

**Exploring Influencing Behaviours in Ransomware Negotiations: Comparing Parties
and Identifying Patterns**

Lilli P. M. Schnirch

Faculty of Behavioural Management and Social Sciences, University of Twente

202000377: BSc Thesis Psychology of Conflict, Risk and Safety

1st Supervisor: Michalis Georgiou MBA, MSc

2nd Supervisor: Prof. Dr. Ellen Giebels

05.07.2024

Abstract

The prevalence of ransomware attacks has increased significantly over the past decade, and there is a notable lack of effective prevention and detection tools. Therefore, the most viable opportunity for victims to reduce financial losses and retrieve their data lies in the negotiation stage of an attack. Here, by effectively employing and responding to influence strategies, victims might be able to persuade the threat actor to decrease the ransom amount. Therefore, insights into influencing behaviours in this context are essential. Hence, this study explored the use of influencing behaviours by both threat actors and victims in ransomware negotiations, building on the Table of Ten influencing strategies by Giebels (2002, as cited in Giebels & Noelanders, 2004). Using a sample of 25 ransomware negotiation logs from an open-source repository, this study conducted a comprehensive analysis of influencing behaviours employed by both parties. Specifically, after coding the logs based on the Table of Ten, the coded data was used to compute the relative frequency of each behaviour, identify the most frequent behaviour (mode) at each utterance level, and conduct sequential pattern analysis using the cSPADE algorithm aiming uncover strategic sequences. The interpretation of the results indicated that there is a special need for threat actors in ransomware negotiations to proof their credibility. Furthermore, it showed that threat actors contribute more to the negotiations in terms of utterances. Lastly, the current study established that there are reoccurring patterns of influencing behaviours in ransomware negotiations. Overall, by applying the Table of Ten to a context of ransomware and producing empirically supported insights, this research bridges a gap in literature, contributing to the understanding of negotiation dynamics in ransomware cases.

Keywords: crisis negotiation, ransomware, social influence, Table of Ten, cSPADE

Table of Content

Introduction.....	4
Theoretical Framework.....	5
Ransomware.....	5
Social Influence in Crisis Negotiations.....	10
Formulation of the Research Questions.....	16
Methods.....	17
Data.....	17
Data Analysis.....	23
Results.....	31
Utterance Ratio.....	31
Frequency Analysis.....	32
Most Frequent Code per Utterance Level (Mode).....	35
Pattern Identification Using cSPADE Algorithm.....	38
Discussion.....	41
Key Findings.....	42
Limitations.....	45
Future Research.....	46
Conclusion.....	48
References.....	49
Appendix A.....	55
Appendix B.....	57

Introduction

Ransomware attacks have become “one of the most devastating threats to organizations” throughout the last decade (Boticiu & Teichmann, 2023). Ransomware is a type of malware (i.e., malicious code) that, upon infiltrating the victim’s system, either encrypts data or denies access to the system in order to subsequently demand a ransom payment in exchange for decryption or restored access (Beaman et al., 2021; Boticiu & Teichmann, 2023; Cartwright et al., 2019; Mijwil et al., 2023; Ryan et al., 2022). According to the United Nations Office on Drugs and Crime (2021) the monetary damage caused through ransomware attacks worldwide amounts to 20 billion US dollars annually.

Despite the high prevalence and significant consequences of ransomware attacks, effective mitigating solutions, such as detection and prevention tools, remain scarce and underdeveloped (Aslan et al., 2023; Beaman et al., 2021; Boticiu & Teichmann, 2023; Cartwright et al., 2019). Furthermore, while technical aspects of ransomware, such as encryption schemes and deployment methods, are continuously studied (see for example: Beaman et al., 2021; Hull et al., 2019), recent research has noted an overall lack of investigation into the negotiation phase of ransomware attacks (Boticiu & Teichmann, 2023). Additionally, a gap in research has been highlighted regarding the underlying dynamics between threat actors and victims (Connolly et al., 2020). Although some emerging insights into these dynamics have been suggested (outlined in the section below), they are primarily theoretical and based on research from traditional crisis contexts (see for example: Ryan et al., 2022; Wade, 2021), rather than being derived or validated by empirical ransomware data.

Within literature on traditional crisis negotiations (e.g., extortion or hostage scenarios) it has been found that influence strategies play a key role (Giebels & Noelanders, 2004; Giebels & Taylor, 2009; Grubb, 2023; Grubb et al., 2019). Furthermore, a widely referenced framework that combines insights into crisis negotiations with theory on social influence

exists; The Table of Ten influencing behaviours (Giebels, 2002 as cited in Giebels & Noelanders, 2004) provides a conceptual foundation for investigating dynamics in crisis negotiations and related fields (see for example Beune et al., 2010, 2011; Giebels et al., 2017; Giebels & Noelanders, 2004; Giebels & Taylor, 2009). Furthermore, the potential for reoccurring patterns of influencing behaviours in crisis negotiations has been outlined and conceptualized as strategic sequences (Beune et al., 2011).

The current study will explore the use of social influencing behaviours within the context of ransomware negotiations. Ultimately, aiming to address the aforementioned gap in research — the lack of insight into the underlying mechanisms of ransomware attacks. Further, this study aims to contribute empirical data to the understanding of ransomware negotiations, seeking to validate theoretical insights from traditional crisis negotiation contexts within the context of ransomware.

Therefore, the following section includes an exploration of existing literature on ransomware attacks, specifically the negotiation stage. Following this, social influence research in crisis negotiations is reviewed, including a detailed description of the Table of Ten. Subsequently, research questions are stated, and the methodology of the current study is explained. Afterwards, the results are reported. The paper concludes with a discussion of the findings, emphasising key insights, addressing limitations, and suggesting avenues for future research.

Theoretical Framework

Ransomware

Cyber security threats aimed at extorting money from victims exist since the 1980s (Beaman et al., 2021). PC Cyborg, the very first ransomware attack occurred in 1989 (Tailor & Patel, 2017). Subsequently, until the early 2000s ransomware attacks were mostly launched by amateur hackers seeking to gain recognition in the emerging cyber community (Beaman et

al., 2021; Srinivasan, 2017). Since then, the threat actors have professionalised their attacks and developed more sophisticated methods (Beaman et al., 2021; Cartwright et al., 2019; Ryan et al., 2022; Srinivasan, 2017). Furthermore, threat actors today are often connected and organized within larger criminal networks, so called ransomware groups (Gray et al., 2022). Srinivasan (2017) described this transition as shift from mere 'cyber-vandals' to 'cyber-criminals' and noted that this development was accompanied by a significant increase in monetary losses.

While the ransom demand in the PC Cyborg attack 35 years ago was 189 US dollars (Tailor & Patel, 2017), it is now estimated to cost between 1.85 million (Beaman et al., 2021) and 4.45 million (Boticiu & Teichmann, 2023) US dollars to recover from a ransomware attack. These costs can include the paid ransom, employee working hours, reputational impact, as well as system downtime (Beaman et al., 2021; Boticiu & Teichmann, 2023).

The increase in monetary losses can be attributed to several factors. The first is an increase in attack frequency (Beaman et al., 2021; Boticiu & Teichmann, 2023; Cartwright et al., 2019; Ryan et al., 2022). In their report, the European Union Agency for Cybersecurity (2022) declared ransomware to be the most prevalent cyber security threat today. Worldwide, a new attack occurs approximately every 11 seconds (United Nations Office on Drugs and Crime, 2021). Furthermore, ransomware groups are highly organized in a business-like structure (Cartwright et al., 2019). They have different 'departments' handling various steps of an attack making them more efficient (Cartwright et al., 2019). Additionally, while threat actors in the past targeted individuals, today they are more likely to focus on large companies with high revenue and significant reputational risk, increasing the potential for a greater payout (Beaman et al., 2021; Boticiu & Teichmann, 2023; Cartwright et al., 2019). Lastly, most ransomware strains today do not allow for reverse-engineering without the original

decryption key held by the threat actor (Cartwright et al., 2019) and detection or prevention tools struggle to match the rapid development of new ransomware variants (Aslan et al., 2023; Beaman et al., 2021). Therefore, now more than ever, victims are compelled to pay if they want a chance at recovering their data (Cartwright et al., 2019). Consequently, a large majority of victims pay the ransom (Boticiu & Teichmann, 2023; Cartwright et al., 2019). In 2023, 73% of ransomware victims globally ultimately paid the threat actors, a figure that exhibited a steady upward trend over the past six years (Statista, 2023, as cited in Boticiu & Teichmann, 2023).

Another aspect of the ‘professionalization’ of ransomware attacks is that today threat actors invest considerable resources in gathering information on their victims to determine an appropriate ransom demand (Boticiu & Teichmann, 2023). According to Cartwright and colleagues (2019) the threat actor would not benefit from issuing a ransom demand that exceeds the financial means of the victim. Therefore, after gaining access to the victims system typically through a phishing campaign (Boticiu & Teichmann, 2023; Manjezi & Botha, 2019), threat actors spend on average 43 days undetected before actively disclosing the attack to the victim (Gerritz, 2019 as cited in Ryan et al., 2022). In this time they broaden their control over the system and comb through confidential information collecting insight on the victim’s security and financial status (Ryan et al., 2022). Afterwards, they disclose the attack usually by posting a ransom note in the system, often in form of a file labelled “ReadMe” (Boticiu & Teichmann, 2023). In most cases, this is followed by a negotiation phase in which threat actor and victim discuss the ransom amount and conditions of a potential deal via chat messages (Boticiu & Teichmann, 2023; Cartwright et al., 2019; Ryan et al., 2022).

In a majority of reported cases victims are able to negotiate a lower ransom than the original demand (Boticiu & Teichmann, 2023; Ryan et al., 2022). According to Boticiu and

Teichmann (2023) the final ransom amount can deviate significantly from the original demand. Threat actors in ransomware negotiations may grant up to 90% discounts (Boticiu & Teichmann, 2023). On average they reduce the original ransom by more than 50% during negotiations (Boticiu & Teichmann, 2023). However, the willingness to negotiate and reduce the ransom varies greatly between threat actors and is likely influenced by the behaviour of the victim (Ryan et al., 2022). Similarly, the way the threat actor behaves throughout the negotiation is expected to have an effect on the victims willingness to cooperate (Ryan et al., 2022). Hence, there seems to be some mutual influence between threat actors and victims within ransomware negotiations. The current state of knowledge on this is summarized below.

Dynamics in Ransomware Negotiations

As stated in the introduction, potential influencing dynamics of ransomware negotiations have not been extensively researched. However, given that ransomware attacks share similarities with crisis scenarios such as extortion, kidnappings, and hostage situations, literature on traditional crisis negotiation has been used to inform the context of ransomware negotiations (see for example: Press et al., 2023; Wade, 2021).

Hence, Ryan and colleagues (2022), in their theoretical game theory model of ransomware negotiations, identified reliability of threat actors as a crucial aspect in persuading victims to pay the ransom. Wade (2021) came to a similar conclusion when drawing on insights from traditional crisis negotiations, stating that trust in the threat actors decryption ability potentially increases the victim's willingness to pay.

Furthermore, Ryan and colleagues (2022) anticipate that victims in ransomware negotiations are at a notable disadvantage in terms of information asymmetry. As explained above, threat actors often prepare negotiations by collecting information on the victim for days sometimes weeks (Gerritz, 2019 as cited in Ryan et al., 2022). At the same time the

victim has very limited ways to gather information on the threat actor (Faivre, 2023; Ryan et al., 2022). Faivre (2023), also described this information asymmetry and further explained that it can create a power imbalance in favour of the threat actor. However, they also noted that the victim holds “the power of financial gains” which is what continuously motivates the threat actor to engage in the negotiation and balances the overall power dynamic (Faivre, 2023).

Furthermore, the dynamics in ransomware cases are often compared to business negotiations, where parties cooperate and are equally engaged by asking questions, making requests, and responding (Boticiu & Teichmann, 2023; Cartwright et al., 2019). This suggest somewhat of a balance between threat actors and victims. Giebels and Noelanders (2004) also noted this almost business-like symmetry in the context of extortion, a scenario similar to ransomware cases.

Another ransomware negotiation dynamic which has been hypothesized is the use of irrational aggression by threat actors to persuade victims into paying the ransom (Cartwright et al., 2019). This insight is grounded in game theory and kidnapping scenarios, suggesting that if threat actors behave aggressive and are perceived as threatening, a fear of data loss is instilled in the victim, thereby increasing the likelihood of them complying with ransom demands (Cartwright et al., 2019).

Additionally, Vakulov (2023, as cited in Boticiu & Teichmann, 2023) stated that it is an effective approach for the victim to reason with the threat actor. Specifically, they advise victims to attempt to persuade threat actors into believing that the victims financial resources are less substantial than assumed (Vakulov, 2023 as cited in Boticiu & Teichmann, 2023). It is important to note that this suggestion appears to rely on anecdotal data.

Lastly, several have noted the role of time restrictions in ransomware negotiations (Boticiu & Teichmann, 2023; Cartwright et al., 2019; Ryan et al., 2022). The threat actor

typically imposes a deadline for the payment of the ransom, threatening to delete or publish the victim's data if the deadline is not met (Cartwright et al., 2019). The timeframes might vary per case, but usually victims are granted 72 hours to comply with the request (Cartwright et al., 2019). Apart from the threat actor introducing this artificial stressor, victims are likely motivated to drive the negotiation forward due to the disruptive impact encrypted files or locked systems may have on their business workflow (Boticiu & Teichmann, 2023; Ryan et al., 2022). While the threat actor is unlikely to experience the same urgency, they are driven by the potential profit, especially after having invested resources into the preparation of the attack (Cartwright et al., 2019). Thus, both parties might employ strategies such as time-bound offers to accelerate the negotiation.

The above summarizes insights into ransomware negotiations that have thus far been made, primarily derived from traditional crisis negotiations and game theory. However, as noted, these hypotheses are yet to be validated through empirical research, highlighting a crucial gap in our understanding of the dynamics at play in ransomware negotiations. To build on this foundation and introduce the frameworks used in the current research, it is essential to first explore the broader role of social influence within crisis negotiation. This will provide a comprehensive basis for the formulation of the research questions and the subsequent description of methodology and results.

Social Influence in Crisis Negotiations

Social Influence Theory suggests that our attitudes and behaviours can be significantly influenced by those around us, even at a subconscious level (Gass & Seiter, 2022; Kelman, 1958). Implicit and explicit forms of social influence manifest through cultural and societal norms, interpersonal interactions, and mass media among other examples (Gass & Seiter, 2022). Furthermore, social influence strategies are frequently employed in various contexts to nudge individuals towards desired changes in behaviour and

or attitude (see for example Gass & Seiter, 2022; Gonçalves et al., 2021; Li et al., 2024; Luther et al., 2023).

Crisis negotiations represent a particularly relevant context for applying these strategies (Giebels & Noelanders, 2004; Grubb, 2023). These types of negotiations are typically intense and emotionally loaded since they take place during crisis incident such as kidnappings, suicide attempts and extortions, and thus involve high stakes (Giebels & Noelanders, 2004; Grubb et al., 2019). Meaning, the consequences of crisis negotiations can be substantial such as potential loss of life, severe psychological distress (Giebels et al., 2005; VasIU & VasIU, 2020), or devastating financial loss (VasIU & VasIU, 2020).

Due to the outlined relevance there is a large body of research on influencing behaviours in crisis negotiations (see for example Beune et al., 2011; Giebels & Noelanders, 2004; Grubb, 2023). The current review focuses on two key findings. First, social influence is inherently present in crisis negotiation as shown by Giebels and Noelanders' (2004) study of 35 crisis negotiations, which found that between 65-70% of all displayed behaviours could be characterized as influencing. Secondly, influencing behaviours may form strategic sequences (Beune et al., 2011). Implying, that there is a potential for patterns of influencing behaviours within crisis negotiations. Investigating whether these findings apply in ransomware negotiations can be done using the frameworks introduced below.

Table of Ten

Throughout the last two decades Giebels and colleagues have researched the use of influencing behaviours in the context of crisis negotiations and similar scenarios, mainly looking at cases involving kidnappings, extortions, sieges or police interviews (see for example Beune et al., 2010, 2011; Giebels et al., 2017; Giebels & Noelanders, 2004; Giebels & Taylor, 2009). Specifically, these studies put into practice and empirically validated the Table of Ten by Giebels (2002, as cited in Giebels & Noelanders, 2004). A framework that

comprises a collection of ten social influence strategies that can be employed in crisis negotiation to alter the behaviour of the other.

Also note that according to the classic tenets of social influence, change in attitude precedes shifts in behaviour (Gass & Seiter, 2022; Kelman, 1958). Therefore, when it comes to influencing in a non-crisis context, the objective is to initially alter the counterparties attitude, thereby facilitating subsequent changes in behaviour (Gass & Seiter, 2022). However, given the urgency of crisis negotiations, strategies in the Table of Ten generally prioritize immediate behavioural change rather than attitude changes associated with long-term behavioural manifestations (Giebels & Noelanders, 2004).

Furthermore, it is important to note that each strategy from the Table of Ten can be implemented by either party in a crisis negotiation, for example both law enforcement and criminals (Giebels & Noelanders, 2004). Research applying the Table of Ten in traditional crisis negotiations has found that there are differences between how parties in crisis negotiations employ these strategies (see for example: Beune et al., 2010; Giebels et al., 2005; Giebels & Noelanders, 2004). However, the Table of Ten has never been applied in the context of ransomware negotiations¹.

To provide more detail on this framework, each influence strategy from the Table of Ten is rooted in a social influence principle (Giebels & Noelanders, 2004). Furthermore, the strategies can be distinguished as being either relational strategies or content strategies (Giebels & Noelanders, 2004). There are three relational strategies in the Table of Ten which emphasize the identity of the sender and their relationship with the recipient (Giebels & Noelanders, 2004). These strategies can be employed to foster a positive relationship and facilitate rapport (Giebels & Noelanders, 2004; Westerveld, 2024), which many negotiation

¹Currently, to the best of my knowledge, there are no published peer-reviewed research articles that have applied the Table of Ten within the context of ransomware negotiations. However, Westerveld (2024) explored this in a master thesis, which investigated “The impact of various negotiation strategies [from the Table of Ten] on ransomware negotiation outcomes from the victim's perspective”.

studies identify as crucial for successful negotiations (see for example (see for example Crighton, 2021; Grubb, 2023). In contrast, within the seven content strategies the subject matter of a message or behaviour is pivotal. The complete Table of Ten, including the strategies and their underlying principles, is presented in Table 1 and discussed below.

Table 1

Table of Ten influencing strategies and underlying principles (Giebels, 2002 as cited in (Giebels & Noelanders, 2004)

Strategy	Underlying Principle	Description Behaviour
Relational Strategies		
Being kind	Sympathy	All friendly, helpful behaviour
Being equal	Similarity	Statements aimed at something the parties have in common
Being credible	Authority	Behaviour showing expertise or proving you are reliable
Content Strategies		
Emotional appeal	Self-image	Playing upon the emotions of the other
Intimidation	Deterrence / Fear	Threatening with punishment or accusing the other personally
Imposing a restriction	Scarcity	Delay behaviour or making something available in a limited way
Direct pressure	Power of Repetition	Exerting pressure on the other in a neutral manner by being firm
Legitimizing	Legitimacy	Referring to what has been agreed upon in society or with others
Exchanging	Reciprocity	Give-and-take behaviour
Rational Persuasion	Consistency	Use persuasive arguments and logic

The first strategy, *being kind*, makes use of the idea that we are more likely to be influenced by individuals we like and feel sympathetic towards (Giebels & Noelanders, 2004), sometimes referred to as the “liking principle” within negotiation literature (Guthrie, 2004; Korobkin, 2024). The second strategy, *being equal*, is based on the social influence principle of similarity, which suggests that we are more likely to be influenced by a person when they establish and refer to shared interests and mutual experiences (i.e., similarities) (Giebels & Noelanders, 2004). The third strategy, *being credible*, combines expertise and reliability, both of which are crucial factors in effectively influencing the opposing party in a negotiation (Giebels & Noelanders, 2004). The three strategies discussed thus far are relational strategies, that foster the interpersonal dynamic of the negotiation parties.

The following seven strategies are content based. The strategy *emotional appeal* leverages the self-image of the counterpart, employing emotional, value-based, and idealistic arguments. Much of this appeal is rooted in the human desire to ‘be good’ (Giebels & Noelanders, 2004). The fifth influence strategy, *intimidation*, straddles the boundary to coercion, as it involves the use of threats, and personal attacks to instil fear in the opposing party (Giebels & Noelanders, 2004). The sixth strategy, *imposing a restriction*, limits the availability of a negotiable. In crisis negotiations this often refers to a time restriction, or scarce responses by either party (Giebels & Noelanders, 2004). The seventh strategy, *direct pressure*, can include behaviours like repeating a request, or giving an instruction. It is similar to *intimidation* but usually less intense and not of a coercive nature (Giebels & Noelanders, 2004). The eighth strategy, *legitimizing*, refers to the utilization of regulations, including legislation as well as societal standards and norms, to justify or validate an argument (Giebels & Noelanders, 2004). The ninth strategy, *exchanging*, is based on the idea that when someone offers us something we are inclined to reciprocate (Giebels & Noelanders, 2004). The tenth strategy, *rational persuasion*, utilizes logical reasoning and factual evidence. With consistency

as underlying principle, rational persuasion specifically makes use of the need to behave in a way that is consistent with earlier decision making (Giebels & Noelanders, 2004). Lastly, next to the influencing strategies discussed above, parties also resort to information exchange (i.e., not an influence strategy), which usually makes up one third of behaviours in crisis negotiations (Giebels & Noelanders, 2004).

How the current study applied the Table of Ten to the context of ransomware negotiations and what findings this generated is explained in later sections.

Strategic Sequences and the Application of cSPADE

As aforementioned, Beune and colleagues (2011) found that influencing behaviours can form strategic sequence. They coined this concept in the context of police interviews, specifically based on considerations regarding the “good cop-bad cop technique” (Beune et al., 2011). A strategic sequence refers to several influencing behaviours being grouped in strategic order to achieve a certain effect (Beune et al., 2011). This sequence can be executed by a single party or several (Beune et al., 2011). To the best of my knowledge the concept of strategic sequence or any comparable approach to identify patterns in communication data has not been applied to ransomware negotiations.

In this study, the cSPADE (constrained Sequential Discovery using Equivalence classes) algorithm (Lesh et al., 2000; Zaki, 2000, 2001) was utilized to identify patterns of influencing behaviours within ransomware negotiations. cSPADE is most commonly applied in market research to identify patterns in customer purchase behaviour (see for example Fernando & Aw, 2023; Liu et al., 2023). Patterns in data mining refer to sequential occurrences of events. In this study, patterns denote successive influencing behaviours (i.e., strategic sequences) observed in ransomware negotiations.

There are various pattern mining tools which allow researchers to extract meaningful patterns from large data (Gupta & Chandra, 2020; Morita et al., 2005; Savaş, 2021).

Generally, they were first developed in the early 1990s (Savaş, 2021) and have since been applied to various different domains, with notable prevalence in economics and biology (Gupta & Chandra, 2020; Savaş, 2021). However, none of them were used in the context of influencing dynamics within ransomware. For the current study, ultimately, the cSPADE algorithm was chosen as analysis tool because it is recognized as one of the most efficient pattern mining algorithms (Aoga et al., 2016). Additionally, several sources commend on its easy implementation (see for example Fernando & Aw, 2023; Zhang & Paquette, 2023) and it is open source (Zaki, 2018/2024). cSPADE is included in the *arulesSequence* R package (Buchta & Hahsler, 2023).

Details on how cSPADE was utilized can be found in the method section.

Formulation of the Research Questions

Based on the exploration of literature on ransomware attacks and social influence, there is a clear need for an empirical investigation into the influencing behaviours present in ransomware negotiations. The review above highlights gaps in understanding the dynamics of ransomware negotiations, particularly the strategies and sequences of behaviours used by both threat actors and victims. To address these gaps, this study aims to explore the influencing behaviours and their patterns within ransomware negotiations using empirical data. Hence, the following research questions guide the current study:

RQ1: What are key influencing behaviours employed by threat actors and victims in ransomware negotiations?

RQ2: Are there frequently occurring patterns of influencing behaviours, across different ransomware cases?

Methods

Before explaining the methodology used in this study, it should be noted that this research was part of a broader project. Together with another bachelor's student from the University of Twente, I collected and coded original ransomware logs, effectively ensuring reliability of the coded material. While we both worked within the context of ransomware negotiations and used the same data, we explored different research goals. When describing the data analysis process below I refer to him as my research partner.

Data

For this study, publicly available data logs were sampled from ransomware.live, a website which hosts a total collection of 147 ransomware logs (Mousqueton, 2024). In this context, a log refers to the written record of an actual ransomware negotiation that occurred in the past. Each log includes messages (i.e., utterances) labelled as having originated from either the victim or the threat actor. These utterances are presented in chronological sequence, mimicking the course of the negotiation. Some logs include timestamps with geographical time zone information, while others lack these contextual details. Additionally, to preserve anonymity, all identifiers of victims (e.g., names, email-addresses, organizations) have been redacted. This is true for almost all logs. According to a statement on ransomware.live, identity markers are only disclosed when the logs document cases where victim information has been publicly disclosed through another source (e.g., the media or threat actor) (Mousqueton, 2024).

The data was selected because it allows for the empirical investigation of influencing behaviours in real-world ransomware negotiation. However, before sampling logs from the total collection, inclusion criteria were determined to ensure relevance of the data in addressing the research questions.

Inclusion Criteria

To determine the inclusion criteria, a cursory examination of five original logs was conducted to gain a preliminary understanding of the dataset. Notably, during this process, me and my coding partner did not discuss our impressions of the data or applied our knowledge of influencing behaviours, to avoid introducing bias. The primary objective at this stage was to anticipate the nature of the data and inform the criteria development.

Sample Size and Number of Utterances. Initially, it was determined that a sample size of $n=25$ logs would be feasible for this research study. This decision was based primarily on time constraints imposed by the University of Twente (2024), necessitating completion of data analysis within one month. Furthermore, this decision was informed by experiences during the coding practice phase, where coding a short log comprising 25-35 utterances took nearly half an hour. Apart from the sample size, it was established that each log must contain a number of utterances between minimum 35 and maximum 80. This range was established through careful consideration of time constraint and the necessity for adequate information density. By establishing this range, we aimed to ensure that the data would enable comprehensive analysis and facilitate meaningful conclusions.

Stratified Sampling. Subsequently, it was determined that stratified sampling based on threat actor groups would add value to this research. By ensuring that the sample reflects the diversity of threat actors within the log repository as closely as possible our aim was to provide an information-rich sample that captures variations in influencing behaviours across different threat actor groups.

Ransom Payment Status. Lastly, the payment status of the ransom was considered, with the objective of achieving a balance between logs that indicated ransom paid and those indicating ransom not paid, adhering to a close to 50:50 ratio (considering $n=25$ this meant a 12:13 ratio). This criterion was established to fulfil the research requirements of another

study, where this “ransom paid versus not paid ratio” was necessary to address research questions. It’s important to note that while achieving a close to 50:50 ratio of logs was advantageous to the other research project, it did not accurately reflect the original dataset. Within the original dataset (no=147), the ratio is 23:77, with 34 logs indicating “paid” ransom and 113 logs indicating ransom “not paid”.

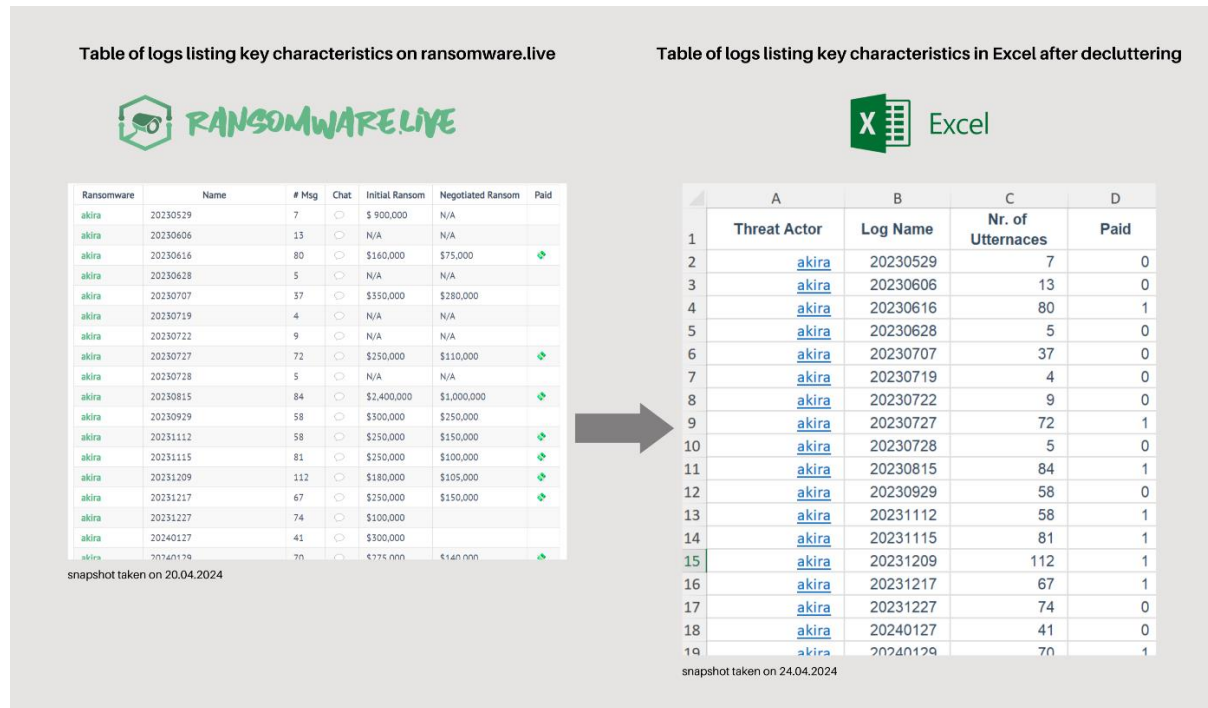
Sample Selection

Having established the above criteria, the sample was determined. For this, the ransomware.live website was utilized (Mousqueton, 2024). Since, it provides a structured overview of the logs; A table summarizing each log based on key characteristics (snapshot can be found in Figure 1). Consulting this table facilitated the sampling based on the criteria determined above, without the need of for an exhaustive review of the logs. This effectively minimized the introduction of additional bias, preserving the randomness of the sample while also ensuring relevance of the data for analysis purposes.

First the table of logs listing their key characteristics was copied into Microsoft Excel. Subsequently the table was cleaned by removing columns which contained information that was not needed for the sampling process. This included hyperlinks to simulated negotiation chats (in the form of speech bubble emojis), as well as information on the initial ransom and the negotiated ransom amount. Additionally, emojis which ransomware.live uses to indicate whether the ransom was paid or not were converted into numerical values of 1 and 0. The result (shown in Figure 1) was a table comprising essential information on each of the 147 logs, including the name of the threat actor, the log ID, number of utterances in the log, and the indication whether the ransom was paid or not.

Figure 1

Key Characteristics Used during Sample Selection in This Study: Snapshot of Original Table from Ransomware.live and Decluttered Table in Excel



Having created the comprised overview of the dataset in Excel, the sampling process continued by applying each of the inclusion criteria listed above.

Number of Utternaces. Initially, compliance with the first inclusion criteria, which necessitated an utterance range between 35 and 80, was ensured. For this the dataset (i.e., Excel table) was sorted based on the number of utterances. Thereafter, logs containing utterances below 35 or above 80 were manually selected and removed, using the “Delete” function in the context menu. This process resulted in a reduced dataset of n=64 logs.

Stratified Sampling. In order to adhere to the second inclusion criteria of producing a stratified sample, the number of logs to be sampled per threat actor was determined as followed: First, the distribution of threat actors in the original dataset (n=147) was established. This was done by dividing the total number of logs (n=147) by 100 and multiplying it by the number of logs per threat actor, previously determined using the

COUNTIFS() function in Excel. Secondly, this distribution was then applied to a sample size of $n=25$ and rounded to the nearest whole number, removing any decimal places. This resulted in a sample of only 22 logs instead of the desired $n=25$ because six threat actors appeared to contribute nothing to the sample. Each of their shares in the original dataset was $\leq 1.4\%$, which was insignificant at a sample size of $n=25$. Consequently, to maintain the distribution of threat actors in the original dataset while sampling a total of 25 logs, a decision was made to randomly select three logs from these six threat actors (i.e., treating them as one group). An overview of the distributions of threat actors in both the original dataset as well as the sample is presented in Figure 2.

Once the distribution was determined, a new Excel table was created for each threat actor, containing only logs with utterances falling between 35 and 80, as identified in the previous step. In total, there were nine tables: eight dedicated to individual threat actors that would contribute to the sample based on the computed distribution and one table comprising logs from the six threat actors that were grouped together (i.e., Avos, Babuk, Blackmatter, Cloak, Mount-Locker, and Ranzy).

Within each table the function `=INDEX(X:X;RANDBETWEEN(1;COUNTA(X:X)))` was used to randomly select the needed number of logs based on the threat actor distribution. This finalized the stratified sampling process and resulted in 25 randomly selected logs that closely resembled the threat actor distribution of the original dataset.

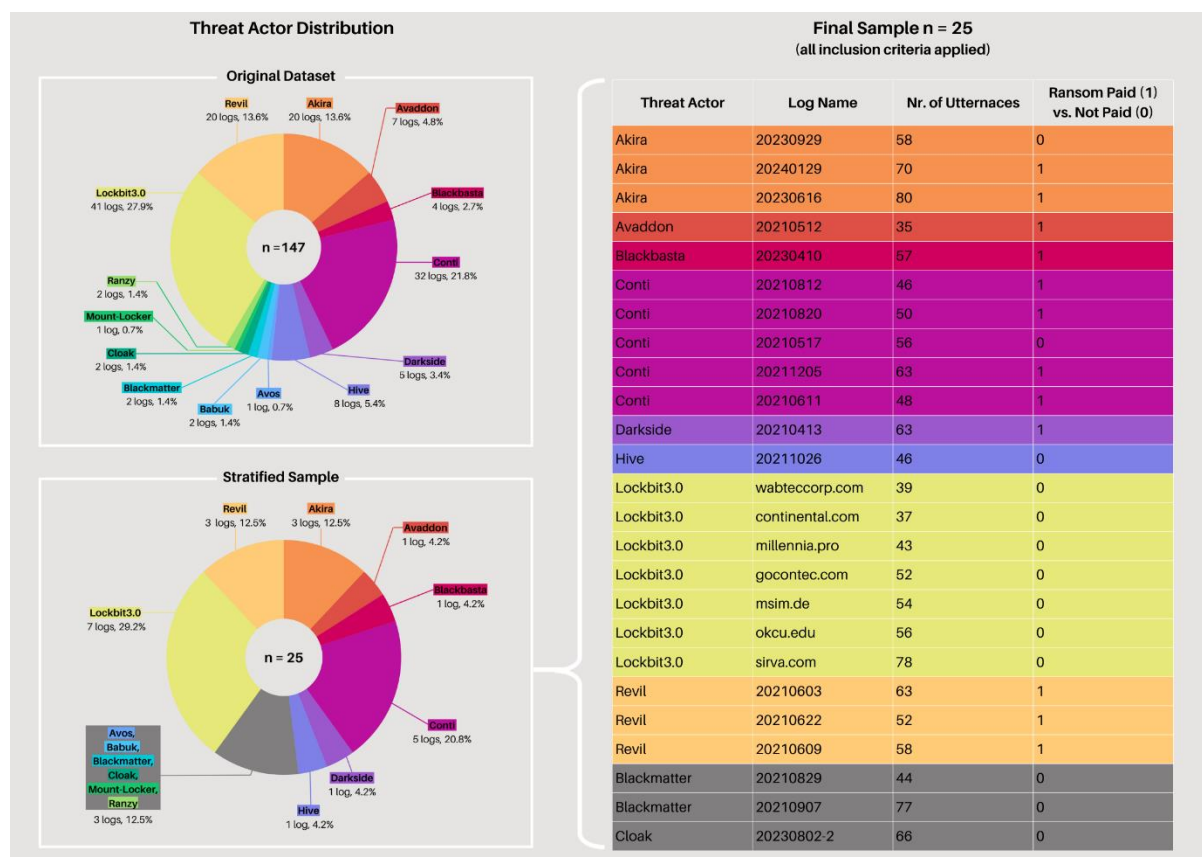
Ransom Payment Status. Within the initial stratified sample, merely eight logs (32%) indicated that the victim paid the ransom. To fulfil requirements of another study using the same data the last step in the sampling process was to rectify the “ransom paid versus not paid ratio” to resemble close to 50:50. This adjustment was accomplished by randomly selecting a log indicating that the ransom was not paid and replacing it with a ransom-paid-log from the same threat actor as to not disturb the stratified distribution.

Final Sample

Through the steps described above, the inclusion criteria were applied to the fullest extent feasible in the context of this study. An overview of the final sample, with each log consisting of 35-80 utterances ($\bar{x} = 55$), the threat actor distribution closely resembling the original dataset and a “ransom paid versus not paid ratio” of 48:52, can be seen in Figure 2.

Figure 2

Threat Actor Distribution and Final Sample



Sample Extraction. After determining the sample using information from the ransomware.live website, an open-source GitHub repository (Rieß-Marchive, 2023/2024) that hosts the same logs was utilized to extract the sample. While the GitHub interface is not as structured and user friendly (e.g., lacking an overview of the key characteristics), the data extraction process was more straightforward on this platform. The logs were downloaded as .json files and converted into Excel files using *JSON to Excel Converter* (n.d.). This

combination of sources allowed for a comprehensive approach to sampling and data retrieval for the current study.

Sample Pre-Processing. Within the excel files, to facilitate the subsequent data analysis, conditional formatting was used to colour all utterances from victims in grey. Constituting a visual help to clearly distinguishing between threat actor and victim utterances. Additionally, the “Wrap Text” function was enabled to ensure that even long utterances were properly displayed in the cells.

Data Analysis

The sample determined in the previous step was analysed as follows: First the content of each log was coded using the official Table of Ten coding guidelines (Giebels, 2023). Subsequently, different data analyses were conducted using the coded material. All steps are described in detail below.

Coding

Me and my research partner underwent training to code ransomware logs based on influencing behaviours outlined in the Table of Ten. The coding guidelines reported by Giebels (2023) formed the basis for this training. An overview of these guidelines can be found in Figure 3, they include eleven codes based on the influence strategies in the Table of Ten and an additional code for “Information exchange”. The latter is only to be coded when no influence strategy is present. As indicated by Giebels (2023) “speaking turn” was determined as level of analysis, meaning that one code would be assigned per utterance.

Figure 3

Coding Guidelines based on the Table of Ten (Giebels, 2023)

Strategy	Underlying Principle	Example Behaviour
1. Being Kind	Sympathy	A. Active listening B. Show empathy C. Kindly offer something
2. Being Equal	Identification	A. Use "We" instead of "I" / "You" B. Stress something you have in common (e.g., background, family circumstances, hobbies) C. Emphasize mutual goal / dependence / enemy
3. Being Credible	Authority	A. Show reliability (do what you say) B. Emphasize your expertise / experience (you know what you are doing) C. Show you are transparent
4. Emotional Appeal	Self-Image (Heart)	A. Touch upon feeling / ask for sympathy (how it affects you / victim) B. Praise other's behaviour C. Boost other's self-respect
5. Intimidation	Deterrence	A. Warnings B. Threats C. Condemn transgression
6. Imposing Restriction	Scarcity	A. Postpone Answer B. Ignore other / not being available C. Offer limited Choice (A or B)
7. Direct Pressure	Power of Fact / Repetition	A. Repeat request (planting the seed) B. Share fact C. Give instruction
8. Legitimizing	Legitimacy (External)	A. Reference to formal rule / the law B. Reference to procedures C. Mentioning of moral / social codes
9. Exchanging	Reciprocity	A. Ask for something in return B. Lower your bid C. Exchange proposal
10. Rational Persuasion	Consistency	A. Use of arguments B. Provide logic C. Confront with inconsistencies
11. Information Exchange		A. Asking information (e.g., What is your name?) B. Giving information (e.g., My name is Pim.)

To aid subsequent data analysis, a distinction was made within the codes between utterances from the threat actor and the victim. Numbers 1-11 were allocated to code victim

utterances, with each number representing a specific influencing behaviour as listed above.

The numbers 12-22 were assigned to code utterances from the threat actor, following the same order (i.e., 12 = being kind, 13 = being equal, 22 = information exchange).

Coding Practice. My research partner and I coded a total of 21 practice logs before proceeding to analyse the actual data sample. For this purpose, AI-generated logs, based on original ransomware negotiations, were utilized.

The practice phase spanned ten hours in total, encompassing individual coding of the logs and subsequent collaborative meetings for comparison. During these discussions, Cohen's Kappa (κ) was calculated to assess intercoder reliability. Initially, the coding of each individual practice log was followed by immediate discussion, to promptly address any discrepancies. Once agreement scores became more stable, a bulk coding approach was adopted, with both of us coding five practice logs consecutively before convening to compare and discuss them.

Throughout the practice phase, additional coding rules (presented in Figure 4) were formulated to enhance intercoder reliability. This was done whenever we were able to identify reoccurring differences, during the discussions following the individual coding. Within these discussions, the underlying principles and example behaviours outlined in regard to the Table of Ten (Giebels, 2023; Giebels & Noelanders, 2004) provided the foundation for any new rule. Overall, the practice and formulation of additional rules contributed to the increase of inter-coder reliability from an initial $\kappa = 0.37$ to an average $\kappa = 0.68$.

Figure 4

Final Coding Scheme: Coding Guidelines based on the Table of Ten (Giebels, 2023) and Additional Rules formulated during the Practice Phase

Strategy	Underlying Principle	Example Behaviour	Additional Coding Rules	General Considerations
1. Being Kind	Sympathy	A. Active listening B. Show empathy C. Kindly offer something	"Thank you" at the end of the transcript is 1 Single word utterance like "ok" or "Hello" etc. are coded 1 for active listening (A) • as long as there is no negative undertone !! • if there is additional information exchange (e.g., "hello what is your name?"), it is 1.1	<p>Consider the reason for the utterance (i.e., intention) rather than focusing on a specific part of an utterance (especially when multiple influencing behaviours seem appropriate). Always highlight the whole utterance in Atlas.ti</p> <p>Redacted messages</p> <ul style="list-style-type: none"> fully redacted messages = not coded partially redacted; message understandable = appropriate influencing behaviour coded partially redacted; message NOT understandable = information exchange (1.1) <p>Difference between being credible (3) and legitimizing (8) Communicating ongoing processes (being transparent) or reference externally dictated expectation or procedures e.g., social norm, the law or corporate governance guidelines (legitimizing):</p> <ul style="list-style-type: none"> "Our stakeholders are involved. This is more complex than a simple transaction." explanation of what is going on and being considered at this point... coded as 3 "We need to ensure that our actions do not set a precedent or encourage further attacks this one is about social rules." this clearly places responsibilities onto an external cause = coded as 8 <p>Difference between intimidation (5) and imposing a restriction (6) Making more time a necessity to continue the negotiation (threatening) or explaining that some time is needed to come to a decision (postponing an answer):</p> <ul style="list-style-type: none"> "We need more time" implies threat to end the negotiation = coded as 5 "We will get back to you" implies willingness to continue after recess = coded as 6 <p>Rules in the context of time limits The first time a time limit is mentioned it is coded as 6 if (after the first occurrence), the time limit is mentioned again unprompted it is coded as 7 if the victim asks for an extension and...:</p> <ul style="list-style-type: none"> the threat actor response is to keep the original time limit it is coded as 6 the threat actor extends the deadline it is coded as 9 <p>Exception for the use of self-image in rational persuasion (10) instead of emotional appeal (4) Most utterances that refer to self-image are coded as 4. Always check if the utterance includes an attempt to arouse an emotional response (e.g., pleading, asking for sympathy, mentioning devastating consequences). However, if the self-image is clearly used in an appeal to logic it is coded as 10. "It seems you are misinformed. We are a small company with only 3 employees. We do not have these kinds of funds" this utterance uses self image in a causal argument = coded as 10</p>
2. Being Equal	Identification	A. Use "We" instead of "I" / "You" B. Stress something you have in common (e.g., background, family circumstances, hobbies) C. Emphasize mutual goal / dependence / onemy		
3. Being Credible	Authority	A. Show reliability (do what you say) B. Emphasize your expertise / experience you know what you are doing C. Show you are transparent	if the reason for the utterance is to prove credibility it is coded as 3 • The threat actor sending of proof (i.e., decrypted files) is coded as 3 (being credible) • The victim sending files that are to be decrypted by the threat actor is coded as 1.1 (information exchange)	
4. Emotional Appeal	Self-Image (Heart)	A. Touch upon feeling / ask for sympathy (how it affects you / victim) B. Praise other's behaviour C. Boost other's self-respect		
5. Intimidation	Deterrence	A. Warnings B. Threats C. Condemn transgression	if victim mentions (threatens with) alternative solution or refusal of payment it is coded as 5 • e.g., mentioning of backups or insurance payments	
6. Imposing Restriction	Scarcity	A. Postpone Answer B. Ignore other / not being available C. Offer limited Choice (A or B)	Only use 6 after checking that the principle of scarcity applies in case of limited choice (C); only code 6 if there is explicit mention of two options	
7. Direct Pressure	Power of Fact / Repetition	A. Repeat request (planting the seed) B. Share fact C. Give instruction	The first time the threat actor states the ransom demands it is coded as 7	
8. Legitimizing	Legitimacy (External)	A. Reference to formal rule / the law B. Reference to procedures C. Mentioning of moral / social codes	Legitimization (8) only applies if the continuation of the negotiation is directly dependent on the referenced framework	
9. Exchanging	Reciprocity	A. Ask for something in return B. Lower your bid C. Exchange proposal	if a question references something as a condition for proceeding it (asking for something in return) is coded as 9 • e.g., "Can you guarantee that paying will prevent future attacks?" When the threat actor accepts a diminished proposal from the victim (lowering your bid) it is coded as 9	
10. Rational Persuasion	Consistency	A. Use of arguments B. Provide logic C. Confront with inconsistencies		
11. Information Exchange		A. Asking information (e.g., What is your name?) B. Giving information (e.g., My name is Fin.)	Given information (B) should be a self-evident fact When deciding between information exchange (1.1) and an influence behaviour; always choose the influencing behaviour	

Coding the Sample. After practicing and agreeing on the rules listed above, a transition from practice data to analysing the determined sample data was made. For this the sample logs were imported into Atlas.ti and a coding environment was set up. Meaning, both coders received identical Atlas.ti project bundles containing the sample logs (n=25) and 22 codes (i.e., eleven codes as formulated by Giebels (2023), each per victim and per threat actor). Subsequently, we started coding individually on our respective devices, working independently to code all 25 logs of ransomware negotiations. Afterwards we convened to determine the intercoder reliability score ($\kappa = 0.66$) and established consensuses for all codes that were coded differently.

After the final codes for all logs were established, I started to analyse them aiming to answer the research question formulated above.

Utterance Ratio

Within the total of coded behaviours there was a potential for an uneven ratio between threat actor and victim utterances. Therefore, the first step was to understand how the utterances were distributed between threat actors and victims within the ransomware logs. Thus, the utterance ratio within each log was calculated by tallying all utterances originating from each party (COUNTIFS() function in Excel) and relating this count to the total number of utterances in the respective log. The average ratio is reported below in Figure 7. Next to it, the overall utterance ratio is reported, showing the total number of utterances each actor has contributed throughout all logs (same Figure 7).

Relative Frequency of Codes

Following this, the relative frequency of each influencing behaviour was computed per party threat actor and victim. This analysis allowed to identify key influencing behaviours for each party and allowed for comparison between them (i.e., addressing RQ1). As previously explained each code (1-22) denoted the party in addition to the influencing behaviour. This facilitated the frequency analysis by allowing to effectively distinguish between utterances of threat actors and victims and thus easily determine the frequency of party-specific behaviours as follows.

First, the raw count (i.e., frequency) of each code was generated within each of the 25 logs using the =COUNTIF function in Excel. After obtaining the frequency of each code per log the total number of utterances per party was determined in the individual logs. This total was then used to calculate the percentage (i.e., relative frequency) representing how extensively each party used each influencing behaviour (i.e., code) within that specific case (i.e., log) relative to the total actions performed by the party in that case.

The average relative frequencies of each code per threat actor and victim are reported below in Table 2 and 3 respectively. Furthermore, additional statistical measures, including minimum and maximum values, and standard deviation, are provided to offer further context.

Most Frequent Code per Utterance Level (Mode)

The subsequent step in this study entailed the computation of the mode at the utterance level. To clarify, the sequence of utterances was numbered based on the order they were sent in each ransomware log. Defining the utterance level like this, allows for meaningful comparisons across different logs by ensuring each level corresponds to a specific point in the negotiation sequence. Thus, after the utterance level were established, the most frequently occurring code at each level across all logs was computed (i.e., the mode of all codes ranging from 1-22). The findings of this analysis are reported below in Figure 8 and Figure 9.

Pattern Identification Using cSPADE Algorithm

As described in the introduction, cSPADE was used to identify frequent patterns of successive influencing behaviours. However, before running the algorithm in R, the dataset had to be formatted to ensure compatibility with cSPADE. The algorithm requires the dataset to include three variables, which must be sorted in vertical order. The first variable is *log ID*, followed by *time*. Since not all logs in the current sample contained timestamps, a sequence of successive dates was randomly generated for each log. It is important to note that this procedural step did not affect the analysis, but rather facilitates the internal processing of cSPADE to sequence the behaviours correctly. To be precise, the algorithm merely utilizes the time variable to later assign an event ID that numerically represents the order of influencing behaviours in each log, commencing from 1. The last variable is *influencing behaviour*. An overview of the dataset as it was loaded in R can be seen in Figure 5.

Figure 5*Snapshot of the Dataset formatted for cSPADE Processing*

	log ID	time	influencing behaviour
	A	B	C
1	customer.identifier	purchase.date	product
2	D1	01.01.2024	L
3	D1	02.01.2024	A
4	D1	03.01.2024	K
5	D1	04.01.2024	R
6	D1	05.01.2024	I
7	D1	06.01.2024	V
8	D1	07.01.2024	C

snapshot taken on 19.06.2024

It can be seen here that, to ensure the R script was running smoothly the variable names (i.e., column headings) were kept as they are in the original R script appropriated from Wan (2022/2022). The figure shows in grey what each column refers to in the context of the current study. Furthermore, cSPADE processes the influencing behaviours using letters instead of numbers. Therefore, as part of the formatting, Excel's 'find and replace' function was employed to convert the coded behaviours (1-22) into corresponding letters (A-V).

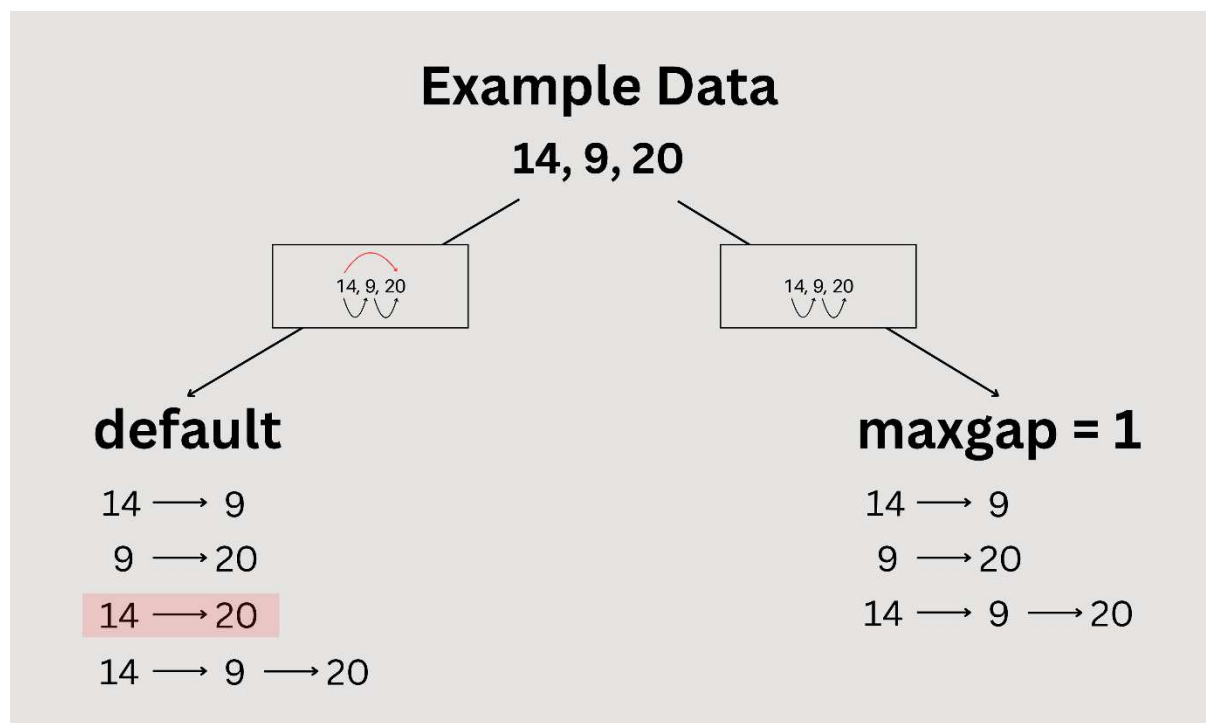
Once formatting was completed the dataset in form of a .csv file was loaded in R. The algorithm allows the application of constraints in the pattern mining process, in order to minimize computational time and produce outcomes that are significant with respect to specific research requirements (Liu et al., 2023). One of the optional constraints involves setting a support threshold, which restricts the algorithm from exploring patterns with low support and high risk of randomness. In cSPADE, support is measured by the frequency of occurrence across sequences (Morita et al., 2005). In the current context, support for a pattern

translates to the number of logs out of the total 25 logs where the pattern occurred at least once. Thus far there are no guidelines for determining an appropriate support threshold when working with cSPADE (Zhang & Paquette, 2023). I decided on 0.2, meaning, that for a pattern to be identified and noted in the cSPADE output it had to occur in at least 5 logs.

Another constraint that was set before running the R script was the maximum gap between sequential behaviours. By default, cSPADE considers sequences of behaviour that are not immediately consecutive when mining for frequent patterns (Buchta & Hahsler, 2007; Zaki, 2000). While this can be useful in the context of (e.g.,) market basket analysis, these non-consecutive patterns are not meaningful in the current context. Therefore, to address this the maxgap parameter was set to 1. An illustration of the difference between the default setting and maxgap parameter equalling one is shown in Figure 6 below.

Figure 6

Illustration of the Difference between the cSPADE Default Setting and Maxgap = 1



Note, that datasets that include original timestamps need to be handled differently since setting maxgap to 1 prevents behaviours occurring on the same day from being grouped

together into patterns. However, with the way the current dataset was formatted this issue does not arise.

After setting the constraining parameters the R script was executed and the results were exported as .csv files. The most significant results are discussed below.

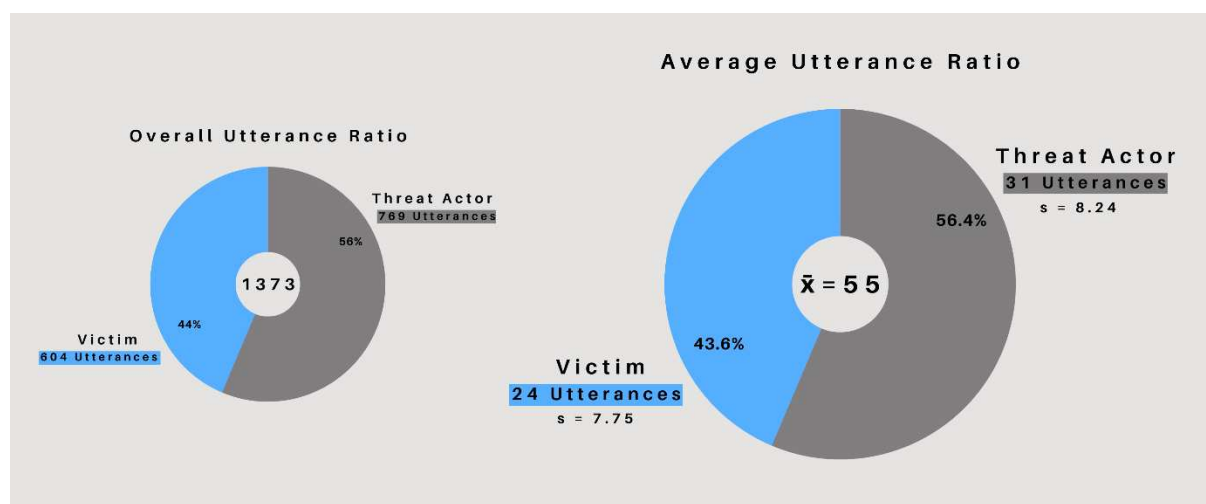
Results

Utterance Ratio

The dataset contained a total of 1,387 utterances. However, 14 of these utterances were fully redacted and therefore excluded from the coding process, resulting in 1,373 coded utterances from both the threat actors and victims. An analysis of the dataset revealed an uneven distribution of utterances. The threat actors contributed 165 more utterances than the victims. Consequently, the overall utterance ratio comprised 44% of messages sent by the victims and 56% by the threat actors. This ratio is closely reflected in the average contributions per log, with victims contributing 43.6% of utterances on average (standard deviation, $s = 7.75$) and threat actors contributing 56.4% on average (standard deviation, $s = 8.24$). A visualization of these results is presented below in Figure 7.

Figure 7

Threat Actor and Victim Utterance Ratio across 25 Ransomware Logs: Overall and Average



Frequency Analysis

Below, the relative frequencies of influencing behaviours (i.e., codes) within the dataset are reported. Table 2 presents the frequencies for the threat actors, while Table 3 provides the frequencies for the victims. As described in the method section, the average relative frequency for each code was calculated individually for each log and then averaged across the logs. This approach addresses the imbalance in the utterance ratio, enabling meaningful comparisons between the two parties. Additionally, the tables include information on the minimum and maximum occurrences, as well as the standard deviation for the relative frequencies.

Table 2

Threat Actors: Average Relative Frequency and Additional Statistical Measures

Influencing Behaviour	Relative Frequency (rounded %)	Min	Max	SD
Being Kind	8%	0	6	1,41
Being Equal	< 1%	0	1	0,20
Being Credible	32%	3	24	6,06
Emotional Appeal	< 1%	0	2	0,44
Intimidation	10 %	0	10	2,96
Imposing a Restriction	4 %	0	4	1,16
Direct Pressure	16 %	0	21	4,24
Legitimizing	< 1%	0	1	0,33

Exchanging	11 %	0	7	2,08
Rational Persuasion	4 %	0	5	1,35
Information Exchange	13 %	0	9	2,19

The table above shows that 13% of the threat actor's utterances were classified as information exchange. The remaining 87% were identified as influencing behaviours. Of these influencing behaviours, 47% were relational strategies (i.e., being kind, being equal, and being credible), and 53% were content strategies.

The most frequent influencing behaviours employed by threat actors were *being credible* ($f_r = 32\%$), *direct pressure* ($f_r = 16\%$) and *exchanging* ($f_r = 11\%$). *Being credible* was the only influencing behaviour found in every log (i.e., $min = 3$). Furthermore, it exhibited the highest standard deviation ($s = 6.06$). Similar to *being credible* ($max = 24$), *direct pressure* showed a high maximum occurrence ($max = 21$) and significant dispersion from the mean ($s = 4.24$). *Exchanging* displayed less variability ($s = 2.19$) and average maximum and minimum occurrences ($min = 0$ and $max = 7$) compared to other influencing behaviours. Additionally, it is noteworthy that *intimidation*, another relatively frequent ($f_r = 10\%$) influencing behaviour employed by the threat actor, shows a comparatively high variability across logs ($s = 2.96$).

In contrast, the least frequent influencing behaviours employed by the threat actors included *being equal* ($f_r = 0.09\%$), *emotional appeal* ($f_r = 0.33\%$) and *legitimizing* ($f_r = 0.34\%$). All of these behaviours exhibited low minimum and maximum occurrence as well as low variability in terms of standard deviation.

Table 3*Victims: Average Relative Frequency and Additional Statistical Measures*

Influencing Behaviour	Relative Frequency (rounded %)	Min	Max	SD
Being Kind	5%	0	6	1,49
Being Equal	< 1%	0	3	0,62
Being Credible	19%	0	13	2,94
Emotional Appeal	2%	0	2	0,77
Intimidation	4%	0	6	1,71
Imposing a Restriction	6%	0	6	1,58
Direct Pressure	2%	0	8	1,68
Legitimizing	2%	0	2	0,71
Exchanging	27%	1	10	2,71
Rational Persuasion	10%	0	7	2,06
Information Exchange	22%	1	12	3,21

The table above shows that 22% of the victims' utterances were classified as information exchange. The remaining 78% were identified as influencing behaviours. Of these influencing behaviours, 30% were relational strategies (i.e., being kind, being equal, and being credible), and 60% were content strategies.

The most frequent influencing behaviours employed by victims were *exchanging* ($f_r = 27\%$), *being credible* ($f_r = 19\%$), and *rational persuasion* ($f_r = 10\%$). *Exchanging* was the only influencing behaviour that the victims employed in each log (i.e., $min = 1$). Additionally, they also made use of information exchange at least once in every log (i.e., $min = 1$). In terms of deviation from the mean influencing behaviours used by the victims showed modest variability. *Being credible* showed the highest standard deviation ($s = 2.94$) out of the influencing behaviours. *Exchanging* exhibited similar values ($s = 2.06$). All other behaviours had less dispersion around the mean.

This was also true for the least frequent influencing behaviours employed by victims, namely *being equal* ($f_r = 0.57\%$), *direct pressure* ($f_r = 2\%$), and *legitimizing* ($f_r = 2\%$). *Emotional appeal* ($f_r = 2\%$), and *intimidation* ($f_r = 4\%$), exhibited similarly low frequencies.

Most Frequent Code per Utterance Level (Mode)

The figure below presents a visualization of the most frequent influencing behaviour per utterance level. All logs begin at the first utterance level, with each log in the sample containing between 35 and 80 utterances, as previously mentioned. Consequently, the number of logs contributing to the mode gradually decreases after the 35th utterance level, as indicated by the lighter shading. Additionally, beyond the 63rd utterance level, there was an insufficient number of logs contributing to the mode resulting in the analysis returning no value, denoted by the fading N/A in the bottom of the figure. Note, that the average number of utterances in the sample is 55.

Figure 8

Most Frequent Code per Utterance Level (Mode)

Utterance Level	Most Frequent Behaviour (Mode)		Support
	Threat Actor	Victim	Percentage of Logs contributing to Mode
1		Information Exchange	52%
2		Information Exchange	24%
3		Information Exchange	40%
4	Being Credible		32%
5	Being Credible		24%
6	Being Credible		28%
7	Being Credible		24%
8	Being Credible		24%
9	Being Credible		28%
10	Being Credible		24%
11	Being Credible		24%
12	Being Credible		28%
13	Direct Pressure		12%
14	Being Credible		28%
15		Information Exchange	20%
16	Direct Pressure		16%
17		Exchanging	20%
18	Being Credible		32%
19		Exchanging	20%
20	Being Credible		24%
21	Being Credible		20%
22		Being Credible	20%
23	Being Credible		16%
24	Being Credible		20%
25		Information Exchange	16%
26	Being Credible		16%
27		Exchanging	28%
28	Direct Pressure		24%
29		Exchanging	16%
30	Direct Pressure		16%
31	Being Credible		16%
32		Exchanging	20%
33		Being Credible	24%
34	Being Credible		24%
35	Direct Pressure		16%
36		Being Credible	12,5
37	Information Exchange		20,83%
38	Being Credible		20,83%
39		Exchanging	27,27%
40	Being Credible		18,18%
41		Being Credible	13,64%
42	Exchanging		22,73%
43	Being Credible		23,81%
44	Information Exchange		19,05%
45		Exchanging	20%
46		Being Credible	20%
47	Being Credible		22,22%
48		Information Exchange	16,67%
49		Being Credible	17,65%
50		Being Credible	23,53%
51	Being Credible		25%
52	Being Credible		40%
53		Being Credible	28,57%
54	Being Credible		23,08%
55		Information Exchange	23,08%
56	Being Credible		27,27%
57		Exchanging	18,18%
58	Being Credible		36,36%
59	Being Credible		37,5%
60	Intimidation		25%
61		Information Exchange	25%
62		Information Exchange	25%
63	Being Credible		37,5%
64	N/A	N/A	N/A
65	N/A	N/A	N/A
66	N/A	N/A	N/A

A notable first observation that can be made from the figure above, is the long sequence of *being credible* by the threat actors from the fourth to the 12th utterance level, with an average support of 26%. Beyond these levels, *being credible*, particularly when employed by the threat actor, appeared frequently throughout the negotiation, often spanning two consecutive utterance levels. Related to this, the figure shows threat actors using *being credible* predominantly, whereas the victim side in the figure shows a more balanced mix of *being credible*, *exchanging* and *information exchange*, without a clear primary behavior.

After threat actor *being credible* in the upper utterance levels, the figure shows them using *direct pressure* sporadically across the utterance levels in the upper half of the figure (above the 35th utterance level). The victim side in this figure does not show *direct pressure* on any utterance level.

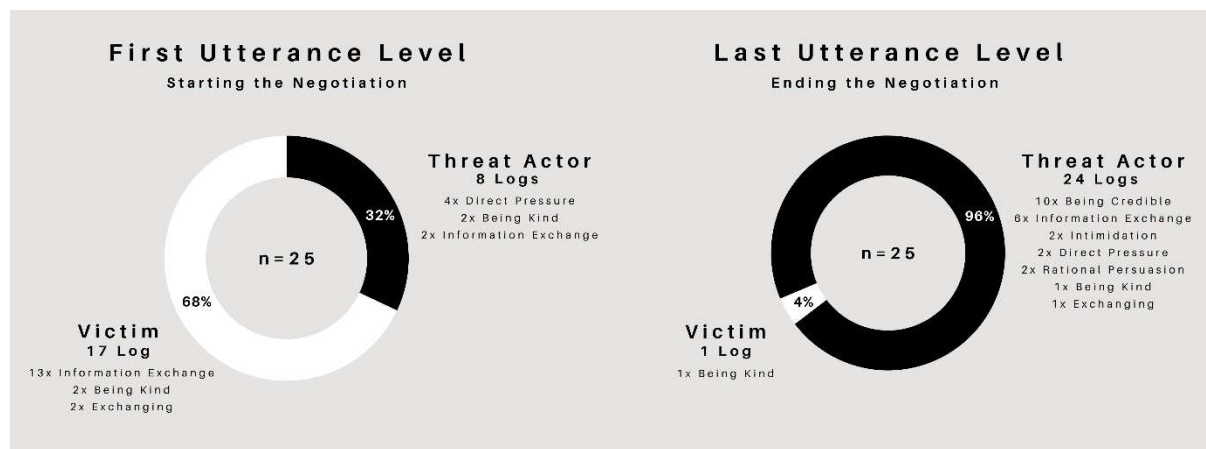
However, victim *exchanging* is depicted relatively consistently across the utterance levels after the initial phase, maintaining stable support (average 21%) until level 57. In contrast, the figure shows that threat actors using *exchanging* being the most frequent behaviour per utterance level occurred only once, at level 42.

Lastly, it is noteworthy that at the very first utterance level, the most frequent behaviour, observed in 13 out of 25 analysed logs (52%), is the victim employing information exchange (not influencing). This behaviour continues for the next two utterance levels with decreasing support (24% and 40%, respectively).

Generally, examining the starting point (first utterance level) of each log is the only directly comparable point across logs of varying lengths. Therefore, Figure 9 summarizes the first utterance level of each log in terms of influencing behaviour. Additionally, the figure provides a summary of the last utterance level in each log.

Figure 9

Codes at the First and Last Utterance Level of Ransomware Logs



Looking at the first utterance level across all logs, there were 17 victim-initiated logs (68%). Among these, as already shown by the mode in Figure 8 (above) 13 were categorized as *information exchange*. Out of the remaining four, two instances involved the victim *being kind* or and two involved *exchanging*. Comparatively, there were 8 logs (32%) where the first utterance came from the threat actor. These can be further divided into four instances of *direct pressure*, two instances of *being kind* and two *information exchange*.

Considering the last utterance level within each log, the majority (96%) of them consisted of threat actor utterances. Among these, there were ten instances of *being credible*, six instances of *information exchange*, as well as two instances each of *intimidation*, *direct pressure* and *rational persuasion*. Furthermore, there was one instance of the threat actor *being kind* and one instance of them *exchanging*. In contrast there was one log (4%) where the victim contributed the last utterance by *being kind*.

Pattern Identification Using cSPADE Algorithm

The cSPADE analysis identified 128 patterns with a support threshold of 20%. These patterns ranged in length from one to four behaviours. The most frequent single-behaviour patterns identified were *threat actor being credible* and *victim exchanging*, both with a support of 100%, indicating that they occur in every log at least once. In the exploration of

identified patterns below, single-behaviour-patterns are excluded, as their frequency has been discussed in detail in the previous analyses. However, it is worth noting that cSPADE corroborated the previous observations regarding the frequency of codes (i.e., single-behaviour patterns). Having established this, the following Figure 10 and related text explores patterns with the highest overall support, beginning with those comprising a minimum of two behaviours.

Figure 10

Patterns with the Highest Overall Support

cSPADE output: Patterns with Highest Overall Support (>50%)		
Pattern		Support
Threat Actor	Victim	Percentage of Logs with at least once occurrence
Being Credible Being Credible		84%
Being Credible	Exchanging	76%
Exchanging	Exchanging	68%
Being Credible	Exchanging	68%
	Information Exchange Information Exchange	60%
Being Credible	Being Credible	60%
Being Credible	Information Exchange	56%
Being Credible	Being Credible	52%
Direct Pressure	Exchanging	52%
Exchanging	Exchanging	52%
Being Credible Being Credible Being Credible		52%

Note. This figure exhaustively lists all patterns (excluding single-behaviour-patterns) that achieved above 50% support. Meaning each pattern displayed above occurred in at least 13 out of the 25 analysed logs at least once. The patterns are ordered by support.

As shown above the most supported pattern (excluding single-behaviour-patterns), occurring in 84% of the logs in this sample, is the *threat actor being credible followed by the threat actor being credible*. Similarly, cSPADE identified a pattern of the *threat actor being credible three times in a row*, which occurred in 52% of the logs.

There was also a pattern identified with the same code being used consecutively by the victim, specifically, the *victim using information exchange twice in a row* (not an influencing behaviour according to the table of ten).

Another interesting observation is that there are three pairs of patterns that include the same behaviours with switched parties. The first pair is the *threat actor being credible followed by the victim being credible* (60% support) and the *victim being credible followed by the threat actor being credible* (52% support). The second pair of patterns is the *victim exchanging followed by the threat actor being credible* occurred in 76% of the logs. The reversed pattern being, the *threat actor being credible first followed by the victim exchanging*, occurred in two fewer logs, exhibiting slightly lower support (68%). Similarly, there is a difference in support between the pattern where the *victim is exchanging and the threat actor respond with exchanging* (68%) and the opposite scenario, where *the threat actor exchanging and the victim responding with the same behaviour* (52%).

Other identified patterns include the *victim using information exchange followed by the threat actor being credible* (56%) and the *threat actor employing direct pressure and the victim reacting by exchanging* (52%).

The patterns with the highest support, as identified above, are all, except for one, two-behaviour patterns. To also include longer patterns, which generally showed lower support,

illustrations of all three-behaviour patterns and four-behaviour patterns identified through cSPADE are shown in Appendix B.

Discussion

The current study set out to explore the use of social influencing behaviors within the context of ransomware negotiations, specifically addressing the following gaps in the existing research: (1) the overall lack of insight into the negotiation stage of ransomware incidents, (2) scarce knowledge regarding the underlying dynamics between threat actors and victims in this context, and (3) the absence of empirical validation in ransomware research.

Additionally, this research makes a novel contribution by applying the Table of Ten by Giebels (2002, as cited in Giebels & Noelanders, 2004) to the context of ransomware negotiations. Furthermore, it utilized a novel approach for pattern analysis by employing the cSPADE algorithm (Zaki, 2000, 2001).

Thus, at the beginning of this paper the following three research questions were formulated to guide the current study:

RQ1: What are key influencing behaviours employed by threat actors and victims in ransomware negotiations?

RQ2: Are there frequently occurring patterns of influencing behaviours, across different ransomware cases?

To answer them 25 original ransomware logs were sourced and coded using the Table of Ten. Subsequently the several analyses were conducted including the computation of the utterance ratio, relative frequency of coded behaviours, the most frequent behaviour per utterance level (mode) and pattern mining. Below the key findings of this study are discussed followed by a section on limitations and recommendations for future research.

Key Findings

The Threat Actor Wants to Seem Credible

The results of the current study highlighted the extensive use of being credible as an influencing behaviour by the threat actor. This was corroborated by all analyses. It is the most frequently used code of threat actors and the mode analysis as well as the pattern mining showed that there are long sequences where the threat actor uses this repetitively.

To give more context, being credible within the sample was mostly observed in the form of file exchanges to provide proof of the threat actor's decryption ability. Usually this meant that the threat actor sends (on average) three decrypted files to the victim (each file coded as 14 Threat Actor Being Credible). This happens in a majority of the logs, usually within the beginning of the negotiation, and it is often unprompted. Meaning, the threat actor initiates this being credible and showing proof before the victim even asks for proof. Likely because their reputation and credibility is one of the most important factors when it comes to convincing the victim to pay (Cartwright et al., 2019). Victims have no other way to assess the threat actor as a 'potential business partner' (Cartwright et al., 2019). Plus, they would never send money if they had doubts about receiving their data after the ransom is paid (Cartwright et al., 2019).

Thus, overall, this observation confirms what has thus far been theorized in literature. As mentioned in the introduction both the game theory study by Ryan and colleagues (2022), and Wade's (2021) inferences based on traditional crisis negotiations suggested that trust in the decryption ability of the threat actor plays a key role in increasing the victim's willingness to pay.

The Threat Actor Sends More Messages

I found that the threat actors in the current sample contributed significantly more to the negotiation (on average 12.8%). This is a steady trend seen in the total utterance ratio and

confirmed by the average utterance ratio (Figure 7). While the average ratio shows significant standard deviation, this may very well be attributed to variability in log length (35-80). The underlying trend persists across the different logs.

There are several possible explanations for this. First, it could imply a certain power asymmetry. In the introduction it was explained that the information asymmetry, manufactured through the attack preparation of the threat actor, is indicative of the threat actor likely dominating the negotiation (Faivre, 2023; Ryan et al., 2022). This has also been found in research on police interviews, which are known to be asymmetric in terms of power dynamic, including the allocation of speaker versus listener role (Momeni, 2011). In this context, the officer assumes a position of power by directing the conversation, including the decision on when the interaction begins and ends (Momeni, 2011). Figure 8 shows that in the current sample of ransomware negotiations the victim was more likely to start however, it could be argued that the initiation of the attack or the posting of the ransom note (i.e., threat actor behaviours) mark the actual start of the negotiation. It was further found that in all but one ransomware case the threat actor ended the negotiation. Comparing this to Momeni's (2011) evaluation of police interviews this too would indicate that the threat actors hold more power as they are the ones directing the negotiation. The frequent use of *direct pressure* might further underline this.

It must be noted though that, as suggested in the introduction, the victim is not completely powerless. As mentioned above they hold "the power of financial gains" (Faivre, 2023). In the current sample this could be occasionally observed when victims used *intimidation*, threatening to leave the negotiation. This connects to what was noted in the discussion ransomware attacks are frequently compared to business transactions (Boticiu & Teichmann, 2023; Cartwright et al., 2019) Similarly, research noted that both threat actors

and victims having the opportunity to influence the outcome of the negotiation is indicative of some possession of power on both sides (Faivre, 2023).

Hence, another explanation for the difference in utterance distribution between threat actor and victim could lie within the communication medium. Ransomware negotiations differ from many other crisis contexts in that they are conducted through text-based communication. Generally, when using text-based communication channels it is common to send several messages without waiting for a response (Alis & Lim, 2013; Gallucci, 2021). Gallucci (2021) argues that this behaviour does not necessarily imply disinterest in receiving a response. Instead, it reflects contemporary texting norms, where extensive information is often segmented across multiple messages. This could be seen in several logs within this sample especially when the threat actor was sending instructions on how to transfer money. Hence, the threat actor contributing more might simply be attributed to the presence of such texting norms.

There are Patterns of Influencing Behaviours that Occur Repeatedly

The current study showed that there are some influencing behaviours which frequently occur in the same order. Thus it seem that what Beune and colleagues (2011) discussed in regard to police interviews, that influencing behaviours can form strategic sequence, also applies for ransomware negotiations. The most frequent patterns identified here mainly included the threat actor *being credible* (which is already addressed in the first key finding above) and instances were the victim employed *exchanging* which the threat actor reacted to by reciprocating the behaviour or by *being credible*. More research is needed to further verify and understand potential strategic sequences in ransomware negotiations. This is outline in more details below.

Limitations

Below the main limitation of this current research are listed. I start by discussing limitations regarding the sample and continue with limitations on the data analysis method employed in this study.

First, it must be noted that the sampling process based on the information provided on the ransomware.live website was flawed. As described, we sampled the logs while adhering to predefined criteria, using the overview of key characteristics on the website. Later, I realized that these characteristics were partially untrue. For example, the log with the ID 20210413 was described to have 63 utterances on the website (Mousqueton, 2024). However, in reality there were 64 utterances. Furthermore, a similar discrepancy was noticed regarding another log. This goes to show that the information that the sampling processes was based on was flawed.

Additionally, the varying log length (35-80 utterances) posed several significant limitations when directly comparing logs. This holds especially true for the analysis of the most frequent code per utterance level (mode) because the context surrounding a specific utterance level can vary significantly. Additionally, important patterns might be diluted when taking a measure of central tendency across logs of varying lengths.

Furthermore, the formatting of the log files might have an effect on how utterances are interpreted. Since the provided logs are only transcripts of ransomware negotiations and do not include snapshots, we cannot be sure the utterances were actually separate messages. This becomes especially noticeable when comparing different logs, like for example D4 and D2. At one point in the latter every single sentence is expressed as part of a new utterance, while in the former similar chunks of information were seemingly send within one message.

Additionally, the intercoder reliability score of $\kappa = 0.66$ suggests that there was an inconsistency in how me and my research partner applied the coding scheme and additional

rules. This indicates that the practice phase was insufficient. However, it must be noted that the AI-generated practice data was largely different from the actual data. It was lacking in depth and nuances. This observation is undermined in that negotiations in the practice data were much shorter (on average 32 utterances) than data in the final sample (on average 55 utterances). Furthermore, the logs within the practice data showed much less variation in terms of the range of scenarios they presented. This became obvious in the pace of the negotiations as well as formulations used and the repetition of sequences of interactions. For an example see the figure below. Meanwhile, during the coding of the actual data, we encountered a diverse range of interactions. Hence, the training might have been relatively ineffective. For future research it is recommended to check whether the practice data closely resembles the actual sample. Although, this might come with an introduction of bias since the actual data would need to be thoroughly viewed before generating practice data.

Future Research

While small suggestions are already made above, I want to outline relevant avenues for future research.

First and foremost, to address the sample limitations stated above, it would be advised to try and collect ransomware log from a primary source. Either, by getting in contact with victims or threat actors. This could be challenging since many victims do not publicly speak out about their involvement in ransomware attacks and finding a threat actor (group) that is willing to share negotiation data is even more unlikely. A different option might be to use police reports on ransomware cases (assuming that they include a full transcript of the negotiation).

Furthermore, primary data would allow for more contextual information to be included in the analysis. Ryan and colleagues (2022) theorized that influencing behaviour in ransomware negotiations is largely dependent on cultural factors. Similarly, research on crisis

negotiation also considered this and found that cultural differences impact crisis negotiations.(see for example Beune et al., 2010; Giebels et al., 2017; Giebels & Taylor, 2009). Hence, investigating this empirically in the context of ransomware might be valuable.

Generally, any empirically supported insight of ransomware negotiations would contribute in addressing the gaps that I outlined in this current paper. To do so novel methods such as pattern mining can and should be utilized as I shown in this current research. However, cSPADE might not be the most appropriate. Sequential pattern mining requires temporal data since it determines which events are likely to occur in successive order (e.g., A is likely to happen before B). However, as in the current sample some ransomware logs might not include timestamps. (As outlined in the method section, I used made up time sequences for cSPADE to work which did not impact the analysis in any way). However, association algorithms (e.g., Apriori) which are a different category of data mining tools might be a better alternative. Different from cSPADE they do not use temporal data and the results merely imply that events occur closely together (e.g., A and B are associated). This could be an alternative analysis for ransomware logs, especially those that do not include timestamps. Similarly, another possible method to explore relations between patterns is the computation of the proximity coefficients as performed by Beune and colleagues (2010).

Whichever, method might be employed continuing to identifying reoccurring patterns of influencing behaviour should remain a research focus since it will help to further understand the influencing dynamics in ransomware negotiations. Similarly, more investigations into the (a)symmetry of power are needed to thoroughly understand ransomware negotiations.

Connected to this I suggest that future research which applies the table of ten to ransomware negotiations refrains from adopting the approach introduced by Beune and colleagues (2010) to code silence. Beune and colleagues (2010) when investigating influence

strategies within police interviews introduced an additional code “Refusing to give information” whenever the threat actor (i.e., suspect in the context of police interviews) remained silent in response to a question or a request. As a result, their approach effectively balanced the ratio of coded behaviours between both parties within their study (Beune et al., 2010), eliminating the need to look for unequal utterance contribution.

They based this approach on the context of police interviews, particularly, considerations regarding the suspect’s right to remain silent (Beune et al., 2010). As noted police interviews are known to be asymmetric in terms of power dynamic, including the allocation of speaker versus listener role (Momeni, 2011). Hence, in the study of Beune and colleagues (2010), their approach, was justified within the specific context (Beune et al., 2010)). However, in the current context of ransomware negotiations it was deemed more appropriate to exclusively code explicit utterances, allowing for the identifications of imbalance in the utterance ratio. This is especially important since utterance ratios can give some indication of who is leading the negotiation.

Conclusion

The current study contributed insights into the dynamics of ransomware negotiations. Specifically, it identified the threat actor being credible as the most frequent influencing behaviour within ransomware negotiations, which empirically validates previous theoretical research. Furthermore, it can be noted that the victim primarily, employs a mix of two influencing behaviours *being credible* and *exchanging*. It was also noted that threat actors contribute more to ransomware negotiations. Additionally, by using cSPADE pattern mining I was able to identify frequently occurring patterns of influencing behaviours.

This research marks a first attempt at understanding influencing behaviours within ransomware negotiations. While, more empirical research is needed, the current study might inspire future explorations both in terms of research focus and method.

References

- Aoga, J. O. R., Guns, T., & Schaus, P. (2016). An Efficient Algorithm for Mining Frequent Sequence with Constraint Programming. In P. Frasconi, N. Landwehr, G. Manco, & J. Vreeken (Eds.), *Machine Learning and Knowledge Discovery in Databases* (pp. 315–330). Springer International Publishing. https://doi.org/10.1007/978-3-319-46227-1_20
- Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A Comprehensive Review of Cyber Security Vulnerabilities, Threats, Attacks, and Solutions. *Electronics*, *12*(6), Article 6. <https://doi.org/10.3390/electronics12061333>
- Beaman, C., Barkworth, A., Akande, T. D., Hakak, S., & Khan, M. K. (2021). Ransomware: Recent advances, analysis, challenges and future research directions. *Computers & Security*, *111*, 102490. <https://doi.org/10.1016/j.cose.2021.102490>
- Beune, K., Giebels, E., Adair, W. L., Fennis, B. M., & Zee, K. I. van der. (2011). Strategic sequences in police interviews and the importance of order and cultural fit. *Criminal Justice and Behavior*, *38*(9), 934–954. <https://doi.org/10.1177/0093854811412170>
- Beune, K., Giebels, E., & Taylor, P. J. (2010). Patterns of interaction in police interviews: The role of cultural dependency. *Criminal Justice and Behavior*, *37*(8), 904–925. <https://doi.org/10.1177/0093854810369623>
- Boticiu, S., & Teichmann, F. (2023). How does one negotiate with ransomware attackers? *International Cybersecurity Law Review*, *5*(1), 55–65. <https://doi.org/10.1365/s43439-023-00106-w>
- Buchta, C., & Hahsler, M. (2023). *arulesSequences: Mining Frequent Sequences* (0.2-30) [Computer software]. <https://cran.r-project.org/web/packages/arulesSequences/index.html>

- Cartwright, E., Hernandez Castro, J., & Cartwright, A. (2019). To pay or not: Game theoretic models of ransomware. *Journal of Cybersecurity*, 5(1), tyz009.
<https://doi.org/10.1093/cybsec/tyz009>
- Connolly, L. Y., Wall, D. S., Lang, M., & Oddson, B. (2020). An empirical study of ransomware attacks on organizations: An assessment of severity and salient factors affecting vulnerability. *Journal of Cybersecurity*, 6(1), tyaa023.
<https://doi.org/10.1093/cybsec/tyaa023>
- Crichton, D. A. (2021). Crisis Negotiation. In *Forensic Psychology* (pp. 350–370). John Wiley & Sons, Ltd. <https://doi.org/10.1002/97811394260669.ch15>
- European Union Agency for Cybersecurity, Lella, I., Tsekmezoglou, E., Naydenov, R. S., Ciobanu, C., Malatras, A., & Theocharidou, M. (2022). *ENISA Threat Landscape 2022* (Annual Report 9). ENISA. <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022>
- Faivre, J. (2023). *Negotiations in Tech: An Analysis of Asymmetric Ransomware Negotiations* (SSRN Scholarly Paper 4530094). <https://doi.org/10.2139/ssrn.4530094>
- Fernando, A. G., & Aw, E. C.-X. (2023). What do consumers want? A methodological framework to identify determinant product attributes from consumers' online questions. *Journal of Retailing and Consumer Services*, 73, 103335.
<https://doi.org/10.1016/j.jretconser.2023.103335>
- Gass, R. H., & Seiter, J. S. (2022). Chapter 1: Why Study Persuasion? In *Persuasion: Social Influence and Compliance Gaining* (7th ed., pp. 1–30). Routledge.
<https://doi.org/10.4324/9781003081388>
- Giebels, E. (2023). *Coding guidelines Tabe of Ten*.
- Giebels, E., & Noelanders, S. (2004). *Research Report Crisis Negotiation presented in "landcommanderij Alden Biesen"*. Universal Press.

https://ris.utwente.nl/ws/portalfiles/portal/321906709/Giebels_Noelanders_crisis_negotiation_English.pdf

Giebels, E., Noelanders, S., & Vervaeke, G. (2005). The hostage experience: Implications for negotiation strategies. *Clinical psychology & psychotherapy*, *12*(3).

<https://doi.org/10.1002/cpp.453>

Giebels, E., Oostinga, M. S. D., Taylor, P. J., & Curtis, J. L. (2017). The cultural dimension of uncertainty avoidance impacts police-civilian interaction. *Law and Human Behavior*, *41*(1), 93–102. <https://doi.org/10.1037/lhb0000227>

Giebels, E., & Taylor, P. J. (2009). Interaction patterns in crisis negotiations: Persuasive arguments and cultural differences. *Journal of Applied Psychology*, *94*(1), 5–19.

<https://doi.org/10.1037/a0012953>

Gonçalves, D., Coelho, P., Martinez, L. F., & Monteiro, P. (2021). Nudging Consumers toward Healthier Food Choices: A Field Study on the Effect of Social Norms.

Sustainability, *13*(4), Article 4. <https://doi.org/10.3390/su13041660>

Gray, I. W., Cable, J., Brown, B., Cuijuclu, V., & McCoy, D. (2022). Money over Morals: A Business Analysis of Conti Ransomware. *Proceedings of the 2022 APWG Symposium on Electronic Crime Research, eCrime 2022*.

<https://doi.org/10.1109/eCrime57793.2022.10142119>

Grubb, A. R. (2023). Effective Police Negotiation: Synthesising the Strategies and Techniques that Promote Success Within Hostage or Crisis Situations. In M. S. Staller, S. Koerner, & B. Zaiser (Eds.), *Police Conflict Management, Volume I: Challenges and Opportunities in the 21st Century* (pp. 285–314). Springer International

Publishing. https://doi.org/10.1007/978-3-031-41096-3_12

Grubb, A. R., Brown, S. J., Hall, P., & Bowen, E. (2019). From “Sad People on Bridges” to “Kidnap and Extortion”: Understanding the Nature and Situational Characteristics of

- Hostage and Crisis Negotiator Deployments. *Negotiation and Conflict Management Research*, 12(1), 41–65. <https://doi.org/10.1111/ncmr.12126>
- Guthrie, C. (2004). Influence: Principles of Influence in Negotiation. *Marquette Law Review*, 87(4). <https://scholarship.law.marquette.edu/mulr/vol87/iss4/20>
- Hull, G., John, H., & Arief, B. (2019). Ransomware deployment methods and analysis: Views from a predictive model and human responses. *Crime Science*, 8(1), 2. <https://doi.org/10.1186/s40163-019-0097-9>
- JSON to Excel Converter*. (n.d.). Conversion Tools. Retrieved 28 June 2024, from <https://conversiontools.io/convert/json-to-excel>
- Kelman, H. C. (1958). Compliance, identification, and internalization three processes of attitude change. *Journal of Conflict Resolution*, 2(1), 51–60. <https://doi.org/10.1177/002200275800200106>
- Korobkin, R. (2024). *Negotiation Theory and Strategy*. Aspen Publishing.
- Lesh, N., Zaki, M. J., & Oglhara, M. (2000). Scalable feature mining for sequential data. *IEEE Intelligent Systems*, 15(2), 48–56. <https://doi.org/10.1109/5254.850827>
- Li, X., Liu, M., Lian, J., & Zhu, Q. (2024). Nudging Away Health Misinformation on Social Media: The Roles of Social Influences and Power Distance. In I. Sserwanga, H. Joho, J. Ma, P. Hansen, D. Wu, M. Koizumi, & A. J. Gilliland (Eds.), *Wisdom, Well-Being, Win-Win* (pp. 268–279). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-57860-1_19
- Liu, H.-W., Wu, J.-Z., & Wang, Y.-H. (2023). Uncovering Insights for New Car Recommendations with Sequence Pattern Mining on Mobile Applications. *Applied Sciences*, 13(11), Article 11. <https://doi.org/10.3390/app13116386>
- Luther, K., Keeping, Z., Snook, B., de Almeida, H., Fahmy, W., Smith, A., & Han, T. (2023). Nudging eyewitnesses: The effect of social influence on recalling witnessed events.

Journal of Criminal Psychology, 14(1), 55–77. <https://doi.org/10.1108/JCP-06-2023-0042>

Manjezi, Z., & Botha, R. A. (2019). Preventing and Mitigating Ransomware. In H. Venter, M. Looek, M. Coetzee, M. Eloff, & J. Eloff (Eds.), *Information Security* (pp. 149–162). Springer International Publishing. https://doi.org/10.1007/978-3-030-11407-7_11

Mijwil, M., Unogwu, O., Filali, Y., Bala, I., & Al-Shahwani, H. (2023). *Exploring the Top Five Evolving Threats in Cybersecurity: An In-Depth Overview*. <https://doi.org/10.58496/MJCS/2023/010>

Momeni, N. (2011). Police Genre: Interruption and its Classification as a Sign of Asymmetry in Police Interview/Interrogation. *International Journal of Criminology and Sociological Theory*, 4(1), Article 1. <https://ijcst.journals.yorku.ca/index.php/ijcst/article/view/32123>

Mousqueton, J. (2024). *Ransomware.live*. Ransomware.Live. <https://www.ransomware.live>

Press, S., Welsh, N., & Schneider, A. K. (2023). Negotiation Theories Engage Hybrid Warfare. *Faculty Scholarship*. <https://open.mitchellhamline.edu/facsch/569>

Rieß-Marchive, V. (2024). *Casualtek/Ransomchats* [Python]. <https://github.com/Casualtek/Ransomchats> (Original work published 2023)

Ryan, P., Fokker, J., Healy, S., & Amann, A. (2022). Dynamics of targeted ransomware negotiation. *IEEE Access*, 10, 32836–32844. <https://doi.org/10.1109/ACCESS.2022.3160748>

Srinivasan, C. (2017). Hobby hackers to billion-dollar industry: The evolution of ransomware. *Computer Fraud & Security*, 2017(11), 7–9. [https://doi.org/10.1016/S1361-3723\(17\)30081-7](https://doi.org/10.1016/S1361-3723(17)30081-7)

- Taylor, J. P., & Patel, A. D. (2017). A Comprehensive Survey: Ransomware Attacks Prevention, Monitoring and Damage Control. *International Journal of Research and Scientific Innovation*, 4.
- United Nations Office on Drugs and Crime. (2021, October). *Ransomware attacks, a growing threat that needs to be countered*. UNODC Regional Office for Southeast Asia and the Pacific. <https://www.unodc.org/roseap/en/2021/10/cybercrime-ransomware-attacks/story.html>
- Vasiu, I., & Vasiu, L. (2020). Forms and Consequences of the Cyber Threats and Extortion Phenomenon. *European Journal of Sustainable Development*, 9(4), Article 4. <https://doi.org/10.14207/ejsd.2020.v9n4p295>
- Wade, M. (2021). Digital hostages: Leveraging ransomware attacks in cyberspace. *Business Horizons*, 64(6), 787–797. <https://doi.org/10.1016/j.bushor.2021.07.014>
- Westerveld, F. (2024). *Navigating the Shadows: Analyzing Negotiation Strategies in Ransomware Incidents and their Impact on Outcome* [Info:eu-repo/semantics/masterThesis, University of Twente]. <https://essay.utwente.nl/98692/>
- Zaki, M. J. (2000). Sequence Mining in Categorical Domains: Incorporating Constraints. *A B*.
- Zaki, M. J. (2001). SPADE: An Efficient Algorithm for Mining Frequent Sequences. *Machine Learning*, 42(1), 31–60. <https://doi.org/10.1023/A:1007652502315>
- Zaki, M. J. (2024). *Zakimjz/cSPADE* [C++]. <https://github.com/zakimjz/cSPADE> (Original work published 2018)
- Zhang, Y., & Paquette, L. (2023). Sequential Pattern Mining in Educational Data: The Application Context, Potential, Strengths, and Limitations. In A. Peña-Ayala (Ed.), *Educational Data Science: Essentials, Approaches, and Tendencies: Proactive Education based on Empirical Big Data Evidence* (pp. 219–254). Springer Nature. https://doi.org/10.1007/978-981-99-0026-8_6

Appendix A

R Markdown for cSPADE Pattern Mining Analysis

Lilli Schnirch

2024-06-26

loading necessary packages

```
library(tidyverse) # data manipulation
library(arulesSequences) # run the sequence mining algorithm

df <- read.csv(file = 'cSPADE_input.csv')
options(scipen=999)
```

data cleaning

```
df1 <- df %>%
  group_by(customer.identifier) %>%
  arrange(purchase.date) %>%
  #Next Line: removes Instances where the same product appears repeatedly
  #distinct(customer.identifier, product, .keep_all = TRUE) %>%
  #Create Item ID Within Customer ID
  mutate(item_id = row_number()) %>%
  select(customer.identifier, purchase.date, item_id, product) %>%
  ungroup() %>%
  #Convert Everything to Factor
  mutate(across(.cols = c("customer.identifier", "product"), .f = as.factor))

df1 <- df1[order(df1$customer.identifier),] # descending order
```

c-spade pre-process

```
sessions <- as(df1 %>% transmute(items = product), "transactions")
transactionInfo(sessions)$sequenceID <- df1$customer.identifier
transactionInfo(sessions)$eventID <- df1$item_id
itemLabels(sessions) <- str_replace_all(itemLabels(sessions), "items=", "" )
inspect(head(sessions,10))
```

cSPADE

```
itemsets <- cspade(sessions,
                  parameter = list(support = 0.2, maxgap = 1),
                  control = list(verbose = FALSE))
inspect((itemsets))
df2 <- itemsets
```

output all results

```
df2 <- as(df2, "data.frame") %>% as_tibble()
df2$pattern <- (str_count(df2$sequence, ",") + 1)
```

```
df2 <- df2[order(-df2$support),] # descending
write.csv(x=df2, file="all_results.csv", row.names=FALSE)
```

output top results

```
c <- df2 %>% group_by(pattern) %>% slice_max(order_by = support, n = 20)
write.csv(x=c, file="top_results.csv", row.names=FALSE)
```

Note. The original R script (Wan, 2022/2022) there is a line of code in the data cleaning section, which removes instances where the same influencing behaviour appears repeatedly (e.g., threat actor being credible followed by threat actor being credible). As reported, the descriptive analysis and face value observations revealed such instances of repeated behaviours in the current dataset.

Initially, I ran the R script including the function to remove repeated values, `distinct(customer.identifier, product, .keep_all = TRUE)`. However, for the results reported here, I chose to exclude this line of code (indicated by the respective line starting with # in the markup above) and allowed repeated behaviours to be identified as patterns. I believe there is value in identifying these repeated behaviours. It should be further noted, though, that omitting this step in the data cleaning process increased the computational time slightly.

Appendix B

Additional cSPADE Output

Figure B.1

cSPADE Output: Most Frequent Three-Behaviour-Patterns













cSPADE Output: 3-Behaviour Patterns with Highest Support (Excluding Previously Listed Pattern with Support >50%)			Support
Threat Actor		Victim	Percentage of Logs with at least once occurrence
Being Credible		Exchanging	40%
Exchanging			
Being Credible		Being Credible	36%
Being Credible			
Exchanging		Exchanging	32%
Exchanging			
Being Credible		Exchanging	32%
Being Credible			
Exchanging		Exchanging	32%
		Exchanging	
Exchanging		Exchanging	32%
		Being Credible	

Figure B.2

cSPADE Output: Most Frequent Two-Behaviour-Patterns

cSPADE Output: 4-Behaviour Patterns with Highest Support		
Pattern		Support
Threat Actor	Victim	Percentage of Logs with at least once occurrence
Being Credible Being Credible Being Credible	Exchanging	24%
Being Credible Being Credible Being Credible	Exchanging	24%
Being Credible Being Credible Being Credible	Being Credible	20%
Being Kind Being Credible Being Credible Being Credible		20%
Being Credible Being Credible Being Credible Being Credible		20%