The interface of the Web Mirror.

By doing this, the tool builds a tracking profile of the user. As shown in Figure 3.1, the profile is presented to the user in a form of a visualisation of the tracking network graph and illustrations of the topics extracted. Through the provision of such a 'Web reflection', the tool educates the user on how trackers are connected to the websites they visit, what information about them might have been collected and what kind of inferences might have been made. The user is able to interact with the information provided by clicking on individual graph nodes to view details about a particular tracker. These include the amount of browsing observed, examples of pages tracked and the inferences drawn. The graph nodes are colour-coded: green nodes are websites that the user has visited, red nodes represent the trackers and purple nodes are the visited websites that also track the user's visits to other websites.

3.2 Motivation

In this study, we explore the relationship between privacy knowledge and privacy choices in online tracking where consent to data collection is often implicit. For example, even after the GDPR came into force, Sanchez-Rola et al. (2019) observed that 92% of the popular websites track users without consent. This happens even before showing a cookie notice and even if the user chooses to opt-out from being tracked. In fact, they also found that only few websites provide an easy way to opt out from tracking: in some cases, users have no choice – continuing to use a service is interpreted as consent to data collection; if a cookie banner has an option to refuse consent, rejecting tracking is often ineffective – the number of cookies set by the server remains the same or increases after refusal. In another study, Matte et al. (2020) found that 7% of the European websites they examined fail to provide any means to refuse consent, 10% store consent as if it

was granted even before a user has made any choice and 5% store a consent as if it was granted even when a user explicitly refuses to grant it. In fact, Degeling et al. (2019) reported that only 62.1–69.9% of European websites have consent notice banners.

Although there exist countermeasures that can effectively prevent online tracking, users tend not to utilise them. Initially, the non-use of countermeasures was associated with usability issues of many of the available anti-tracking tools (Leon et al., 2012). However, Shirazi and Volkamer (2014) later identified seven more detailed explanations:

1. '*People primarily correlate privacy issues with issues not related to identification and tracking (and thus corresponding countermeasures do not help to protect against identification and tracking on the Internet)*';
2. '*People are not aware that meta information is actually sent with each webpage request*';
3. '*People who are aware of meta data being transmitted tend not to be aware that such information can be used to identify and/or track them (or how)*';
4. '*People who are aware that identification and tracking is possible using their meta data are not concerned for various reasons. The first is because they feel that they are not important enough or have nothing to hide. The second is that they are not aware of who and why different entities would want to identify and track them other than to produce personalized advertisements. Third, they are not aware of actual consequences (other than personalized advertisements)*';
5. '*People are not aware of existing effective countermeasures*';
6. '*People are not able to use them properly*';
7. '*People are able to use them, but become side tracked for other reasons*'.

What can be observed is that all of these explanations are related to the lack of privacy knowledge. As detailed in Section 2.2.3, there are three dimensions of privacy knowledge: factual knowledge, procedural knowledge and experiential knowledge. In particular, procedural knowledge means understanding how to apply privacy-enhancing strategies and tools protect personal data. In context of online tracking, this refers to the tracking countermeasures listed in Section 3.1.2. Furthermore, factual knowledge refers to the awareness of corporate practices such as online tracking, the risks to user's privacy that these practices pose and the extent to which users realise when they share personal data. Tools such as the Web Mirror described in Section 3.1.3 provide users with factual knowledge on online tracking. In contrast, experiential knowledge refers to the general familiarity with online technology and any first-hand experience with privacy violations. Since such familiarity and experience is usually acquired by users over time, in this

study, we focus specifically on procedural and factual knowledge as our first step towards understanding how privacy knowledge affects privacy behaviour.

With regard to that, McDonald and Cranor (2010) found that the majority of American online users hold misconceptions about online tracking. Importantly, they identified a gap between the knowledge users currently have and the knowledge they would need to have to make informed privacy choices. To that end, they argued that 'consumers cannot protect themselves from risks they do not understand'. Therefore, in our study, we hypothesise that when users obtain procedural knowledge on privacy protection, they are able to make more informed privacy choices and, thus, some of them may decide to adopt privacy-protective behaviour. Stated formally:

> ***H1:*** *The provision of procedural privacy knowledge has an effect on users' privacy behaviour.*

However, some users who are familiar with ways to protect their privacy may not realise of the extent of data collection if they are not provided with factual knowledge on privacy. For example, Schaub et al. (2016) conducted a 24-participant lab study evaluating three anti-tracking browser extensions. Before the study, many of their participants were in the abstract aware of online tracking taking place, yet were unsure of who the tracking companies are, what is collected and for what purpose. With regard to that, they reported that the use of extensions on its own provided limited insight and at the end of the experiment their participants remained confused about many aspects of online tracking. Hence, we hypothesise that if users obtain both procedural knowledge on privacy protection and factual knowledge on online tracking, they are able to make even more informed privacy choices and, thus, more of them may decide to adopt privacy-protective behaviour. Formally, we propose the following:

> ***H2:*** *The provision of factual privacy knowledge in addition to procedural knowledge has a larger effect on users' privacy behaviour than the provision of procedural knowledge only.*

In addition, given that users in general have a low level of privacy knowledge and, consequently, are surprised when they learn about online tracking (Ur et al., 2012), it is currently unclear whether the discrepancies between their intentions and actual behavior (i.e. privacy paradox) translate directly into the online tracking context where data collection is often hidden. In fact, psychological reactance theory posits that whenever people perceive that their freedom to act or decide is threatened, restricted or eliminated, they tend to undergo a motivational reaction that intends to reestablish the affected freedom (Brehm, 1966; Brehm and Brehm, 2013). Therefore, users who learn about the extent of tracking may perceive that their choices are affected and may want to

reestablish them by using their knowledge about tracking prevention. Based on this reasoning, we expect that:

> ***H3:*** *If privacy knowledge is provided, intent to withdraw implicit consent leads to actual withdrawal of implicit consent.*

In order to test these hypotheses, we design an intervention where participants are provided with actionable information about selected tracking countermeasures: how they work and what actions a user can take to adopt them. In addition, some of the participants in our experiment are provided with access to the Web Mirror.

The connection between privacy knowledge and users' *intent* to adopt tracking countermeasures has been previously explored by Gomer (2018). However, users' intentions often do not to align with their actual behaviour (cf. privacy paradox). Therefore, in this chapter, we extend that work with the following contributions:

1. We investigate the impact of privacy knowledge on users' actual privacy behaviour in context of online tracking. Specifically, we explore:
   (a) the effect of procedural privacy knowledge on users' privacy behaviour,
   (b) the effect of procedural and factual privacy knowledge on users' privacy behaviour,
   (c) the connection between users' intent to adopt online tracking countermeasures and the actual adoption of online tracking countermeasures.
2. We provide a browser extension to log users' countermeasure adoption behaviour.

3.3 Experiment

We designed and conducted a randomised between-subjects experiment where participants were exposed to factual and procedural privacy knowledge on online tracking. In this section, we first report on the software equipment used in the study. Second, we present the methodology of the experiment. Third, we describe the participant recruitment criteria and demographics of the recruited sample. After that, we report on the results of the experiment with regard to the proposed hypotheses. Then, we discuss the implications of our results in the general context of privacy knowledge and privacy behaviour. Finally, we provide a brief summary of the contributions presented in this chapter.

3.3.1 Apparatus

To provide participants with privacy knowledge about the extent of the tracking network that is collecting information about their own online activity, we used the personalised visualisation delivered through Web Mirror. Therefore, to integrate the tool into our survey, Web Mirror was configured such that the participant spent at least three minutes interacting with their tracking network visualisation. After that time, a button appeared allowing the participant to continue.

In order to make this integration possible, we developed an extension to Google Chrome browser with the following functionalities:

- testing that there are no countermeasures installed in the browser at the beginning of the study;
- retrieving a sample of the participant's browsing history and submitting it to Web Mirror[9];
- logging the use of online tracking countermeasures for the duration of eight hours.

When a participant reviewed the information about the study and signed the consent form, they were asked to proceed to downloading our Chrome extension from Google Play store. Once connected to the Chrome browser, the extension was designed to automatically confirm that no countermeasures were installed – otherwise, interrupt the study. If the check was successful, it collected the participant's browsing history, uploaded it into Web Mirror and redirected the user back to our survey website. For the eight ours of Web browsing, the extension allowed us to log the adoption and deletion of the following countermeasures:

- manual cookie deletion,
- AdBlock Plus,
- AdBlock,
- Ghostery,
- SuperBlock Adblocker,
- Ad Remover,
- uBlock,

[9] After the History page in Google Chrome was updated with version 55.0.2883 in December 2016, it was no longer possible for a participant to copy the browsing history directly from the browser to paste into the Web Mirror.

- Disconnect,
- AdGuard,
- Baycloud Bouncer.

Afterwards, the extension automatically uninstalled itself from the browser and redirected the participant to the second part of the questionnaire. The extension was developed in JavaScript. The code has been published as part of the dataset.

3.3.2 Methodology

The experiment was designed to be conducted remotely. It consisted of two randomised treatments: participants in one group were provided with actionable information on anti-tracking protection (T1); in the other group, participants were additionally exposed to the visualisation of their tracking network (T2).

The experimental procedure started with a brief description of the study. As part of it, participants were informed about the requirement to install our browser extension for the duration of the experiment. This was communicated as follows:

> *As part of the study we will ask you to install an extension for Google Chrome and keep it enabled for a period of up to 8 hours – please use your browser as normal during that time.*

Following participants' consent, both treatments completed an entry questionnaire about their demographics and Web browsing habits (see Appendix A). In addition, we used an instrument developed by Chellappa and Sin (2005) to collect information about participants' privacy concerns and attitudes towards online services personalised through online tracking. We were particularly interested in the following statements included in the survey instrument:

> PRI1. *I am sensitive about giving out information regarding my preferences.*
>
> PRI2. *I am concerned about anonymous information (information collected automatically but cannot be used to identify me, such as my computer, network information, operating system, etc.) that is collected about me.*
>
> PRI3. *I am concerned about how my personally unidentifiable information (information that I have voluntarily given out but cannot be used to identify me, e.g., zip code, age range, sex, etc.) will be used by the website.*
>
> PRI4. *I am concerned about how my personally identifiable information (information that I have voluntarily given out AND can be used to identify me*

> *as an individual, e.g., name, shipping address, credit card or bank account information, social security number, etc.) will be used by the website.*

Then, they were redirected to Google Chrome Web Store to download the browser extension we developed for the study. Once the installation was complete, the extension automatically directed them back to our survey website.

From there, participants in the T2 treatment were asked for a permission for the browser extension to upload a sample of their Web browsing data to the Web Mirror through the browser extension. After a brief waiting time while their data was analyzed, they were shown their Web Mirror 'reflection'. To ensure that the participants engaged with the tracking network information provided, they were asked to interact with the visualisation for at least 3 minutes. After that, a 'continue' button appeared in a banner on top of the page, allowing them to move to the next part of the experiment.

Differently, participants in the T1 treatment were not exposed to the Web Mirror at any point. Instead, they were presented with a brief description of online tracking, which read as follows:

> *As we browse the web our web browsers connect to numerous different websites, including many that operate "behind the scenes". Behind the scenes websites deliver content and adverts to the "first party" website that we are visiting. These behind the scenes websites are called "third parties", and some of them deliver content and adverts to more than one first party. For instance, some large advertising companies deliver adverts across millions of different web pages. When we visit a first party site that is connected to third parties, the third parties can record our visits and build up a log of many different websites that we've visited. That information can be used to choose which adverts we see, based on the things that we might be interested in. This is why it can sometimes feel like a particular advert is "following" us!*

After that, participants in both treatments were directed to the intent questionnaire. In order to simplify the countermeasure information, we limited it to only three countermeasures: one ad blocker (AdBlock Plus), one tracker blocker (Disconnect) and manual cookie deletion. The participants were briefly informed about the effect of each of them and potential disadvantages, and given the exact instructions on how to adopt them. The step-by-step guidance included links to AdBlock Plus and Disconnect on the Google Chrome Web Store, where they could download the countermeasure in one click. However, participants were not requested to adopt the countermeasures. After each countermeasure was presented, the participants were asked about their intent to adopt it.

Next, participants in both treatments were asked to use the Chrome browser as usual until the extension was automatically uninstalled. During that time, the Chrome extension detected and logged their actual adoption of the countermeasures. The minimum time of logging was set to eight hours. Participants were allowed to turn the browser and their device off during that period, in which case the completion of the study took longer.

After at least eight hours, a new tab automatically opened in their browsers with the final part of the questionnaire. If any countermeasure adoption was detected during the eight-hour period, participants were asked for some reasons why they adopted the particular countermeasure. Those who earlier declared intent to adopt one or more countermeasures but decided not to do it were instead asked to select the reasons why they did not do it.

Before conducting the experiment, we carried out a week-long pilot study with 10 participants to validate the methodology and survey instruments, as well as to test the reliability of the apparatus and other aspects of the study such as its duration. During the pilot study, we observed that all of the participants who decided to adopt any countermeasures, did so on the same day. We therefore decided to limit the duration of the study to one working day, i.e. eight hours. The feedback from our participants indicated that 3 minutes was sufficient time for them to explore their Web Mirror reflection. All of the pilot study participants received £5 Amazon vouchers independently of the treatment allocation.

3.3.3 Participants

Following approval of the study by the Faculty Ethics Committee (ref. ERGO/FEPS/24293), we recruited 50 participants through social media posts and posters spread around the university campus. Having recognised that the recruited sample might have already had advanced privacy knowledge, we designed our participant selection criteria accordingly. Namely, instead of asking potential participants whether they were aware of online tracking (in particular, because users tend to have misconceptions about tracking and those who are aware of it are not necessarily aware of the large extent of it), we selected a number of factors that could contribute to a participant having decreased privacy concerns. Moreover, for technical reasons related to the accuracy of their Web Mirror 'reflection', participants were required to be regular Chrome browser users and to have sufficient browsing history stored in Chrome. More specifically, we were looking for participants who met the following criteria:

- were 18+ years old,
- used Google Chrome on a regular basis,

- never used browser extensions such as Ghostery, Disconnect or any kind of ad blockers,
- did not regularly clear their browsing history,
- had not participated in any previous user studies on online tracking countermeasures.

To double-check, the non-use of tracking countermeasures was confirmed in two ways at the beginning of the study. First, participants were asked to select from list the software tools they have used. Specifically, countermeasures were included in the list along with popular software such as Facebook, Twitter and Spotify. Then, the Chrome extension verified that no countermeasures were installed in the browser.

Out of those who responded to our study advertisement, 40 (80%) identified as women and 10 (20%) identified as men. 35 of the participants (70%) were between 18 and 40 years old, whereas the remaining 15 participants (30%) were over 40. Furthermore, 33 of the participants (66%) were British nationals. The remaining 17 participants (34%) were nationals of other countries such as Poland (5; 10%), Italy (3; 6%) and the USA (2; 4%). 20 participants (40%) reported having no undergraduate or professional degree, 17 (34%) had an undergraduate degree, and 7 (14%) had a postgraduate degree: master's (5; 10%) or Ph.D. (2; 4%). On average, 21 participants (42%) were spending 4 or less hours per day browsing the Web, 19 (38%) were spending between 5 and 8 hours, and 10

TABLE 3.1: Demographics of the study participants.

Variables		**Experimental Group**	**Control Group**	**Chi-Square Significance Level**
Gender	Male	2	8	$p = 0.0339$
	Female	23	17	
Age	18-25	5	8	$p = 0.2941$
	26-30	6	5	
	31-40	4	7	
	41-50	7	5	
	51+	3	0	
Degree	No degree	11	9	$p = 0.72$
	Bachelor's degree	6	11	
	Master's degree	3	2	
	Doctoral degree	1	1	
	Other degree	3	1	
	Prefer not to say	1	1	
Nationality	UK national	18	15	$p = 0.3705$
	Other	7	10	
Web browsing (h/day)	4 or less	9	12	$p = 0.3423$
	5-8	12	7	
	more than 8	4	6	

(20%) were spending more than 8 hours. In general, our population was notably young and over-represented women.

Subsequently, 25 of the participants (50%) were assigned to treatment T1 and the other 25 (50%) – to treatment T2. Table 3.1 presents in detail the differences in demographics between the experimental treatments, with all differences being statistically non-significant ($p > 0.05$, chi-squared test) apart from gender ($p = 0.0339$, chi-squared test). At the end, all participants who completed the study were rewarded with a £5 Amazon voucher independently of the treatment allocation.

3.3.4 Results

Overall, we found that 11 participants adopted any countermeasures, which is 22% of the total sample (5 in treatment T1 and 6 in T2). This finding suggests that the provision of procedural privacy knowledge has an effect on users' privacy behaviour. Therefore, hypothesis **H1** proved to be *true*. However, no significant difference was found between treatments T1 and T2 in terms of adoption of the countermeasures. In particular, exposure to user's tracking information did not significantly affect their actual countermeasure adoption. Therefore, hypothesis **H2** is *false*.

Interestingly, we found that 33 (66%) of the participants declared intent to adopt at least one of the countermeasures they were informed about. 16 (32%) of the participants intended to install AdBlock Plus, 15 (30%) – Disconnect, and 25 (50%) – to clear cookies. We also found that only 11 participants out of the 33 who declared the intent (6 out of 16 in the experimental treatment and 5 out of 17 in the control treatment) actually adopted any countermeasures. Thus, while majority of the participants intend to adopt countermeasures when told how to do so, our results show that a minority of them actually takes action to adopt them. This outcome suggests little correlation between the intent to adopt countermeasures and the actual adoption even when users are provided with privacy knowledge. Thus, hypothesis **H3** is *false*.

Furthermore, we compared the concern for privacy of participants who adopted any of the countermeasures and those who declared intent to adopt but did not take action. In particular, we looked at the responses to the statements from the survey instrument of Chellappa and Sin (2005). We found that a majority of the participants agree[10] with each of the above statements (PRI1: 36 (72%), PRI2: 31 (62%), PRI3: 29 (58%), PRI4: 42 (84%)). No significant difference in concern was found between the participants who adopted the countermeasures and those who declared the intent but did not take action.

In the final questionnaire, those who adopted countermeasures were asked to provide some reasons that they did so in a free-text response. The reasons were very diverse,

[10]They rated the statement as 5, 6 or 7 on a 7-point Likert scale (1 – strongly disagree, 7 – strongly agree).

such as *avoiding ads* (2 out of 11), *faster browsing* (2 out of 11), and *privacy* (2 out of 11). For instance, participant 30 said: "*after the information I got yesterday, I decided to take action to safeguard my privacy*". Those who intended to adopt a countermeasure, but did not do it, were asked to select one or more reasons from the list, constructed based on the qualitative responses from Experiment 2. In this case, the reasons were mostly usability related – the most commonly selected reason for not installing AdBlock Plus was "*It slows down my browsing*" (4 out of 11 participants), for Disconnect and cookie deletion – "*It's too much hassle*" (3 out of 9 for Disconnect, 5 out of 17 for cookie deletion). Interestingly, 3 out of 17 participants as a reason for not clearing cookies picked option 'other' and in the short-text response, reported lack of time to do it.

Our Chrome extension detected only one case of removal of the countermeasures during the study. When asked what were some reasons why they installed AdBlock Plus and Disconnect, participant 43 (in treatment T2) answered: "*I was intrigued to see how these add-ons worked but I did not keep them as I don't fully understand the purpose of them and so was a bit worried as to what they would do to my computer*".

3.4 Discussion

Making consent *informed* is a broad topic. In this chapter, we provide new results towards addressing research question **RQ1** on the impact of privacy knowledge on users' disclosure behaviour. To this end, we specifically focus on the provision of procedural and factual privacy. In this section, we discuss the implications of our findings that contribute towards that goal. We also compare our findings to those of previous work, to identify where they confirm existing knowledge, and where they represent new contributions.

Our results show that most of the participants intended to adopt additional countermeasures to stop tracking, regardless of whether they were exposed to the tracking visualisation or not (**H1**). This finding suggests that the provision of information about countermeasures on its own seems to provide an adequate stimulus to motivate intent to adopt countermeasures and any effect from the additional information about the tracking network is rendered inconsequential. When those measures and instructions how to adopt them are clearly described to them, the participants' willingness to consider tracking countermeasures makes us believe that providing readily-available means of acting with respect to privacy could be effective even without finding ways to make detailed information about privacy-invasive processes themselves easily consumable. This is an important outcome for **RQ1** and for the design of informed consent mechanisms, because it suggests that perhaps procedural information is more important for the user to make informed decisions than detailed factual information. Future work should further explore this research direction.

In addition, given that more detailed information about the trackers did not lead to our participants' stronger motivation to prevent tracking or greater adoption in practice (**H2**), we recommend that HCI designers and data protection regulators working to encourage users to increase their level of privacy should focus their efforts on actionable information about how users can prevent tracking, which as our study has shown, is enough to motivate both willingness and action to counteract tracking. Referring back to **RQ1**, this finding suggests that different kinds of privacy knowledge impact privacy behaviour differently. It also suggests that educational programmes where users learn about privacy online should focus on actionable information how users can choose to opt out of online tracking and protect their privacy in general.

Based on the results of the previous study by Gomer (2018) involving the Web Mirror, we expected that exposure to the tracking information would lead to a significant countermeasure adoption effect. Instead, our results suggest that advanced comprehension of tracking is not what users need to be informed about in order to be motivated to act (**H3**). This is an important information for efforts towards answering **RQ1**, because it strongly suggests that research on the impact of privacy knowledge on privacy behaviour should not rely on self-reported intent. From an empowerment perspective, this raises a question of whether users are informed about data processing practices purely to enable them to make own decisions, or because this information is valuable in and of itself. If the former, then identifying the minimum amount of information necessary for them to reach a stable decision (i.e., one that is not affected by further information) appears to have some interaction benefits in terms of lowering cognitive effort and decreasing the amount of time required to engage with privacy-protection mechanisms.

Considering the participants' ratings of the statements from the Chellappa and Sin's instrument, it is clear that the participants were concerned about their privacy before being told about the countermeasures and/or seeing the Web Mirror. However, we discovered that although majority of the participants declared the intent to adopt countermeasures, 11 out of 50 of them actually did so. While this is not a small number given the artificial study setup, this result reflects the reality of the *privacy paradox* present in the tracking context even when users are clearly instructed how they can protect themselves against being tracked. This finding suggests that although the privacy-unaware users report that they are concerned about the misuse of their data, their informed actions rarely follow their attitudes.

Nonetheless, we do not interpret this response to mean that people are not interested in protection against tracking, but that the perceived cost, including the effort or complexity of acting, is higher than the perceived benefit. Our qualitative results show that the actions (or lack of them) were taken based on rational decisions with concrete reasons such as that the countermeasure would slow down their browsing, takes too much time to install or '*it's too much hassle*'. This finding speaks in favour of the 'privacy calculus' theory and raises a question of whether the privacy paradox phenomenon is

indeed a paradox or instead, a rational judgement of costs (inconvenience, time, effort) and benefits (protection against tracking). In other words, a user might be concerned about the collection or use of their personal information in general, when it comes to a specific privacy-related situation, their decision on whether to take action to protect it would depend on their cost-benefit assessment at that moment. Even if factors such as misperceptions, social norms, emotions, and heuristics affect what those costs and benefits are perceived to be, it seems that the *perceived* costs and benefits play an important role in users' decision-making. This question opens up a line of further investigation of the topic.

Importantly, if the cost-benefit trade-off is expressed as a comparison between the payoff from adopting a certain countermeasure minus the cost of adopting it, and the payoff and cost of *not* using it (Acquisti, 2002), then our findings suggests that the cost of adoption and/or the perceived benefit does not encourage users to take action to protect their privacy. Since a minority of our participants adopted any countermeasures in practice, in this study we provide more evidence that requiring users to take action to protect themselves, even with access to current state-of-the-art transparency tools, is ineffective. Hence, instead of expecting people to take action and adopt countermeasures manually, we encourage designers to consider the importance of privacy by design (Hustinx, 2010) and by default (Tschersich and Niekamp, 2015; Mathur et al., 2018).

3.5 Limitations

Our study is not without limitations. Firstly, the small sample size (N=50) and demographics of the participants could have introduced a bias to the results. Given that a high proportion of the participants was young and educated, it is likely that they were also more tech-savvy and more likely to adopt countermeasures that the general population. In particular, for technical reasons of running the experiment, our participants were required to be regular Chrome users. Since Chrome is developed by an advertising company (Google), Chrome users may have different levels of comfort with ads and different attitudes towards ad blocking than the general population.

Secondly, the minimum duration of the experiment was only 8 hours. It would be interesting to see whether the results were different if the participants' activity was observed for longer and they were sent reminders with additional procedural and factual information during that time. In particular, the situational context in which the participants were provided privacy knowledge is unknown to us. It is possible that those participants who saw some unexpected results in their 'Web reflections' became more privacy-concerned and, thus, decided to adopt countermeasures. Future research should also examine the provision of privacy knowledge in situations that provoke privacy concerns, e.g. seeing unexpected advertisement.

Finally, our experiment was designed in a way that participants in both treatments were provided with procedural information. Given the results we observed, the next logical step would be to consider an experimental design where a treatment with procedural information only can be compared to a treatment with factual information only.

3.6 Summary

In this chapter, we focused on the role of privacy knowledge in making consent *informed*. Specifically, we investigated the impact of procedural and factual privacy knowledge on privacy behaviour. This is especially important for automated consent mechanisms, because as users gain more privacy knowledge, their privacy preferences may change.

To that end, we conducted a user study, where we observed how the provision of procedural and factual privacy information affected the participants' privacy behaviour. Our findings suggest that the procedural information acted as a stimulus that made 11 out of 50 (22%) of our participants take action to protect their privacy. This result shows that knowledge on how users can control their privacy has impact on their consent decisions.

Additionally, we compared the participants' privacy intentions to their actual privacy behaviour. In general, we observed that even when users are given detailed information on the actions they can take to protect their privacy, their stated intent rarely follows privacy behaviour. While this is consistent with the 'privacy paradox' phenomenon in other privacy behaviours, our results speak in favour of the 'privacy calculus' theory where users' behaviours are results of privacy-utility tradeoffs.

Our findings have valuable implications for the development of automated consent mechanisms. In particular, they highlight the importance of articulating available control options in consent mechanisms. Moreover, they suggest that there are concrete reasons why users may consent to data collection even when they recognise the privacy risk, which should be taken into account for consent mechanisms to be functional and usable. Last but not least, the study found many promising avenues for further exploration.

Chapter 4

Negotiable Consent

While some of the data that service providers collect is necessary for them to be able to provide the expected functionality to the user, other information can be shared by the user voluntarily. However, the user may be more willing to share their non-essential personal information if they receive an incentive in return. Thus, as discussed in Chapter 2, there is an opportunity for negotiating consent to collection, use and processing of non-essential personal data for an incentive to achieve outcomes beneficial for both the user and the service provider.

In this chapter, we illustrate how agent-based negotiation can support the process of reaching a privacy agreement between the user and the service provider, and how a user can be represented by an agent that learns their privacy preferences. Firstly, we propose a novel framework for automated negotiation of privacy agreements. As part of the framework, we introduce an alternating-offers, multi-issue negotiation protocol, in which an agent representing the user can propose partial offers by specifying values for some issues of the agreement, and the service provider's agent completes those offers. We argue that such a framework is more suitable for this negotiation domain, allows for more collaborative exploration of the negotiation space to find mutually beneficial agreements and avoids distributive negotiation on single issues such as price.

Furthermore, in an empirical study, we look at two different ways of modelling the preference profile of the user in such negotiation. Specifically, we compare an approach where the preference profile is personalised for each user based on their previous decisions (accepted or rejected offers), to the approach of Baarslag et al. (2017), where the agent classifies a user as being one of a three different privacy types later adjusted based on their privacy behaviour. The results of our user study show that offers proposed by the latter are more accurate than those of the former. However, we can observe a rising trend of the accuracy of the new agent, which in the end, exceeds the accuracy of the agent of Baarslag et al. (2017).

The remainder of this Chapter is structured as follows. In Section 4.2, we motivate the development of our new framework. Next, in Section 4.3, we formally propose our framework for such negotiation, with a protocol for bilateral multi-issue negotiation. After that, in Section 4.4, we use the agent design of Baarslag et al. (2017) to demonstrate how the general framework can be implemented. Next, in Section 4.5, we apply the framework and its proposed implementation to the privacy permission negotiation domain. Furthermore, in Section 4.6, we describe the apparatus, methodology, results and limitations of our experimental evaluation. Ultimately, in Section 4.7, we discuss the implications of our results for future design choices for privacy management and automated negotiation. In Section 4.8 we mention the limitations of our experiments. Lastly, in Section 4.9, we provide a brief summary of our work.

4.1 Preliminaries

Before we move on to the framework for privacy negotiations, in this section, we introduce the topic of automated negotiation under preference uncertainty. Firstly, we explain the general concepts behind automated negotiation. Secondly, we discuss the common strategies used in user preference elicitation. Finally, we briefly summarise the recent work on an automated negotiation agent for permission management that this chapter builds upon.

4.1.1 Automated Negotiation

In general, negotiation is about a joint exploration of outcomes in search for mutual gains. In doing so, the ultimate goal of negotiation is to resolve the conflict of interest present among different parties. Since negotiation covers so many aspects of people's lives this has led to an increasing focus on the design of automated negotiators, i.e., autonomous agents capable of negotiating with other agents in a specific environment (Jennings et al., 2001; Kraus, 2001).

There is a significant body of literature that deals with automated negotiation. This interest has been growing since the beginning of the 1980s with the work of early adopters such as Smith's Contract Net Protocol (Smith, 1980), Sycara's persuader (Sycara-Cyranski, 1985; Sycara, 1988), Robinson's oz (Robinson, 1990), as well as the work by Rosenschein (1986), and by Klein and Lu (1989). Broadly, there are two approaches: those using protocols based on the alternating offers approach where only offers are exchanged, and argumentation-based approaches where additional information is conveyed in an attempt to convince their counter part (Karunatillake, 2006). In the former the agent preferences are typically modelled using utility theory and techniques such as game theory and decision theory are applied, whereas the latter is mostly based on logical inference. Our approach in this chapter is based on the former.

In more detail, the alternating offers protocol (Osborne and Rubinstein, 1994; Rubinstein, 1982) is the best-known and most widely studied model for bargaining (Fatima and Wooldridge, 2014). Following this protocol, offers are exchanged between two agents over a sequence of rounds. When one agent (the proposer) submits an offer, the other one (the responder) can either accept it or reject it. If the offer is rejected, the responder proposes an alternative offer which the proposer can accept or reject. The negotiation continues until one of the following conditions is satisfied: an offer is accepted, a deadline is reached or one of the parties terminates the negotiation. While there also exist other negotiation protocols such as the monotonic concession protocol (Rosenschein and Zlotkin, 1994), one of the main advantages of following this approach is that an abundance of agents have already been formulated for the alternating offers protocol (e.g. Aydoğan et al. (2017); Baarslag et al. (2015); Chen et al. (2013); Fatima et al. (2002); Hao and Leung (2014); Ilany and Gal (2014); Kawaguchi et al. (2012); Kraus (2001); Williams et al. (2012)) which could be easily adapted to our model in the privacy context.

More specifically, the protocol we propose in this chapter is based on the multi-issue bilateral alternating-offers protocol where two agents negotiate offers with regard to not just a single item or a single bundle of items, but on many issues. In general, there are two approaches to such negotiations: one way is to negotiate all the issues together and the other is to negotiate them one by one (issue-by-issue) (Fatima et al., 2003). As users' privacy preferences are influenced by the context (Nissenbaum, 2009, 2011), in our work, we follow the former approach.

4.1.2 User Preference Elicitation

Despite the benefits that negotiation can provide, the negotiation process is also a time-consuming and expensive activity that people often find challenging and stressful (Fatima et al., 2014). To alleviate these difficulties, in automated negotiation, users are represented by agents that are familiar with the users' individual preferences. To do so effectively, the agent needs to obtain an accurate model of the user's preferences through the process of *preference elicitation* before it starts negotiating on the user's behalf.

In general, while research has focused on opponent preference modelling, user preference elicitation in automated negotiation has received little attention. A large family of strategies proposed for user preference modelling are the UTA methods (UTilités Additives) which obtain a ranked set of outcomes as input and formulate a linear utility function through the use of linear programming (Jacquet-Lagreze and Siskos, 1982; Greco et al., 2012; Van Nguyen, 2013; Roszkowska et al., 2016). CP-nets, which provide a qualitative representation of preferences that reflects conditional dependence (Boutilier et al., 2004), were also studied in the automated negotiation context (Baarslag and Gerding, 2015a; Mohammad and Nakadai, 2018). Furthermore, Tsimpoukis et al. (2018) proposed a decision model that uses linear optimization to translate partial information

into utility estimates based on a set of ranked outcomes. However, as we explain later in Section 4.4.2, when the input data stems from possibly erroneous information collected from a human user, the problem quickly becomes over-constrained and linear constraint solvers can no longer be applied. Instead, in this chapter, we propose a preference elicitation method based on approximation techniques studied by Popescu (2003) to find the most preferable outcomes.

When it comes to modelling the user's privacy preferences specifically, a number of studies have looked these issues. In particular, early work has focused on modelling users' location sharing preferences (Benisch et al., 2011; Cranshaw et al., 2011; Ravichandran et al., 2009; Sadeh et al., 2009). For instance, personas and incremental suggestions were used to learn users' location privacy rules, resulting in users sharing significantly more without a substantial difference in comfort (Kelley et al., 2008; Mugan et al., 2011; Wilson et al., 2013). However, as users have different sensitivity regarding different types of personal data and their willingness to provide information varies depending on the perceived risk of sharing the particular information (Milne et al., 2017), location sharing preferences do not easily translate to other types of personal data. Thus, a more general approach is needed.

To understand users' privacy preferences for other types of data, researchers surveyed users' expectations and subjective feelings about different kinds of sensitive data, and identified key factors that influence them (Lin et al., 2012). Then, clustering techniques were used to define a set of *privacy profiles* based on users' self-reported privacy preferences (Lin et al., 2014; Liu et al., 2014). However, privacy profiles do not take individual user differences into account. To address this issue, in our study, we compare the profile-based approach to one where the prediction outcomes are based on the behaviour of an individual user.

4.1.3 Automated Negotiation Agent for Permission Management

In 2017, Baarslag et al. (2017) proposed a user agent for an automated negotiation of privacy permissions. In general, the agent was able to model the permission preferences of different user profiles and apply the right model when it represented a user assigned to one of these profiles.

More specifically, they conducted a user study with a mobile app, where participants were asked to select the resources (*contacts*, *messages*, *the list of applications installed*, *photos* and *browsing history*) they would allow the app to access. In the first phase of the study, the participants initially completed the three-question Westin Segmentation Index survey which classified them into one of the three privacy profiles: *Fundamentalist*, *Pragmatist* or *Unconcerned*. Then, they were proposed a 'default setting' of the permissions, which was a randomly selected proposal of enabled permissions. The user

was able to modify this setting by allowing or disallowing each of the permissions individually before granting them in exchange for points. This At the end, the final number of points collected was converted to a monetary reward. In Section 4.6.3, we refer to this treatment as the Manual Negotiation (MN).

Doing this allowed the researchers to collect data on the permission preferences of users assigned to the three privacy profiles. Therefore, they conducted the second phase of the user study where after the user was classified into a privacy profile, instead of the 'default setting', the agent used the preference model of the group of users previously assigned to the particular privacy profile. While the opponent agent was not developed, its behaviour was mocked by rewarding the user with a uniformly random number of points, depending on the number of permissions enabled. We explain this in detail in Section 4.5, referring to this agent as Agent 2.

4.2 Motivation

In the classic multi-issue alternating offer protocol, offers are exchanged between two parties that specify values for each of the negotiable issues. As detailed in Section 4.1.1, the proposer submits a fully specified offer, which the responder can either accept or reject. If the offer is rejected, the responder proposes an alternative offer, which the proposer can accept or reject, and the negotiation continues until an offer is accepted, a deadline is reached or one of the parties terminates the negotiation.

However, in practice, parties typically have asymmetric roles such as user and service provider. Often, a party can find it difficult to determine the value of some issues, which are more naturally determined by the counter party. For example, when negotiating a software development contract with negotiable issues such as the functionality requirements, maintenance level, hardware infrastructure and delivery timeframe, it is more natural for the customer to solely specify the functionality requirements, whereas the developer may choose how much infrastructure they can make available for the project given the requirements. In a similar way, during a negotiation of an insurance policy, it is usually the buyer who specifies the conditions for the extent of the cover – the seller then completes the contract by proposing a price.

In particular, the ability of a party to leave some issues unspecified is important for privacy negotiations. In such negotiations, some essential issues need to be controlled by the user, e.g. what kind of data access the user is happy to consent to, while others must stay under control of the service provider, e.g. what reward or extra functionality can be offered in exchange for access to the data. Thus, the user may want to initiate the negotiation, specifying the requirements for a subset of important issues (e.g. *I am happy to share my contacts but not my browsing history; what will I get for sharing this?*). The service provider is then able to submit a complete counter-offer based on

the proposed offer (e.g. *without your browsing history, I miss out on ad revenue; given that, I can offer you £2.99*). The user can either accept the proposal or submit a new enquiry (e.g. *What if I choose to share my contacts and location, but not my browsing history?*).

To model this kind of automated negotiation, our premise is that, before the negotiation takes place, the proposer and the responder prescribe which issues they specify values for. This makes sure that, during the process, essential issues are under control of the party exchanging them, whereas others are left open. In Section 4.3, we do so by introducing *partial offers*. More specifically, as the first contribution of this chapter, we formalise the process of privacy negotiations by proposing a novel negotiation framework for multi-issue bilateral negotiations where offers of the proposer allow unspecified values.

Furthermore, ensuring that users are fairly compensated for the privacy loss suffered as a result of sharing personal information is a non-trivial problem. This is particularly challenging when the user is represented by an automated agent, because the user's views of fairness are subjective within the privacy context (Campbell, 1997; Wilkinson et al., 2018). In fact, the *effortless privacy negotiations* envisioned by Krol and Preibusch (2015) specifically apply to negotiations where agreements are reached between the service provider and individual users with heterogeneous privacy preferences. Thus, an important part of our work is modelling and learning the user's privacy preferences such that the agreement can be personalised according to each individual user's needs. For this reason, the second contribution of this chapter is the new preference elicitation method presented in Section 4.4.

Last but not least, to make sure that our theoretical results can be successfully applied in practice, it is important that they are evaluated with human users. This is because, as mentioned earlier in Section 4.1.2, input data that stems from interactions with a human user may be erroneous. Therefore, the third contribution presented in this chapter is a user study, where we compare our new preference elicitation method personalised to the individual user to the method of Baarslag et al. (2017) based on privacy profiles. In Section 4.6, we refer to these treatments as A1 and A2 accordingly. We also compare our negotiation approach to two treatments without a negotiation agent where the offers presented to the human user are chosen randomly. Specifically, in treatment MN, users are able to communicate their preferences in the same way as in A1 and A2, whereas treatment TIOLI represents the *take-it-or-leave-it* approach with no tunable control over the privacy trade-offs. To do so, we recruit 66 study participants for treatments A1 and TIOLI, and present a new analysis of data collected by Baarslag et al. (2017) for treatments A2 and MN[1].

More specifically, this chapter provides the following original contributions vs. Baarslag et al. (2017):

[1] In their paper, Baarslag et al. (2017) refer to these treatments as 'Agent' and 'Random' accordingly.

1. We generalise the agent from Baarslag et al. (2017) to propose a novel negotiation framework for multi-issue bilateral negotiations where the roles of the proposer and the responder are asymmetric, such as a buyer and a seller, or a user and a service provider. As part of the framework, we propose a new alternating offers protocol with costly quoting, in which one agent can propose partial offers by specifying values for some issues, and the other agent completes those offers.

2. We propose a new user preference elicitation method personalised to the needs of an individual user, while the method published by Baarslag et al. (2017) relied on privacy profile classification.

3. We present results of a new user study with 66 human users. We also re-analyse the data collected during the study published by Baarslag et al. (2017) and compare it to the data from our new experiment.

4.3 Negotiation Framework

In this chapter, we propose a negotiation framework with a novel protocol we refer to as the partial-complete offer protocol, which is a variant of the multi-issue alternating-offers protocol. We first introduce the negotiation setting. After that, we describe the partial-complete offer protocol, where two parties propose offers consisting of values of the negotiable issues. Finally, we explain how the utility of each of the two agents is calculated when the negotiation concludes.

4.3.1 Negotiation Setting

In this chapter, we focus on bilateral multi-issue negotiations. In more detail, a negotiation domain is specified by m issues, where $m \in \mathbb{N}$ and $m > 1$. To reach an agreement, the agents must settle on an outcome that is accepted by both parties of the form $\vec{\omega} = (\omega_1, \ldots, \omega_m)$, where ω_i denotes a value associated with the i-th negotiable issue for $i \in \{1, \ldots, m\}$. Then, $\Omega = \prod_{i=1}^{m} \Omega_i$, where Ω_i represents the set of possible values of ω_i. For example, in the context of privacy permissions, if the issues Ω_i correspond to *Shared data*, *Purpose of sharing*, *Retention policy*, *Price discount*, then an example agreement in $\vec{\omega} \in \Omega$ is (*GPS location*, *Targeted ads*, *Shared with third parties only*, *£0.20*).

In this framework, we allow a value of an issue to be undefined. We denote so by a special character $\perp \in \Omega_i$ for each $i \in \{1, \ldots, m\}$. Moreover, we define a subset $S \subseteq \Omega$ containing all offers where all values are defined. Formally, for each offer $\vec{\omega} \in S$, $\omega_i \neq \perp$ for any $i \in \{1, \ldots, m\}$. Because of this property, the elements of S are called the *complete offers*. On the contrary, the elements of the complement of S, $\overline{S} \subseteq \Omega$, are called the *partial offers*.

While the domain is common knowledge, the preferences of each agent are its private information. Therefore, in addition to its own preferences, every agent also has an *opponent model*, which is an abstract description of the opponent's preferences. This allows the agent to employ a *negotiation strategy* to determine its optimal action in a given state of the negotiation.

Since we focus on bilateral negotiation, there are two negotiating agents involved. We call them the *proposer* and the *responder*, and present the protocol that dictates their moves.

4.3.2 Negotiation Protocol

In the Partial-Complete Offer Protocol, the proposer submits a partial offer specifying the requirements for a subset of issues. Formally, the proposer offers $\vec{\omega}^p = (\omega_1^p, \ldots, \omega_m^p)$ such that there exists (possibly more than one) $i \in \{1, \ldots, m\}$ for which the value remains unspecified, i.e. $\omega_i^p = \perp$. The responder is then able to complete the offer, taking into account the proposed partial offer. That is, the responder replies with a complete offer $\vec{\omega}^r = (\omega_1^r, \ldots, \omega_m^r)$ such that $\omega_i^r = \omega_i^p$ for all $\omega_i^p \neq \perp$. This ends the first *negotiation round*. From here, the proposer can either accept a complete offer and end the negotiation, reject it and submit a new partial one starting the next negotiation round, or break off the negotiation.

As the negotiation continues, we assume that previous offers remain valid. That is, the complete offers returned by the responder generate a growing set $Q \subseteq S$ of possible outcomes that the proposer can agree to. Then, the negotiation ends when the proposer either accepts an offer $\omega \in \Omega$ or actively ends the negotiation by signalling a break off.

In addition, negotiations often involve some form of time pressure to ensure that they finish in a timely manner. In order to avoid the proposer exploring all possible partial offers, the proposer incurs a bargaining cost $c^p(n) \in \mathbb{R}$ where n is the total number of negotiation rounds. Similarly, to make sure that the responder's offers are more likely to be accepted by the proposer, additional bargaining cost $c^r(n) \in \mathbb{R}$ is levied on the responder.

To summarise, this the step-by-step process is presented in Protocol 1. Additionally, the protocol is illustrated in Figure 4.1 as a sequence diagram. Next, we explain how the negotiation outcome and costs affect the utility of each of the negotiating agents.

4.3.3 Utility

When negotiation ends, the utility of each agent is updated. This depends on the cost the agent incurred and, if an offer is agreed, the agent's valuation of the offer. Furthermore, each agent has a *reservation value*, which is the utility of a disagreement.

Protocol 1 Partial-Complete Offer Protocol

Inputs. The number of issues m; for each $i \in \{1, \ldots, m\}$, a set of values Ω_i associated with the ith issue; the set of complete offers S; the set of partial offers $\overline{S}$; the cost per n rounds for the proposer $c^p(n) \in \mathbb{R}$ and the responder $c^r(n) \in \mathbb{R}$.
Goal. The proposer and the responder agree on the outcome of the negotiation.

Initialisation. The number of completed rounds is $n = 0$.

The protocol:

1. The proposer submits a partial offer $\vec{\omega^p} = (\omega_1^p, \ldots, \omega_m^p) \in \overline{S}$ such that there exists $i \in \{1, \ldots, m\}$ for which $\omega_i^p = \perp$.
2. The responder replies with a complete offer $\vec{\omega^r} = (\omega_1^r, \ldots, \omega_n^r) \in S$, where $\omega_i^r = \omega_i^p$ for all $\omega_i^p \neq \perp$.
3. The number of completed rounds is incremented: $n = n + 1$.
4. The proposer performs one of the following actions:
   (a) rejects $\vec{\omega}^r \rightarrow$ go to 1.
   (b) accepts $\vec{\omega}^r \rightarrow$ the negotiation ends with outcome $\vec{\omega}^r$, and with costs $c^p(n)$ for the proposer and $c^r(n)$ for the responder.
   (c) breaks off the negotiation $\rightarrow$ the negotiation ends with no outcome, and with costs $c^p(n)$ for the proposer and $c^r(n)$ for the responder.

Protocol 1: The Partial-Complete Offer Protocol.

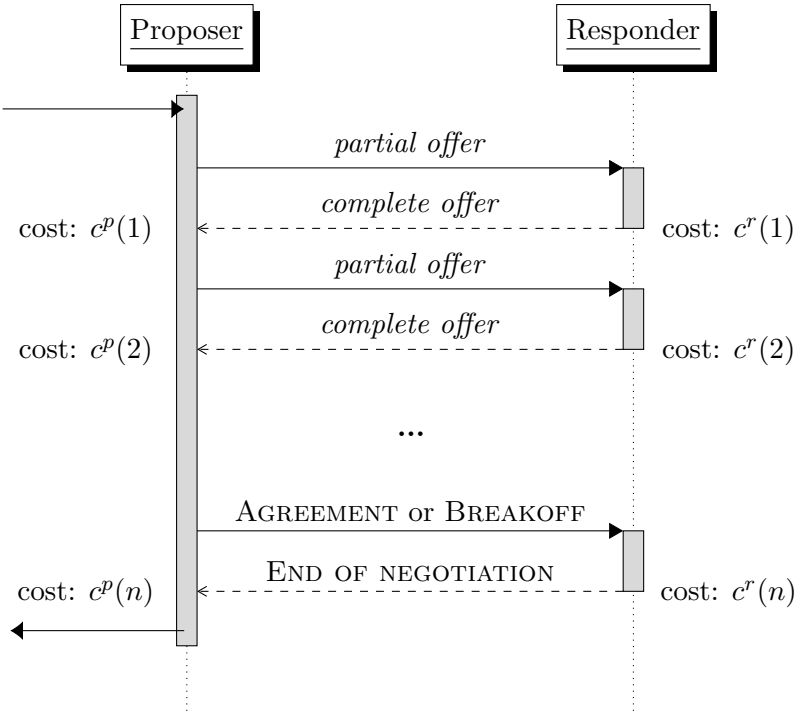


Figure 4.1: Sequence diagram of a negotiation that ended after n rounds.

Specifically, the negotiation either ends with an outcome $\vec{\omega} \in Q$ or with no outcome. Then, if the proposer's valuation is defined by a valuation function $v^p : \Omega \longrightarrow [0,1]$ and the proposer's reservation value is $r^p \in [0,1]$, the utility of the proposer when the negotiation stops is given by:

$$U^p(Q) = \begin{cases} v^p(\vec{\omega}) - c^p(|Q|) & \text{if } \vec{\omega} \in Q \text{ is the agreed outcome,} \\ r^p - c^p(|Q|) & \text{if no agreement is reached.} \end{cases} \tag{4.1}$$

Consequently, if the responder's valuation is encoded by $v^r : \Omega \longrightarrow [0,1]$ and the responder's reservation value is $r^r \in [0,1]$, then the utility of the responder is defined as:

$$U^r(Q) = \begin{cases} v^r(\vec{\omega}) - c^r(|Q|) & \text{if } \vec{\omega} \in Q \text{ is the agreed outcome,} \\ r^r - c^r(|Q|) & \text{if no agreement is reached.} \end{cases} \tag{4.2}$$

Given that the valuation function of one agent is not necessarily know to the other one, each of the agents aims to maximise their expectation over utility. Thus, the challenge the agent faces lies in employing an adequate negotiation strategy to determine the optimal sequence of actions. Furthermore, in some situations where human users are involved, the exact valuation function may not necessarily be defined for an agent due to the uncertainty over the user's preferences. To address this, we show how such negotiations can be modelled using this framework.

4.4 Negotiation Agent

In this section, we use the agent design of Baarslag et al. (2017) to demonstrate how the framework can be implemented to account for situations where a user is negotiating with a service provider. Specifically, we consider a setting, where the proposer is negotiating on behalf of a user who is unwilling or unable to fully specify the valuation function. This is particularly the case in many realistic scenarios, where collecting the necessary preference information to define the valuation function is time-consuming or costly, or communicating them is difficult for the user. To address the challenge that such preference uncertainty brings into the negotiation, we first discuss how the agent builds the user and the opponent models. Second, we explain the method of preference elicitation which allows the agent to learn the user's valuation. Finally, we present the negotiation strategy which allows the agent to decide what offer, if any, to send to the opponent and when to terminate the negotiation.

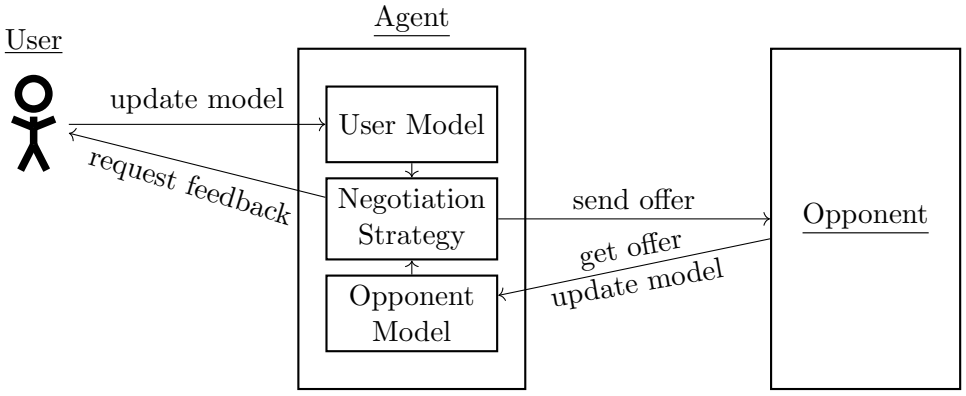


FIGURE 4.2: The interaction model between the user, the agent and the opponent.

4.4.1 Models of Uncertainty

In order for an agent to faithfully represent the user, it is important that the agent's offers are aligned with the user's preferences. In our setup, we assume that the proposer agent first negotiates the offer with the responder (its opponent) on behalf of the user, and then proactively interacts with the user to establish whether the negotiated offer aligns with their preferences. Rather than asking the user directly, the agent derives the user's preferences from feedback on the negotiation. This way, the agent is able to incrementally collect information on user's preferences for future negotiations at the most relevant time (that is, *during* the negotiation process), while constantly keeping the user in the negotiation loop, as explained in detail in Section 4.5.3. Similarly, information on the opponent's valuation can be updated through the offers that are exchanged with the opponent.

To this end, we can identify two kinds of preference uncertainty in this negotiation:

- the user's preferences regarding the possible negotiation outcomes,
- the opponent's reactions to the offers.

In order to model this uncertainty, the agent builds the *user model* and the *opponent model*. Specifically, the user model consists of the agent's beliefs about the user's preferences. This can be elicited from the agent's interactions with the user. Conversely, the opponent model, which reflects the agent's beliefs about the opponent's reactions, depends on the negotiation strategy used by the service provider. Such a model could be constructed from prior knowledge or previous interactions with the opponent, and can be based on the relative likelihood of the counter-proposal from the service provider (for an overview, see Baarslag et al. (2012, 2016)). This model of uncertainty has been illustrated in Figure 4.2.

4.4.2 Privacy Preference Elicitation

In a negotiation, a user has a specific set of preferences regarding the possible outcomes. The so-called *preference profile* is given by an ordinal ranking over the set of the outcomes: an outcome ω is said to be weakly preferred over an outcome ω' if $\omega \succeq \omega'$ where $\omega, \omega' \in \Omega$ or strictly preferred if $\omega \succ \omega'$ (Tsimpoukis et al., 2018). Given the outcome ranking, the agent's goal is to formulate its estimated the valuation function v^p that approximates the real user's valuation as much as possible so that the preferences are expressed in a cardinal way:

$$\omega \succeq \omega' \iff v^p(\omega) \geqslant v^p(\omega'). \tag{4.3}$$

Following the literature on multi-issue negotiation (Keeney and Raiffa, 1976), we make the common assumption that the agent's valuation function is linearly additive. That is, it has the following form:

$$v^p(\vec{\omega}) = \omega_1 w_1 + \cdots + \omega_m w_m, \tag{4.4}$$

where w_i is a weight indicating the importance of the i-th negotiable issue, and the weights are normalized such that:

$$\sum_{i=1}^{m} w_i = 1. \tag{4.5}$$

Therefore, deriving the valuation for the issues specified in the partial offer means deriving the weights $w_1, \ldots, w_m$. This can be performed with feedback from the human user on the previous negotiation outcomes as explained in detail in Section 4.5. That is, if a user previously disliked a negotiated outcome ω (including the case of granting consent to share some personal data, as well as the case of sharing nothing) in favour of approving an outcome ω', we can assume that $v^p(\omega') \geqslant v^p(\omega)$. This can be written as:

$$(\omega'_1 - \omega_1)w_1 + \cdots + (\omega'_m - \omega_m)w_m \leqslant v^p(\vec{\omega'}) - v^p(\vec{\omega}). \tag{4.6}$$

If we do this for all previously approved and disapproved complete offers, we obtain a set of inequalities, from which we can deduce the most appropriate overall weights $w_1, \ldots, w_m$. To this end, note that this procedure transposes the problem into a set of linear inequalities of the form:

$$Aw \leqslant b, \tag{4.7}$$

where the entries of A and b correspond, for each equation, to the values of $(\omega'_i - \omega_i) \in \{-1, 0, 1\}$ and $b = v^p(\vec{\omega'}) - v^p(\vec{\omega})$ respectively. We refer to these combinations of A and b as *constraints*.

However, as this data stems from human interaction and the problem quickly becomes over-constrained, we find that these inequalities are typically not consistent. Therefore, we cannot simply use standard linear constraint solvers. To address this, instead, we can find a solution that best satisfies the constraints following the techniques described by Popescu (2003). Specifically, we determine the weights w^* that minimize the least squares norm:

$$w^* = \arg\min_w ||(Aw - b)_+||^2, \tag{4.8}$$

where $(Aw - b)_+$ is the vector whose i-th component equals $\max\{(Aw - b), 0\}$.

The weights w^* can be then plugged to Equation 4.4 to derive the valuation function, which the negotiation strategy relies on to select the most preferable offers.

4.4.3 Negotiation Strategy

Using the valuation function, the negotiation strategy needs to determine which of the partial offers, if any, the agent should propose to the opponent. What makes this problem non-trivial is its sequential nature: whether or not to propose a partial offer depends on the offers proposed so far. Therefore, the goal is to find an optimal *sequence* of offers to propose and a corresponding strategy which specifies when to conclude the negotiation process. To find such optimal negotiation strategy, we propose a similar approach to the one used in related work for preference elicitation (cf. Baarslag and Gerding (2015b)).

To evaluate partial offers, we assume that the agent has a model of the likelihood of receiving a certain complete offer from the opponent. That is, the probability of a complete offer given a partial offer $\vec{\omega}_p$ is given by a stochastic variable X_{ω_p}, with a cumulative distribution function $G_{\omega_p}(x)$ known to the agent. From this, the expected value of a partial offer can be derived. Specifically, the valuation of a complete offer returned by the opponent is described by a stochastic variable $Y_{\omega_p} = v^p(X_{\omega_p}) \sim [0, 1]$ with a corresponding cumulative distribution function $H_{\omega_p}(y)$. Additionally, we assume that the total cost of the negotiation for the agent after n rounds is defined as $c^p(n) = C^p n$.

With these assumptions, our goal is to formulate a negotiation policy π which, given a state $\mathcal{E}$, decides whether to continue or stop the negotiation. Since the costs of offer proposals are sunk, it is easily verified that the negotiation strategy depends on the offers proposed, $P \subseteq \overline{S}$, and the offers received, $Q \subseteq S$ by the agent so far. Thus, the current negotiation state can be summarised by $\mathcal{E} = \langle P, Q \rangle$.

Given the state $\mathcal{E}$, the negotiation policy π should either stop the negotiation (with an outcome in Q or with no outcome) and obtain utility $U^p(Q)$ (defined in Equation 4.1), or continue the negotiation by proposing a new partial offer $\vec{\omega}^p \in \{\overline{S} \backslash P\}$ at cost C^p. If $\vec{\omega}^p$ is proposed and a new complete offer $\vec{\omega}^r \in S$ received, the negotiation enters a new state which can be described as $\mathcal{E}' = \langle P \cup \{\vec{\omega}^p\}, Q \cup \{\vec{\omega}^r\} \rangle$. Since at that point the future utility in state $\mathcal{E}'$ cannot be observed, the expectation over utility can be represented by the expected value $\mathbb{E}_{\vec{\omega}^r|\vec{\omega}^p}\{U(\pi, \mathcal{E}')\}$.

Therefore, the utility of a policy π given the state $\mathcal{E}$ can be computed as follows:

$$U(\pi, \mathcal{E}) = \begin{cases} U^p(Q) & \text{if the negotiation stops,} \\ \mathbb{E}_{\vec{\omega}^r|\vec{\omega}^p}\{U(\pi, \mathcal{E}')\} & \text{otherwise.} \end{cases} \tag{4.9}$$

Now, given the state $\mathcal{E}$, we are looking for the optimal negotiation policy $\pi^* = \arg\max_\pi U(\pi, \mathcal{E})$. Note that when all offers are observed, $U(\pi^*, \langle \overline{S}, S \rangle) = U^p(S)$; otherwise, the agent may choose to propose one or more partial offers $\vec{\omega}^p \in \{\overline{S} \backslash P\}$. Given this, the optimal negotiation policy π^* should consider the following: either stop the negotiation and obtain $U^p(Q)$, or propose a new partial offer $\vec{\omega}^p \in \{\overline{S} \backslash P\}$ which maximises the expected value.

More formally, $U(\pi^*, \mathcal{E})$ must satisfy the following recursive relation:

$$U(\pi^*, \mathcal{E}) = \max \left\{ U^p(Q), \max_{\vec{\omega}^p \in \overline{S} \backslash P} \left\{ \mathbb{E}_{\vec{\omega}^r|\vec{\omega}^p}\{U(\pi, \mathcal{E}')\} \right\} \right\}. \tag{4.10}$$

The relation for π^* given in Equation 4.10 is essentially a Bellman equation which, in principle, could be solved by backward induction. However, even for a moderate-size negotiation space, this approach quickly becomes intractable. Instead, we use a simple index-based alternative method to decide which partial offers to propose and whether to break off the negotiation.

Specifically, the negotiation strategy of the agent can be mapped onto a variant of the so-called Pandora's Problem (Weitzman, 1979): a search problem involving boxes that contain a stochastic reward. As such, each partial offer $\vec{\omega}_p \in \{\overline{S} \backslash P\}$ can be regarded as a *closed* box with stochastic reward $Y_{\vec{\omega_p}}$ that can be opened at cost C^p, while every partial offer $\vec{\omega_p} \in P$ can be represented by an *open* box with a known reward $v(\vec{\omega_r})$ where $\vec{\omega_r}$ is the complete offer observed after proposing $\vec{\omega_p}$. As a consequence of Pandora's Rule (Weitzman, 1979), we can assign an index $z_{\vec{\omega}_p}$ for every partial offer $\vec{\omega}_p \in \{\overline{S} \backslash P\}$,

satisfying:

$$\int_{y=z_{\vec{\omega}_p}}^{1} (y - z_{\vec{\omega}_p})\, \mathrm{d}H_{\vec{\omega}_p}(y) = C^p. \tag{4.11}$$

After identifying the offer $\vec{\omega}_p^* \in \{\overline{S} \backslash P\}$ with the highest index $z^*_{\vec{\omega}_p}$, we apply the following negotiation strategy:

> Selection Rule: *If an offer is proposed, it should be* $\vec{\omega}_p^*$.
>
> Stopping Rule: *Terminate the negotiation whenever the reservation value* r^p *or the highest valuation* $\max_{\vec{\omega}_r \in Q}(v^p(\vec{\omega}_r))$ *exceeds* $z^*_{\vec{\omega}_p}$. *Choose the negotiation outcome as follows:*
>
> - *If* $\max\limits_{\vec{\omega}_r \in Q}(v^p(\vec{\omega}_r)) > z^*_{\vec{\omega}_p}$, *then* $\arg\max\limits_{\vec{\omega}_r \in Q}(v^p(\vec{\omega}_r))$ *is the outcome.*
> - *If* $r^p > z^*_{\vec{\omega}_p}$, *the negotiation ends with no outcome.*

This negotiation strategy completely characterises the optimal policy π^*, as it has been proved to be optimal in terms of maximizing expected utility (Weitzman, 1979). In practice, the effectiveness of the optimal quoting strategy depends on the accuracy of the valuation function and the opponent model. However, with a faithful model, the agent's strategy is optimal in a non-myopic sense: it will negotiate taking into account not only the costs, but also the incremental effect of any subsequent rounds.

4.5 Negotiation of Privacy

In this section, the theoretical framework described in Sections 4.3 and 4.4 is applied to privacy negotiations, where the proposer represents interests of the user and the responder – the service provider. First, the negotiation domain is formally defined. Second, we provide a top-level overview of how the negotiation of privacy permissions is performed with human users in the loop. Finally, since the utility of both agents is similarly based on the valuation function and therefore, on the preference profile, as the first step, we focus specifically on modelling the user's agent, and present two variants of how such preference profile can be created.

4.5.1 Negotiation Domain

In privacy permission management, the negotiation domain may consist of a set of permissions to access user's device resources such as contacts or text messages, as well as other negotiable issues such as access to certain features of the service. For the purpose of our proof of concept, we consider a case where the permissions to access resources are

granted in exchange for a monetary reward. This monetary reward is a total value a user can receive for granting permissions to the specific set of resources, and we refer to it as the *quote*. We assume that the user's agent is in control of the values of issues related to the set of permissions and the service provider's agent controls the value of the quote. For instance, if the permissions grant access to the user's *Contacts*, *Messages*, *Apps List*, *Photo Gallery* and *Browsing History*, an example partial offer proposed by the user's agent may be: (*permission granted*, *no permission*, *permission granted*, *no permission*, *permission granted*, $\perp$). The service provider's agent may complete the offer and respond with: (*permission granted*, *no permission*, *permission granted*, *no permission*, *permission granted*, *£0.55*).

Formally, following the notation from Section 4.3, we define the negotiation domain such that $\omega_i \in \{0, 1\}$ for $i \in \{1, \ldots, m-1\}$ is a permission with binary values 1 and 0 indicating whether the permission is granted or not, and ω_m is a continuous issue representing the quote. This value is normalized such that $\omega_m \in [0, 1]$. Hence, the user's agent proposes a partial offer providing values for $\omega_1, \ldots, \omega_{m-1}$, whereas the service provider's agent responds with a complete offer, including a value for ω_m. This way, the example offer returned by the service provider's agent can be represented as: $(1, 0, 1, 0, 1, 0.55)$.

4.5.2 Overview of the Negotiation

To demonstrate the applicability of the framework, we consider a specific bargaining situation where the user's agent takes the role of the proposer by sending partial offers to the service provider's agent. The offers indicate the permissions the user is willing to grant access to. Next, the service provider's agent returns a complete offer, specifying the quote the service provider can offer for the permissions proposed. Consequently, the user's agent negotiates the best possible offer (given the current user and opponent models) with the service provider's agent on behalf of the human, and then proactively interacts with the human user to establish whether the negotiated offer meets the user's expectations. Although the agent performs the negotiation autonomously, it is the human user who has the final say on whether the permissions are granted. If the user refuses granting the permissions in exchange for the given quote, the agent automatically starts a new negotiation and continues to do so until the user eventually approves a negotiated offer – we refer to this as the negotiation *scenario*.

This way of negotiating privacy permissions provides two benefits: the user has full control over their privacy and the agent can use the feedback for constructing the user's preference profile. However, in order for the agent to accurately represent the user, establishing an accurate preference profile is crucial. This is because the weights in the valuation function depend on the data present in the preference profile (see Equation 4.4). To this end, in this chapter, we compare two ways of using users' feedback to

derive the user's valuation. Since they are two different variants of the user's agent, we refer to them as Agent 1 and Agent 2.

4.5.3 Agent 1: Individual Preferences

Agent 1 recomputes the weights after each negotiation based on an individual user's decisions in the previous negotiations. At the beginning, to deal with the "cold-start" problem (i.e. where no information is available about the past negotiations), the agent applies the weight derivation method on data collected from other users' negotiations. In more detail, data from previous user studies is aggregated to generate a set of constraints according to Inequality 4.6 which are, in turn, used to derive an initial set of weights through Equation 4.8. Thus, when no information about an individual user is known, the initial weights w_i are set the same for all users.

Then, whenever a user rejects an offer $\vec{\omega}$ in favour of accepting $\vec{\omega}'$, a new inequality of the form of Inequality 4.6 is added to the set of already existing inequalities. However, if there is a conflict between constraints, the user's own constraints are always prioritised over the ones from the manual negotiation trials. Then the new weights are again derived and updated using Equation 4.8 for use in the next negotiation.

4.5.4 Agent 2: Type-Based Preferences

Agent 2 is the agent previously published by (Baarslag et al., 2017). It assigns users into categories before the start of their negotiations. We define these categories as follows:

- *Fundamentalists* – users generally granting less than 33% permissions.
- *Pragmatists* – users generally granting between 33% and 66% permissions.
- *Unconcerned* – users generally to granting more than 66% permissions.

Initially, the preferences of a newly added user are based on the negotiation data from users previously classified into the same category. In more detail, the agent applies the weight derivation method on negotiation data from the previous studies where participants were categorised in the same way to determine the weights offline.

After each negotiation, the cluster classification is updated according to the percentage of permissions that were actually granted. Specifically, if it is less than 33%, they are re-classified as *Fundamentalists*, between 33% and 66% – *Pragmatists*, and *Unconcerned* otherwise. This way, even though the weights are learnt offline, the classification of the user is performed online.

4.6 Experimental Evaluation

To evaluate the suitability of the agent-based negotiation framework for consent management, we conduct a user study with participants negotiating consent to their actual personal data stored on mobile phones in exchange for points converted to a monetary value at the end of the experiment. In this section, we provide a detailed description of the assumptions made about the service provider's agent, the apparatus, the methodology used and the findings of our experimental evaluation. In particular, we compare the negotiation approach to the prevailing take-it-or-leave-it approach, which we use as a benchmark for our experiments. The studies have been approved by the Faculty Ethics Committee (ref.: ERGO/FPSE/18082; ERGO/FEPS/58090).

4.6.1 Offer Completion

In this chapter, in order to conduct the experimental comparison of the user's agent, we abstract away from designing the actual strategy of the service provider's agent and instead, we design a stochastic way of completing the partial offers. In more detail, we first assume that the order in which partial offers are made does not impact the complete offers, i.e. the complete offers are not different depending on the negotiation round. However, we assume that the complete offers depend on the number of granted permissions $N = \sum_{i=1}^{m-1} \omega_i$. That is, the agent completes every partial offer with a uniformly random quote $\omega_m \in [\max(0, N - 1/m),\ N/m)$. Using the previous example, if the partial offer received from the user's agent is: $(1, 0, 1, 0, 1, \perp)$, then some of the possible complete offers are: $(1, 0, 1, 0, 1, 0.4)$, $(1, 0, 1, 0, 1, 0.45)$ and $(1, 0, 1, 0, 1, 0.55)$. This approach of completing the offers ensures that granting more permissions results in a higher quote, while different combinations of permissions result in a different (not necessarily linearly additive) quotes. This reflects situations where some data types could complement each other (i.e. when a specific combination allows the service provider to derive more relevant information) or substitute each other (in which case the added data provides little additional benefit to the service provider).

4.6.2 Apparatus

In our experiments, we use a tool developed and previously used in user studies by Baarslag et al. (2017) which allows participants to negotiate combinations of selected permissions on their smartphones in exchange for points, which map directly to a monetary reward. In the setting screen, the user can see which permissions are going to be granted, the number of points they receive as a reward, and the total number of points collected. The more data a user shares, the more points they receive. In this section, we describe the tool and the ways users interact with it.

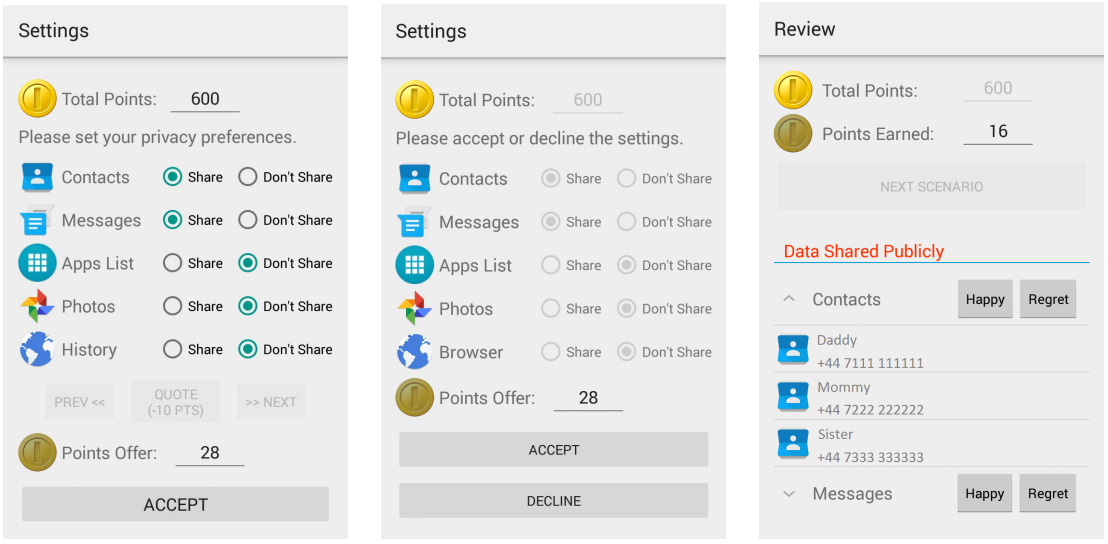
In order to avoid any personal preference towards the service and to focus on users' permission preferences, the tool we developed has no intrinsic functionality other than capturing and displaying selected data points, and the testbed for negotiation. The advantage of using a mobile app as our experimental environment is that it allows us to use participants' own data that they personally care about in a real, privacy-sensitive situation. This level of realism is particularly important in privacy-related experiments where past studies (Taddicken, 2014; Barth and De Jong, 2017) have consistently shown a discrepancy between a person's stated preferences from surveys and actual disclosure decisions (i.e. the so-called privacy paradox).

The app requests access permissions to the following data:

- the list of *contacts*,
- the text *messages*,
- the list of installed *apps*,
- the gallery of *photos*, and
- the browsing *history*.

The reason why the above permissions were selected is that they are among the most often used, as ranked by (Liccardi et al., 2014), and can be acquired, mined from users' smartphones and quantified. It is important to note that, although we decided to use the mobile app as our experimental environment, the aim is to test the negotiation framework in practice and to compare different privacy preference learning approaches. We present two designs of this interface: the negotiation screen allows the user to request a new negotiation of the permissions and the "Take It Or Leave It" screen that represents the *take-it-or-leave-it* approach that is adopted by most services requiring users to share data.

In the negotiation screen, the user is presented with the outcome of the negotiation. When the negotiation between the user's agent and the service provider's agent ends, the agreed permission settings (*Share* or *Don't Share*) as well as the agreed quote (a number of points received in exchange) are displayed. If the user is happy with the offer, they can press the *Accept* button. In that case, the offered number of points is added to their total budget. Otherwise, the user can communicate their permission preferences by selecting *Share* or *Don't Share* to grant or refuse granting a permission. An example configuration is presented in Figure 4.3(a), where a user is offered 28 points for access to their contacts and messages only, but can change these settings. By pressing the *Quote* button, the user can then request a new quote. To prevent the user from constantly doing so, 10 points are subtracted from their accumulated budget every time they request the offer to be renegotiated. The user can also freely switch between the



(a) The negotiation screen. (b) Take-it-or-leave-it screen. (c) The review screen. screen.

FIGURE 4.3: The interface of the experimental tool.

received offers using the *Prev* and *Next* buttons. Once they are happy with one of the offers, they can press *Accept* to approve it.

The *take-it-or-leave-it* approach represents the situation where the data sharing terms lack any kind of tunable control over the privacy trade-offs. For example, prior to Android 6.0 users were required to accept all data access permissions requested by mobile apps in order to proceed with the installation on their smartphones. Moreover, the take-it-or-leave-it approach is typical in many exchanges and not limited to apps. This approach is reflected in the differences between the negotiation screen and the "take it or leave it" one. Specifically, the "take it or leave it" screen does not allow the user to modify the permissions. The user can only accept or decline an offer by pressing the *Accept* or *Decline* buttons accordingly. The *Decline* option is equivalent to selecting *Don't Share* for all data types. For example, in Figure 4.3(b) the user is able to accept or decline access to contacts and messages in return for 28 points.

Following each sharing decision (through either negotiation or *take-it-or-leave-it*), the app randomly collects three data points of each shared permission type from the user's device. The user is then presented with a review screen showing those data points (which the users are made to believe is made public) and asked to retrospectively express whether they are *Happy* about or *Regret* granting the permission. Figure 4.3(c) presents an example review screen displaying three sample contacts and messages of the user.

The purpose of this screen is to collect data on regret of granting permissions, which is one of the success measures we use. This allows us to assess whether retrospectively the agent has made the right decision and how this compares to the *take-it-or-leave-it* approach. Note that users cannot revoke their decision at this point.

TABLE 4.1: A summary of the treatments used in the experimental evaluation.

Treatment	Summary
TIOLI	*Take-it-or-leave-it* (no agent)
MN	Manual negotiation (no agent)
A1	Agent 1 negotiating
A2	Agent 2 negotiating

4.6.3 Methodology

Experiments were conducted through lab studies to allow for more control of the conditions, and using the participants' own mobile phones so that they truly care about the data being shared. In this section, we report on the experimental design, the procedure of the lab study and the participant recruitment.

In order to avoid bias caused by the learning effect, we decided to employ the between-subject experimental design with four treatments summarised in Table 4.1. We aimed to test if the proposed agent designs can facilitate data sharing better than manual negotiation and to evaluate the performance of Agents 1 and 2, comparing them to the manual negotiation and to each other. Our treatment where Agent 1 was used is referred to as treatment Agent 1 (A1). The treatment where secondary data from the agent experiments of Baarslag et al. (2017) was re-analysed is referred to as treatment Agent 2 (A2).

To learn how a human negotiates without the help of an agent, in addition to the treatments with the agents, we introduced two control treatments: Take-It-Or-Leave-It (TIOLI) and Manual Negotiation (MN). In both of these treatments, the final offers that participants could see on the screen were pre-defined beforehand through a uniformly random selection. Specifically, we used the dataset published by Baarslag et al. (2017) at `http://doi.org/10.5258/SOTON/405394`. The difference between the two treatments was in what happened after the offer was presented to the user:

- in the TIOLI treatment, the *take-it-or-leave-it* screen was used – the users were able to either *Accept* or *Reject* the offer,
- in the MN treatment, the negotiation screen was used – participants were able to interact with the negotiation screen after seeing the offer in the same way as in A1 and A2.

The lab study procedure involved four parts: the initial survey, the main experiment, a post-study questionnaire and a debrief. Table 4.2 summarises the user study procedure as it was experienced by the participant. At the beginning, participants were asked to download the mobile app from the Play Store. They entered their demographics (age,

Table 4.2: A summary of the user study procedure as experienced by a participant.

Part of experiment	Summary
Initial survey	Demographics (age, gender, nationality, university course) Westin's Privacy Segmentation Index
Main experiment	"Take it or leave it" or negotiation screen + review screen repeated eight times
Post-study questionnaire	Data sharing sensitivity survey NASA-TLX

gender, nationality, university course) into the app, and completed Westin's Privacy Segmentation Index survey (Kumaraguru and Cranor, 2005). Westin's Privacy Segmentation Index is a widely used tool for measuring privacy attitudes (Woodruff et al., 2014) and categorising individuals into three privacy types: Fundamentalists, Pragmatists, and Unconcerned. As part of the Index, they indicated on a 4-point Likert scale (1 – strongly disagree, 4 – strongly agree) the extent to which they agreed with the following statements:

1. Consumers have lost all control over how personal information is collected and used by companies.
2. Most businesses handle the personal information they collect about consumers in a proper and confidential way.
3. Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.

They were informed that their monetary reward will be based on the total number of points earned and (to elicit a genuine response) that any data shared during the experiment will be made available on a public website. They were not informed about the agent negotiation taking place in this experiment.

After reading the app manual, participants were asked to interact with either the negotiation or take-it-or-leave-it screen, depending on the treatment. In order to control the conditions, the offers for treatments with no agent negotiation (TIOLI and MN) were pre-defined using random sampling prior to the experiment. Once a participant accepted or declined an offer, they proceeded to the review screen. As the experiment continued, these interactions were repeated such that each participant engaged in eight negotiation scenarios in total. To explore varying reward levels, those interactions differed in the maximum possible number of points to be gained by a participant: 25, 50 or 100. To cancel out possible interaction effects, this maximum reward was set to 50 in the first and last interaction for all participants, and a balanced Latin square design was used to determine the order of maximum rewards in all others. Consequently, the

TABLE 4.3: The number of participants of each privacy type per treatment.

Privacy Type / Treatment	**TIOLI**	**MN**	**A1**	**A2**	**Total**
Fundamentalists	5	5	7	7	18.18%
Pragmatists	26	26	23	24	75%
Unconcerned	2	2	3	2	6.82%

number of points a participant gained during a single negotiation or a take-it-or-leave it interaction was the maximum reward level multiplied by the negotiated quote ω_m in the given scenario. In the end, the total number of points collected by a participant was the sum of points collected in all scenarios.

After that, they completed a questionnaire about their data sharing sensitivity. Specifically, they were asked to rate the following statements on a 7-point Likert scale (1 – strongly disagree, 7 – strongly agree):

1. I am sensitive about sharing the contacts stored on my phone.
2. I am sensitive about sharing the text messages stored on my phone.
3. I am sensitive about sharing the apps stored on my phone.
4. I am sensitive about sharing the photos stored on my phone.
5. I am sensitive about sharing the browsing history stored on my phone.

Additionally, participants were asked to complete NASA Task Load Index (NASA-TLX), which is a widely used, multidimensional assessment tool that rates perceived workload (Hart and Staveland, 1988). As part of this survey, they were asked to rate the effort they had to put into this experiment on a 20-point Likert scale.

Finally, the participants were debriefed about the purpose of the study in more detail and informed that their data was never made publicly available on any website, despite the initial claim. All participants received a cash payment of between £5 and £10, directly depending on the number of points accumulated during the experiment, regardless of treatment allocation, e.g. if they collected 658 points, they received £6.60; if they collected less than 500 points, they received £5.

4.6.4 Participants

For treatments MN and A1, we recruited 66 participants from the University of Southampton. At the recruitment stage, they were informed that, during the experiment, they will be asked to download an Android application and make privacy-related decisions

to earn between £5 and £10. Additionally, we use secondary data from the experiment conducted by Baarslag et al. (2017) collected from another 66 participants.

In total, our sample consisted of 132 participants. The participants were undergraduate, Master's or Ph.D. students from a variety of disciplines (e.g. Engineering, Medicine, Law). Since university students typically have a good level of digital literacy and a variety of attitudes towards privacy, such sample was suitable for the purpose of evaluating agent-based negotiation. 37.12% of them identified as women and 62.88% identified as men. 45.45% of the sample was British; others were nationals of 32 different countries such as Romania (7.58%), Malaysia (6.06%) and India (5.3%). Their age ranged from 17 to 43 (mean: 21.69, median: 21, st. dev.: 3.78). The participant poll consisted of 18.18% Fundamentalists, 75% Pragmatists and 6.82% Unconcerned (as defined by Westin's Privacy Segmentation Index) which is broadly consistent with the overall American population (Kumaraguru and Cranor, 2005).

The participants were randomly allocated into the treatments, i.e. there were 33 participants in each treatment. As in the experiment conducted by Baarslag et al. (2017), the allocations were performed such that any differences in privacy attitudes between the treatments were non-significant. To illustrate these differences, Table 4.3 presents the number of participants of each privacy type per treatment.

4.6.5 Results

In this section, we present the results of our data analysis. In particular, we report on the impact of automated negotiation on data sharing and user's post-sharing regret, aligning users' decisions with their self-reported data sharing sensitivity, the effort required from the users and the accuracy of the proposed agent variants.

One aim of our research was to investigate how agent negotiation may influence users' data sharing behavior. The results, based on participants' own private data show that, on average, participants allowed access to data of the five data types over 2.5 times more often when they were able to negotiate. Figure 4.4 shows the percentages of how many times the participants allowed access to each of the data types in all scenarios. In particular, participants in treatment MN decided to share their list of installed applications 3.5 times more often than those in treatment TIOLI, and those in treatments A1 and A2 – nearly four times more often. The messages stored on the participants' mobile phones in treatment MN were shared almost twice more often than in treatment TIOLI, and in treatments A1 and A2 – nearly three times more often. These findings suggest that negotiation leads to a win-win situation, for both the participants: the user received higher payoffs from sharing more data, and our hypothetical service provider received more data from them.

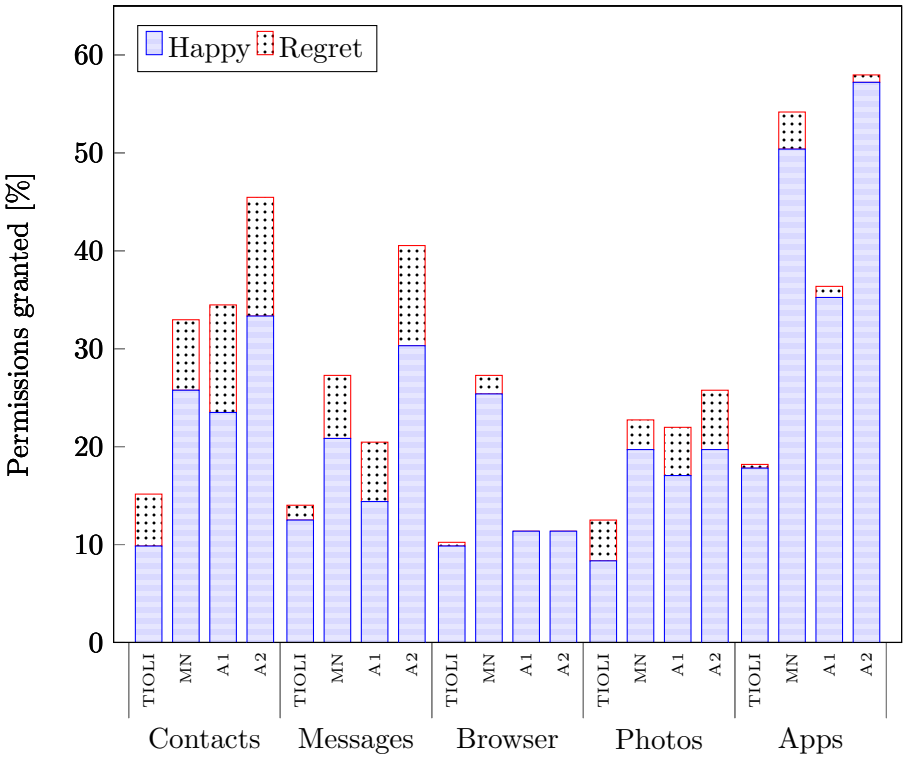


FIGURE 4.4: Number of times the participants granted permissions to each of the resources.

Based on the results, we can see that people are happily sharing certain kinds of data, and regret having shared others. However, the regret rate does not change when the negotiation agent is introduced. On average, the participants expressed regret allowing access to their data in 15.96% cases. This is consistent with findings from related work where in 10% of decisions, users were granting permissions reluctantly (Bonné et al., 2017).

Figure 4.4 illustrates how many times the participants were happy having shared their data of each type and how many times they regretted their decisions. Most often that was access to their contacts (27.78%), least often – their list of installed applications (3.55%). Nonetheless, there were no significant differences between the regret rates in the treatments. We consider this a positive outcome for the potential of automated negotiation in this area. That is, even though users grant access to their data more often when a negotiation agent is involved, there is no significant increase in regret.

Our results show that, when users are allowed to negotiate, not only does their sharing behavior change radically, but their choices also better reflect their privacy preferences. For each treatment, Figure 4.5 shows the means of the users' self-reported data sharing sensitivity for granting each permission on 7-point Likert scale (see Section 4.6.3 for details). As expected, there are no statistical differences found between the treatments.

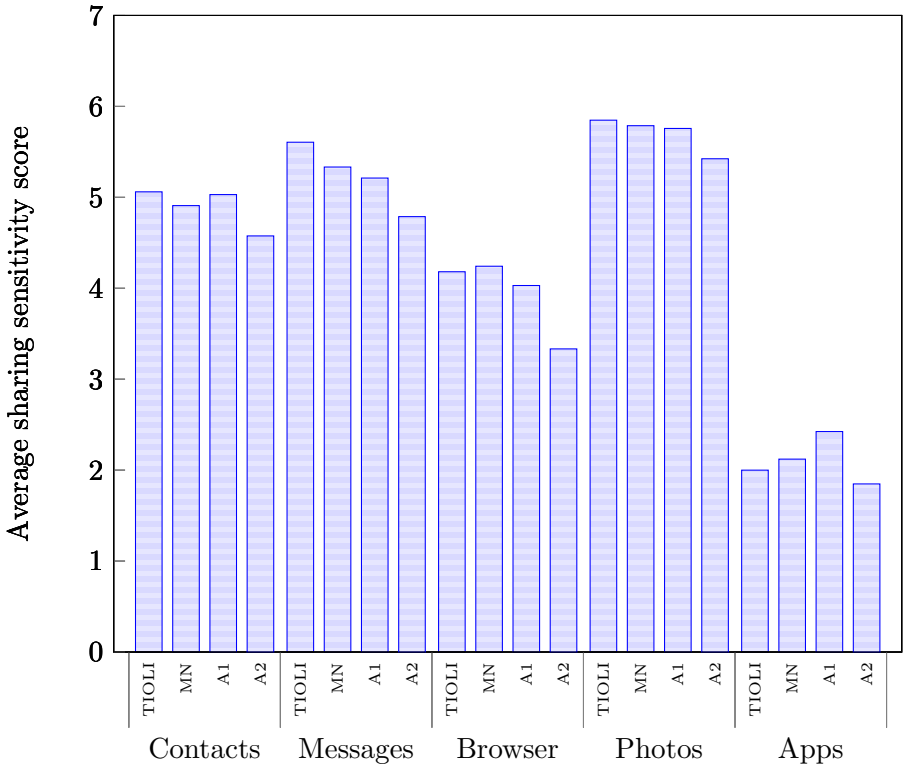


FIGURE 4.5: Means of self-reported data sensitivity scores on a 7-point Likert scale.

When we compare the reported sensitivity of each data type with users' actual sharing decisions (also illustrated in Figure 4.5), we observe a more marked correspondence for the negotiation settings. Although the average scores hide a one-to-one correspondence between sensitivity scores and sharing actions, we can clearly observe how often-shared data in these three treatments correspond with permissions of markedly lower sensitivity, such as apps, while access to photos and messages, which are both highly sensitive, is permitted far more sporadically than other permissions. The only outlier in this order seems to be the browsing history, which we believe is due to the fact that a number of participants did not have any browsing history available, independent of the sensitivity.

The last aim of our research was to examine the accuracy of the proposed negotiation agents. To do so, we calculated the accuracy by comparing the number of changes that the users made to the negotiated outcome.

Specifically, we define the accuracy as the difference between the accepted offer and the offer negotiated by the agent. Building on the notation from Section 4.5.1, if the initial offer is $\vec{\omega^1} = (\omega_1^1, \ldots, \omega_{m-1}^1, \omega_m^1)$ and the final offer is $\vec{\omega^k} = (\omega_1^k, \ldots, \omega_{m-1}^k, \omega_m^k)$, the difference between them, δ, is calculated as:

$$\delta = \sum_{i=1}^{m-1} |\omega_i^k - \omega_i^1|$$

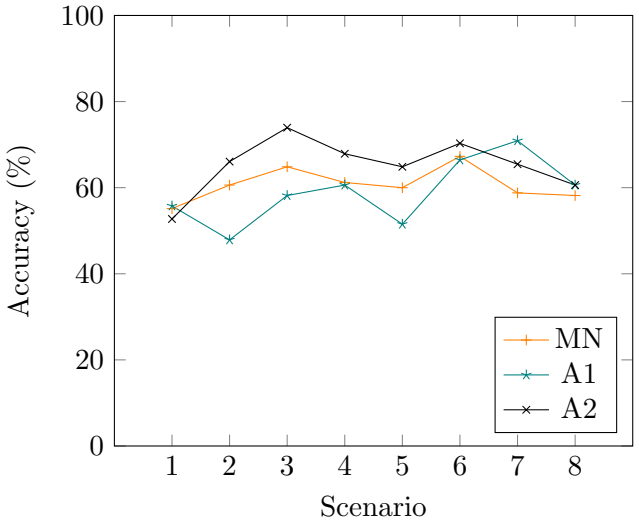


FIGURE 4.6: The accuracy of Agent 1 (A1), Agent 2 (A2) and the manual negotiation (MN) in each scenario.

where the first $m-1$ issues are the permissions (i.e., contacts, messages, browsing history, photos and list of applications installed) and the m^{th} issue is the number of points received in exchange for them, which is ignored here since this is set by the service provider and not the user. Hence, the *accuracy* of the negotiated outcome is calculated as:

$$1 - \frac{\delta}{m-1}$$

Figure 4.6 presents the accuracy of Agent 1, Agent 2 and the manual negotiation in each scenario. Results show that, in all scenarios except the first one (when the agent is still relying on Westin's Privacy Segmentation Index categorisation), the users made the least changes to the default settings when Agent 2 was negotiating on their behalf. On average, offers proposed by Agent 2 were the most accurate (65.23%). After the first scenario, the choices of Agent 2 for the default settings were more accurate at accommodating the users' privacy preferences than the manual negotiation.

Although, on average, the accuracy of Agent 1 (58.86%) is lower than the accuracy of manual negotiation (60.76%), overall, we can observe a rising trend of the accuracy of Agent 1. In particular, in the penultimate scenario, it exceeds the accuracy of both Agent 2 and the default settings. This suggests that, with more learning, the individual approach could eventually outperform an agent which bases the preferences on a limited number of profiles. This opens up a number of potential lines of further investigation.

Our experimental setup, in which participants were not aware of the agent, allows us to be confident that this accuracy is the result of correctly predicting preference, rather than a tendency of participants to "go along with" suggestions that they know are made by an agent. Although we detect some bias resulting from the defaults, this is apparent in all three conditions – since defaults in the first scenario were set randomly, we expected

Table 4.4: Mean, median and standard deviation of the perceived effort of negotiation in each treatment on a scale from 0 to 20.

Effort / Treatment	**TIOLI**	**MN**	**A1**	**A2**
Mean	4.19	5.33	5.88	3.55
St. dev.	4.09	3.54	3.82	2.88
Median	3	5	5	3

them to be aligned with user preferences 50% of the time; in fact, the slightly higher percentage in all treatments (MN: 55.15%; A1: 55.76%; A2: 52.73%) show that the defaults exert some influence and act as a means to promote exploration of the different options.

Lastly, we measured the perceived effort of negotiation via the NASA-TLX questionnaire. Table 4.4 presents the mean, median and standard deviation of the results in each treatment. We can see that the effort required from the user supported by Agent 2 is not only less than in the *take-it-or-leave-it* approach but also during manual negotiation. This finding shows a potential for automated negotiation to be less demanding than the manual negotiation.

4.7 Discussion

In this chapter, we explore research question **RQ2** on privacy negotiations that automatically take into account the user's individual preferences. To this end, we propose a novel multi-issue negotiation framework, where two agents exchange partial and complete offers, bargaining over a number of issues in bundle.

An advantage of this approach is that it prevents competitive, zero-sum negotiations on isolated issues and, instead, promotes mutually beneficial deals. Moreover, the protocol allows users to focus on issues that are important to them and leave out issues for which users find it difficult to determine a precise value and are more naturally determined by the counter party. This is especially important in negotiating privacy permissions, because the benefits of privacy protection are often uncertain and intangible (Acquisti and Grossklags, 2005b) and, as a result, users find it difficult to express the exact willingness to pay for revealing certain information. It is easier for users to decide, through relative comparisons, which of the complete offers they prefer in order to assess the value of protecting their privacy (Tsai et al., 2011). In this way, users can easily explore the set of possible agreements, while the service provider, provided with information about monetizing the data (e.g. through advertising), has the ability to exercise the final say. However, such negotiations often occur in practice in a number of settings not necessarily limited to permissions management. For example, when negotiating insurance policies,

buyers often specify certain conditions for the extent of cover, for which the seller completes the possible contract by proposing a price. Other examples include negotiating mortgages and broadband packages. For this reason, we believe that this contribution generalises to other negotiation domains with similar individual-vendor relationships.

Furthermore, we demonstrate the applicability of the framework in a specific bargaining situation, which is the negotiation of permissions between a user and a service provider. In doing so, we assume that it is the user's agent who starts the negotiation, specifying the requirements for a subset of important issues. The service provider's agent is then able to submit a complete counter-offer based on the proposed offer. The user's agent can either accept the complete offer or submit a new partial one. It can also break off the negotiation. Although this ensures that the service provider's agent can always 'price out' any undesired partial offers, an equally valid use of this framework could be where the negotiation is initiated by the service provider's agent. Then, the user's agent could specify the conditions, under which they would agree to the service provider's conditions or 'price out' any undesired proposals. Future work on this topic should investigate how this setting impacts user's data sharing, regret, preferences and effort, comparing to the setting we used in this chapter.

In addition, we show how the framework can be used in practical context through a user study with human participants and their private data. The findings from our user study and secondary data analysis suggest that users can be incentivised to share much more data when they are able to negotiate, with no increase in regret about their decisions. We also show that negotiation enables users to align their privacy choices more closely with their preferences. In particular, we found that the deals negotiated by the agents are more accurate than the baseline in that the resulting agreements are better aligned with the user's actual preferences. These outcomes suggest that negotiation is a powerful interaction mechanism for achieving mutually beneficial data sharing agreements.

Moreover, we compare two variants of the user agent, which differ in the way the user's preferences are elicited. Our results indicate that, with limited data, a profile-based variant might be better where users can be categorised into a limited number of types, but with more interactions, this can be further personalised. This is an important insight towards answering **RQ2**, because suggests that the profile-based variant could effectively used at the beginning when the number of previous negotiations is very small. However, as the number of negotiations increases, the agent can be more effective personalising offers based on s single user's negotiation data. Alternatively, the negotiation could benefit from a variant that incorporates the two approaches at the same time, potentially supported by the use of machine learning techniques. Future work should further explore different options of implicit preference elicitation.

Lastly, while the proposed variants of the user's agent are both quite general, our experimental setting is limited to reasoning about data types. Future work should consider

other factors beyond data types, for example, the recipient, retention period, purpose, quality, and privacy risks. In addition, we noticed that, when expressing regret, users often did so for specific data points (e.g. specific contacts or photos). Therefore, it is clear that a model based on permissions alone is too coarse to accurately capture the privacy preferences. Combining a semi-autonomous agent with a more meaningful classification of data (perhaps using signals such as location, time of the day and relations to other users) is another avenue that warrants further exploration. In our future work, we plan to study various designs of privacy agents that can learn to negotiate on users' behalf, and engage users directly only when in doubt. Additionally, the agent's user model was based on types derived from results of a short survey before the negotiations started. A more personalised model, derived from e.g. apps installed on the phone and other factors, could increase the accuracy of the deal negotiated by the agent even further.

4.8 Limitations

As the main limitation of our study we consider the small number of participants in each treatment and, therefore, a limited amount of data. Although this sample size allowed us to proof the concept and make conclusions from our observations, it is hard to generalise findings based on a study of data from only 132 participants.

Another limitation of the study was that, instead of trading their personal data for access to the service, the participants were receiving cash payments. Whereas in our study participants were told that their data will be posted on a public website, in a real-life scenario, personal information shared with the service provider might have been monetized to generate profit, e.g. by selling the data.

4.9 Summary

In this chapter, we explore multi-issue negotiation as a mechanism that can make consent *negotiable*. Specifically, we presented a novel multi-issue negotiation framework with a new variant of the alternating-offers protocol, based on exchanges of partial and complete counter-offers. Moreover, we demonstrate how this framework can be used in bilateral negotiations of privacy permissions between users and service providers.

The results of our evaluation provide evidence that agents are able to automatically negotiate consent on behalf of users. In fact, we find that users decide to consent 2.5 times more often when they are able to negotiate while maintaining the same level of decision regret. Moreover, we observe that negotiation can be less mentally demanding than the take-it-or-leave-it approach and that it enables people to better align their privacy choices with their actual preferences. We discuss our findings, which point

to several avenues of future work on automated and negotiable privacy management mechanisms.

In order for an agent to faithfully represent the user, we compare two approaches to implicit preference elicitation: one approach personalised to each individual user and one personalised depending on the user's privacy profile classification. Furthermore, we find that the latter agent variant performed better in the initial negotiation rounds. However, there is potential that the former variant results in more accurate offers in the further rounds.

This work sits within the wider agenda of privacy engineering that has received renewed momentum with the introduction of the GDPR, requiring greater transparency and user empowerment, and with opportunities for multi-agent systems to provide technological solutions. Our ultimate aim is to enable automated negotiation of consent, where both the agent and the service provider are represented by autonomous agents. The work described in this chapter is an essential step towards addressing this broader vision.

Chapter 5

Implementable Consent

After a privacy agreement is negotiated and consent is granted, the agreement must be stored and honoured. With regard to that, related work presented in Chapter 2 suggests several ways of representing the privacy agreement and implementing it in all aspects of data processing. For example, the creators of P3P proposed that data use practices are described as machine-readable statements and the user expresses their privacy preferences as a set of rules on those statements. However, the implementation of P3P turned out challenging for usability reasons: data practices were complex and users had little experience expressing their privacy preferences. As the scale of data processing is rapidly increasing and, thus, becoming even more complex, there is a need for ways of expressing privacy agreements that can make consent *implementable*.

In this chapter, we take the first step towards addressing this gap. In particular, we focus on the problem of identifying consented data processing based on the privacy agreement between the service provider and an individual user. More specifically, if the privacy agreement constrains the flow of information that is processed, then by applying privacy agreement constraints on the data processing model, we can establish which parts of the data processing the user has consented to.

Consequently, we propose a new data processing model and formalise a problem of applying such constraints on it. Moreover, we formally prove that the problem in general is $\mathcal{NP}$-hard. However, if we are able to make certain assumptions about the valuation of the data by the service provider, we are able to provide algorithms that can effectively find a representation of the consented data processing practices, given the service provider's data processing model and the user's constraints. Our contribution in this chapter includes the comparison of the algorithms in terms of the runtime and accuracy.

The remainder of this chapter is structured as follows. Firstly, in Section 5.1, we introduce the theoretical concepts our research is based on: the computational complexity theory, graph theory and network flow theory. Secondly, in Section 5.2, we discuss the

motivation of this study and present an example scenario where our solution finds a use case. Then, in Section 5.3, we present our new data processing model and state the problem of identifying the consented data processing given the user's consent constraints. Next, in Section 5.4, we present theoretical results regarding the complexity of the problem. Subsequently, in Section 5.5, we present a simplified instance of the problem where the utility is linearly additive. In Section 5.6, we propose five algorithms that can help service providers to implement the privacy constraints into data workflows. Then, we analyse the (non-)optimality of these algorithms with respect to their computational complexity in Section 5.7. In Section 5.8, we report on the results of our experimental comparison of the proposed algorithms. In Section 5.9, we discuss the broader context of our findings in light of privacy in algorithmic data processing. In Section 5.10, we list the limitations of our research. Finally, in Section 5.11, we provide a brief summary of the work presented in this chapter.

5.1 Preliminaries

In this section, we introduce terminology from the computational complexity theory and graph theory used in this chapter. We also give a brief introduction to the classic graph-cutting problems.

5.1.1 Computational Complexity Theory

Computational complexity theory focuses on classifying computational problems into classes according to their inherent difficulty, and relating these classes to each other. Such problems can be expressed as relations $P \subseteq I \times \mathcal{S}$, where I is the set of problem instances and $\mathcal{S}$ is the set of problem solutions (Ausiello et al., 2012). Especially, if P reduces to a function $f : I \to \mathcal{S}$, where $\mathcal{S}$ is the binary set $\mathcal{S} = \{\text{YES}, \text{NO}\}$ (or $\mathcal{S} = \{0, 1\}$), the problem is denoted as a *decision problem* (Ausiello et al., 2012). Differently, if given an instance $x \in I$, the problem asks for the "best" solution y^* among all solutions $y \in \mathcal{S}$ such that $(x, y) \in P$ is verified, a problem of this kind is called an *optimisation problem* (Ausiello et al., 2012).

In particular, computational complexity theory is the study of how much computational resources are required to solve a given problem (Arora and Barak, 2009). To this end, a *complexity class* is a set of problems that can be solved within a given resource (Arora and Barak, 2009). The most fundamental examples of complexity classes are **P** and **NP**. While **P** is the class of decision problems that can be efficiently solved, **NP** is the class of decision problems whose solutions can be efficiently verified (Arora and Barak, 2009).

Some problems, which we call **NP**-hard, have the property of being at least as hard as the hardest problems in **NP** (Arora and Barak, 2009). This property of being "at least as hard as" is formalised through the notion of reduction (Kleinberg and Tardos, 2006):

Definition 5.1. *Polynomial-time reduction.* If arbitrary instances of problem Y can be solved using a polynomial number of computational steps, plus a polynomial number of calls to an algorithm that solves problem X, then Y is *polynomial-time reducible* to X, denoted as $Y \leqslant_p X$.

Formally, **NP**-hardness is defined as follows:

Definition 5.2. *NP-hardness.* A problem H is **NP**-hard if for every problem $L \in \mathbf{NP}$ there is a polynomial-time reduction from L to H.

Markedly, if an **NP**-hard problem can be solved using a polynomial number of computational steps, i.e. it has a polynomial-time algorithm, then so do all the problems in **NP**. However, assuming that $\mathbf{P} \neq \mathbf{NP}$, the fact that a problem is **NP**-hard can be viewed as evidence that it cannot be solved in polynomial time (Arora and Barak, 2009).

Note that although **NP** is a class of decision problems, the class **NP**-hard is not restricted to decision problems. In particular, there are several graph-theoretic optimisation problems that are **NP**-hard.

5.1.2 Graph Theory

Graph theory is the study of graphs, which are mathematical structures used to model pairwise relations between objects (Bagdasar, 2013). Formally, a *graph* $G = (V, E)$ consists of a set V of objects called *vertices* and another set E of objects called *edges*, such that each edge $e \in E$ is identified with an unordered pair (v_i, v_j) of vertices $v_i, v_j \in V$ (Diestel, 2000). A specific graph where $V = \varnothing$ and $E = \varnothing$ is called the *empty graph* (Diestel, 2000). Furthermore, any graph $G' = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E$ is a *subgraph* of G (Diestel, 2000).

A special kind of a graph is the *directed graph*, where the pair (v_i, v_j) is ordered (Diestel, 2000). In such a graph, there are two maps: $init : E \to V$ and $ter : E \to V$ (Diestel, 2000). If $e \in E$ is identified with the pair (v_i, v_j) (denoted as $e = (v_i, v_j)$), then $init(e) = v_i$ and $ter(e) = v_j$. For any vertex $v \in V$, if $init(e) = v$, then e is an *outgoing edge* of v, and if $ter(e) = v$, then e is an *incoming edge* of v. We define functions $in : V \to \mathcal{P}(E)$ and $out : V \to \mathcal{P}(E)$ to denote the sets of all incoming and outgoing edges a vertex. That is, if $v = ter(e)$, then $e \in in(v)$, and accordingly, if $v = init(e)$, then $e \in out(v)$. Moreover, the number of incoming edges at v is the *in-degree* of v, denoted as $deg^+(v)$. Consequently, the number of outgoing edges at v is the *out-degree* of v, denoted as $deg^-(v)$. In more detail, $deg^+(v) = |in(v)|$ and $deg^-(v) = |out(v)|$.

In a graph, a *path* is a non-empty subgraph $P = (V_P, E_P)$ of the form $V_P = \{v_1, v_2, \ldots, v_k\}$, $E_P = \{(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k)\}$, where all v_i are distinct for $i \in \{1, \ldots, k\}$ (Diestel, 2000). We say that vertices v_i and v_j are *linked* by P if $v_i \in V_P$ and $v_j \in V_P$. A non-empty graph G is called *connected* if any two of its vertices are linked by a path in G (Diestel, 2000). In addition, a vertex $v_j \in V$ is *reachable* from a vertex $v_i \in V$ if there exists a path in G such that $v_1 = v_i$ and $v_k = v_j$. If $k \geqslant 3$, then the graph $C = (V_P, E_P \cup \{(v_k, v_1)\})$ is called a *cycle* (Diestel, 2000). Thus, a directed graph that does not contain any cycles is a *directed acyclic graph* (DAG).

Another special kind of a graph is the *weighted graph* (also called a *network*), in which a number called its *weight*, is assigned to each edge $e \in E$ (Fletcher et al., 1991). That is, in a weighted graph, there exists a function $w : E \to \mathbb{R}$. For a vertex $v \in V$ in such a graph, we define the *weighted in-degree* as the sum of the weights of all incoming edges of v:

$$wdeg^{+}(v) = \sum_{e \in in(v)} w(e).$$

Consequently, the *weighted out-degree* of v is the sum of the weights of all outgoing edges of v. We defined it as follows:

$$wdeg^{-}(v) = \sum_{e \in out(v)} w(e).$$

5.1.3 Graph-cutting Problems

From set theory, a set $S = \{S_1, \ldots, S_k\}$ of disjoint subsets of a set S is called a *partition* of S if $S = \bigcup_{i=1}^{k} S_i$ and $S_i \neq \varnothing$ for every i (Diestel, 2000). Then, if $\{V_1, V_2\}$ is a partition of the set of vertices V, the set of all edges $(v_1, v_2) \in E$ such that $v_1 \in V_1, v_2 \in V_2$ is called a *cut*. Specifically, when there is a pair of vertices $s \in V_1$ and $t \in V_2$, we can say that the cut is an *s-t cut* for (s, t). The vertex s is then a *source* and t – the *sink*.

While there can be many s-t cuts that partition a graph such that s and t belong to different partitions, a classic graph-cutting problem is to find the s-t cut where the sum of weights of the included edges is minimal. More formally, the so-called Minimum s-t Cut Problem (MinCut) is defined as follows:

Definition 5.3. Minimum s-t Cut Problem (MinCut)

Instance: a graph $G = (V, E)$, a weight function $w : E \to \mathbb{N}^*$ and a pair (source s, sink t) of terminal vertices of G.

Question: find a set of edges of G, E_{MinCut} whose removal leaves no directed path from s to t such that:

$$E_{MinCut} = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.1}$$

The method of solving the MinCut problem was first proposed by Ford and Fulkerson (1956). Later, the implementation of the Ford-Fulkerson method was independently published by Dinic (Dinic, 1970; Dinitz, 2006) and by Edmonds and Karp (1972). In particular, Dinic's algorithm includes techniques that reduce its running time to $\mathcal{O}(|V^2||E|)$ (Dinitz, 2006).

Furthermore, a cut that partitions the graph such that sinks and their respective sources from multiple pairs of vertices are in different partitions is described as a *multicut*. With regard to that, a classic problem in graph theory and combinatorial optimisation is the Minimum Multicut Problem (MinMC) (Costa et al., 2005), defined as follows:

Definition 5.4. Minimum Multicut Problem (MinMC)

Instance: a graph $G = (V, E)$, a weight function $w : E \to \mathbb{N}^*$ and a set $\mathcal{N}$ of pairs (source s, sink t) of terminal vertices of G.

Question: find a set of edges of G, E_{MinMC} whose removal leaves no directed path from s to t for each $(s, t) \in \mathcal{N}$ such that:

$$E_{MinMC} = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.2}$$

MinMC is known to be $\mathcal{NP}$-hard (Dahlhaus et al., 1994). Therefore, unless $\mathbf{P} = \mathbf{NP}$, it cannot be solved in polynomial time.

In this chapter, we specifically consider the MinMC problem in DAGs, which is defined as follows:

Definition 5.5. Minimum Multicut Problem in DAGs (MinMC-DAG)

Instance: a directed acyclic graph $G = (V, E)$, a weight function $w : E \to \mathbb{N}^*$ and a set $\mathcal{N}$ of pairs (s, t) of terminal vertices of G.

Question: find a set of edges of G, $E_{MinMC-DAG}$, whose removal leaves no directed path from s to t for each $(s, t) \in \mathcal{N}$ such that:

$$E_{MinMC-DAG} = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.3}$$

Importantly, MinMC-DAG has also been proven $\mathcal{NP}$-hard (Bentz, 2011).

5.2 Motivation

In early computing systems, personal data was collected from the user by the service provider, occasionally passed on to a small number of other parties and processed, often manually, for few specific purposes. Gradually, this has been moving towards automated data processing, where personal information provided by the user is processed by automated algorithms. As a result, modern computing has created a proliferation of large volumes of user data across thousands of companies. When user data enters such a system, it is automatically processed, possibly by several service providers. Consequently, this processing creates new information, often conveying predictions and inferences about the user, that is used to finally fulfill a certain purpose.

As part of this data processing, different kinds of personal data of one or more users can be combined to learn more information about the users. By doing so, the service provider may gain utility (such as monetary benefits) directly e.g. through selling the data or indirectly e.g. through providing personalised features. This is especially the case when the processing involves several different parties. For instance, consider the following scenario:

> *Mississippi is an online book store which needs to obtain the user's name, shipping address, and credit card number to complete a purchase transaction. After that, Mississippi uses the purchase history of customers to offer book recommendations on its site. It also partners with Nile – a social media platform that gathers virtual book club communities. As part of this partnership, Mississippi shares information about the user's book recommendations and shipping address with Nile in return for a premium, which Nile uses to send the user suggestions on the potential communities they may like to join. However, to support their business model, Nile also uses the data to serve the user personalised advertising on their website*[1].

In this scenario, the user's data, including the name, credit card number and shipping address, enters the system and is processed by an aggregation algorithm into purchase details. Then, this new information is used for completing the purchase transactions and processed further by a prediction algorithm. The output of the prediction algorithm is subsequently used for serving book recommendations and processed even further by a clustering algorithm to serve the purposes of suggesting book clubs and personalised advertising. In this chapter, we refer to this sequence of processing tasks as the *data workflow*. The workflow of the data in our scenario is illustrated in Figure 5.1.

However, the user may refuse to consent to some of the data processing. Particularly, in certain regulatory frameworks such as the GDPR, unless there exists another legal

[1] Our running scenario is a revised version of the scenario proposed by Agrawal et al. (2002).

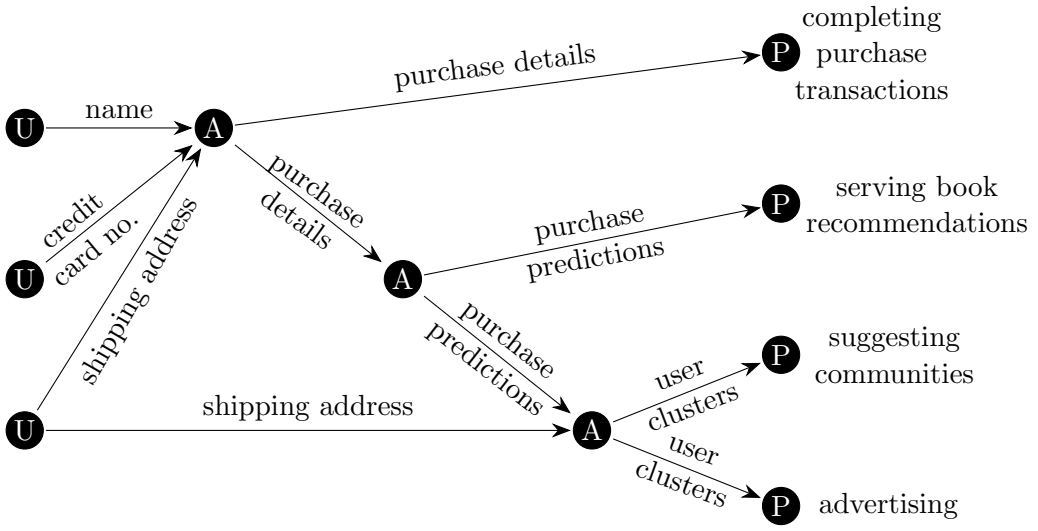


FIGURE 5.1: Data workflow in the system described in the motivating scenario (U – user vertex, A – algorithm vertex, P – purpose vertex).

basis, user's consent is necessary to legally allow the data to be processed. In the given scenario, they the user may, e.g. be happy for their purchase information to be used for suggesting book clubs to join, but may not wish to be subject to personalised advertising based on it. To satisfy these *privacy constraints* imposed on the data processing by the user, Mississippi could refuse to share the user's purchase prediction data with Nile, which is created based on the purchase details. In that case, if the user consents to receive personalised advertising based on their address, Mississippi could share the user's address with Nile to make community suggestions and advertising based on the address only (option 1). Alternatively, Mississippi could share the information about the user's purchase predictions with Nile but ask Nile not to use this data for advertising purposes (option 2).

Nonetheless, privacy constraints may affect the utility of the service providers. In the first case (option 1), if both advertising and community suggestions are personalised based on the user's address only, Nile may earn smaller profit overall and offer Mississippi a smaller premium in exchange for the data. In the second case (option 2), if Nile uses the purchase prediction data for community suggestions but respects Mississippi's request not to use it for advertising purposes, they may make a way smaller profit from advertising and may have to offer Mississippi a way smaller premium relative to their profit. Thus, Mississippi and Nile need to carefully consider how they adapt to the user's privacy constraints, optimising the utility they expect to receive. What complicates the task even more is its large scale: there may be many further stages of data processing, in a workflow that involves hundreds of nodes. For instance, the purchase history can be combined with information from other sources and used for predicting the user's general interests, which can be processed by other partner organisations for other purposes.

Additionally, when the number of users involved is large, each one with different privacy preferences and constraints, these decisions must be automated.

To date, no approach exists which can help the service provider decide how to satisfy privacy constraints of an individual user optimally. Existing approaches to data privacy such as k-anonimity (Sweeney, 2002), l-diversity (Machanavajjhala et al., 2007), t-closeness (Li et al., 2007) and differential privacy (Dwork, 2008) protect personal information attributes by masking or altering the data or the processing output. However, these methods do not take into account the custom constraints of an individual user and the purpose of data processing. Instead, access control methods, including purpose-based access access control (Byun et al., 2005; Byun and Li, 2008), Hippocratic databases (Agrawal et al., 2002) as well as machine-readable privacy policies, such as those enabled by P3P (Cranor, 2002), were designed for use cases where data processing is relatively limited: there is no way to describe the workflow of the data within the system or a conglomeration of systems in a way that benefits both users and service providers. While there is work on implementing privacy constraints in individual algorithms (Debruyne et al., 2019, 2020), no one has looked at implementing them within networks of algorithmic nodes.

In this chapter, we propose a novel approach to finding the most optimal ways of satisfying the user's privacy constraints. Specifically, we model the data workflow as a graph and privacy constraints as pairs of vertices of the graph. We formulate the problem as an optimisation problem where pairs of graph vertices must be disconnected such that utility is maximised. Furthermore, through a polynomial-time reduction from the minimum multicut problem (MinMC) which is $\mathcal{NP}$-hard, we prove that our problem in general is $\mathcal{NP}$-hard. Nonetheless, certain assumptions on the valuation of the user's data types and the expected utility can help us simplify the problem. Thus, we present a simplified instance of the problem where the utility is linearly additive. Moreover, we propose five algorithms that can satisfy the privacy constraints in data workflows. Then, we analyse the (non-)optimality of these algorithms with respect to their computational complexity. We evaluate and compare the algorithms in terms of accuracy and performance. Notably, we show that, although computing the optimal solution can be very time-consuming, the proposed approximations can provide very accurate and efficient alternatives.

5.3 Consented Data Workflow Problem

We consider consent in data processing as a graph theory problem. First, we propose a new data processing model, which describes the workflow of personal information. Second, we present a formal formulation of the problem of identifying the parts of data

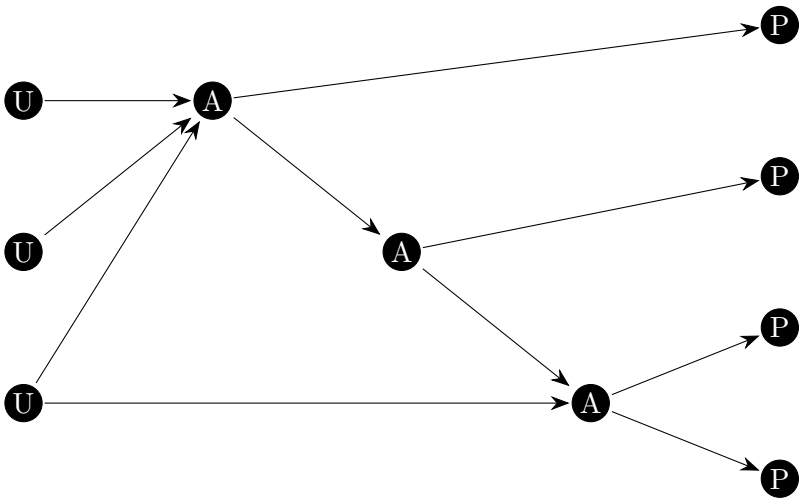


FIGURE 5.2: Illustration of the data workflow (U – user vertex, A – algorithm vertex, P – purpose vertex).

processing workflows that align with the user's constraints, and bring the most benefits to the service provider.

5.3.1 Model

In order to describe the data workflow in the system, we formulate our data processing model as a directed graph (di-graph) $G = (V, E)$ with a set of edges E representing the data flow and a set of vertices V representing the stages of data processing. These stages refer to three activities: data collection from the user (start), algorithmic data processing (possibly at multiple stages) and satisfying the purpose of processing (end). For this reason, we distinguish three kinds of vertices, i.e. $V = V^U \cup V^A \cup V^P$, where:

- V^U is a set of user data vertices, which represent the types of data collected directly from the user, such as name, shipping address, credit card number;
- V^A is a set of algorithm vertices, which represent data processing algorithms that take one or more data types as input, e.g. purchase predictions and location, and output a new data type, e.g. predicted clusters of users;
- V^P is a set of purpose vertices, which represent the end goals of data processing, e.g. serving personalised advertising or suggesting communities to join.

To illustrate this with regard to our running scenario, Figure 5.2 shows how different points from our data flow diagram in Figure 5.1 map to these vertex categories. Importantly, we view data processing algorithms as 'black boxes'. That is, our model does not describe their internal workings, only the input and output represented by the incoming and outgoing edges. Thus, in our model, vertices in V^A (marked with 'A') have at least

one incoming and at least one outgoing edge. Differently, vertices in V^U (marked with 'U') have no incoming edges and those in V^P (marked with 'P') – no outgoing edges.

Furthermore, satisfying the given purposes is what brings service providers utility. For this reason, for any vertex in V^P, we will have an associated utility that reflects the value processing for this purpose brings to service providers. In practice, the service providers' valuation depends on factors such as the accuracy of the datasets used as the input to the processing algorithms (Li et al., 2014; Ghorbani and Zou, 2019). In particular, where data processing is a multi-stage process, the utility is affected by all stages and all datasets processed for the purpose.

Therefore, in order to calculate the utility of data processing for a given purpose, in our model we look for all vertices and edges that carry the data workflow to the given purpose vertex. Formally, we say that a vertex $v_i \in V$ is *reachable* from a vertex $v_j \in V$ if there exists a path in G defined as a graph $(\{v_1, v_2, \ldots, v_k\}, \{(v_1, v_2),(v_2, v_3),\ldots\ (v_{k-1}, v_k)\})$ such that $v_1 = v_j$ and $v_k = v_i$. Now, we consider the so-called *reachability subgraph* of a purpose vertex. For each purpose vertex $p \in V^P$, the reachability subgraph of p is the graph $G_p = (V_p, E_p)$ where $V_p \subseteq V$ is the set of the vertices that p is reachable from and $E_p \subseteq E$ is the set of edges that connect them. If an edge is removed from G, the reachability subgraph of one or more purpose vertices is affected. In general, we write $\mathcal{R}(G_p)$ to denote the set of all possible reachability subgraphs of p in G. Then, to calculate the utility of fulfilling a purpose, for each purpose vertex $p \in V^P$ we define a utility function $u_p : \mathcal{R}(G_p) \to \mathbb{R}_0^+$, which is a function of the reachability subgraph of p.

While u_p can be an arbitrary function dependent on the valuation of datasets in the corresponding reachability subgraph, the valuations of some datasets may influence the valuations of others. For example, when users are clustered based on their shipping address to be served personal advertising, the accuracy and, thus, valuation of the address may impact the accuracy, i.e. valuation of the clustering. To describe the relationships between these valuations in our model, we define a valuation function $\pi : E \to \mathbb{R}_0^+$, representing the valuation of the data propagating through the edge in the data processing system. As we later show in Section 5.5, given the reachability subgraph $G_p = (V_p, E_p)$ of a vertex $p \in V^P$, the utility function $u_p(G_p)$ at p can be defined as a function of the valuations of edges in E_p.

5.3.2 Problem Formulation

The presented data processing model describes the workflow of data in the system. However, the user may refuse consent to some of this processing. For example, one of the user's constraints may be: *I'm happy for my shipping address to be used for suggesting me book clubs to join, but I don't want to be served personalised advertising based on it.* To formulate this specific constraint, the user does not need to have an expert

understanding of *how* their personal data is processed. Whereas, in order to satisfy this constraint, the system should be able to analyse the data workflow and make sure that there is no connection between the vertex where the shipping address enters the workflow and the vertex representing advertising. If there is no such connection, we call this data workflow *consented*.

Thus, our goal is to find the *consented data workflow* under certain constraints. Specifically, we focus on what we call the Consented Data Workflow problem (CDW): given user's constraints expressed in terms of the vertices that they do not wish to be connected, find a configuration of the data workflow where these constraints are satisfied. However, if there is more than one way of disconnecting the given vertices, the optimal solution should minimise the utility loss for the service provider from applying the constraints. In other words, we are looking for the utility-maximising solution subject to the users' privacy constraints.

Formally, users' constraints can be expressed as a set of pairs of vertices. In this chapter we focus on personal consent against a set of purposes, thus we will want to disconnect particular user vertices in V^U with purpose vertices in V^P. Formally, our set of constraints is a set $\mathcal{N} = \{(v_s, v_t) \mid v_s \in V^U, v_t \in V^P\}$. In order to satisfy the constraints, the initial graph G needs to be modified by removing one or more edges that belong to the paths between pairs (v_s, v_t), such that the utilities u_p are maximised. In essence, our problem is a multi-objective optimisation problem, where the objectives are to maximise u_p for all $p \in V^P$. The most common approach to multi-objective optimization is to turn the problem into a single-objective optimization using a weighted sum (Marler and Arora, 2004). This allows us to define the utility of the system G as:

$$U(G) = \sum_{p \in V^P} w_p u_p(G_p), \tag{5.4}$$

where w_p is the weight of the purpose corresponding to vertex p and G_p is the reachability subgraph of p. Therefore, given $\mathcal{N}$, the objective of CDW is to find the *consented subgraph* of G:

$$G^* = arg \max_{G'} U(G') \tag{5.5}$$

such that $G' = (V, E')$ is a subgraph of G where $E' \subseteq E$ and there is no path from s to t for each $(s, t) \in \mathcal{N}$.

5.4 Complexity Analysis

In this section, we study the complexity of our problem. Specifically, we show that the CDW problem is $\mathcal{NP}$-hard. In order to do so, we reduce from the minimum multicut (MinMC) problem which is $\mathcal{NP}$-hard in directed acyclic graphs (DAGs) (Bentz, 2011).

Given a DAG, $G = (V, E, w)$, where V is a set of vertices, E is a set of edges and $w : E \to \mathbb{N}^*$ an edge weight function, as well as a set $\mathcal{N}$ of pairs (source s, sink t) of terminal vertices of G, the objective of MinMC is to find a set of edges of G, E_{MinMC}, whose removal leaves no directed path from s to t for each $(s, t) \in \mathcal{N}$ such that:

$$E_{MinMC} = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.6}$$

To prove the complexity of CDW, let us consider an instance $I_{MinMC} = (G, \mathcal{N})$ of MinMC in DAGs and translate it into an instance of CDW. Let $G = (V, E)$. We are going to construct an instance of CDW on the same DAG G and the same constraints $\mathcal{N}$ such that, for a set of edges E^{MinMC}, the subgraph of G, $(V, E \backslash E^{MinMC})$ is a solution to the CDW instance if and only if E^{MinMC} is a solution to I_{MinMC}.

For simplicity, for any vertex $v \in V$, we use $in(v)$ to denote a set of incoming edges of v and $out(v)$ to denote outgoing edges in G. We first construct the set $V^U \subseteq V$ such that for any vertex $v \in V$ $in(v) = \varnothing$ if and only if $v \in V^U$. Similarly, we construct the set $V^P \subseteq V$ such that, for any vertex v, $out(v) = \varnothing$ if and only if $v \in V^P$. Note that V^U and V^P cannot be empty sets, because G is acyclic. Then, we construct a set $V^A = V \backslash (V^U \cup V^P)$. Moreover, we construct a purpose-reachability function $r : V^U \cup V^A \to \mathcal{P}(V^P)$, which for any vertex $v \in V^U \cup V^A$ returns a set of vertices in V^P such that $p \in r(v)$ iff p is reachable from v. Note that this construction happens in polynomial time. Then, for each edge $e = (v, v') \in E$, we construct a valuation function $\pi : E \to \mathbb{R}$ such that $\pi(e) = \frac{w(e)}{|r(v)|}$. In addition, for each $p \in V^P$, we have a utility function $u_p(G_p) = \sum_{e \in E_p} \pi(e)$, where $G_p = (V, E_p)$ is a reachability subgraph of p. Lastly, we construct a utility function of G as $U(G) = \sum_{p \in V^P} u_p(G_p)$, which implies that:

$$U(G) = \sum_{e \in E} w(e) \tag{5.7}$$

Note that each edge contributes to $U(G)$ exactly its original weight $w(e)$ because of the way $\pi(e)$ is constructed.

This concludes the construction of instance I_{CDW}, which is a CDW instance with the objectives weighted equally, i.e. for all $p \in V^P$, w_p is equal. Note that since all pairs $(s, t) \in$ consist of terminal vertices, $\mathcal{N}$ is the same set for I_{CDW}. Now, let us prove the following:

Lemma 5.6. *Given an instance I_{MinMC} of MinMC, there is a polynomial time reduction to an instance I_{CDW} of* CDW *such that a graph $(V, E \backslash E_{MinMC})$ is the solution to I_{CDW} iff E_{MinMC} is the solution to I_{MinMC}.*

Proof. ($\Leftarrow$) Let E_{MinMC} be a solution to I_{MinMC}. Since E_{MinMC} is a multicut of G given $\mathcal{N}$, this guarantees that when the edges in E_{MinMC} are removed from G, there is no directed path from s to t for each $(s,t) \in \mathcal{N}$. We show that the removal of set E_{MinMC} maximises the utility $U(G^*)$. Firstly, the set of edges after removal of E_{MinMC} from G can be expressed as:

$$E^* = E \backslash E_{MinMC}. \tag{5.8}$$

Then, if we plug Equation 5.6 to Equation 5.8, we have:

$$E^* = E \backslash \{arg \min_{E'} \sum_{e \in E'} w(e)\}. \tag{5.9}$$

Since E^* is a difference between E and the subset of E whose sum of edge weights is minimal, the sum of edge weights of E^* is maximal. Therefore, Equation 5.9 is equivalent to:

$$E^* = arg \max_{E'} \sum_{e \in E'} w(e). \tag{5.10}$$

Thus, if we plug Equation 5.7 into Equation 5.10, we have:

$$G^* = arg \max_{G'} U(G). \tag{5.11}$$

As this is exactly Equation 5.5, the graph $G^* = (V, E \backslash E_{MinMC})$ is the solution to I_{CDW}. Notably, this transition is performed in polynomial time.

($\Rightarrow$) Conversely, let $G^* = (V, E \backslash E_{MinMC})$ be a solution to I_{CDW}. By definition of the consented subgraph, in graph G^* there is no directed path from s to t for each $(s,t) \in \mathcal{N}$ and the utility is maximised as per Equation 5.5. If we plug Equation 5.7 into Equation 5.5, we have:

$$E^* = arg \max_{E'} \sum_{e \in E'} w(e). \tag{5.12}$$

Since $E^* \subseteq E$, there exists a set of edges $E \backslash E^*$ such that:

$$E \backslash E^* = E \backslash \{arg \max_{E'} \sum_{e \in E'} w(e)\}. \tag{5.13}$$

This is equivalent to:

$$E \backslash E^* = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.14}$$

If we call this set E_{MinMC}, i.e. $E \backslash E^* = E_{MinMC}$, then:

$$E_{MinMC} = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.15}$$

As this is exactly Equation 5.6, E_{MinMC} is the minimum multicut of G given $\mathcal{N}$. Therefore, E_{MinMC} is the solution to I_{MinMC}. Notably, this is achieved in polynomial time. □

Given this, we can now show that even for a small number of user's constraints our problem is $\mathcal{NP}$-hard, since MinMC is an $\mathcal{NP}$-hard problem in di-graphs:

Theorem 5.7. CDW *is $\mathcal{NP}$-hard, even if* $|\mathcal{N}| = 2$.

Proof. By Lemma 5.6, an instance of MinMC in DAGs can be converted to CDW in polynomial time. Moreover, solving CDW yields in polynomial time a solution to MinMC. Since MinMC in DAGs is known to be an $\mathcal{NP}$-hard problem for any $|\mathcal{N}| > 1$ (Bentz, 2011), there exists a polynomial-time reduction from a known $\mathcal{NP}$-hard problem to CDW. Therefore, the CDW problem is $\mathcal{NP}$-hard for any $|\mathcal{N}| > 1$, which concludes the proof. □

Thus, assuming that $\mathbf{P} \neq \mathbf{NP}$, we can conclude that the CDW problem cannot be solved in polynomial time. However, we can find efficient algorithms to solve some specific instances of the problem.

5.5 Additive Model

As shown in Section 5.4, CDW in general is $\mathcal{NP}$-hard, which makes it difficult to expect a relatively efficient algorithm. Moreover, the valuations and utilities defined in Section 5.3.1 can be arbitrary complex functions. In this section, we focus on a simple but practical instance of the problem, where these functions are linearly additive.

In practice, data valuation is determined by a complex interaction of multiple factors including its age, accuracy and reliability (Heckman et al., 2015). Here, we choose a linear valuation function as a natural choice of a function. In particular, we assume that the valuation function of the data type going out of a vertex is linearly additive with respect to the importance of the data types on the incoming edges. In more detail, consider an instance of CDW, where for each edge $e = (v, v') \in E$, the valuation is defined recursively as follows:

$$\pi(e) = \sum_{e' \in in(v)} \pi(e'). \tag{5.16}$$

Similarly, we model the utility gained from processing the data for a purpose as a linearly additive function with respect to the valuation of the data types on the incoming edges.

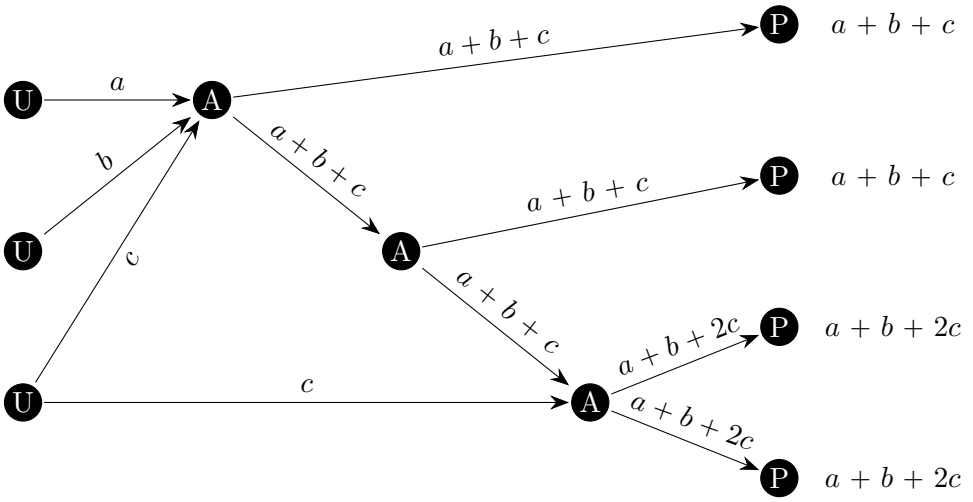


FIGURE 5.3: Example importance and utility values in the special instance of the data processing system (U – user vertex, A – algorithm vertex, P – purpose vertex).

That is, for each purpose vertex $p \in V_P$ and its reachability subgraph G_p, we define a utility function as follows:

$$u_p(G_p) = \sum_{e \in in(p)} \pi(e). \tag{5.17}$$

Since the valuation is defined recursively, we also assume that our model has no cycles. That is, the original graph G is a directed acyclic graph (DAG).

To illustrate this instance, Figure 5.3 shows how the valuation values propagate within the graph from our running scenario in Figure 5.1. For example, if the importance of the edge representing the data type 'purchase predictions' in our data processing model is $a + b + c$ and the importance of the edge representing 'shipping address' is c, then we assume that the importance of the edge representing 'user clusters' is $a + b + 2c$. In this case, since the 'user clusters' edge is the only edge incoming to the 'advertising' vertex, the utility gained from serving personalised advertising is also $a + b + 2c$.

Formally, the objective of our linearly additive instance of the CDW problem, called CDW-LA, is to find a the consented subgraph of G, given: a DAG $G = (V^U \cup V^A \cup V^P, E)$, a valuation function $\pi(e) = \sum_{e' \in in(v)} \pi(e')$ for each $e = (v, v') \in E$, a utility function $u_p(G_p) = \sum_{e \in in(p)} \pi(e)$ for each reachability subgraph G_p of each $p \in V_P$, a weight $w_p = 1$ of the purpose represented by vertex p and a set $\mathcal{N}$ of pairs (s, t) of terminal vertices of G such that $s \in V_U$ and $t \in V_P$.

5.6 Algorithms

In this section, we devise a range of new algorithmic approaches that can solve the CDW-LA problem. Although there might exist multiple optimal solutions, we design our algorithms looking for a single solution $G^* = (V^*, E^*)$. While some of them offer optimal solutions to CDW-LA, others serve as viable heuristics. Note that, even though the algorithms may not be optimal (i.e. utility maximising), all five algorithms always return a *feasible* solution, which is any subgraph of G with no path between each $(s_i, t_i) \in \mathcal{N}$ for $i \in \{1, \ldots, |\mathcal{N}|\}$.

Firstly, a simple heuristic for finding a feasible solution is an algorithm that removes a random edge from each of the paths connecting $(s, t) \in \mathcal{N}$. In more detail, Algorithm 1 RemoveRandomEdge finds all paths from s to t (in lines 1 - 2) and from each of the paths selects a random edge to remove (in lines 3 - 4). Then, before the edge is removed (in line 7), the other edges whose valuation depends on the presence of the given edge in the graph must be updated. This is done by the `updateDependencies` function (in line 6). In particular, if the valuation of an edge after the update is 0, such edge must also be removed, e.g. if edge (s, v_1) in Figure 5.4 is removed, edges (v_1, t) and (v_1, t') also require removal. Although the solution has a high variance, the run time of this algorithm is polynomial.

Algorithm 1 RemoveRandomEdge

Input: A graph G and a set of constraints $\mathcal{N}$.
Output: A graph G.

```
1: for all (s,t) ∈ N do
2:     for all p ∈ getAllEdgePaths(G,s,t) do
3:         edgeIndex ← getRandomInteger(1,|p|)
4:         e ← p[edgeIndex]
5:         if hasEdge(G,e) then
6:             updateDependencies(G,e)
7:             removeEdge(G,e)
8:         end if
9:     end for
10: end for
```

Secondly, as the valuation function of the edges is additive, and because the valuation of the incoming edge of an algorithm vertex is always greater or equal than the outgoing one, the removal of the first edge of each path from s to t can serve as another trivial heuristic. Specifically, Algorithm 2 RemoveFirstEdge is very similar to RemoveRandomEdge, except that, instead of selecting a random edge, it removes the first edge from each path (in line 3). This algorithm reflects an approach whereby the user's data type is removed entirely and not even collected by the system. Similarly to RemoveRandomEdge, the runtime of this algorithm is polynomial.

Algorithm 2 RemoveFirstEdge

Input: A graph G and a set of constraints $\mathcal{N}$.
Output: A graph G.

```
1: for all (s, t) ∈ N do
2:     for all path ∈ getAllEdgePaths(G, s, t) do
3:         e ← getFirstEdge(path)
4:         if hasEdge(G, e) then
5:             updateDependencies(G, e)
6:             removeEdge(G, e)
7:         end if
8:     end for
9: end for
```

Next, we look for algorithms that can provide more accurate solutions. In particular, we propose a greedy algorithm that can provide a feasible solution in polynomial time. This algorithm follows the heuristic of making locally optimal choices for each constraint. To do so, it uses a polynomial-time algorithm solving the Minimum Cut problem (MinCut) (Dinitz, 2006; Edmonds and Karp, 1972), defined as follows: given a graph $G = (V, E)$, a weight function $w : E \to \mathbb{N}^*$ and a single pair (source s, sink t) of terminal vertices of G, find a set of edges of G, E_{MinCut} whose removal leaves no directed path from s to t for each $(s, t) \in \mathcal{N}$ such that:

$$E_{MinCut} = arg \min_{E'} \sum_{e \in E'} w(e). \tag{5.18}$$

To design a greedy algorithm, we can use algorithms solving MinCut to find a minimum cut of G for each $(s, t) \in \mathcal{N}$. Consequently, we remove the minimum cut before moving on to the next constraint, which results in a partial solution. In more detail, Algorithm 3 RemoveMinCuts starts from initialising the weights $w(e) = \pi(e) \sum_{p \in r(v)} w_p$ for all edges $e \in E$ (in lines 1 - 4). Then, for each constraint $(s, t) \in \mathcal{N}$, it finds the minimum cut that solves MinCut for vertices s and t in G with weights w (in line 6). For each edge in the minimum cut, it uses the `updateDependencies` function to update the valuations of the consecutive edges (in line 8) before removing the given edge (in line 9). Given that MinCut is known to be solvable in polynomial time (Dinitz, 2006), the outcome of this heuristic can also be found in polynomial time.

Another way of approximating the solution is by converting our problem to MinMC, defined in Section 5.4. That is, we can solve MinMC with weights $w(e) = \pi(e) \sum_{p \in r(v)} w_p$ for all edges $e \in E$ and then use the MinMC solution to find a solution to CDW-LA. In the same way as RemoveMinCuts, Algorithm 4 RemoveMinMC starts from initialising the weights w (in lines 1 - 4). Then, it finds the minimum multicut of graph G for constraints $\mathcal{N}$ by executing the algorithm solving MinMC for input $(G, \mathcal{N}, w)$ (in line 5). Subsequently, for each edge in the minimum multicut, it uses the `updateDependencies`

Algorithm 3 RemoveMinCuts

Input: A graph $G = (V, E)$ and a set of constraints $\mathcal{N}$.
Output: A graph G.

```
1: w ← ∅
2: for all e ∈ E do
3:     w(e) ← π(e) Σ_{p∈r(v)} w_p
4: end for
5: for all (s,t) ∈ N do
6:     for all e ∈ MinCut(G, w, s, t) do
7:         if hasEdge(G, e) then
8:             updateDependencies(G, e)
9:             removeEdge(G, e)
10:        end if
11:    end for
12: end for
```

function to update the valuations of the consecutive edges (in line 8) before removing the given edge (in line 9).

Algorithm 4 RemoveMinMC

Input: A graph $G = (V, E)$, a set of constraints $\mathcal{N}$.
Output: A graph G.

```
1: w ← ∅
2: for all e ∈ E do
3:     w(e) ← π(e) Σ_{p∈r(v)} w_p
4: end for
5: multicut ← MinMC(G, N, w)
6: for all e ∈ multicut do
7:     if hasEdge(G, e) then
8:         updateDependencies(G, e)
9:         removeEdge(G, e)
10:    end if
11: end for
```

Finally, we propose an algorithm that can guarantee achieving an optimal solution. That is, Algorithm 5 BruteForce is an exhaustive search algorithm that enumerates all feasible candidates for the solution and compares them to eventually output the one that maximises the utility. More specifically, BruteForce starts from finding the set of all paths $\mathcal{A}$ from s to t for all $(s, t) \in \mathcal{N}$, which need to be broken (in lines 1 - 4). In order to list all feasible multicuts of G for the given $\mathcal{N}$, the Cartesian product of $\mathcal{A}$ is computed (in line 5). Then, the algorithm systematically checks the utility of G after the removal of each multicut (in lines 8 - 27). Importantly, at the beginning of the multicut check, copies are made of the valuation values π' of each edge and the number of paths p' the edge belongs to in G, as well as of the graph G itself (in lines 9 - 14). Before an edge of the feasible multicut is removed from the copy of G (in line

18), the valuation π' and the number of paths p' in the copy of G are updated for its dependencies (in line 17). At the end of the multicut check, the utility of the copy of G is compared to the utility of the most optimal solution found so far (in lines 21 - 25). This way, given that all possible solutions that satisfy the constraints are checked, the algorithm can guarantee eventually finding the optimal solution. However, the runtime of this algorithm is exponential even in the best case.

Algorithm 5 BRUTEFORCE

Input: A graph $G = (V, E)$ and a set of constraints $\mathcal{N}$.
Output: A graph G^*.

```
1: A ← ∅
2: for all (s,t) ∈ N do
3:     A ← A ∪ getAllEdgePaths(G,s,t)
4: end for
5: multicuts ← cartesianProduct(A)
6: maxUtility ← 0
7: G* ← G
8: for all multicut ∈ multicuts do
9:     G' ← G
10:    π', p ← ∅, ∅
11:    for all e ∈ E do
12:        π'(e) ← π(e)
13:        p(e) ← Σ_{p∈r(v)} w_p
14:    end for
15:    for all e ∈ multicut do
16:        if hasEdge(G',e) then
17:            updateDependencies(G',e,π',p)
18:            removeEdge(G',e)
19:        end if
20:    end for
21:    utility ← U(G')
22:    if utility > maxUtility then
23:        maxUtility ← utility
24:        G* ← G'
25:    end if
26: end for
27: return G*
```

While Algorithms 1, 2 and 5 are designed to work on models with arbitrary valuation and purpose utility functions, specifying these functions is needed to calculate the weights $w(e)$ in Algorithms 3 and 4.

5.7 Optimality of Solutions

Out of five algorithms proposed in Section 5.6, only BRUTEFORCE guarantees an optimal solution to CDW-LA. In contrast, it is clear that REMOVERANDOMEDGE does not

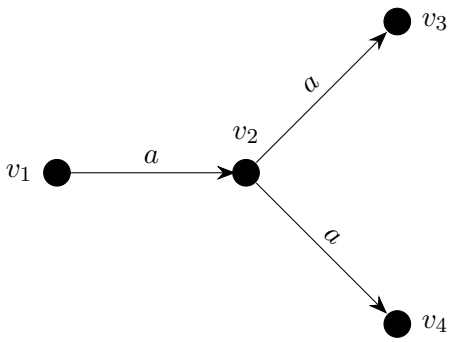


FIGURE 5.4: A data processing model where $V^U = \{v_1\}$, $V^A = \{v_2\}$, $V^P = \{v_3, v_4\}$ and $E = \{(v_1, v_2), (v_2, v_3), (v_2, v_4)\}$.

guarantee an optimal solution – we use it as a benchmark for our evaluation. In this section, we analyse the properties of the solutions returned by our three remaining heuristics and prove that none of them can guarantee finding an optimal solution even for the linear setting.

Firstly, we show that a simple removal of the first edge of each path proposed in Algorithm 2 REMOVEFIRSTEDGE does not guarantee an optimal solution. In more detail, for each $(s_i, t_i) \in \mathcal{N}$, there is at least one path $P = (V_P, E_P) \in \mathcal{A}$ of the form $V_P = \{v_1, v_2, \ldots, v_k\}$, $E_P = \{(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k)\}$ where $v_1 = s_i$ and $v_k = t_i$. From each such path $P \in \mathcal{A}$, we could remove edge (v_1, v_2). We refer to (v_1, v_2) as the *first edge*. We show that the removal of the first edge from each P does not always result in an optimal solution to CDW-LA by the following example.

Let G be a data processing model where $V^U = \{v_1\}$, $V^A = \{v_2\}$, $V^P = \{v_3, v_4\}$, $E = \{(v_1, v_2), (v_2, v_3), (v_2, v_4)\}$ and for each $p \in V^P$, $w_p = 1$. In addition, assume that for edge $e_1 = (v_1, v_2)$, $\pi(e_1) = a$ where $a \in \mathbb{R}_0^+$ and that $\mathcal{N} = \{(v_1, v_3)\}$. This model is illustrated in Figure 5.4.

In such case, we use Equation 5.16 to calculate the valuation of edges $e_2 = (v_2, v_3)$ and $e_3 = (v_2, v_4)$, which is $\pi(e_2) = \pi(e_3) = a$. We also use Equation 5.4 to calculate the initial utility of G, which is $U(G) = 2a$. Given that $\mathcal{N} = \{(v_1, v_3)\}$, we establish that there is one path that needs to be disconnected in order to satisfy the constraints, i.e. $\mathcal{A} = \{P\}$ where $P = (\{v_1, v_2, v_3\}, \{(v_1, v_2), (v_2, v_3)\})$.

Then, we remove the first edge (v_1, v_2) from P. The utility of the resulting graph G'_1 is $U(G'_1) = 0$, since purpose vertices v_3 and v_4 are now not linked to any user vertex. However, if instead we removed the alternative edge (v_2, v_3), vertex v_4 would still be linked to v_1 and therefore the utility of the resulting graph G'_2 would be $U(G'_2) = a$. Thus, in this case the removal of the first edge does not provide us with an optimal solution.

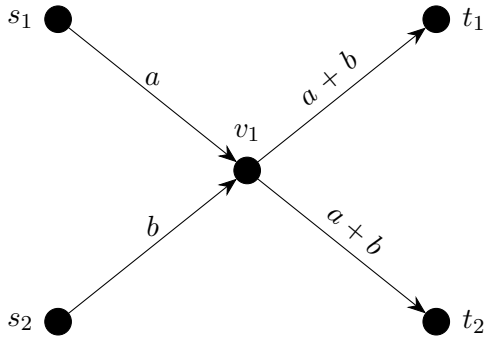


FIGURE 5.5: A data processing model where $V^U = \{s_1, s_2\}$, $V^A = \{v_1\}$, $V^P = \{t_1, t_2\}$ and $E = \{(s_1, v_1), (s_2, v_1), (v_1, t_1), (v_1, t_2)\}$.

In a similar way, we can prove that removing the last edge (v_{k-1}, v_k) from each path in $\mathcal{A}$ also does not guarantee the optimal solution. However, since for all $(s, t) \in \mathcal{N}$ all paths from s to t are broken, it similarly provides a feasible solution.

Furthermore, we show that solving the problem in a greedy way, i.e. where we apply the constraints one at a time as proposed in Algorithm 3 RemoveMinCuts, also does not lead to an optimal solution. Specifically, consider a series of graphs $G^0, G^1, \ldots, G^{|\mathcal{N}|}$. These graphs are computed recursively such that $G^0 = G = (V, E)$ and for all $i \in \{1, \ldots, |\mathcal{N}|\}$, $G^i = (V, E^i)$ where $E^i = E^{i-1} \backslash \text{MinCut}(G^{i-1}, s_i, t_i, w)$ corresponds to i-th pair of vertices in $\mathcal{N}$. Given a solution to MinCut, one could transform CDW-LA into a repeated MinCut problem looking for $G^{|\mathcal{N}|}$. While this approach leads to a feasible solution, $G^{|\mathcal{N}|}$ is not necessarily an optimal solution.

We show this by the following example. Let G be a data processing model where $V^U = \{s_1, s_2\}$, $V^A = \{v_1\}$, $V^P = \{t_1, t_2\}$, $E = \{(s_1, v_1), (s_2, v_1), (v_1, t_1), (v_1, t_2)\}$ and for each $p \in V^P$, $w_p = 1$. Assume that for $e_1 = (s_1, v_1)$, $\pi(e_1) = a$ and for $e_2 = (s_2, v_1)$, $\pi(e_2) = b$, where $a, b \in \mathbb{R}_0^+$ and $a > b$. This model is illustrated in Figure 5.5. In addition, there are two constraints: $\mathcal{N} = \{(s_1, t_1), (s_1, t_2)\}$.

In such case, $i = 2$. We look for $G^2 = (V, E \backslash \text{MinCut}(G^1, s_1, t_1, w))$. Thus, we first calculate $G^1 = (V, E \backslash \text{MinCut}(G, s_1, t_1, w))$. We observe that there is one path $P = (\{s_1, v_1, t_1\}, \{(s_1, v_1), (v_1, t_1)\})$ between vertices s_1 and t_1. Because $a > b$, $w((v_1, t_1)) = a + b < w((s_1, v_1)) = 2a$. Thus, $G^1 = (V, E \backslash \{(v_1, t_1)\})$. With this information, we return to looking for G^2. We observe that there is one path $P = (\{s_1, v_1, t_2\}, \{(s_1, v_1), (v_1, t_2)\})$ between vertices s_1 and t_2. However, now $w((s_1, v_1)) = a$ and $w((v_1, t_2)) = a + b$. So, $w((s_1, v_1)) < w((v_1, t_2))$ and $G^2 = (V, E \backslash \{(s_1, v_1), (v_1, t_1)\})$.

After that, we calculate the utility $U(G^2) = b$. However, we can see that in order to optimally solve CDW-LA, it is sufficient to remove edge (s_1, v_1) only. That is, the optimal solution to CDW-LA in this case is $G^* = (V, E \backslash \{(s_1, v_1)\})$, because its utility is $U(G^*) = 2b$. Therefore, G^2 is not an optimal solution to CDW-LA.

Intuitively, it is reasonable to assume that the optimal solution requires removing no more than one edge per path between s and t for each $(s,t) \in \mathcal{N}$. In what follows, we first prove that, for settings where this is indeed the case, our Algorithm 4 RemoveMinMC finds the optimal set of edges to remove. However, we then show that there are settings where this assumption does not hold, and where the optimal solution requires removing more than one edge from the same path. However, in Section 5.8 we show that these cases are rare and, in most cases, the algorithm does return the optimal solution. Hence showing that the algorithm guarantees the optimal solution in restricted settings is useful.

Theorem 5.8. *If for each path $P = (V_P, E_P) \in \mathcal{A}$ there is exactly one edge $e \in E_P$ such that $e \in E_{MinMC}$, then there exists a solution $G^* = (V, E \backslash E_{MinMC})$.*

Proof. Let $T \subseteq V^P$ be a set of purpose vertices such that for all $(s,t) \in \mathcal{N}$, $t \in T$. If for each path $P = (V_P, E_P) \in \mathcal{A}$ there is exactly one edge $e \in E_P$ such that $e \in E_{MinMC}$, then the removal of a set of edges E_{MinMC} reduces the utility of a purpose vertex $t \in T$ by $\sum_{e \in E_t} \pi(e)$ where $E_t \subseteq E_{MinMC}$ is a set of those edges in E_{MinMC} that are within the reachability subgraph of t, G_t. Thus, if the resulting graph after the removal of set E_{MinMC} from G is $G' = (V, E')$, then using Equation 5.4, the total loss of the utility can be calculated as follows:

$$U(G) - U(G') = \sum_{t \in T} \sum_{e \in E_t} w_t \pi(e). \tag{5.19}$$

This is equivalent to the following equation:

$$U(G') = U(G) - \sum_{e \in E \backslash E'} \pi(e) \sum_{t \in T} w_t. \tag{5.20}$$

We are looking for the consented subgraph $G^* = (V, E^*)$. In fact, if we plug Equation 5.20 into Equation 5.5, we have:

$$G^* = arg \max_{G'} \{U(G) - \sum_{e \in E \backslash E'} \pi(e) \sum_{t \in T} w_t\}. \tag{5.21}$$

Equivalently, we are looking for a subgraph $G^* = (V, E^*)$ where:

$$E^* = E \setminus \{arg \min_{E \backslash E'} \sum_{e \in E \backslash E'} \pi(e) \sum_{t \in T} w_t\}. \tag{5.22}$$

Thus, by Equation 5.6, E^* is the set difference of E and the minimum multicut of G given $\mathcal{N}$, where the edge weight is $w(e) = \pi(e) \sum_{t \in T} w_t$. In more detail, the minimum

multicut with edge weights $w(e) = \pi(e) \sum_{t \in T} w_t$ can be expressed as:

$$E_{MinMC} = arg \min_{E'} \sum_{e \in E'} \pi(e) \sum_{t \in T} w_t. \tag{5.23}$$

If we plug Equation 5.22 into Equation 5.23, then what we are looking for is $G^* = (V, E^*)$ where:

$$E^* = E \setminus E_{MinMC}. \tag{5.24}$$

Since E_{MinMC} is a solution to MinMC, there is $G^* = (V, E \setminus E_{MinMC})$. □

However, removing a single edge from each path does not always result in an optimal solution. We prove this by the following example. Consider a graph G where $V^U = \{s_1, s_2\}$, $V^A = \{v_1\}$, $V^P = \{t_1, t_2\}$, $E = \{(s_1, v_1), (s_2, v_1), (v_1, t_1), (v_1, t_2)\}$ and for each $p \in V^P$, $w_p = 1$. In addition, assume that for $e_1 = (s_1, v_1)$, $\pi(e_1) = a$ and for $e_2 = (s_2, v_1)$, $\pi(e_2) = b$, where $a, b \in \mathbb{R}_0^+$ and $a > b$. This graph is illustrated in Figure 5.5. Now, let the set of constraints be as follows: $\mathcal{N} = \{(s_1, t_1), (s_1, t_2), (s_2, t_1)\}$. We can see that the optimal solution in this case is $G^* = (V, \{(s_2, v_1), (v_1, t_2)\})$. However, since $(s_1, t_1) \in \mathcal{N}$ and there is a path from s_1 to t_1 in the original graph G, we can observe that the optimal solution G^* does not contain two edges (s_1, v_1) and (v_1, t_2) from that path.Therefore, in general, it is not true that Algorithm 4 RemoveMinMC can guarantee finding an optimal solution to CDW-LA by removing only one edge from each path.

5.8 Experimental Evaluation

In this section, we evaluate the performance of the proposed algorithms empirically. First, we discuss the experimental setup, including details on the algorithm implementation and graph data generation. Then, we present results of the experiments focusing on the runtime and accuracy of each algorithm.

5.8.1 Methodology

We implement the proposed algorithms using the NetworkX[2] library. In particular, this library provides a method to solve the MinCut problem in Algorithm 3. In addition, we implement Algorithm 4 using the PICOS[3] API for optimization solvers. Specifically,

[2]NetworkX, `https://networkx.org/`.
[3]PICOS, `https://picos-api.gitlab.io/picos/`.

we use the GLPK (GNU Linear Programming Kit)[4] package for solving the MinMC problem.

We compare the different algorithms by measuring their performance on synthetic data. This choice allows us to test the algorithms when all the assumptions of the CDW-LA instance are met. To do so, our graph generation method includes the following parameters:

- number of constraints $|\mathcal{N}|$;
- number of vertices $|V|$;
- path length k – for any $(s,t) \in \mathcal{N}$, if there is a path $P = ((v_1, v_2, \ldots, v_k), ((v_1, v_2) \ldots (v_{k-1}, v_k)))$ such that $v_1 = s$ and $v_k = t$, then k defines the number of workflow stages, i.e. data that 'flows' from s to t though $k-2$ algorithm nodes;
- vertex distribution vector X_k – proportions of vertices at workflow stages, e.g. a setting $X_k = (50\%, 25\%, 10\%, 10\%, 5\%)$ represents a scenario for $k = 5$ where half of the vertices are the user data vertices, 35% are the algorithm vertices and 5% are the number of the purpose vertices;
- minimum density d – the proportion of initially generated edges between any two workflow stages.

To generate a graph, we distribute $|V|$ vertices onto k workflow stages as per vector X_k. For any two workflow stages, the initial d of all possible edges are generated through a pseudo-random[5] selection of vertices. Then, to ensure that all vertices in V^U and V^A have at least one outgoing edge, and that all vertices in V^A and V^P have at least one incoming edge, we add additional edges with one of the vertices selected randomly. All edges going out of the user vertices are assigned integer valuations $\pi(e)$ through a uniform selection from a range of 1–100. Furthermore, for each purpose vertex $p \in V_p$, the purpose weight introduced in Equation 5.4 is set to $w_p = 1$. Then, we generate the set of constraints by selecting $|\mathcal{N}|$ distinct pairs of randomly selected user data vertices and purpose vertices, ensuring that for any $(s,t) \in \mathcal{N}$, there exists at least one path from s to t.

This way, we prepare three datasets with different configurations of the above parameters. To make the dataset reflect data processing scenarios similar to the one described in Section 5.2, we ensure that in addition to the user vertices and purpose vertices, the generated graphs have more than one layer of algorithm vertices. Firstly, to observe how

[4]GLPK (GNU Linear Programming Kit), `https://www.gnu.org/software/glpk/`.

[5]For details, see the Python Standard Library documentation: `https://docs.python.org/3/library/random.html`.

TABLE 5.1: Parameter configurations for datasets 1, 2 and 3.

	Dataset 1			**Dataset 2**	**Dataset 3**
	a	**b**	**c**		
$\|\mathcal{N}\|$	1 – 50	1 – 50	1 – 50	10	5
$\|V\|$	100	1000	100	150 – 5000	100 – 10000
k	5	5	5	3 – 50	5
X_k	NUnif	NUnif	Unif	Unif	NUnif
d	0	0	20%	0	0

the number of constraints affects the runtime of algorithms and graph utility, we generate dataset 1. We create this dataset in different variants: a variant with 100 vertices (1a), a variant with 1000 vertices (1b) and a variant with a minimum density of 20% (1c). To keep the number of data processing stages constant in the first experiments, we choose to generate graphs with three algorithm stages, which allows us to focus the available computing resources on scaling the number of constraints. To observe how the shape of the graph affects the runtime and utility, we apply a non-uniform vertex distribution ($X_k = (50\%, 25\%, 10\%, 10\%, 5\%)$ abbreviated as 'NUnif') in variants 1a and 1b, and a uniform vertex distribution ($X_k = (20\%, 20\%, 20\%, 20\%, 20\%)$ abbreviated as 'Unif') in variant 1c. Secondly, to observe the impact of the path length on the runtime and utility, we generate dataset 2 with graphs that have a constant number of paths, but differ in the number of data processing stages from relatively small (i.e. one layer of algorithm vertices linking the user and purpose vertices) to relatively large (i.e. 48 stages of data processing, such that the total length of the longest path is 50). Specifically, we first generate graphs with $|V| = 150$, $k = 3$, and vertices distributed uniformly such that $|V^U| = 50$, $|V^A| = 50$ and $|V^P| = 50$. Then, we keep generating new graphs by adding 50 additional vertices to the previous graph and connecting each vertex to the graph with a single edge. This way we extend the length of each path in the previous graph while keeping the number of paths constant. We also adjust the constraints such that they relate to the same paths as for the previously generated graph. Finally, to observe how the size of the graph affects the runtime and utility, we generate graphs of 100–10,000 vertices with a constant number of constraints. The exact parameter configurations are specified in Table 5.1.

We perform our experiments on the University of Southampton High Performance Computing service Iridis 4[6] which offers 750 compute nodes in total with dual 2.6 GHz Intel Sandybridge processors. Each compute node has 16 CPUs per node with 64 GB of memory and the maximum runtime of a job is 60 hours. In order to ensure that the average result has low variance, we repeat the experiments until we have at least 30 runs with a sufficiently low standard error (SE).

[6]The Iridis Compute Cluster, `https://cmg.soton.ac.uk/iridis`.

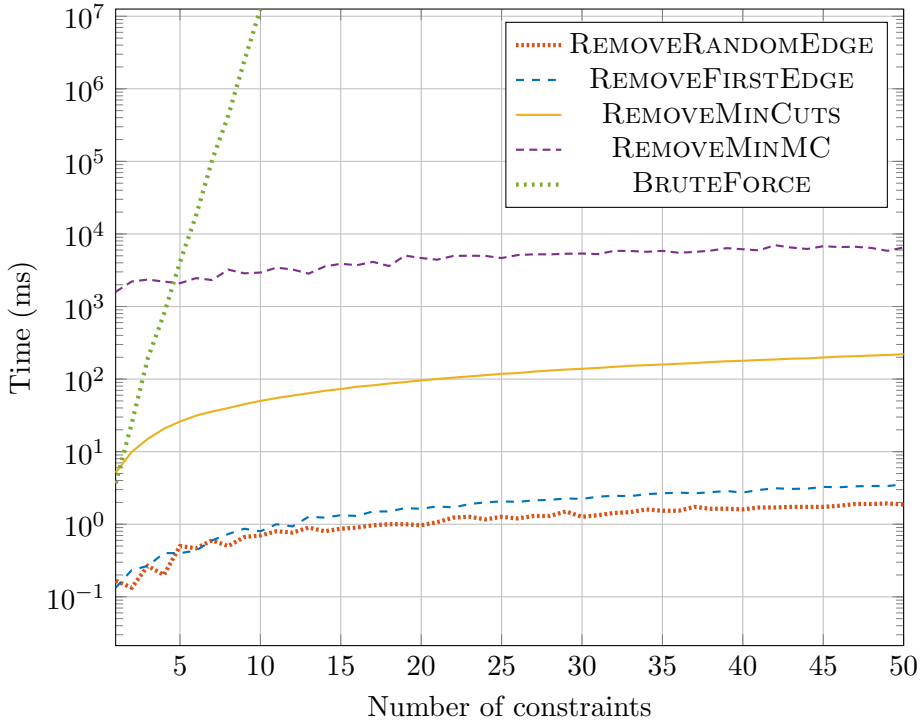


FIGURE 5.6: The number of constraints vs. the runtime of the algorithms in graphs (dataset 1a).

5.8.2 Results

We apply the algorithms from Section 5.6 to datasets 1, 2 and 3. In this section, we report on the runtime of the algorithms and changes in the graph's utility with respect to the number of privacy constraints, number of data processing stages and the size of the workflow.

We first observe the runtime of the algorithms as the number of privacy constraints grows. We measure their performance on 100-vertex graphs (dataset 1a) and compare it to the performance on 10 times larger, 1000-vertex graphs (dataset 1b). Since in datasets 1a and 1b the number of paths between the constraints is equal or close to the number of constraints given, we also consider slightly denser 100-vertex graphs, where the number of edges between each level of data processing is at least 20% of all possible edges (dataset 1c). In general, when we compare the average runtime on datasets 1a (Figure 5.6), 1b (Figure 5.7) and 1c (Figure 5.8), we observe very similar trends. As the graph size increases 10 times, the average runtimes of BruteForce, RemoveMinMC and RemoveMinCuts also increases approximately 10 times. This suggests that, as expected, the execution time of these three algorithms depends on the size of the graph and the number of constraints given as input.

In particular, we observe that the runtime of BruteForce increases rapidly with the increasing number of constraints, reaching an average time of 14838508.46 ms (i.e. over

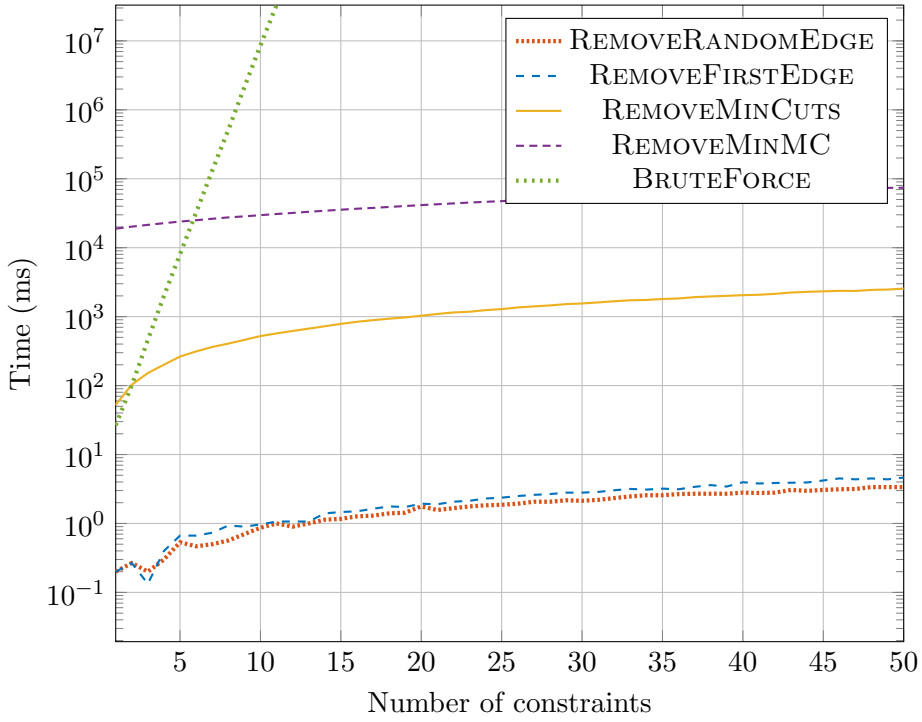


FIGURE 5.7: The number of constraints vs. the runtime of the algorithms in graphs (dataset 1b).

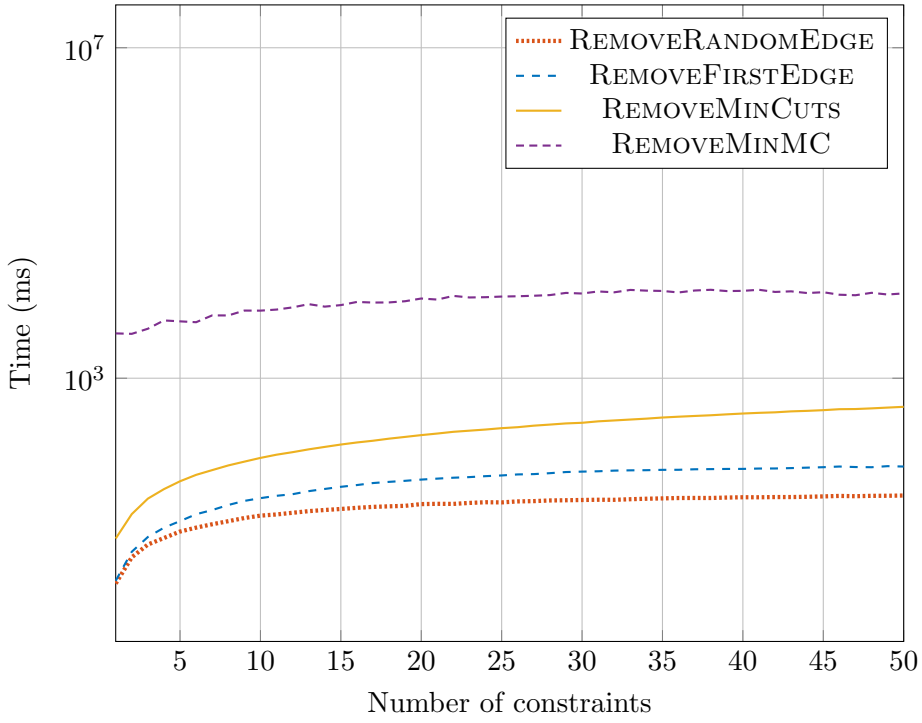


FIGURE 5.8: The number of constraints vs. the runtime of the algorithms in graphs (dataset 1c).

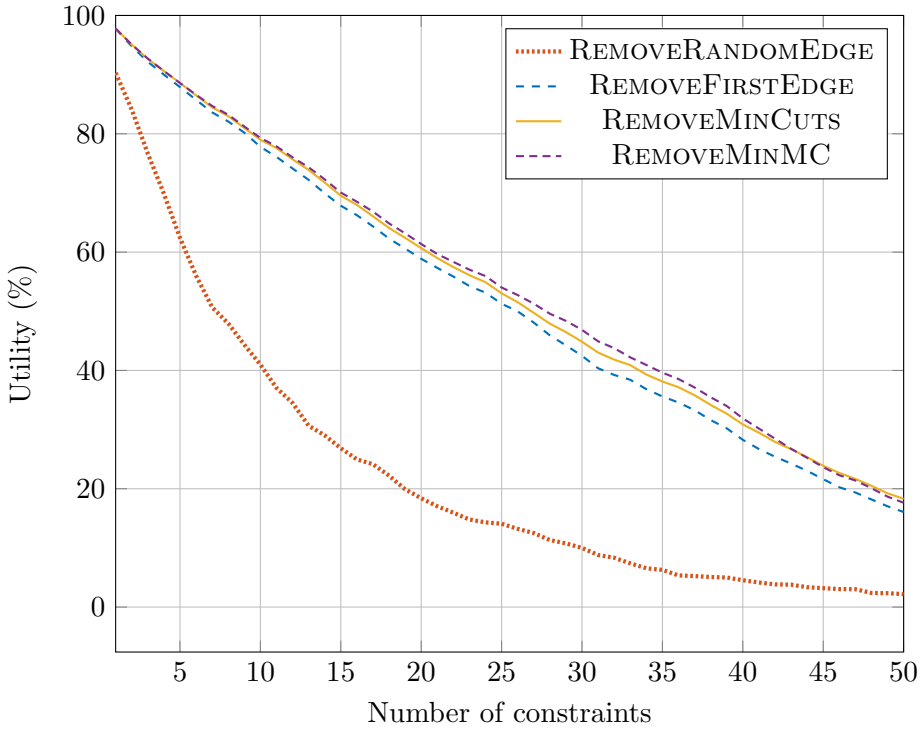


Figure 5.9: The number of constraints vs. graph utility after applying the algorithms on graphs (dataset 1a).

4 h) for just 10 constraints on dataset 1a and 8563968 ms (i.e. over 2 h; SE: 1058979.73 ms) on dataset 1b. For dataset 1c, in most cases, BruteForce is unable to return a result in 60 hours even for just a single pair of constraints. Thus, although BruteForce guarantees finding an optimal solution, its average runtime on such a small and sparse graphs makes this algorithm impractical. At the same time, the solver-based RemoveMinMC can reach an approximate solution for even 50 constraints on average in 6.51 seconds (SE: 405.02 ms) on dataset 1a, 74.1 seconds (SE: 439.96 ms) on dataset 1b and 11 seconds (SE: 290.47 ms) on dataset 1c. RemoveMinCuts can on average find an approximate solution for 50 constraints in 220.2 milliseconds (SE: 1.1 ms) on dataset 1a, 2.55 seconds (SE: 17.27 ms) on dataset 1b and 450.57 ms (SE: 3.17 ms) on dataset 1c.

Moreover, we observe the change in the graph's utility as the number of privacy constraints grows. Specifically, in Figure 5.9, we consider sparse graphs from dataset 1a. We can see that for RemoveMinMC, RemoveMinCuts and RemoveFirstEdge the utility of the graph after applying the algorithm decreases almost linearly as the number of constraints grows. Out of these three, RemoveMinMC tends to provide the most accurate solutions, reducing the utility down to 17.63% on average (SE: 1.15%) for 50 constraints. Although the exponential runtime of the BruteForce algorithm means we cannot run the experiments for more than 10 constraints, we still compare the results to RemoveMinMC for this limited setting. Results are presented in Table 5.2 and show that the utility using RemoveMinMC is nearly optimal in this case. In Section 5.7,

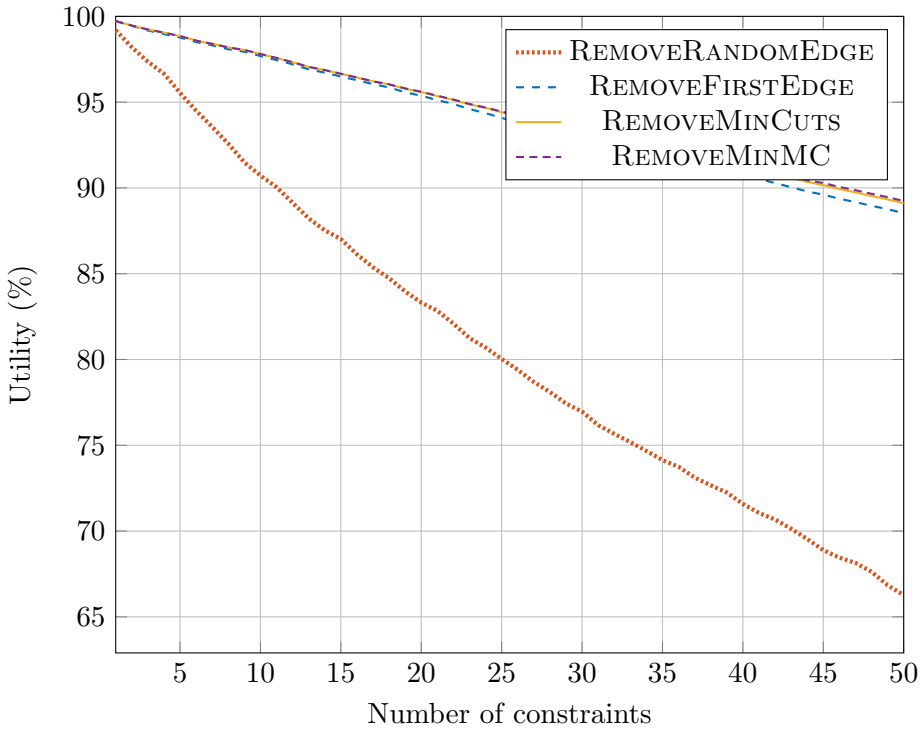


FIGURE 5.10: The number of constraints vs. graph utility after applying the algorithms on graphs (dataset 1b).

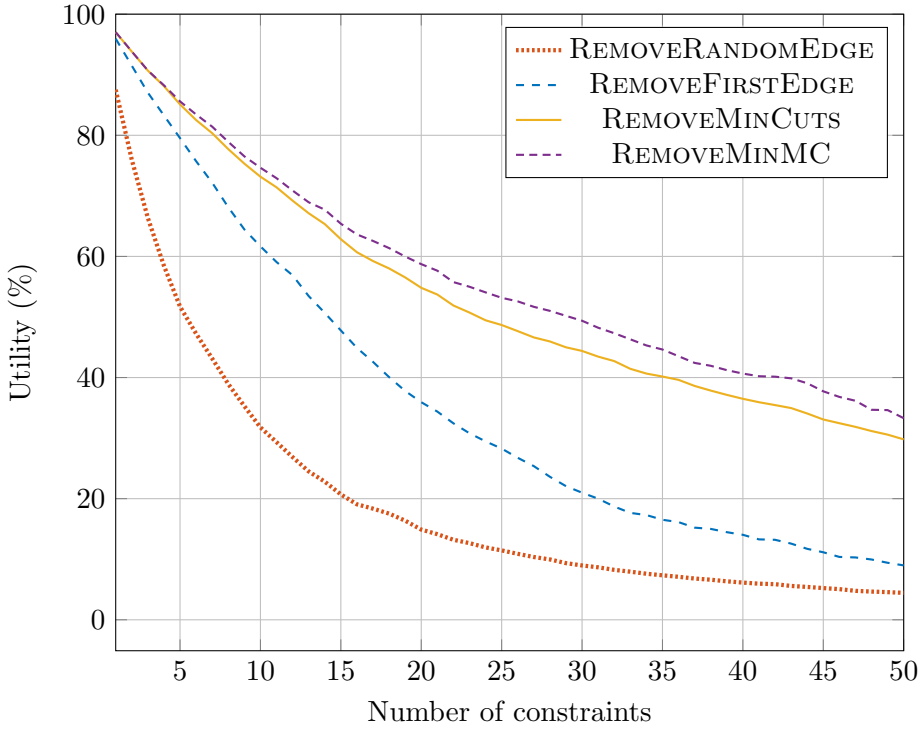


FIGURE 5.11: The number of constraints vs. graph utility after applying the algorithms on graphs (dataset 1c).

Table 5.2: Comparison of the graph's utility after applying RemoveMinMC and BruteForce.

Number of constraints	RemoveMinMC		BruteForce	
	% of original	SE	% of original	SE
1	97.79	0.32	97.79	0.32
2	95.08	0.35	95.08	0.35
3	92.71	0.58	92.71	0.58
4	90.62	0.57	90.63	0.57
5	88.59	0.75	88.65	0.75
6	86.59	0.72	86.66	0.72
7	84.71	0.72	84.77	0.71
8	83.22	0.70	83.28	0.70
9	81.24	0.69	81.33	0.69
10	79.30	0.69	79.39	0.68

we have shown that the algorithm is only guaranteed to be optimal for specific settings where the optimal solutions consists of only a single edge being removed from each path. Nevertheless, this empirical outcome suggests that, for graphs with a relatively small number of constraints RemoveMinMC is likely to provide very accurate solutions.

In Figure 5.10, we consider the utility changes in sparse graphs with 1000 vertices distributed non-uniformly (dataset 1b). As in Figure 5.9, RemoveMinMC provides solutions with the highest utility, i.e. on average 89.24% (SE: 0.19%) given 50 constraints. As expected, the utility here is higher than in Figure 5.9. After applying the algorithms with same number of constraints, proportionally less paths are broken in the graphs with 1000 vertices than with 100 vertices. At the same time, this results suggests that, for very large graphs, faster algorithms such as RemoveMinCuts or even RemoveFirstEdge may be able to provide sufficiently accurate solutions.

In Figure 5.11 we consider denser graphs with 100 vertices. We observe that the differences in utility between algorithms is more evident when the graphs are denser, resulting in significantly poorer performance especially for RemoveMinCuts, RemoveFirstEdge and RemoveRandomEdge. This is because denser graphs have more paths that need to be broken. Nonetheless, as previously, RemoveMinMC provides the best solution with the average utility being 33.29% (SE: 1.09%) of the original utility of the graph.

Next, we observe how the execution time depends on the number of paths between pairs of vertices that connect the constraints. Figure 5.12 presents a scatter plot of the runtime of the algorithms and distribution of utility in dense graphs (dataset 1c). In particular, we can see that, in case of RemoveMinCuts and RemoveMinMC, the runtimes increase almost linearly with respect to the number of paths. However, the execution times for these two algorithms differ significantly. For example, for a graph where 822 paths are required to be broken, RemoveMinMC takes 12116 ms to return a solution, whereas RemoveMinCuts can provide one in only 472 ms. Similarly, we can

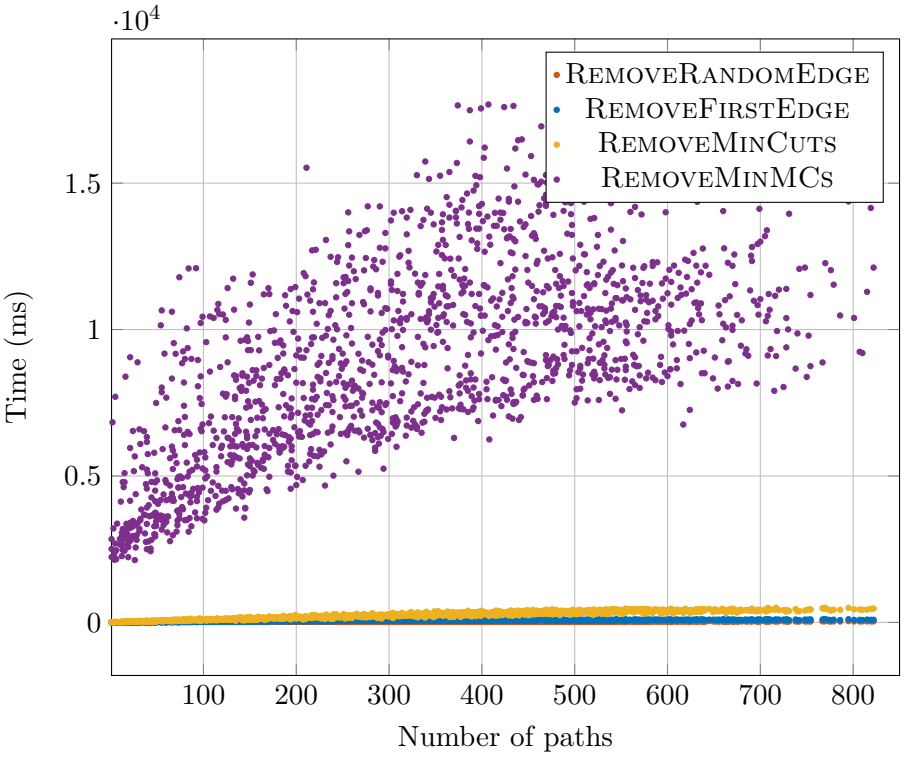


FIGURE 5.12: Number of paths vs. runtime (dataset 1c).

see that the utility decreases as the number of paths connecting constraints increases. Yet again, the utility after executing RemoveMinMC tends to decrease the slowest, reaching on average the utility of 32.27% of the original utility for the graph with 822 paths to be broken. For the same graph, the next best solution is RemoveMinCuts with an accuracy of 24.58%. For comparison, RemoveFirstEdge achieves on average a utility of 15.73% .

Next, we apply the algorithms to graphs with a constant number of paths. Since, in the linear model, only edges connected to the purpose vertices affect the utility of the graph, increasing the lengths of the paths on its own does not affect the utility. Thus, in this experiment, we focus on the execution time of the algorithms as the length of the paths grows. In Figure 5.14, we consider sparse graphs with vertices distributed uniformly with the same number of user data vertices as purpose vertices (dataset 2). We can see that as the path length grows, the runtime in case of BruteForce increases faster than the others.

Lastly, we analyse how the number of vertices in the graph impacts the runtime and the utility of the graph after applying the algorithms. To do this, we run the algorithms on sparse graphs of sizes between 100 and 10000 vertices and corresponding sets of 10 constraints (dataset 3). As the number of paths between the constraints and their length are equal for these graphs, in Figure 5.15 we can see that the size of the graph has only a slight impact on the execution time for BruteForce. In addition, RemoveMinCuts

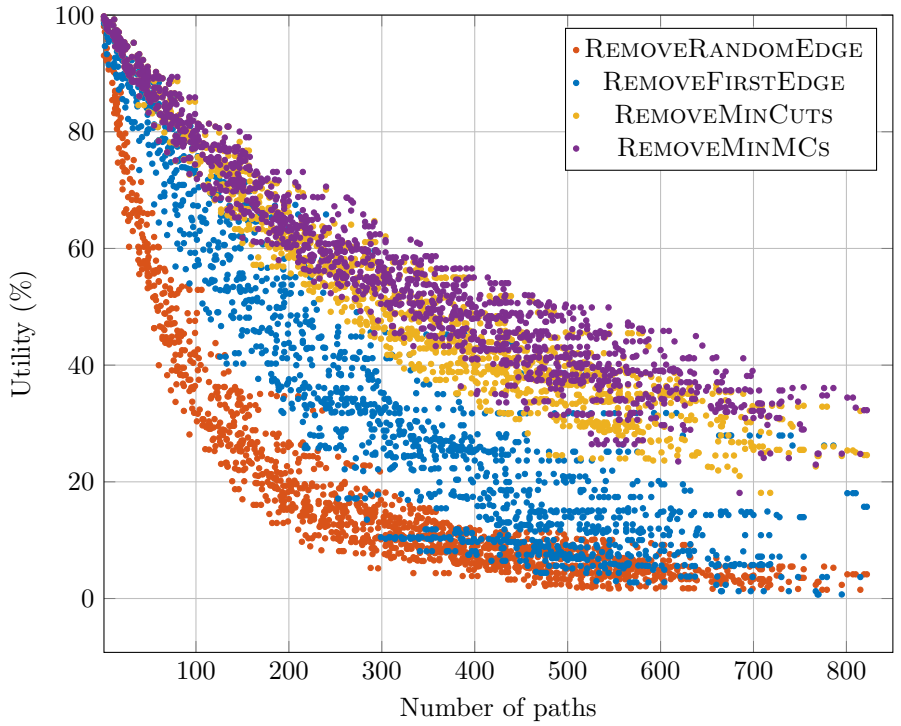


FIGURE 5.13: Number of paths vs. utility (dataset 1c).

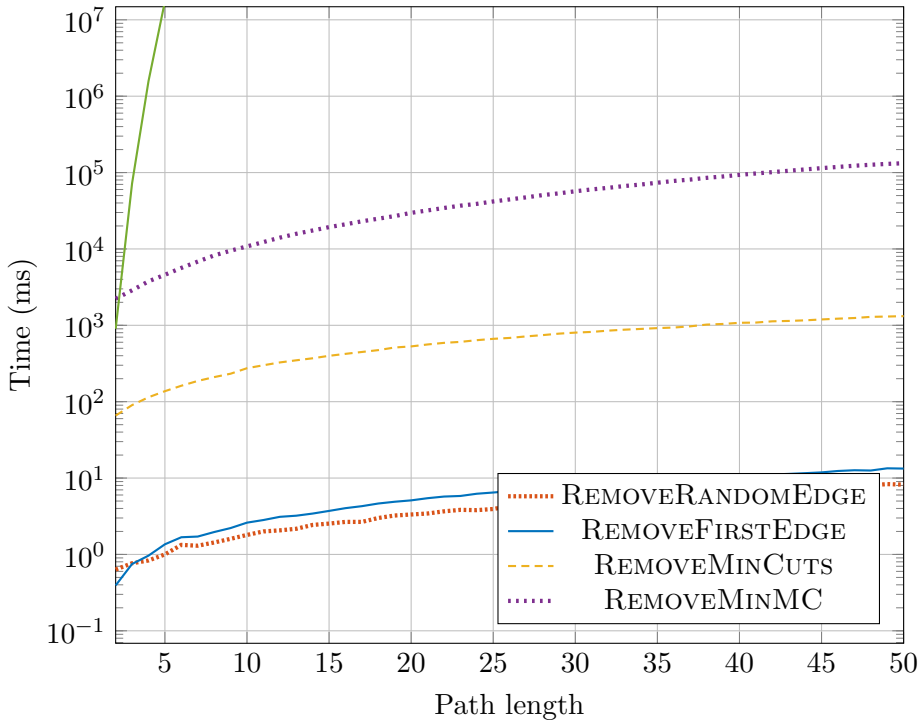


FIGURE 5.14: Path length vs. time in sparse graphs (dataset 2).

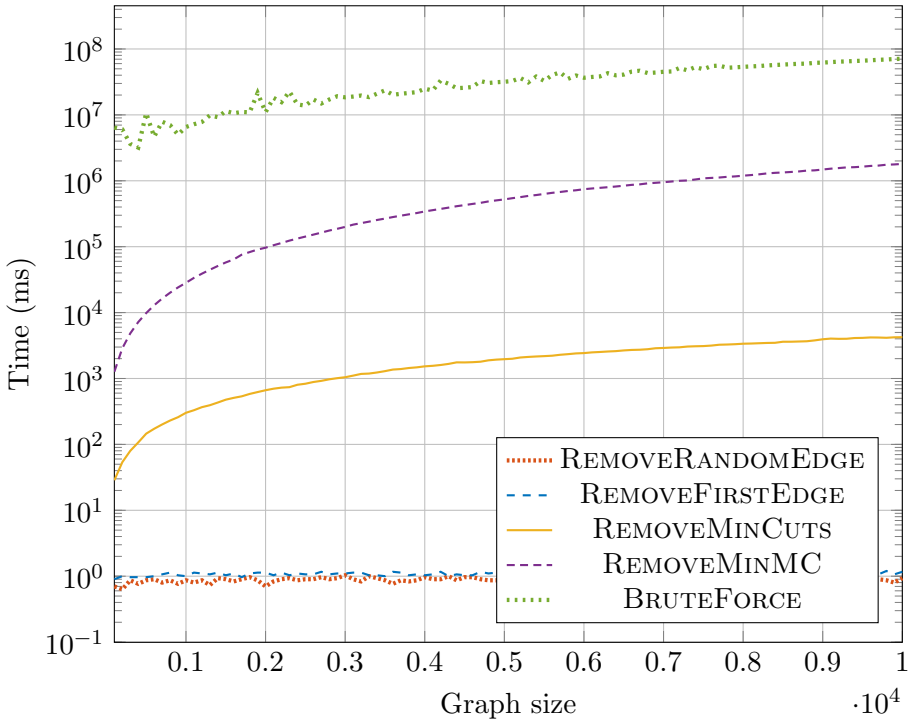


FIGURE 5.15: Graph size vs. runtime (dataset 3).

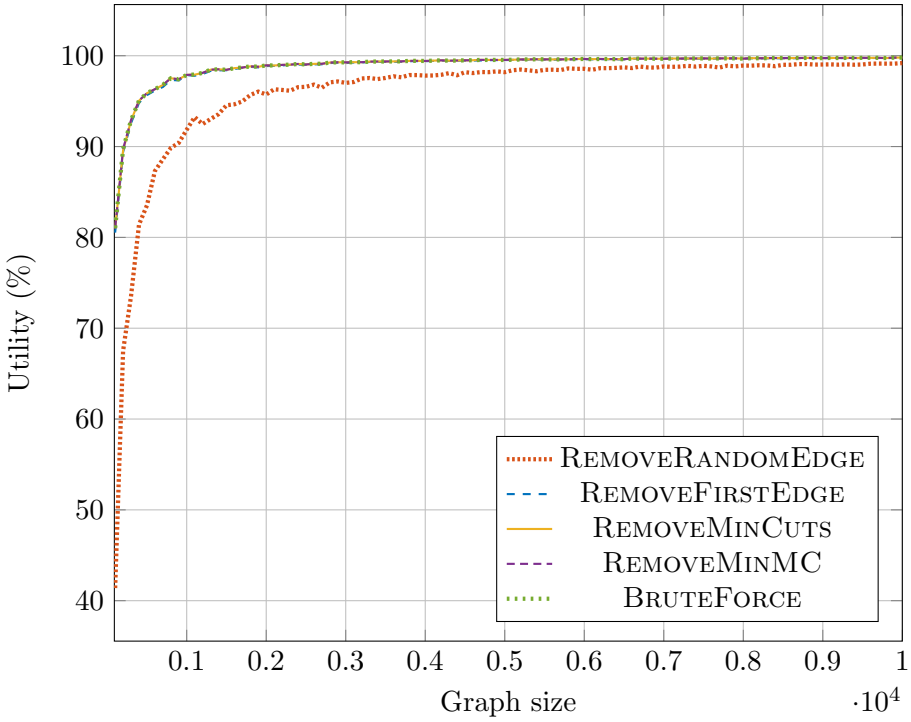


FIGURE 5.16: Graph size vs. utility (dataset 3).

is faster on average compared to BruteForce and RemoveMinMC. Considering the utility, Figure 5.16 shows that the graph size does not have a significant impact on the utility when the number of paths between the constraints and their length remain equal for the graphs.

5.9 Discussion

As the scale of data processing is rapidly increasing, we must ensure that consent is *implementable*. To that end, we focus on the problem of identifying consented data processing based on the privacy agreement between the service provider and an individual user. More specifically, if the privacy agreement constrains the flow of information that is processed, then by applying privacy constraints on the data processing model, we can establish which parts of the data processing the user has consented to.

In this chapter, we address research question **RQ3**, exploring methods that the service provider can apply to satisfy the user's privacy constraints optimally. To that end, we propose a new data processing model and formalise a problem of applying such constraints on it. Notably, the novelty of our approach is that the consent mechanism proposed benefits both the user and the service providers. This is especially important for automated consent, because as more users gain privacy knowledge, more and more privacy constraints may be placed by them on the processing of their personal data.

Interestingly, there is an extensive collection of open problems and challenges around such workflows. First, our theoretical results show that the problem in general is NP-hard. This result provides us with the lower bound on its complexity. The upper bound, however, depends largely on the complexity of the selected valuation and utility functions. Future work should investigate the upper bound of the problem with different functions that depend on the application.

Second, we focused on a specific instance of the problem, where the importance of an outputted data type is linearly additive with respect to the input. In addition, we assumed that the utility functions of specific data processing purposes are also linearly additive with respect to the importance of the data processed for the purpose. For this instance of our problem, our algorithm RemoveMinMC can find very accurate approximate solutions in seconds even for large workflow graphs. Nonetheless, the complexity of the problem in this additive case remains unknown. Further investigation is needed not just to study even more efficient and optimal algorithms, but also to establish the bounds on its complexity.

Third, some of our heuristic algorithms rely on the simplifying assumption that the value of different information sources is additive. While this can be a reasonable approximation in some settings, in practice the value may be subadditive (e.g. in the case

of redundant data) or superadditive (when data complements each other). At the moment, finding realistic data to create large real-world models and design algorithms for them is challenging. However, as data processing systems keep expanding, future work should focus on more realistic workflow models.

In addition, there are several open problems regarding scalability of the solution. Currently, the solution needs to be recomputed every time a new user enters the system or when an existing user updates their constraints. What is more, every time a change is made, some of algorithms that process the data would need to be re-run as well, which could be costly. At the same time, there could be many users of the same type, i.e. with similar privacy constraints, and a limited number of different user types which can be known in advance. To take advantage of this, users of the same type could e.g. be treated as a single user to enable the system to cope with thousands and even millions of users. This way, if new users enter the system, a new solution can be found quickly. Generally, as there are more and more users with privacy constraints, new methods are needed that take into account scalability by re-using some of the computation performed for the previous solution, as well as the costs of making changes.

Furthermore, new user interfaces are needed to collect privacy constraints from users. As data processing systems become more complex, explaining the complexity of data workflows to the user in a way that is usable, transparent and empowers them to make consentful decisions regarding the processing of their personal data becomes more and more challenging.

Finally, there are plenty of opportunities to consider richer types of privacy constraints and user preferences. For example, users may have constraints on combinations of different data types for a specific purpose (e.g. a user may say '*I'm okay with you using my data for advertising, but don't combine my location with my purchase history*'), processing the data types by specific service providers (e.g. '*I'm okay with you sharing my purchase history with anyone but Nile*') or time restrictions on data processing (*'I'm ok with you sharing my purchase history with Nile, but I don't want them to keep it for more than 30 days*'). For such constraints, future work should formulate new problems around consented data workflows.

5.10 Limitations

Our research is not without limitations. Firstly, the additive version of the problem we are focusing on is a simplistic model. In practice, the importance and utility functions are likely to be more advanced. Moreover, the importance and utility functions may differ across edges and purpose vertices. While some of the algorithms we propose may still be applicable to other importance and utility functions with minor modifications, more sophisticated approaches to the problem are needed when assumptions are relaxed.

Secondly, we run our experiments on Iridis 4, the High-Performance Computing service that is available to us. In particular, Iridis 4 imposes a limit on the run time and memory resources a process can use, which is reflected in our experiments. To provide the exact comparison of the execution times of Algorithm 5 CheckFeasibleMCs to the other algorithms, longer running time is required.

Finally, the problem we focus on in this chapter is only one of the problems within this space of consented data processing. Other problems include constraints on combining different data types for a specific purpose (e.g. a user may say '*I'm okay with you using my data for advertising, but don't combine my location with my purchase history*'), processing the data types by specific service providers (e.g. '*I'm okay with you sharing my purchase history with anyone but Nile*') which would require labelling the vertices with the service provider or time restrictions on data processing ('*I'm ok with you sharing my purchase history with Nile, but I don't want them to keep it for more than 30 days*').

5.11 Summary

In this chapter, we address the problem of making consent *implementable*. In particular, this means that when the privacy agreement is reached and the user refuses to consent to some forms of data processing, that refusal must be implemented by the service provider(s) to ensure consented information flow in the data processing system.

To that end, we designed a mechanism where the data flow in the processing system is represented as a directed graph and privacy constraints – as pairs of vertices that must be disconnected. Furthermore, we modelled our problem as an optimisation problem where the constraints must be satisfied. However, our theoretical results show that the problem is NP-hard.

Then, we focused on a specific instance of the problem, where the importance of an outputted data type is linearly additive with respect to the input. In addition, we assumed that the utility functions of specific data processing purposes are also linearly additive with respect to the importance of the data processed for the purpose. For this instance of our problem, we proposed five different algorithms that can provide feasible and optimal solutions.

Our findings have valuable implications for the service providers implementing automated consent mechanisms. In particular, we show that although computing the optimal solution can be very time-consuming, the proposed approximations can provide very accurate alternatives. Moreover, our theoretical results suggest that once the user's constraints are satisfied, when an update is requested with additional constraints, it is more beneficial for the service provider to re-compute the solution for all the constraints

than to add new constraints to the existing model. In general, establishing the importance of the data type helps significantly to effectively determine the point in the data processing system where the user's privacy constraints should be applied to gain the most utility. With regard to that, the study opened up new problems for further exploration of consented data processing.

Chapter 6

Conclusions

Consent is more than merely clicking 'I agree'. It is a *process* that defines the relationship between the user and the service provider. As part of this process, they reach a privacy agreement which should thereafter be honored in all aspects of data processing. Although concerns and preferences of both are equally important, the two sides are 'informationally asymmetric': the service provider can only gain access to the data through consent and the user may not be aware of the consequences they consent to.

Therefore, in this thesis, we propose that privacy engineering mechanisms that empower consent should be *informed*, *negotiable* and *implementable*. However, doing so requires solving three main problems: understanding how to meaningfully inform about the consequences of consent, how to enable negotiation of consent and how to implement consent to make sure it is complied with.

As we discuss in Chapter 2, previous research partially addressed some of these problems. There, we provide a brief background on the legal requirement of consent, as well as consent decision-making, its automation, negotiation and enterprise infrastructure. Yet, the existing body of knowledge is incomplete: we do not have a full understanding of how privacy knowledge affects consent decisions, we do not have sufficiently accurate privacy preference models to automate consent negotiations and we do not have ways to ensure that consent (and, in particular, the refusal of thereof) is implemented accordingly.

Since filling these research gaps is not trivial, we approach each of these problems separately. To this end, in Section 1.2 we posed three questions that guided the research presented in this thesis. Firstly, in Chapter 3, to answer **RQ1**, we focused on the role of privacy knowledge in users' privacy behaviour. In this regard, we studied the impact of procedural and factual privacy knowledge on the adoption of online tracking countermeasures. Our findings suggest that procedural knowledge on how users can control their privacy has impact on their consent decisions. Not only does this highlight the importance of articulating available control options in consent mechanisms, but also

helps us better understand what factors motivate updates to users' privacy preferences and consent. We also found that provision of factual privacy knowledge in addition to procedural knowledge does not significantly increase the impact on users' privacy behaviour and that the intent to adopt countermeasures does not necessarily lead to their adoption.

Secondly, to address **RQ2**, we explore privacy mechanisms that enable the negotiation of consent where the preferences of the user are automatically taken into account. That is, in Chapter 4, we extend the work of Baarslag et al. (2017) with a theoretical framework for automated consent negotiations between the user and the service provider. In addition, we compare the agent of Baarslag et al. (2017) to a new variant which personalises the agent's privacy preference profile to the individual user. We find that while this form of negotiation is a powerful mechanism in general, when the number of historical negotiations is small, classifying users into privacy profiles allows the agent to represent the user more accurately than personalising the preference profile based on data from a single user. Importantly, our framework generalises to other kinds of negotiation with the asymmetry of power between the negotiating parties, beyond privacy and consent.

Finally, to address **RQ3**, in Chapter 5, we take the first step towards enabling the service provider to implement consent. To that end, we model the flow of information in the service provider's data processing system and propose algorithms that allow service providers to satisfy constraints of the user. Specifically, given consent constraints from the user, the algorithms determine the stages of data processing that cannot be performed on the user data. In general, we find that the problem is $\mathcal{NP}$-hard. Nonetheless, for specific instances of the problem it is possible to design efficient and very accurate approximation methods. Apart from the regulatory compliance, our approach can also serve as a catalyst for transparency and explainability of data processing. Notably, our graph-theoretic contributions generalise to other graph-cutting problems with a similar additive-weight model.

More specifically, in this thesis we make the following contributions:

1. We show that privacy knowledge affects users' actual privacy behaviour in context of online tracking (Chapter 3). Building upon research of Gomer (2018) on the *intent* of countermeasure adoption, our research indicates that:

   (a) The provision of procedural privacy knowledge has an effect on users' actual privacy behaviour,

   (b) The provision of factual privacy knowledge in addition to procedural knowledge does not have a larger effect on users' actual privacy behaviour than the provision of procedural knowledge only.

   (c) If privacy knowledge is provided, intent to withdraw implicit consent does not necessarily lead to actual withdrawal of implicit consent.

2. We provide a browser extension to log users' countermeasure adoption behaviour (Chapter 3).

3. We generalise the agent of Baarslag et al. (2017) to propose a novel negotiation framework for multi-issue bilateral negotiations where the roles of the proposer and the responder are asymmetric, such as a buyer and a seller, or a user and a service provider (Chapter 4). As part of the framework, we propose a new alternating offers protocol with costly quoting, in which one agent can propose partial offers by specifying values for some issues, and the other agent completes those offers.

4. We propose a new user preference elicitation method personalised to the needs of an individual user, while the method published by Baarslag et al. (2017) relied on privacy profile classification (Chapter 4).

5. We present results of a new user study with 66 participants. We also re-analyse the data collected during the study published by Baarslag et al. (2017) and compare it to the data from our new experiment (Chapter 4).

6. We propose a novel approach to finding the most optimal ways of satisfying the user's privacy constraints in large-scale data workflows (Chapter 5). To this end, we formulate the consent problem as an optimisation problem where pairs of graph vertices must be disconnected such that utility is maximised. Furthermore, we prove that this problem is $\mathcal{NP}$-hard.

7. We propose a simplified instance of the Consented Data Workflow problem where the utility is linearly additive and five algorithms that can satisfy the privacy constraints (Chapter 5).

8. We evaluate our algorithms using synthetically generated data and show that our algorithms can provide a nearly optimal solution in the face of tens of constraints and graphs of thousands of nodes in a few seconds (Chapter 5).

Overall, we make the first step in advancing the technological aspects of consent mechanisms online. While we recognise that privacy is a complex issue and technology on its own cannot address all concerns around consent in handling privacy-sensitive data, we hope that these advances in privacy engineering, together with appropriate regulations, societal norms and corporate practices, can help building trust between users and service providers. However, despite these advances, there are many remaining challenges in this area. In Section 6.1, we outline some exciting directions for future work that our research opens.

6.1 Future Work

Throughout this thesis, we focus on making consent *informed*, *negotiable* and *implementable*. In that regard, our work found many promising avenues for further exploration. Specifically, we envision the following potentially fruitful research directions:

- **Privacy knowledge and consent.** Our study in Chapter 3 focuses on factual and procedural privacy knowledge. Nonetheless, another interesting dimension of privacy knowledge is experiential privacy knowledge, which refers to the users' general familiarity with technology and first-hand experience with privacy violations. In particular, as the number and scale of known privacy breaches increases and so increases the number of daily online users, this kind of knowledge may have increasingly greater impact on consent decisions. However, since users acquire experiential knowledge over a period of time, longitudinal studies are necessary to adequately assess its impact on users' consent (or the refusal of thereof). In fact, to explore this, one could conduct multiple Westin-style studies, c.f. Kumaraguru and Cranor (2005). Notably, in order to do so, relevant instruments need to be developed in order to perform a detailed assessment of users' privacy knowledge. Given that online privacy is a broad topic and spans over all of our daily data-sharing activities, this is particularly challenging. Yet, contributions in this area could provide insights crucial for engineering consent mechanisms – not just ones that provide greater opportunities for informed choice, but also for a more accurate automation of that choice.

- **Consent decision predictions.** With an ever-growing number of consent requests, detailed insights on how exactly users consent could help us develop models for consent decision predictions. To that end, machine learning methods could play a crucial role. Thus, apart from understanding the exact factors that come into play during consent decision-making such as privacy knowledge researched in Chapter 3, it would be interesting to see whether users could potentially be clustered according to their consent behaviour. If so, such *consent behaviour types* could provide a powerful instrument for building accurate privacy preference profiles of users for consent negotiations with service providers in a similar way as we use the Westin Segmentation Index in Chapter 4.

- **Service provider's agent for consent negotiations.** In Chapter 4, we proposed a framework for consent negotiations. To begin our exploration, we chose to focus on the user's agent and make certain assumptions about the service provider. However, designing the agent for the service provider is another interesting research direction. What is particularly exciting here is the preference profile of the service provider, because their valuation of specific data types is not necessarily equal or linearly additive. For example, if the user is willing to share two data types,

the service provider may not consider them both equally important. In addition, their valuation is not necessarily equal to the sum of the valuation of each of them shared individually (e.g. a permission to access browsing history may be valuable to the service provider, but additional access to text messages may not increase the valuation much, even if text messages themselves with the browsing history may also be very valuable). In general, the valuation of personal data on the service provider's side is a broad topic.

- **Data processing models for consent implementation.** Accurate the valuation methods for types of users' personal data could also be used to develop more accurate data processing models for consent implementation. To begin our exploration, in Chapter 5, we assume that the valuation is linearly additive, and design our model and algorithms accordingly. However, in practice, finding efficient and accurate algorithms for optimal consent satisfaction is not trivial – in fact, we prove that when this assumption is relaxed completely, the problem becomes NP-hard. In particular, privacy constraints can be expressed in different ways which open new opportunities for algorithmic contributions (e.g. consent expressed by the user as: '*I'm ok with you using my data for advertising, but don't combine my location with my purchase history*'). Moreover, automated creation of such data processing models for existing services that are already on the market is another interesting challenge with opportunities for multiple algorithmic contributions.

- **Articulating privacy constraints and preferences.** Last but not least, new user-centred interfaces are needed to enable users to articulate consent in a fine-grained way. In particular, when consent is negotiated, novel ways of communicating privacy preferences can improve the agent's model of the user's privacy preference profile. However, as discussed in Chapter 5, there may be many stages of data processing and many possible consequences of granting consent. Therefore, research within the area of human-computer interaction is required to analyse how the constraints and preferences of the user should be communicated to the agent and through the agent to the service provider.

These research directions complete the vision of making consent *informed*, *negotiable* and *implementable*. However, apart from these characteristics, there are also other important aspects of engineering mechanisms for consent that call for further exploration. For example, one of them is about making consent *time-limited*. In that line of research, future work should explore mechanisms for ensuring that consent is granted not just for a specific purpose to a specific service provider, but also for a specific period of time.

Appendix A

Questionnaire

1. How old are you?
   (a) 18-25,
   (b) 26-30,
   (c) 31-40,
   (d) 41-50,
   (e) 51-60,
   (f) 61+.
2. What is your gender?
   (a) female,
   (b) male,
   (c) I think of myself in another way,
   (d) prefer not to say.
3. What is your nationality? (Select from list)
4. What is the highest degree of education you have completed?
   (a) None,
   (b) Pre-university Qualification (A-level/AS-level/IB or equivalent),
   (c) Foundation Year,
   (d) Bachelor's Degree,
   (e) Master's Degree,
   (f) PhD,
   (g) Trade/Technical/Vocational Training,
   (h) Prefer not to say.

5. How many hours a day do you spend browsing websites?
6. From the stand point of personal privacy, please indicate the extent to which you, as an individual, agree or disagree with each statement by circling the appropriate number (Chellapa and Sin's survey instrument).
7. Have you used* the following:
   (a) Facebook,
   (b) Twitter,
   (c) Ghostery,
   (d) WhatsApp,
   (e) Adblock Plus,
   (f) Instagram,
   (g) Disconnect,
   (h) Spotify,
   (i) Tor.

   * – for each one, select:

   - No, I never heard of it.
   - No, but I heard of it.
   - Yes, I tried it.
   - Yes, I use it regularly.
   - Yes, I used to use it regularly before, but I don't use it now.
8. Have you ever cleared browsing history?**
9. Have you ever deleted browser cookies?**
10. Have you disabled third party cookies in your browser?**
11. Have you enabled opt-out in your browser?**

    ** – select:

    - No, I never heard of it.
    - No – I heard of it, but I don't know why I would need to do it.
    - No, but I heard of it and understand why others do it.
    - Yes, I've done it once in the past year.
    - Yes, I've done it before, but I don't do it now.
    - Yes, I have.

Appendix D. Experiment 3 Intent Questionnaire

The information below describes ways of preventing online tracking. Please read it carefully and answer all the questions.

Ad blockers. Ad blockers are extensions for Web browsers such as Google Chrome that block adverts on the Web pages you visit. They prevent a lot of the tracking that is conducted by advertising networks. An example of an ad blocker is AdBlock Plus.

Some websites that rely on advertising revenue block ad blocker users from viewing their content. You might be asked to disable Adblock Plus when you visit those websites.

To install AdBlock Plus, you can:

- Visit AdBlock Plus on the Chrome Web Store (link to AdBlock Plus on the Chrome Web Store included here) and click "Install". Adblock Plus will be installed, and ready to use shortly.

1. Do you intend to install Adblock Plus? (Yes/No)
2. What are some reasons?

Disconnect. Disconnect is a Web browser extension that blocks many 'requests' (connections) that your browser makes to known tracking websites, preventing them from tracking you.

Some websites that rely on advertising revenue block Disconnect users from viewing their content. You might be asked to disable Disconnect when you visit those websites.

To install Disconnect, you can:

- Visit Disconnect on the Chrome Web Store (link to Disconnect on the Chrome Web Store included here) and click "Install". Disconnect will be installed, and ready to use shortly.

3. Do you intend to install Disconnect? (Yes/No)
4. What are some reasons?

Clearing cookies. Clearing all of the cookies in your Web browser removes the tracking cookies that are used to identify you each time you visit a website. It will cause most trackers to "forget" about your computer.

Some trackers use other means to identify you, such as fingerprinting, so clearing cookies might not be 100% effective at preventing tracking.

Clearing cookies will cause you to be logged out of all websites and have to sign in again.

Some websites might forget personalisation settings.

Clearning your cookies does not prevent tracking from taking place again in the future.

To clear your browser cookies, you can:

- In the top right of Google Chrome, click “More”.
- Select “More Tools” and then click on “Clear Browsing Data”.
- In the ”Clear browsing data” box, click the checkboxes for “Cookies and other site data” and “Cached images and files”.
- Use the menu at the top to select the amount of data that you want to delete. Choose “beginning of time” to delete everything.
- Click “Clear browsing data”.

5. Do you intend to clear browser cookies? (Yes/No)

6. What are some reasons?

Bibliography

Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. The web never forgets: Persistent tracking mechanisms in the wild. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 674–689. ACM, 2014.

Alessandro Acquisti. Protecting privacy with economics: Economic incentives for preventive technologies in ubiquitous computing environments. In *Proceedings of Workshop on Socially-informed Design of Privacy-enhancing Solutions, 4th International Conference on Ubiquitous Computing (UBICOMP 02)*, 2002.

Alessandro Acquisti. Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, pages 21–29. ACM, 2004.

Alessandro Acquisti. Nudging privacy: The behavioral economics of personal information. *IEEE Security & Privacy*, 7(6), 2009.

Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Privacy and human behavior in the age of information. *Science*, 347(6221):509–514, 2015.

Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *International Workshop on Privacy Enhancing Technologies*, pages 36–58. Springer, 2006.

Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 2(2005):24–30, 2005a.

Alessandro Acquisti and Jens Grossklags. Uncertainty, ambiguity and privacy. In *Fourth Workshop on the Economics of Information Security (WEIS)*, pages 2–3, 2005b.

Nabil R. Adam and John C. Worthmann. Security-control methods for statistical databases: a comparative study. *ACM Computing Surveys*, 21(4):515–556, 1989.

Idris Adjerid, Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Sleights of privacy: Framing, disclosures, and the limits of transparency. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, pages 1–11. ACM, 2013.

Yuvraj Agarwal and Malcolm Hall. Protectmyprivacy: detecting and mitigating privacy leaks on ios devices using crowdsourcing. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 97–110. ACM, 2013.

Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. Hippocratic databases. In *Proceedings of the 28th International Conference on Very Large Databases (VLDB'02)*, pages 143–154. Elsevier, 2002.

Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. Implementing p3p using database technology. In *Proceedings 19th International Conference on Data Engineering (Cat. No. 03CH37405)*, pages 595–606. IEEE, 2003a.

Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. An xpath-based preference language for p3p. In *Proceedings of the 12th international conference on World Wide Web*, pages 629–639. ACM, 2003b.

Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. XPref: a preference language for P3P. *Computer Networks*, 48(5):809–827, 2005.

Reza Ghaiumy Anaraky, David Cherry, Marie Jarrell, and Bart Knijnenburg. Testing a comic-based privacy policy. In *Proceedings of the 15th Symposium on Usable Privacy and Security*. USENIX, 2019.

Julia Angwin. The web's new gold mine: Your secrets. In Hesse Pesta, Julia Angwin, Scott Thurm, Steve Yoder, and Mitch Pacelle, editors, *What They Know: The Business of Tracking You on the Internet. A Wall Street Journal Investigation*, pages 7–14. The Wall Street Journal, 2010. [cited: 24 Mar 2021]. Available from: `http://www.cs.cornell.edu/~shmat/courses/cs5436/whattheyknow.pdf`.

Pauline Anthonysamy, Awais Rashid, and Ruzanna Chitchyan. Privacy requirements: present & future. In *2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering in Society Track (ICSE-SEIS)*, pages 13–22. IEEE, 2017.

Sanjeev Arora and Boaz Barak. *Computational Complexity: a Modern Approach*. Cambridge University Press, 2009.

Paul Ashley, Satoshi Hada, Günter Karjoth, Calvin Powers, and Matthias Schunter. Enterprise Privacy Authorization Language (EPAL). Technical report, IBM Research, 2003.

Paul Ashley, Satoshi Hada, Günter Karjoth, and Matthias Schunter. E-p3p privacy policies and privacy authorization. In *Proceedings of the 2002 ACM Workshop on Privacy in the Electronic Society*, pages 103–109. ACM, 2002a.

Paul Ashley, Calvin Powers, and Matthias Schunter. From privacy promises to privacy management: a new approach for enforcing privacy throughout an enterprise. In *Proceedings of the 2002 Workshop on New Security Paradigms*, pages 43–50. ACM, 2002b.

Giorgio Ausiello, Pierluigi Crescenzi, Giorgio Gambosi, Viggo Kann, Alberto Marchetti-Spaccamela, and Marco Protasi. *Complexity and approximation: Combinatorial optimization problems and their approximability properties*. Springer Science & Business Media, 2012.

Brooke Auxier, Lee Rainie, Monica Andreson, Andrew Perrin, Madhu Kumar, and Erica Turner. Americans and privacy: Concerned, confused and feeling lack of control over their personal information. Technical report, Pew Research Center, Nov 2019. [cited 31 Jan 2021]. Available from: `https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/11/Pew-Research-Center_PI_2019.11.15_Privacy_FINAL.pdf`.

Reyhan Aydoğan, David Festen, Koen V Hindriks, and Catholijn M Jonker. Alternating offers protocols for multilateral negotiation. In *Modern approaches to agent-based complex automated negotiation*, pages 153–167. Springer, 2017.

Tim Baarslag, Alan Alper, Richard Gomer, Muddasser Alam, Perera Charith, Enrico Gerding, et al. An automated negotiation agent for permission management. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)*. ACM, 2017.

Tim Baarslag and Enrico H Gerding. Optimal incremental preference elicitation during negotiation. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015a.

Tim Baarslag and Enrico H. Gerding. Optimal incremental preference elicitation during negotiation. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 3–9, 2015b.

Tim Baarslag, Enrico H. Gerding, Reyhan Aydoğan, and m. c. schraefel. Optimal negotiation decision functions in time-sensitive domains. In *Proceedings of the 2015 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 2, pages 190–197. IEEE, Dec 2015.

Tim Baarslag, Mark J.C. Hendrikx, Koen V. Hindriks, and Catholijn M. Jonker. Measuring the performance of online opponent models in automated bilateral negotiation. In Michael Thielscher and Dongmo Zhang, editors, *AI 2012: Advances in Artificial Intelligence*, volume 7691 of *Lecture Notes in Computer Science*, pages 1–14. Springer Berlin Heidelberg, 2012.

Tim Baarslag, Mark J.C. Hendrikx, Koen V. Hindriks, and Catholijn M. Jonker. Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems*, 30 (5):849–898, 2016.

Ovidiu Bagdasar. Elements of graph theory. In *Concise Computer Mathematics*, pages 73–80. Springer, 2013.

Rebecca Balebako, Jaeyeon Jung, Wei Lu, Lorrie Faith Cranor, and Carolyn Nguyen. Little brothers watching you: Raising awareness of data leaks on smartphones. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, page 12. ACM, 2013.

Rebecca Balebako, Pedro G Leon, Hazim Almuhimedi, Patrick Gage Kelley, Jonathan Mugan, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. Nudging users towards privacy on mobile devices. In *Proc. CHI 2011 Workshop on Persuasion, Nudge, Influence and Coercion*, 2011.

Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Cranor, Shomir Wilson, et al. Finding a choice in a haystack: automatic extraction of opt-out statements from privacy policy text. In *Proceedings of the 29th International Conference on World Wide Web*, pages 1943–1954. ACM, 2020.

Gaurav Bansal, David Gefen, et al. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems*, 49(2):138–150, 2010.

Susanne Barth and Menno D. T. De Jong. The privacy paradox–investigating discrepancies between expressed privacy concerns and actual online behavior–a systematic literature review. *Telematics and Informatics*, 34(7):1038–1058, 2017.

Miriam Bartsch and Tobias Dienlin. Control your facebook: An analysis of online privacy literacy. *Computers in Human Behavior*, 56:147–154, 2016.

Masooda Bashir, Carol Hayes, April D Lambert, and Jay P Kesan. Online privacy and informed consent: The dilemma of information asymmetry. *Proceedings of the Association for Information Science and Technology*, 52(1):1–10, 2015.

Stefano Belloro and Alexios Mylonas. I know what you did last summer: New persistent tracking mechanisms in the wild. *IEEE Access*, 6:52779–52792, 2018.

Scott Bender. Privacy in the cloud frontier: Abandoning the take it or leave it approach. *Drexel Law Review*, 4:487, 2011.

Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing*, 15(7):679–694, 2011.

Colin J Bennett. *Regulating privacy: Data protection and public policy in Europe and the United States*. Cornell University Press, 1992.

Marcel Bennicke et al. Towards automatic negotiation of privacy contracts for internet services. In *Networks, 2003. ICON2003. The 11th IEEE International Conference on*, pages 319–324. IEEE, 2003.

Cédric Bentz. On the hardness of finding near-optimal multicuts in directed acyclic graphs. *Theoretical computer science*, 412(39):5325–5332, 2011.

Bettina Berendt, Oliver Günther, and Sarah Spiekermann. Privacy in e-commerce: Stated preferences vs. actual behavior. *Communications of the ACM*, 48(4):101–106, 2005.

Alastair R Beresford, Andrew Rice, Nicholas Skehin, and Ripduman Sohan. Mockdroid: trading privacy for application functionality on smartphones. In *Proceedings of the 12th workshop on mobile computing systems and applications*, pages 49–54. ACM, 2011.

Hal Berghel. Malice domestic: The Cambridge analytica dystopia. *Computer*, 51(5): 84–89, 2018.

Sophie C. Boerman, Sanne Kruikemeier, and Frederik J. Zuiderveen Borgesius. Online behavioral advertising: A literature review and research agenda. *Journal of Advertising*, 46(3):363–376, 2017.

Bram Bonné, Sai Teja Peddinti, Igor Bilogrevic, and Nina Taft. Exploring decision making with android's runtime permission dialogs using in-context surveys. In *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*. USENIX Association, 2017.

Craig Boutilier, Ronen I. Brafman, Carmel Domshlak, Holger H. Hoos, and David Poole. Cp-nets: A tool for representing and reasoning withconditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research*, 21:135–191, 2004.

Laura Brandimarte and Alessandro Acquisti. The economics of privacy. In Martin Peitz and Joel Waldfogel, editors, *The Oxford Handbook of the Digital Economy*, chapter 20, pages 547–571. Citeseer, 2012.

Jack W Brehm. A theory of psychological reactance. 1966.

Sharon S Brehm and Jack W Brehm. *Psychological reactance: A theory of freedom and control*. Academic Press, 2013.

Aaron R Brough and Kelly D Martin. Critical roles of knowledge and motivation in privacy research. *Current Opinion in Psychology*, 31:11–15, 2020.

Lee A. Bygrave. Privacy protection in a global context–a comparative overview. *Scandinavian Studies in Law*, 47:319–348, 2004.

Ji-Won Byun, Elisa Bertino, and Ninghui Li. Purpose based access control of complex data for privacy protection. In *Proceedings of the 10th ACM Symposium on Access Control Models and Technologies*, pages 102–110. ACM, 2005.

Ji-Won Byun and Ninghui Li. Purpose based access control for privacy protection in relational database systems. *The VLDB Journal*, 17(4):603–619, 2008.

Alexandra J Campbell. Relationship marketing in consumer markets: A comparison of managerial and consumer attitudes about information privacy. *Journal of Direct Marketing*, 11(3):44–57, 1997.

Fred H. Cate. The failure of fair information practice principles. In Jane K. Winn, editor, *Consumer Protection in the Age of the Information Economy*, chapter 13, pages 343–379. Ashgate, 2006.

Fred H. Cate. The limits of notice and choice. *IEEE Security & Privacy*, 8(2):59–62, 2010.

Ann Cavoukian. Privacy by design: The 7 foundational principles. Technical report, Information and privacy commissioner of Ontario, Canada, 2009. [cited 15 Mar 2021]. Available from: `http://dataprotection.industries/wp-content/uploads/2017/10/privacy-by-design.pdf`.

Ramnath K. Chellappa and Raymond G. Sin. Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management*, 6(2):181–202, 2005.

Siqi Chen, Haitham Bou Ammar, Karl Tuyls, and Gerhard Weiss. Optimizing complex automated negotiation using sparse pseudo-input gaussian processes. In *Proceedings of the 12th International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pages 707–714, Richland, SC, 2013. International Foundation for Autonomous Agents and Multiagent Systems.

Vivying SY Cheng, Patrick CK Hung, and Dickson KW Chiu. Enabling Web services policy negotiation with privacy preserved using XACML. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, pages 33–33. IEEE, 2007.

Hanbyul Choi, Jonghwa Park, and Yoonhyuk Jung. The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior*, 81:42–51, 2018.

Lizzie Coles-Kemp and Elahe Kani-Zabihi. On-line privacy and consent: a dialogue, not a monologue. In *Proceedings of the 2010 New Security Paradigms Workshop*, pages 95–106, 2010.

Sunny Consolvo, Ian E Smith, Tara Matthews, Anthony LaMarca, Jason Tabert, and Pauline Powledge. Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the 2005 CHI Conference on Human Factors in Computing Systems*, pages 81–90, 2005.

Marie-Christine Costa, Lucas Létocart, and Frédéric Roupin. Minimal multicut and maximal integer multiflow: a survey. *European Journal of Operational Research*, 162 (1):55–69, 2005.

Lorrie Faith Cranor. *Web privacy with P3P*. O'Reilly Media, Inc., 2002.

Lorrie Faith Cranor. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *Journal on Telecommunication & High Technology Law*, 10:273, 2012.

Justin Cranshaw, Jonathan Mugan, and Norman Sadeh. User-controllable learning of location privacy policies with gaussian mixture models. In *AAAI*, 2011.

Mary J. Culnan and Pamela K. Armstrong. Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science*, 10 (1):104–115, 1999.

Elias Dahlhaus, David S. Johnson, Christos H. Papadimitriou, Paul D. Seymour, and Mihalis Yannakakis. The complexity of multiterminal cuts. *SIAM Journal on Computing*, 23(4):864–894, 1994.

Christophe Debruyne, Harshvardhan J Pandit, Dave Lewis, and Declan O'Sullivan. Towards generating policy-compliant datasets. In *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, pages 199–203. IEEE, 2019.

Christophe Debruyne, Harshvardhan J. Pandit, Dave Lewis, and Declan O'Sullivan. "just-in-time" generation of datasets by considering structured representations of given consent for gdpr compliance. *Knowledge and Information Systems*, 62(9):3615–3640, 2020.

Martin Degeling, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz. We Value Your Privacy... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. *Informatik Spektrum*, 42:345–346, 2019.

Department of Justice, Canada. Personal Information Protection and Electronic Documents Act, April 2000. [cited 20 Jan 2021]. Available from: `https://laws-lois.justice.gc.ca/eng/acts/P-8.6/FullText.html`.

Reinhard Diestel. Graph theory. *Graduate Texts in Mathematics*, 173:7, 2000.

Efim A. Dinic. Algorithm for solution of a problem of maximum flow in networks with power estimation. In *Soviet Mathematics. Doklady*, volume 11, pages 1277–1280. Doklady, 1970.

Yefim Dinitz. Dinitz'algorithm: The original version and Even's version. In Oded Goldreich, Arnold L. Rosenberg, and Alan L. Selman, editors, *Theoretical Computer Science: Essays in Memory of Shimon Even*, pages 218–240. Springer, 2006.

Curt J. Dommeyer and Barbara L. Gross. What consumers know and what they do: An investigation of consumer knowledge, awareness, and use of privacy protection strategies. *Journal of Interactive Marketing*, 17(2):34–51, 2003.

Cynthia Dwork. Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation*, pages 1–19. Springer, 2008.

Jack Edmonds and Richard M. Karp. Theoretical improvements in algorithmic efficiency for network flow problems. *Journal of the ACM*, 19(2):248–264, 1972.

Manuel Egele, Christopher Kruegel, Engin Kirda, and Giovanni Vigna. Pios: Detecting privacy leaks in ios applications. In *NDSS*, pages 177–183, 2011.

Isioma Elueze and Anabel Quan-Haase. Privacy attitudes and concerns in the digital lives of older adults: Westin's privacy attitude typology revisited. *American Behavioral Scientist*, 62(10):1372–1391, 2018.

William Enck, Peter Gilbert, Seungyeop Han, Vasant Tendulkar, Byung-Gon Chun, Landon P Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N Sheth. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems (TOCS)*, 32(2):5, 2014.

European Parliament and the Council of the European Union. Directive 95/46/ec of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal of the European Union*, 281:31–50, Oct 1995.

European Parliament and the Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, 119:1–88, May 2016.

Lujun Fang, Heedo Kim, Kristen LeFevre, and Aaron Tami. A privacy recommendation wizard for users of social networking sites. In *Proceedings of the 17th ACM Conference on Computer and Communications Security*, pages 630–632. ACM, 2010.

Lujun Fang and Kristen LeFevre. Privacy wizards for social networking sites. In *Proceedings of the 19th International Conference on World Wide Web*, pages 351–360, 2010.

Kaniz Fatema, Ensar Hadziselimovic, Harshvardhan J Pandit, Christophe Debruyne, Dave Lewis, and Declan O'Sullivan. Compliance through informed consent: Semantic based consent permission and data management model. In *PrivOn@ ISWC*, 2017.

S. Fatima and Michael Wooldridge. The negotiation game. *Intelligent systems*, 29:57–61, 10 2014.

Shaheen Fatima, Sarit Kraus, and Michael Wooldridge. *Principles of automated negotiation*. Cambridge University Press, 2014.

Shaheen Fatima, Michael Wooldridge, and Nicholas R Jennings. Optimal Agendas for Multi-Issue Negotiation. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 129–136. ACM, 2003.

Shaheen S. Fatima, Michael J. Wooldridge, and Nicholas R. Jennings. Multi-issue negotiation under time constraints. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 143–150, New York, NY, USA, 2002. ACM.

Federal Trade Commission. Protecting consumer privacy in an era of rapid change: Recommendations for businesses and policymakers. Technical report, Federal Trade Commission, March 2012. [cited 31 May 2020]. Available from: `http://ftc.gov/os/2012/03/120326privacyreport.pdf`.

Adrienne Porter Felt, Erika Chin, Steve Hanna, Dawn Song, and David Wagner. Android permissions demystified. In *Proceedings of the 18th ACM Conference on Computer and Communications Security*, pages 627–638. ACM, 2011.

Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the eighth symposium on usable privacy and security*, page 3. ACM, 2012.

Edward W. Felten and Michael A. Schneider. Timing attacks on web privacy. In *Proceedings of the 7th ACM Conference on Computer and Communications Security*, pages 25–32. ACM, 2000.

Martin Fishbein and Icek Ajzen. *Belief, attitude, intention and behavior: An introduction to theory and research*. Addison-Wesley, Reading, MA, 1975.

Peter Fletcher, Hughes B. Hoyle, and C. Wayne Patty. *Foundations of discrete mathematics*. Boston: PWS-KENT Publishing Co., 1991.

Joshua Fogel and Elham Nehmad. Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, 25(1):153–160, 2009.

Lester Randolph Ford and Delbert R. Fulkerson. Maximal flow through a network. *Canadian journal of Mathematics*, 8:399–404, 1956.

Charles Fried. Privacy. *The Yale Law Journal*, 77(3):393–422, 1968.

Christopher S Gates, Jing Chen, Ninghui Li, and Robert W Proctor. Effective risk communication for android apps. *IEEE Transactions on dependable and secure computing*, 11(3):252–265, 2014.

Robert Gellman. Fair Information Practices: a Basic History - Version 2.19, October 2019. [cited 15 January 2021]. Available from: `https://ssrn.com/abstract=2415020`.

Nina Gerber, Paul Gerber, and Melanie Volkamer. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77:226–261, 2018.

Amirata Ghorbani and James Zou. Data Shapley: Equitable valuation of data for machine learning. In *International Conference on Machine Learning*, pages 2242–2251. PMLR, 2019.

Ian Goldberg, David Wagner, and Eric Brewer. Privacy-enhancing technologies for the internet. In *Proceedings IEEE COMPCON 97. Digest of Papers*, pages 103–109. IEEE, 1997.

Richard C. Gomer. *Values in Technology and Practice: Using Activity Theory to consider the role of values and technology in everyday activities.* PhD thesis, University of Southampton, 2018.

Salvatore Greco, Miłosz Kadziński, Vincent Mousseau, and Roman Słowiński. Robust ordinal regression for multiple criteria group decision: Utagms-group and utadisgms-group. *Decision Support Systems*, 52(3):549–561, 2012.

Graham Greenleaf. Sheherezade and the 101 data privacy laws: Origins, significance and global trajectories. *Journal of Law Information & Science*, 23:4, 2014.

Paolo Guarda and Nicola Zannone. Towards the development of privacy-aware systems. *Information and Software Technology*, 51(2):337–350, 2009.

Seda Gürses, Carmela Troncoso, and Claudia Diaz. Engineering privacy by design. *Computers, Privacy & Data Protection*, 14(3):25, 2011.

Jianye Hao and Ho-fung Leung. CUHK agent: an adaptive negotiation strategy for bilateral negotiations over multiple items. In *Novel Insights in Agent-based Complex Automated Negotiation*, volume 535 of *Studies in Computational Intelligence*, pages 171–179. Springer, Japan, 2014.

Marian Harbach, Markus Hettig, Susanne Weber, and Matthew Smith. Using personal examples to improve risk communication for security & privacy decisions. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2647–2656. ACM, 2014.

Eszter Hargittai and Alice Marwick. "What can I really do?" Explaining the privacy paradox with online apathy. *International Journal of Communication*, 10:21, 2016.

Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

Judd Randolph Heckman, Erin Laurel Boehmer, Elizabeth Hope Peters, Milad Davaloo, and Nikhil Gopinath Kurup. A pricing model for data markets. *iConference 2015 Proceedings*, 2015.

Maximilian Hils, Daniel W. Woods, and Rainer Böhme. Measuring the emergence of consent management on the web. In *Proceedings of the ACM Internet Measurement Conference*, pages 317–332, 2020.

Harry Hochheiser. The platform for privacy preference as a social protocol: An examination within the us policy context. *ACM Transactions on Internet Technology*, 2(4): 276–306, 2002.

Kimberly A. Houser and W. Gregory Voss. GDPR: the end of Google and Facebook or a new paradigm in data privacy. *Richmond Journal of Law & Technology*, 25:1–109, 2018.

Peter Hustinx. Privacy by design: delivering the promises. *Identity in the Information Society*, 3(2):253–255, 2010.

L. Ilany and Y. (K.) Gal. The simple-meta agent. In *Novel Insights in Agent-based Complex Automated Negotiation*, volume 535 of *Studies in Computational Intelligence*, pages 197–200. Springer, Japan, 2014.

Jim Isaak and Mina J. Hanna. User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, 51(8):56–59, 2018.

Eric Jacquet-Lagreze and Jean Siskos. Assessing a set of additive utility functions for multicriteria decision-making, the uta method. *European journal of operational research*, 10(2):151–164, 1982.

In Joo Jang, Wenbo Shi, and Hyeon Seon Yoo. Policy negotiation system architecture for privacy protection. In *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, volume 2, pages 592–597. IEEE, 2008.

Nicholas R. Jennings, Peyman Faratin, Alessio R. Lomuscio, Simon Parsons, Michael J. Wooldridge, and Carles Sierra. Automated negotiation: Prospects, methods and challenges. *Group Decision and Negotiation*, 10(2):199–215, 2001.

Carlos Jensen, Colin Potts, and Christian Jensen. Privacy practices of Internet users: self-reports versus observed behavior. *International Journal of Human-Computer Studies*, 63(1):203–227, 2005.

Adam N. Joinson, Ulf-Dietrich Reips, Tom Buchanan, and Carina B. Paine Schofield. Privacy, trust, and self-disclosure online. *Human–Computer Interaction*, 25(1):1–24, 2010.

Yogesh Kalyani and Carlisle Adams. Privacy negotiation using a mobile agent. In *2006 Canadian Conference on Electrical and Computer Engineering*, pages 628–633. IEEE, May 2006.

Günter Karjoth, Matthias Schunter, and Michael Waidner. Platform for enterprise privacy practices: Privacy-enabled management of customer data. In *2nd Workshop on Privacy-Enhancing Technologies. Lecture Notes in Computer Science*, pages 69–84. Springer, 2002.

Nishan C. Karunatillake. *Argumentation–Based Negotiation in a Social Context*. PhD thesis, University of Southampton, 2006.

Shogo Kawaguchi, Katsuhide Fujita, and Takayuki Ito. Compromising strategy based on estimated maximum utility for automated negotiating agents. In *New Trends in Agent-based Complex Automated Negotiations, Series of Studies in Computational Intelligence*, pages 137–144, Berlin, Heidelberg, 2012. Springer-Verlag.

Ralph L. Keeney and Howard Raiffa. *Decisions with Multiple Objectives*. Cambridge University Press, 1976.

Dilara Kekulluoglu, Nadin Kokciyan, and Pinar Yolum. Preserving privacy as social responsibility in online social networks. *ACM Transactions on Internet Technology (TOIT)*, 18(4):1–22, 2018.

Patrick Kelley, Sunny Consolvo, Lorrie Cranor, Jaeyeon Jung, Norman Sadeh, and David Wetherall. A conundrum of permissions: installing applications on an android smartphone. *Financial cryptography and data security*, pages 68–79, 2012.

Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12. ACM, 2009.

Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the 2010 SIGCHI Conference on Human factors in Computing Systems*, pages 1573–1582. ACM, 2010.

Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3393–3402. ACM, 2013.

Patrick Gage Kelley, Paul Hankes Drielsma, Norman Sadeh, and Lorrie Faith Cranor. User-controllable learning of security and privacy policies. In *Proceedings of the 1st ACM Workshop on Workshop on AISec*, pages 11–18. ACM, 2008.

Sabrina Kirrane, Javier D. Fernández, Wouter Dullaert, Uros Milosevic, Axel Polleres, Piero A. Bonatti, Rigo Wenning, Olha Drozd, and Philip Raschke. A scalable consent, transparency and compliance architecture. In Aldo Gangemi, Anna Lisa Gentile, Andrea Giovanni Nuzzolese, Sebastian Rudolph, Maria Maleshkova, Heiko Paulheim, Jeff Z. Pan, and Mehwish Alam, editors, *The Semantic Web: ESWC 2018 Satellite Events. ESWC 2018. Lecture Notes in Computer Science*, volume 11155, pages 131–136. Springer, 2018.

Mark Klein and Stephen C-Y Lu. Conflict resolution in cooperative design. *Artificial Intelligence in Engineering*, 4(4):168–180, 1989.

Jon Kleinberg and Eva Tardos. *Algorithm design*. Pearson Education, 2006.

Bart Knijnenburg and David Cherry. Comics as a medium for privacy notices. In *Proceedings of the 12th Symposium on Usable Privacy and Security*. USENIX, 2016.

Nadin Kökciyan, Nefise Yaglikci, and Pinar Yolum. An argumentation approach for resolving privacy disputes in online social networks. *ACM Transactions on Internet Technology (TOIT)*, 17(3):1–22, 2017.

Nadin Kökciyan and Pınar Yolum. Priguard: A semantic approach to detect privacy violations in online social networks. *IEEE Transactions on Knowledge and Data Engineering*, 28(10):2724–2737, 2016.

Spyros Kokolakis. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 64:122–134, 2017.

Jan Kolter, Rolf Schillinger, and Günther Pernul. A privacy-enhanced attribute-based access control system. In *IFIP Annual Conference on Data and Applications Security and Privacy*, pages 129–143. Springer, 2007.

Stefan Korff and Rainer Böhme. Too much choice: End-user privacy decisions in the context of choice proliferation. In *Proceedings of the 10th Symposium On Usable Privacy and Security*, pages 69–87. USENIX Association, 2014.

Gina Kounga, Marco Casassa Mont, and Pete Bramhall. Extending xacml access control architecture for allowing preference-based authorisation. In *International Conference on Trust, Privacy and Security in Digital Business*, pages 153–164. Springer, 2010.

Sarit Kraus. *Strategic Negotiation in Multiagent Environments*. MIT press, 2001.

Balachander Krishnamurthy and Craig Wills. Privacy diffusion on the web: a longitudinal perspective. In *Proceedings of the 18th International Conference on World Wide Web*, pages 541–550. ACM, 2009.

Kat Krol and Sören Preibusch. Effortless privacy negotiations. *IEEE Security & Privacy*, 13(3):88–91, 2015.

Kat Krol and Sören Preibusch. Control versus effort in privacy warnings for webforms. In *Proceedings of the 2016 ACM Workshop on Privacy in the Electronic Society*, pages 13–23. ACM, 2016.

Ponnurangam Kumaraguru, Lorrie Cranor, Jorge Lobo, and Seraphin Calo. A survey of privacy policy languages. In *Workshop on Usable IT Security Management (USM 07): Proceedings of the 3rd Symposium on Usable Privacy and Security*. ACM, 2007.

Ponnurangam Kumaraguru and Lorrie Faith Cranor. Privacy indexes: a survey of Westin's studies. Technical report, Carnegie Mellon University, School of Computer Science, 2005.

Kristen LeFevre, Rakesh Agrawal, Vuk Ercegovac, Raghu Ramakrishnan, Yirong Xu, and David DeWitt. Limiting disclosure in Hippocratic databases. In *Proceedings of the 30th International Conference on Very Large Databases*, pages 108–119. VLDB Endowment, 2004.

Pedro Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako, and Lorrie Cranor. Why johnny can't opt out: a usability evaluation of tools to limit online behavioral advertising. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 589–598, 2012.

Chao Li, Daniel Yang Li, Gerome Miklau, and Dan Suciu. A theory of pricing private data. *ACM Transactions on Database Systems*, 39(4):1–28, 2014.

Han Li, Rathindra Sarathy, and Heng Xu. The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors. *Decision Support Systems*, 51(3):434–445, 2011.

Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *2007 IEEE 23rd International Conference on Data Engineering*, pages 106–115. IEEE, 2007.

Ninghui Li, Ting Yu, and Annie I. Antón. A semantics based approach to privacy languages. *Computer Systems Science and Engineering*, 21(5):339, 2006.

Ilaria Liccardi, Joseph Pato, and Daniel J Weitzner. Improving user choice through better mobile apps transparency and permissions analysis. *Journal of Privacy and Confidentiality*, 5(2):1, 2014.

Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 501–510. ACM, 2012.

Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I. Hong. Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *Proceedings*

of the 10th Symposium On Usable Privacy and Security (SOUPS 2014). USENIX Association, 2014.

Thomas Linden, Rishabh Khandelwal, Hamza Harkous, and Kassem Fawaz. The privacy policy landscape after the GDPR. *Proceedings on Privacy Enhancing Technologies*, 2020(1):47–64, 2020.

Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Proceedings of the 12th Symposium on Usable Privacy and Security*, pages 27–41. USENIX Association, 2016.

Bin Liu, Jialiu Lin, and Norman Sadeh. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd international conference on World wide web*, pages 201–212. ACM, 2014.

Yue Liu. User control of personal information concerning mobile-app: Notice and consent? *Computer Law & Security Review*, 30(5):521–529, 2014.

Mark Lizar and David Turner. Consent Receipt Specification. Technical report, Kantara Initiative, February 2018. [cited 22 Feb 2021]. Available from: `https://kantarainitiative.org/file-downloads/consent-receipt-specification-v1-1-0/`.

Ewa Luger, Stuart Moran, and Tom Rodden. Consent for all: revealing the hidden complexity of terms and conditions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 2687–2696, 2013.

Ewa Luger and Tom Rodden. Terms of agreement: Rethinking consent for pervasive computing. *Interacting with Computers*, 25(3):229–241, 2013.

M Maaser, Steffen Ortmann, and Peter Langendörfer. Nepp: Negotiation enhancements for privacy policies. In *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement*, 2006.

Michael Maaser and Peter Langendoerfer. Automated negotiation of privacy contracts. In *Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International*, volume 1, pages 505–510. IEEE, 2005.

Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):1–52, 2007.

Michelle Madejski, Maritza Johnson, and Steven M. Bellovin. The failure of online social network privacy settings. Technical report, Columbia University, 8 2011. [cited

11 Feb 2021]. Available from: https://academiccommons.columbia.edu/doi/10.7916/D8NG4ZJ1.

Miguel Malheiros, Sören Preibusch, and M Angela Sasse. "fairly truthful": The impact of perceived effort, fairness, relevance, and sensitivity on personal data disclosure. In *International Conference on Trust and Trustworthy Computing*, pages 250–266. Springer, 2013.

Naresh K. Malhotra, Sung S. Kim, and James Agarwal. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4):336–355, 2004.

R. Timothy Marler and Jasbir S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, 2004.

Kirsten Martin. Do privacy notices matter? comparing the impact of violating formal privacy notices and informal privacy norms on consumer trust online. *The Journal of Legal Studies*, 45(2):191–215, 2016.

Fabio Massacci, John Mylopoulos, and Nicola Zannone. Hierarchical Hippocratic databases with minimal disclosure for virtual organizations. *The VLDB Journal*, 15(4):370–387, 2006.

Arunesh Mathur, Jessica Vitak, Arvind Narayanan, and Marshini Chetty. Characterizing the use of browser-based blocking extensions to prevent online tracking. In *Proceedings of the 14th Symposium on Usable Privacy and Security*, pages 103–116. USENIX, 2018.

Célestin Matte, Nataliia Bielova, and Cristiana Santos. Do cookie banners respect my choice?: Measuring legal compliance of banners from iab europe's transparency and consent framework. In *Proceedings of the 41st IEEE Symposium on Security and Privacy*, pages 791–809. IEEE, 2020.

Jonathan R. Mayer and John C. Mitchell. Third-party web tracking: Policy and technology. In *2012 IEEE Symposium on Security and Privacy*, pages 413–427. ACM, 2012.

Aleecia M. McDonald and Lorrie Faith Cranor. Americans' attitudes about internet behavioral advertising practices. In *Proceedings of the 9th Annual ACM Workshop on Privacy in the Electronic Society*, pages 63–72. ACM, 2010.

Yavuz Mester, Nadin Kökciyan, and Pınar Yolum. Negotiating privacy constraints in online social networks. In *Advances in Social Computing and Multiagent Systems*, pages 112–129. Springer, 2015.

George R. Milne and Shalini Bahl. Are there differences between consumers' and marketers' privacy expectations? a segment-and technology-level analysis. *Journal of Public Policy & Marketing*, 29(1):138–149, 2010.

George R Milne and Maria-Eugenia Boza. Trust and concern in consumers' perceptions of marketing information management practices. *Journal of interactive Marketing*, 13 (1):5–24, 1999.

George R. Milne, Mary J. Culnan, and Henry Greene. A longitudinal assessment of online privacy notice readability. *Journal of Public Policy & Marketing*, 25(2):238–249, 2006.

George R. Milne and Mary Ellen Gordon. Direct mail privacy-efficiency trade-offs within an implied social contract framework. *Journal of Public Policy & Marketing*, 12(2): 206–215, 1993.

George R Milne, George Pettinico, Fatima M Hajjat, and Ereni Markos. Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs*, 51(1):133–161, 2017.

Anthony D. Miyazaki and Ana Fernandez. Consumer perceptions of privacy and security risks for online shopping. *Journal of Consumer Affairs*, 35(1):27–44, 2001.

Yasser Mohammad and Shinji Nakadai. Fastvoi: Efficient utility elicitation during negotiations. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 560–567. Springer, 2018.

Jonathan Mugan, Tarun Sharma, and Norman Sadeh. Understandable learning of privacy preferences through default personas and suggestions, 2011.

Toru Nakamura, Shinsaku Kiyomoto, Welderufael B Tesfay, and Jetzabel Serna. Easing the burden of setting privacy preferences: A machine learning approach. In *International Conference on Information Systems Security and Privacy*, pages 44–63. Springer, 2016.

Qun Ni, Elisa Bertino, Jorge Lobo, Carolyn Brodie, Clare-Marie Karat, John Karat, and Alberto Trombeta. Privacy-aware role-based access control. *ACM Transactions on Information and System Security (TISSEC)*, 13(3):1–31, 2010.

Helen Nissenbaum. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press, Stanford, CA, 2009.

Helen Nissenbaum. A contextual approach to privacy online. *Daedalus*, 140(4):32–48, 2011.

Patricia A. Norberg, Daniel R. Horne, and David A. Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41(1):100–126, 2007.

Amy Nordrum. Popular Internet of Things forecast of 50 billion devices by 2020 is outdated. *IEEE Spectrum*, 18(3), 2016.

Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

Glen J. Nowak and Joseph E. Phelps. Understanding privacy concerns: An assessment of consumers' information-related knowledge and beliefs. *Journal of Direct Marketing*, 6(4):28–39, 1992.

OASIS. eXtensible Access Control Markup Language (XACML). Version 3.0., January 2013. [cited 19 Feb 2021]. Available from: `http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html`.

Hendrik J. G. Oberholzer and Martin S. Olivier. Privacy contracts incorporated in a privacy protection framework. *Computer Systems Science and Engineering*, 21(1): 5–16, 2006.

Shintaro Okazaki, Hairong Li, and Morikazu Hirose. Consumer privacy concerns and preference for degree of regulatory control. *Journal of Advertising*, 38(4):63–77, 2009.

Organisation for Economic Cooperation and Development. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, September 1980. [cited 20 Jan 2021]. Available from: `http://www.oecd.org/digital/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm`.

Organisation for Economic Cooperation and Development. The OECD Privacy Framework, 2013. [cited 20 January 2021]. Available from: `http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf`.

Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. The MIT Press, 1st edition, 1994.

Dara O'Neil. Analysis of internet users' level of online privacy concerns. *Social Science Computer Review*, 19(1):17–31, 2001.

Carina Paine, Ulf-Dietrich Reips, Stefan Stieger, Adam Joinson, and Tom Buchanan. Internet users' perceptions of 'privacy concerns' and 'privacy actions'. *International Journal of Human-Computer Studies*, 65(6):526–536, 2007.

Harshvardhan J. Pandit, Christophe Debruyne, Declan O'Sullivan, and Dave Lewis. GConsent-a consent ontology based on the GDPR. In *European Semantic Web Conference*, pages 270–282. Springer, 2019a.

Harshvardhan J. Pandit, Declan O'Sullivan, and Dave Lewis. Test-driven approach towards GDPR compliance. In *International Conference on Semantic Systems*, pages 19–33. Springer, 2019b.

Yong Jin Park. Digital literacy and privacy behavior online. *Communication Research*, 40(2):215–236, 2013.

Yong Jin Park, Scott W Campbell, and Nojin Kwak. Affect, cognition and reward: Predictors of privacy protection online. *Computers in Human Behavior*, 28(3):1019–1027, 2012.

Mario Pascalev. Privacy exchanges: restoring consent in privacy self-management. *Ethics and Information Technology*, 19(1):39–48, 2017.

Emmanuel Pernot-Leplay. China's Approach on Data Privacy Law: A Third Way Between the US and the EU? *Penn State Journal of Law & International Affairs*, 8(1), 2020.

Milan Petković, Davide Prandi, and Nicola Zannone. Purpose control: Did you process the data for the intended purpose? In *Workshop on Secure Data Management*, pages 145–168. Springer, 2011.

Spyros E. Polykalas. Assessing general data protection regulation for personal data privacy: is the end of "take it or leave it" approach for downloading apps? In *Proceedings of the Seventh International Conference on Social Media Technologies, Communication, and Informatics*, 2017.

Elena Popescu. On the approximation of inconsistent inequality systems. *Analele Ştiinţifice ale Universităţii Ovidius*, 11(2):109–118, 2003.

Richard A. Posner. The right of privacy. *Georgia Law Review*, 12(3):393–422, 1978.

Calvin S. Powers, Paul Ashley, and Matthias Schunter. Privacy promises, access control, and privacy management. enforcing privacy throughout an enterprise by extending access control. In *Proceedings of the Third International Symposium on Electronic Commerce*, pages 13–21. IEEE, 2002.

Sören Preibusch. Implementing privacy negotiations in e-commerce. In *Asia-Pacific Web Conference*, pages 604–615. Springer, 2006.

Sören Preibusch. Guide to measuring privacy concern: Review of survey and observational instruments. *International Journal of Human-Computer Studies*, 71(12):1133–1143, 2013.

Sören Preibusch, Kat Krol, and Alastair R Beresford. The privacy economics of voluntary over-disclosure in web forms. In *The Economics of Information Security and Privacy*, pages 183–209. Springer, 2013.

Mian Qin, Scott Buffett, and Michael W. Fleming. Predicting user preferences via similarity-based clustering. In *Proceedings of the Canadian Society for Computational Studies of Intelligence, 21st Conference on Advances in Artifical Intelligence (Canadian AI'08)*, pages 222–233. Springer, 2008.

Margaret Jane Radin. *Boilerplate: The fine print, vanishing rights, and the rule of law.* Princeton University Press, 2012.

Raghu Ramakrishnan and Johannes Gehrke. *Database management systems.* McGraw-Hill, 2000.

Neil Randall. The new cookie monster. *PC Magazine*, 16(8):211–214, 1997.

Ramprasad Ravichandran, Michael Benisch, Patrick Gage Kelley, and Norman M Sadeh. Capturing social networking privacy preferences. In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 1–18. Springer, 2009.

Ian K Reay, Patricia Beatty, Scott Dick, and James Miller. A survey and analysis of the p3p protocol's agents, adoption, maintenance, and future. *IEEE Transactions on Dependable and Secure Computing*, 4(2):151–164, 2007.

Priscilla M. Regan. *Legislating privacy: Technology, social values, and public policy.* University of North Carolina Press, 2000.

Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, and Rohan Ramanath. Disagreeable privacy policies: Mismatches between meaning and users' understanding. *Berkeley Technology Law Journal*, 30:39, 2015.

Tim Ring. Keeping tabs on tracking technology. *Network Security*, 2015(6):5–8, 2015.

William N. Robinson. Negotiation behavior during requirement specification. In *Proceedings of the 12th International Conference on Software Engineering*, pages 268–276. IEEE, 1990.

Jeffrey S. Rosenschein and Gilad Zlotkin. *Rules of encounter: designing conventions for automated negotiation among computers.* MIT Press, 1994.

Jeffrey Solomon Rosenschein. Rational interaction: cooperation among intelligent agents. Stanford, CA, United States, 1986. Stanford University.

Max Roser, Hannah Ritchie, and Esteban Ortiz-Ospina. Internet. Technical report, Our World in Data, 2015. [cited 15 Mar 2021]. Available from: `https://ourworldindata.org/internet`.

Ewa Roszkowska et al. The application of uta method for support evaluation negotiation offers. *Optimum. Economic Studies*, 80(2):144–162, 2016.

Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109, 1982.

Norman Sadeh, Alessandro Acquisti, Travis D. Breaux, Lorrie Faith Cranor, Aleecia M. McDonald, Joel R. Reidenberg, Noah A. Smith, Fei Liu, N. Cameron Russell, Florian

Schaub, et al. The usable privacy policy project. Technical report, Carnegie Mellon University, 2013.

Norman Sadeh, Jason Hong, Lorrie Cranor, Ian Fette, Patrick Kelley, Madhu Prabaker, and Jinghai Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(6):401–412, 2009.

Iskander Sanchez-Rola, Matteo Dell'Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos. Can I opt out yet? GDPR and the global illusion of cookie control. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, pages 340–351, 2019.

Ravi S. Sandhu, Edward J. Coyne, Hal L. Feinstein, and Charles E. Youman. Role-based access control. *IEE Computer*, 29(2):39–47, February 1996.

Cristiana Santos, Nataliia Bielova, and Célestin Matte. Are cookie banners indeed compliant with the law? *Technology and Regulation*, pages 91–135, 2020.

Peter Schaar. Privacy by design. *Identity in the Information Society*, 3(2):267–274, 2010.

Florian Schaub, Rebecca Balebako, and Lorrie Faith Cranor. Designing effective privacy notices and controls. *IEEE Internet Computing*, 2017.

Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. A design space for effective privacy notices. In *Proceedings of the 11th Symposium On Usable Privacy and Security*, pages 1–17. USENIX, 2015.

Florian Schaub, Aditya Marella, Pranshu Kalvani, Blase Ur, Chao Pan, Emily Forney, and Lorrie Faith Cranor. Watching them watching me: Browser extensions' impact on user privacy awareness and concern. In *Proceedings of the Network and Distributed System Security (NDSS) Symposium 2016. NDSS Workshop on Usable Security (USEC)*, pages 1–10. Internet Society, 2016.

Bart W Schermer, Bart Custers, and Simone van der Hof. The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology*, 16(2):171–182, 2014.

Paul M. Schwartz. Internet privacy and the state. *Connecticut Law Review*, 32:815, 1999.

Seagate. Rethink data: Put more of your business data to work – from edge to cloud. Technical report, International Data Corporation (IDC), March 2012. [cited 15 Mar 2021]. Available from: `https://www.seagate.com/files/www-content/our-story/rethink-data/files/Rethink_Data_Report_2020.pdf`.

Kim Bartel Sheehan. Toward a typology of Internet users and online privacy concerns. *The Information Society*, 18(1):21–32, 2002.

Kim Bartel Sheehan and Mariea Grubbs Hoy. Dimensions of privacy concern among online consumers. *Journal of Public Policy & Marketing*, 19(1):62–73, 2000.

Fatemeh Shirazi and Melanie Volkamer. What deters Jane from preventing identification and tracking on the Web? In *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, pages 107–116. ACM, 2014.

Janice C. Sipior, Burke T. Ward, and Ruben A. Mendoza. Online privacy concerns associated with cookies, flash cookies, and web beacons. *Journal of Internet Commerce*, 10(1):1–16, 2011.

Robert H. Sloan and Richard Warner. Beyond notice and choice: Privacy, norms, and consent. *Journal of High Technology Law*, 14(2):370–414, 2014.

H. Jeff Smith, Tamara Dinev, and Heng Xu. Information privacy research: an interdisciplinary review. *MIS Quarterly*, 35(4):989–1015, 2011.

H. Jeff Smith, Sandra J. Milberg, and Sandra J. Burke. Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 20(2):167–196, 1996.

Reid G. Smith. The contract net protocol: High-level communication and control in a distributed problem solver. *IEEE Transactions on Computers*, 29(12):1104–1113, 1980.

Daniel J. Solove. Introduction: Privacy Self-Management and the Consent Dilemma. *Harvard Law Review*, 126:1880, 2013.

Ashkan Soltani, Shannon Canty, Quentin Mayo, Lauren Thomas, and Chris Jay Hoofnagle. Flash cookies and privacy. Technical report, Summer Undergraduate Program in Engineering Research at Berkeley (SUPERB), Aug 2009. [cited 25 Mar 2021]. Available from: `https://ssrn.com/abstract=1446862`.

Signe Sophus Lai and Sofie Flensburg. A proxy for privacy uncovering the surveillance ecology of mobile apps. *Big Data & Society*, 7(2):1–20, 2020.

Sarah Spiekermann and Lorrie Faith Cranor. Engineering privacy. *IEEE Transactions on Software Engineering*, 35(1):67–82, 2008.

Sarah Spiekermann, Jens Grossklags, and Bettina Berendt. E-privacy in 2nd generation e-commerce: privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM Conference on Electronic Commerce*, pages 38–47. ACM, 2001.

Anna Cinzia Squicciarini, Elisa Bertino, Elena Ferrari, and Indrakshi Ray. Achieving privacy in trust negotiations with an ontology-based approach. *IEEE Transactions on Dependable and Secure Computing*, 3(1):13–30, 2006.

Alexa Stein, Norman Makoto Su, and Xinru Page. Learning through videos: Uncovering approaches to educating people about facebook privacy. In *Proceedings of the 16th Symposium on Usable Privacy and Security*. USENIX, 2020.

Kathy A. Stewart and Albert H. Segars. An empirical examination of the concern for information privacy instrument. *Information Systems Research*, 13(1):36–49, 2002.

Jose M Such and Natalia Criado. Resolving multi-party privacy conflicts in social media. *IEEE Transactions on Knowledge and Data Engineering*, 28(7):1851–1863, 2016.

Daniel Susser. Notice after notice-and-consent: Why privacy disclosures are valuable even if consent frameworks aren't. *Journal of Information Policy*, 9:148–173, 2019.

Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.

Katia Sycara. Resolving goal conflicts via negotiation. In *Proceedings of the Seventh AAAI National Conference on Artificial Intelligence*, volume 88, pages 245–250, 1988.

Katia Sycara-Cyranski. Arguments of persuasion in labour mediation. In *Proceedings of the 9th international joint conference on Artificial intelligence*, pages 294–296, San Francisco, CA, USA, 1985. Morgan Kaufmann Publishers Inc.

Monika Taddicken. The 'privacy paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, 19(2):248–273, 2014.

Darren Thomson, Peter Cochrane, Sian John, Udo Helmbrecht, Ilias Chantzos, Philip Carter, and Steward Room. State of privacy report 2015. Technical report, Symantec, 2015. [cited 8 Feb 2021]. Available from: `https://www.nortonlifelock.com/content/dam/nortonlifelock/pdfs/reports/state-of-privacy-report-en-2015.pdf`.

Janice Y. Tsai, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, 22(2):254–268, 2011.

Markus Tschersich and Michael Niekamp. Pros and cons of privacy by default: Investigating the impact on users and providers of social network sites. In *2015 48th Hawaii International Conference on System Sciences (HICSS)*, pages 1750–1756. IEEE, 2015.

Dimitrios Tsimpoukis, Tim Baarslag, Michael Kaisers, and Nikolaos G Paterakis. Automated negotiations under user preference uncertainty: A linear programming approach. In *International Conference on Agreement Technologies*, pages 115–129. Springer, 2018.

Joseph Turow. Americans and online privacy: The system is broken. Technical report, Annenberg Public Policy Center of the University of Pennsylvania, 2003.

Joseph Turow, Lauren Feldman, and Kimberly Meltzer. Open to exploitation: America's shoppers online and offline. Technical report, Annenberg Public Policy Center of the University of Pennsylvania, 2005.

Joseph Turow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy. Americans reject tailored advertising and three activities that enable it. Technical report, Annenberg School for Communication (University of Pennsylvania) and Berkeley School of Law (University of California, Berkeley), 2009. [cited 26 Mar 2021]. Available from: `https://repository.upenn.edu/asc_papers/524`.

United Nations General Assembly. *Universal Declaration of Human Rights*, volume 3381. Department of State, United States of America, 1948.

Blase Ur, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. Smart, useful, scary, creepy: perceptions of online behavioral advertising. In *Proceedings of the 8th Symposium on Usable Privacy and Security*, pages 1–15. ACM, 2012.

US Department of Health, Education, and Welfare. Records, computers and the rights of citizens. *Report of the Secretary's Advisory Committee on Automated Personal Data Systems*, 1973. [cited 15 Jan 2021]. Available from: `https://epic.org/privacy/hew1973report/default.html`.

Brendan van Alsenoy. *Data Protection Law in the EU: Roles, Responsibilities and Liability*, volume 6. Intersentia, 2019.

G. W. van Blarkom, John J. Borking, and P. Verhaar. PET. In G. W. van Blarkom, John J. Borking, and J. G. Eddy Olk, editors, *Handbook of privacy and privacy-enhancing technologies. The case of Intelligent Software Agents*, pages 33–54. CBP (Dutch Data Protection Authority), The Hague, 2003.

Thea Van Der Geest, Willem Pieterson, and Peter De Vries. Informed consent to address trust, control, and privacy concerns in user profiling. *Privacy Enhanced Personalisation, PEP*, pages 23–34, 2005.

Max Van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J Weitzner, and Nigel Shadbolt. Better the devil you know: Exposing the data sharing practices of smartphone apps. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5208–5220. ACM, 2017.

Duy Van Nguyen. Global maximization of uta functions in multi-objective optimization. *European Journal of Operational Research*, 228(2):397–404, 2013.

Sandra Wachter and Brent Mittelstadt. A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Columbia Bussines Law Review*, pages 494–620, 2019.

Ari Ezra Waldman. Cognitive biases, dark patterns, and the 'privacy paradox'. *Current opinion in psychology*, 31:105–109, 2020.

Yang Wang and Alfred Kobsa. Respecting users' individual privacy constraints in web personalization. In *International Conference on User Modeling*, pages 157–166. Springer, 2007.

Samuel D. Warren and Louis D. Brandeis. The right to privacy. *Harvard Law Review*, 4(5):193–220, December 1890.

Martin L Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3): 641–654, 1979.

Alan F. Westin. Privacy and freedom. *Washington and Lee Law Review*, 25(1):166, 1968.

Daricia Wilkinson, Moses Namara, Karla Badillo-Urquiola, Pamela J Wisniewski, Bart P Knijnenburg, Xinru Page, Eran Toch, and Jen Romano-Bergstrom. Moving beyond a" one-size fits all" exploring individual differences in privacy. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–8. ACM, 2018.

Colin R. Williams, Valentin Robu, Enrico H. Gerding, and Nicholas R. Jennings. Iamhaggler: A negotiation agent for complex environments. In *New Trends in Agent-based Complex Automated Negotiations*, Studies in Computational Intelligence, pages 151–158, Berlin, Heidelberg, 2012. Springer-Verlag.

Craig E. Wills and Mihajlo Zeljkovic. A personalized approach to web privacy: awareness, attitudes and actions. *Information Management & Computer Security*, 2011.

Shomir Wilson, Justin Cranshaw, Norman Sadeh, Alessandro Acquisti, Lorrie Faith Cranor, Jay Springfield, Sae Young Jeong, and Arun Balasubramanian. Privacy manipulation and acclimation in a location sharing application. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 549–558. ACM, 2013.

Shomir Wilson, Florian Schaub, Rohan Ramanath, Norman Sadeh, Fei Liu, Noah A. Smith, and Frederick Liu. Crowdsourcing annotations for websites' privacy policies: Can it really work? In *Proceedings of the 25th International Conference on World Wide Web*, pages 133–143. ACM, 2016.

Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Laura Brandimarte, and Alessandro Acquisti. Would a privacy fundamentalist sell their ${DNA}$ for $1000... if nothing bad happened as a result? the Westin categories, behavioral intentions, and consequences. In *Proceedings of the 10th Symposium On Usable Privacy and Security*, pages 1–18. USENIX Association, 2014.

Kuang-Wen Wu, Shaio Yan Huang, David C Yen, and Irina Popova. The effect of online privacy policy on consumer privacy concern and trust. *Computers in human behavior*, 28(3):889–897, 2012.

Mike Z Yao, Ronald E Rice, and Kier Wallis. Predicting user concerns about online privacy. *Journal of the American Society for Information Science and Technology*, 58 (5):710–722, 2007.

Abdulsalam Yassine and Shervin Shirmohammadi. Measuring users' privacy payoff using intelligent agents. In *International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA'09)*, pages 169–174. IEEE, 2009.

Seounmi Youn. Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents. *Journal of Consumer affairs*, 43(3): 389–418, 2009.

Aristea M. Zafeiropoulou, David E. Millard, Craig Webber, and Kieron O'Hara. Unpicking the privacy paradox: can structuration theory help to explain location-based privacy decisions? In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 463–472. ACM, 2013.

Bo Zhang and Heng Xu. Privacy nudges for mobile applications: Effects on the creepiness emotion and privacy attitudes. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1676–1690. ACM, 2016.

Yajin Zhou, Xinwen Zhang, Xuxian Jiang, and Vincent W Freeh. Taming information-stealing smartphone applications (on Android). In *International Conference on Trust and Trustworthy Computing*, pages 93–107. Springer, 2011.

Yixin Zou, Abraham H. Mhaidli, Austin McCall, and Florian Schaub. "i've got nothing to lose": Consumers' risk perceptions and protective actions after the equifax data breach. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 197–216. USENIX, 2018.