Forensic Linguistics and Cybersecurity Combined: Investigating E-Fraud through Language

دمج اللّسانيّات الجنائيّة بالأمن السّيبرانيّ: التّحقيق اللّغويّ في الاحتيال الالكترونيّ

Nidaa Mercel Al Saifi[1]

نداء مرسل الصّيفي

تاريخ الاستلام 3/ 10/ 2025 **تاريخ القبول 2025/11/30**

Abstract

Within the cyberworld, language is a crucial component through which the entire communicative process is executed. However, in the cyberspace, the function of language is not only communicative, but it also relates to privacy and security matters defining cybersecurity. Grounded in this critical role of language, this study approached the context of scam job offers as a significant manifestation of e-fraud, even regarded as a form of language crimes. Particularly, this study blended computationally-driven methods with forensic linguistic tools to uncover deception indicators distinguishing scams form genuine job offers. It employed text mining techniques and stylometric analysis to investigate the linguistic patterns and style differentiating scam job offers from authentic ones. Besides, this study adopted the **Aristotelian Triad** as a powerful theoretical framework to explore how scammers deploy rhetorical elements as a persuasive strategy shaping their victims' behavior. Through its mixed-methods design, the research revealed that informality, substandard professionalism, privacy abuse, and exaggerated promises are evident-based indicators of deception distinguishing scams from genuine offers. Also, the forensic stylometric analysis showed that scams often lack the homogeneity and consistency in sentence length as well as the accuracy in punctuation patterns. Finally,

[1] 1- A PhD student in English Linguistics at the Lebanese University- Doctoral School of Literature, Humanities & Social Sciences

the theoretical analysis grounded in the adopted **Aristotelian Triad** revealed that scammers heavily rely on emotional appeals and often refer to fabrication and impersonation to benefit from the power of ethos and logos in achieving persuasion and manipulation.

Keywords: E-fraud, cybersecurity, text mining, stylometry, Aristotelian Triad

الملخص

في إطار العالم السّيبرانيّ، تعدّ اللّغة مكوّنًا حيويًّا إذ من خلالها تتم العملية التّواصليّة. غير أنّ وظيفة اللّغة في الفضاء الرّقميّ لا تقتصر على كونها تواصليّة، بل إنّها تتداخل أيضًا في قضايا الأمن والحماية التي تشكّل جوهر الأمن السّيبرانيّ. انطلاقًا من هذا الدّور الفعال، يتطرّق هذا البحث الى سياق عروض العمل الاحتياليّة كتجلٍ بارزٍ للاحتيال الرّقميّ، وكما أنّه يصنّف كونه أحد أشكال الجرائم اللّغويّة. على وجه الخصوص، يدمج هذا البحث الطّرق القائمة على الحوسبة مع اللّغويّات الجنائيّة بهدف كشف مؤشّرات الاحتيال التي تميز عروض العمل الوهميّة عن الاصليّة منها. تستعين هذه الدّراسة بتقنيّات التّنقيب عن النّصوص والتّحليل الأسلوبيّ الرّقميّ لكشف تباين الأنماط والأساليب اللّغويّة بين العروض الوهميّة والأصليّة. إضافة الى ذلك، تعتمد هذه الدّراسة على الثّالوث الأرسطيّ كإطار نظريّ لاستكشاف كيفية استخدام محتالي الانترنت للعناصر البلاغيّة وسيلة إقناعيّة للتّحكم في سلوك ضحاياهم والإيقاع بهم. باعتمادها تصميم الطّرق المختلطة، أظهرت هذه الدّراسة أنّ عروض العمل الوهميّة تفتقر الى الرّسميّة، وتتميز باحترافيّة منخفضة المستوى إضافة الى انتهاك الخصوصيّة والوعود المضخمة كمؤشرات حيويّة تميّزها عن عروض العمل الأصلية. فضلًا عن ذلك، أظهر التّحليل الأسلوبيّ الجنائيّ أنّ العروض الوهميّة تفتقر إلى التّجانس والاتّساق في طول الجملة كما أنّها تفتقد الدّقة في علامات التّرقيم. أخيرًا، أكّد التّحليل النّظريّ المبني على الثّالوث الأرسطي أن مرتكبي الاحتيال الإلكترونيّ يعتمدون على الاستمالة العاطفيّة وكما يلجؤون الى الفبركة والانتحال لتوظيف الايثوس واللوغوس في سبيل تحقيق الإقناع والتّحايل.

الكلمات المفتاحيّة: الاحتيال الرّقميّ، الامن السّيبرانيّ، التّنقيب عن النّصوص، والتّحليل الأسلوبيّ الرّقميّ، الثّالوث الأرسطيّ.

Introduction

Amidst the escalating involvement of individuals in the digital world, the threats of privacy attack, safety breach, and e-fraud witnessed a marked increase. In light of this phenomenon, cybersecurity is inevitably compromised, hindering digital users' engagement in the cyberspace and disrupting their virtual experience. In this context, e-fraud practiced through scam job offers is a valid demonstration depicting the criticality of these fraudulent activities and highlighting the alarming need to closely consider the strategies deployed in undergoing deception and manipulation as a serious form of cybercrime. Blended with the scope of language and communication, Williams (2001) emphasizes the interplay between e-fraud and language stating that "most forms of abuse online manifest textually" (p. 164). This shows that cybercrime heavily relies on language as a means of communication through which deception is performed. This interconnection portrays the significance of linguistic components and their invisible yet effective role in the execution of e-fraud. Intersecting at a single point, cybersecurity, language, and e-fraud manifest as a fertile focus of Forensic Linguistics (FL) as a newly emerging field embarking on the avenues of justice, language, and law (Dusane, 2021).

Given the identified interface, the conducted research ventured into the realm of FL and cybersecurity through investigating e-fraud practiced in scam job offers. Its analytical approach heavily relied on digitalized analysis, mainly computational text mining and forensic corpus-based stylometric analysis. This provided insights into the deceptive strategies practiced through language in scam recruitment offers as a valid manifestation of linguistically committed cybercrimes. Besides, the comparative method was employed attempting to distinguish the style and content of scam job offers from authentic ones. In addition to the digital quantitative approach, the study also embarked on qualitative analysis with the lens of the Aristotelian Triad through which the use of rhetoric as a persuasive strategy underpinning deception in the cyberworld was portrayed.

Statement of the Problem

According to the Verizon Business' **2024 Data Breach Investigations Report**, the rate of cybercrimes is rising steadily mainly as the world is increasingly orienting toward digitalization. Statistically, the report shows an alarming increase in ransomware and phishing attacks during 2023, with the former accounting for 23% of the total analyzed breaches and the latter for 20% (Verizon Business, 2024). Kuzior et al. (2024) claim that since cybercrimes result in critical consequences hindering users' financial and personal data, "protecting this data from breaches is crucial for maintaining privacy and trust" (p. 222). In light of the stated problem, it is crystallized that privacy and security are pivotal matters challenged by the expanding threat imposed by cyber fraud. In this context, Kolupuri et al. (2025) argue that scams are a common type of digital fraud schemes that have severe effects on individuals and businesses. These fraudulent activities are mainly perpetrated through fake news and unsolicited mails such as those impersonating a company representative to suggest a job offer. Thus, the spread of scam job offers is a pressing matter challenging cybersecurity.

Considering the seriousness of this issue, this study approaches the context of scam job offers through an innovative forensic linguistic approach. Being regarded as a form of cybercrime, scam job offers deploy deceptive strategies in victimizing users in the digital world. Since language lies at the center of the communicative process in the cyber world, investigating linguistic deception indicators assists in distinguishing scam from authentic recruitment offers. Despite the significance of this matter, the scope of cybersecurity and e-fraud, particularly in the context of scam job offers, has been scarcely explored. Therefore, this study mingles computational textual analysis with stylometric investigation and rhetorical interpretation to bridge the gap explored in the realm of cybersecurity.

Purpose of the Study

The ultimate purpose of this study lies in addressing the stated research problem through an interdisciplinary perspective. The conducted research primarily employs a forensic corpus-based approach blended with computational linguistic methods and the Aristotelian framework to investigate cyber fraud through language. Particularly, it intends to foster cybersecurity through uncovering deception indicators in the context of scam job offers communicated via email systems. Moreover, this study subjects the investigated corpus to forensic stylometric analysis. This contributes to revealing stylometric features such as sentence length and punctuation patterns distinguishing scam from authentic job offers. Besides, to deepen the linguistic dimension of the conducted analysis, the researcher approached the selected scam offers through the lens of the **Aristotelian Triad**. This reflects how fraudsters harness rhetorical appeals (ethos, pathos, logos) in scam job offers to persuade their victims and manipulate their behavior. Therefore, this study mingles computational linguistic analysis with a forensic corpus-based approach and an Aristotelian rhetorical framework to investigate e-fraud as one common form of cybercrimes.

Significance of the Study

The relevance of this study lies in its potential to investigate cyber fraud through an interdisciplinary linguistic approach, fostering the functionality of language in diverse domains. It embarks on a scarcely explored subject manifested in the synchronization of Forensic Linguistics and Cybersecurity as two booming disciplines in the digital age. In line with its novel approach, this study touches on a critical phenomenon hindering users' privacy and digital security. While research on the cybersecurity mainly centers on software-based algorithms, log analysis, and encryption, this study delves beyond conventional boundaries, emphasizing the power of language in supporting cybersecurity and detecting cybercrimes. This is portrayed in the intersection of computational linguistic methods and forensic corpus-based approaches supported by rhetorical analysis of the invisible persuasive strategies characterizing the context of scam job mails. Besides, combining stylometric analysis and computational methods with a rhetorical approach provides a blended sense of quantitative and qualitative interpretation of the explored context. Accordingly, this study positions language not only as a communicative tool, but also as a means through which forensic evidence can be explored. Thus, it reflects the functionality of linguistic analysis in uncovering criminal practices and assisting the judicial system. This contributes to promoting cybersecurity and protecting users from cyber attacks that threaten their privacy in the digital world.

Research Questions

Considering the illustrated problem, the researcher posed the following research questions:

1. How do computational text mining tools (concordances, n-grams, collocates) unveil deception indicators embedded in scam job offers compared to genuine ones?
2. How does forensic corpus-based stylometric analysis of sentence length and punctuation frequencies distinguish authentic recruitment offers from scams?
3. To what extent do scam job offers deploy the Aristotelian Triad to deceive their victims and manipulate their behavior?

Literature Review

Forensic Linguistics: The Intersection of Language and Law

Forensic Linguistics (FL) is a subfield of applied linguistics and an emerging branch of the overarching discipline of forensic science (Ahmed, 2021; Dusane, 2021). It is characterized by its heterogenous nature that is structured on an interdisciplinary basis, primarily aiming to associate language with legal concerns as a means of assisting the judicial system (Momeni, 2012). In other words, Fromkin et al. (2014) argue that Forensic Linguistics operates at the level of employing language as a tool for facilitating legal cases through its integration in the realm of investigative purposes. This integration is displayed in diverse practical applications including plagiarism detection, authorship attribution, fraud identification, lip-reading, document examination, and courtroom interaction (Ariani et al., 2014; Sakakini, 2020).

In describing this field, linguists and pioneers agreed on multiple definitions of FL. For Shuy (2006), FL is the study of language which contributes to facilitating legal cases and enhancing the demonstration of justice. This definition is further asserted by Olsson (2012) who states that "forensic linguistics began life as an instrument to correct miscarriages of justice" (p. 5). Furthermore, FL is also considered a subfield of forensics that delves into investigating the invisible factors that effectively yet implicitly contribute to addressing forensic issues. Particularly, it functions at the level of detecting and analyzing authentic linguistic evidence, employing diverse analytical tools (MacLeod & Wright, 2020; Syam, 2018). These tools heavily rely on digital methods, mirroring the interplay between this field and digital textual analysis that is booming in the technological era (MacLeod & Wright, 2020). In illustrating the operational procedure of FL, MacLeod and Wright (2020) maintain that in scrutinizing linguistic evidence, FL mingles three major elements, namely: "(i) the (written) language of the law, (ii) the language of (spoken) legal processes, and (iii) language analysis as evidence or as an investigative tool" (p. 360). Through these components, FL uncovers the invisible realities through decoding beyond-line clues (Dusane, 2021).

In tracing its foundation, FL, like many other disciplines has roots in Greek philosophies mainly as Greek authors explored the authenticity of texts and highlighted plagiarism as a form of fraud (Ramezani et al., 2016). Besides, McMenamin (2002) claims that the field underwent different stages of progression, particularly during the 19th century, manifested in the evolution of multiple methods that fostered authorship attribution. Nevertheless, it was not until 1963 that FL appeared as a separate discipline in The United States of America. Its evolution was brought forth by the case of Ernesto Miranda who was arrested by the police. Miranda's case inspired the development of a formal legal text titled **Miranda Warnings** which fostered the need "to concentrate on interviewing of witness rather than on cop's statements" (Dusane, 2021, p. 5174). Due to the explored significance of interviewing the witness, Miranda's case foregrounded forensic linguists' development of **Miranda Rights** which mirrored how the meaning of a text is shaped by how the content is communicated, interpreted, and understood. This revealed that any misinterpretation of the criminal's message can result in systemic injustice (Dusane, 2021). In 1968, FL was first coined as a distinct science by the Swedish linguist Jan Svartvik in his book titled **The Evans Statements: A Case for Forensic Linguistics** (Ali, 2020; Ariani et al., 2014).

Cybersecurity: A Comprehensive Overview

One of the earliest definitions of cybersecurity dates back to the 1980s when Steve Lipner developed the **CIA triad** accounting for defining cybersecurity in terms of: Confidentiality, Integrity, and Availability. However, the progression of technology and networking systems has weakened this triad and its functionality, driving the evolution of updated approaches that offer a more comprehensive definition of this notion (Ham, 2021). Among these approaches, the NIST cybersecurity framework is perceived as a generic approach that complements the CIA triad, aiming at enhancing the infrastructure of cybersecurity (National Institute of Standards and Technology, 2018). In this regard, the NIST framework is a collection of the following activities:

(1) Identify: identify the assets that must be secured and the context that they are in

(2) Protect: define and implement protective measures for the assets

(3) Detect: put sensors and processes in place to detect when protection has been breached

(4) Respond: define response processes for when an incident has been detected

(5) Recover: develop plans for resilience in the organisation, as well as recovery mechanisms. (Ham, 2021, pp. 2-3)

Based on the mentioned frameworks, cybersecurity is a process through which digital devices, accounts, and networks are protected against electronic threats such as theft and cyber fraud (Mijwil et al., 2023). It operates at the level of implementing practical tactics that help protecting the digital environment from electronic threats that are becoming popular due to the expansion of the digital world (Kuzior et al., 2024). Therefore, cybersecurity aims at mitigating cybercrimes through monitoring and controlling the communicative process in cyberspace as a means of maintaining security properties of users' assets in the cyber environment (Galinec et al., 2017).

E-Fraud as a Form of Cybercrime

According to Kuzior et al. (2024), the growing orientation of humans toward the digital world increased the risk of encountering cybercrimes. This phenomenon is highlighted and reported in various reports and by diverse statisticians who "demonstrate cybercrime's increasing frequency and sophistication, including hacking, phishing, ransomware, and other malicious activities" (Kuzior et al., 2024, p. 222). In this regard, Mijwil et al. (2023) state that cyberattacks serve as a form of digital fraud, resulting in serious losses for individuals in the cyber environment. In this context, Dzomira (2014) identifies the multiple types of cyber fraud among which identity theft is a critical category. In addition, another frequent fraud scheme is manifested in scam-based and spam-related operations. This often occurs when fraudsters assume the false identity of a company or a financial institution as a means of deceiving the receivers and obtaining their personal assets (Dzomira, 2014; Geeta, 2011). Kolupuri et al. (2025) shed light on the subtypes of scam attacks, outlining five categories that fall under this umbrella:

1. Phishing Attacks
2. Click Fraud
3. Content based SMS scam
4. Fraud in advertising
5. Online Social Network Scam (p.2)

Kolupuri et al. (2025) consider these practices as manifestations of cybercrimes since they pose serious threat to individuals and organizations. Phishing attacks, for example, employ different deceptive strategies to manipulate the victims' behavior and drive them to share sensitive information. Similarly, click fraud results in critical financial damages particularly for online businesses as it “requires the generation of counterfeit clicks on online advertisements to fraudulently inflate ad revenue or deplete a competitor's ad budget” (Kolupuri et al., 2025, p. 3). Likewise, other types of scams extensively rely on the power of anonymity to deceive the users through mimicking legitimate figures' styles and content. Thus, this reflects the seriousness of the cybercriminal practices as they tend to be real-life crimes, mandating immediate consideration and legal adjustments (Sousa-Silva, 2024).

Methods

Research Design

According to Kumar (2011), a research design is a strategic plan outlining the research investigation and process followed by the researcher to reliably respond to the posed research questions and accurately address the stated problem. Within this scope, the present study adopted a mixed-methods approach, combing qualitative and quantitative analysis in exploring the studied context. Obeyd (2021) distinguishes quantitative and qualitative research claiming that:

> Quantitative research is characterized with the use of numbers and they are enlivened to meaning when they are contextually backed up … it is systematic, controlled, involving exact measurements resulting in reliable and generalized results … [However,] qualitative research is effective in exploring new areas by studying and explaining in details a phenomenon. (pp. 56, 58)

Therefore, the mixed-methods approach is considered a third approach combining techniques of both qualitative and quantitative research in a way that strengthens the study and contributes to eliminating the weaknesses of a single design (Creswell, 2014; Obeyd, 2021). In the context of this study, the quantitative approach was manifested in deploying computational techniques through text mining tools that allowed for uncovering deception indicators embedded in scam job offers. Besides, the corpus-based stylometric analysis also brought forth significant numerical data differentiating scam offers from genuine ones. At the level of qualitative data, this was primarily obtained through the adoption of the Aristotelian Triad to investigate scammers adherence to ethos, pathos, and logos to manipulate their victims' behavior. Moreover, both qualitative and quantitative analysis were displayed in terms of exploring collocates and analyzing keywords in context.

Corpus of the Study

As this study is corps-based, the corpora were selected, compiled, and cleaned before being set into analysis. Particularly, samples of genuine and scam job offers were compiled and balanced according to the criteria of genre and size. In terms of genre, all selected samples were samples of recruitment offers, representatively involving both authentic and scam offers. As for the corpus size, each of the selected corpora adhered to a range of 2650-2675 tokens/corpus. Thus, the number of tokens was 2673 and 2652 for genuine and scam job offers respectively, resulting in a total of 5325 tokens. In terms of the corpus selection and compilation, the scam job offers were retrieved from the **Nigerian scams** website available at http://www.419scam.org./. The researcher particularly visited the **Company representative scam** section and randomly selected scam job offers presented in this section. As for the authentic offers, the researcher contacted the Human Resources (HR) directors at different reputable institutions to obtain some samples of the recruitment offers they usually use. This provided a valid and reliable access to both genuine and scam job offers.

To ensure the confidentiality of participants mentioned in the compiled corpus, the researcher followed a strategic anonymization criterion through which personal information was replaced by coded identifiers such as [RECIPIENT_NAME], [EMAIL_ADDRESS], [SENDER_NAME], [COMPANY_NAME], etc. These anonymization tokens enhanced the objectivity of the study and fostered the researcher's ethical approach to sensitive data. Besides, consistency was maintained across all datasets to avoid bias and to maintain a reliable preservation of any structural or linguistic features. In addition to anonymization, corpus cleaning was performed prior to the analytical process at some levels of the study. While answering RQ#1 and RQ# 3 required using full email samples (including greeting, content, closing, and other components), responding to RQ# 2 rather necessitated the preservation of the email body and the elimination of other components. Cleaning the corpus at the level of RQ#2 was due to the researcher's focus on sentence length and punctuation patterns which might be misinterpreted in the presence of line breaks separating the email content from other components. In other words, the algorithm of the used software might treat a line break as new sentence, although it does not correspond to a true linguistic sentence, resulting in misinterpretations in sentence length and number. Thus, only the email body was consistently analyzed in responding to RQ#2, while the full email was considered in the context of RQ#1 and RQ#3.

Tools/Instruments of the Study

1. AntConc (3.5.9)

In the realm of digital analysis and corpus linguistics, **AntConc** emerges as a freeware corpus analysis toolkit encompassing diverse analytical tools (Anthony, 2004). To carry out text mining procedures, the researcher employed **AntConc (3.5.9)** as a corpus analysis software through which different textual features are investigated. Its computerized algorithm provides accurate and authentic representation of linguistic features, mainly in terms of numerical data such as concordances, n-grams, and frequency of collocation. Through its KWIC (Key Word In Context) feature, this software presents results in terms of "the hit number, KWIC line, and the file name" (Anthony, 2004, p. 9). This mirrors its functionality in assisting the computational analytical process in this study through which deception indicators in scam job offers are explored and, accordingly, scams are distinguished from genuine offers.

2. Signature Stylometric System (1.0)

Signature Stylometric System (1.0) is a freeware computational stylistics tool designed by Peter Millican to facilitate stylometric analysis as a central aspect of forensic linguistic investigation (Guillen-Nieto et al., 2008). It operates through displaying two-dimensional and three-dimensional graphs with numerical values measuring style markers including sentence length, word length, paragraph length, and punctuation patterns. In highlighting its functionality in the scope of forensic linguistic investigation, Guillen-Nieto et al. (2008, p. 14) argue that "the tool provides the forensic language researcher with a remarkable sample of disputed texts for forensic linguistic analysis". They also added that the analytical approach of this tool "provides full comparison of all the results obtained and graphic output" (Guillen-Nieto et al., 2008, p. 14). This mirrors the practicality of using this software tool to conduct forensic corpus-based stylometric analysis within which the content of scam job offers is juxtaposed with that of authentic ones in terms of sentence length and punctuation patterns as two main features central to this study.

Adopted Framework: Aristotle's Rhetorical Triangle

In retrospect, the concept of rhetoric was first introduced by Aristotle in the 4th century B.C. when it was coined as "Rhetorica" in its Greek origin which means the "art of persuasion" (Mshvenieradze, 2013). Being integrated into the discourse context, Aristotle's rhetoric operates at three major levels triangulating the argumentative context and constituting the Aristotelian Triad. These levels are: Logos, Ethos, and Pathos (Murthy & Ghosal, 2014). While these components complement each other is fostering persuasion and manipulation within a given scope, each element functions at a different level.

According to Nurrosyidah (2016), logos is a pivotal component of rhetoric within which persuasion and manipulation are established through reasoning, logical arguments, and factual data. It contributes to influencing the audience's behavior through strengthening the speaker's argument and fostering credibility. However, Aristotle's ethos operates at the level of trustworthiness, emphasizing the power of the speaker's identity and position in gaining the audience's trust, and accordingly, establishing persuasion and manipulation (Perloff, 2003). Finally, "Aristotle's 'Rhetoric' Pathos is the power with which the writer's (speaker's) message moves the audience to his or her desirable emotional action … It is important to know not only how the orator can express but how he or she can by help of discourse cause favorable emotions, like anger, insult, empathy, fear, confusion, etc." (Amossy, 2000, p. 178). All in all, combining these components under the umbrella of Aristotle's Triad means that the writer/speaker strategically varies the use of logic, emotions, and trustworthiness to demonstrate rhetorical influence and manifest the power of persuasion.

In light of the illustrated theoretical framework, this study adopted the Aristotelian Triad to investigate how scammers deploy its components to deceive their victims. Through this study, the researcher approached the studied corpus through a critical analytical lens attempting to highlight the major rhetoric device underlying deception and fraud in scam job offers. This provided a robust theoretical background portraying the linguistic and communicative strategies underlying e-fraud as a serious matter hindering cybersecurity.

Results and Discussion

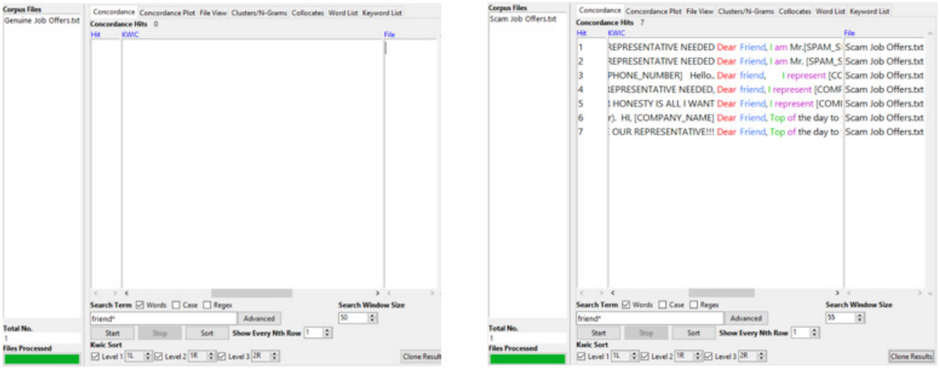
Identifying Deception Indicators in Scam Job Offers Compared to Genuine Ones

To uncover the diverse deception indicators characterizing e-fraud in scam job offers, the researcher conducted a comparative digital textual analysis in which scams and authentic offers were explored. The analytical process primarily involved examining word occurrences through investigating frequency hits and concordance lines. Besides, n-grams and collocates in each corpus were studied and juxtaposed to illustrate the lexical and semantic features distinguishing scam job offers from genuine ones.

In this context, the researcher randomly selected some keywords and investigated their frequency of occurrence in each corpus. One search term was 'friend*' as it provides insights into the degree of formality/informality in each of the studied job offers. As displayed in **Table 1**, the term 'friend*' yielded 0 hits in the corpus of genuine job offers and 7 hits in the context of scam job offers. This shows that the degree of formality is poor in scams as scammers lack the skills needed to professionally communicate with recipients regarding recruitment matters.

Table 1

Concordance hits of ‘friend*’ in genuine and scam job offers



Corpus Files: Genuine Job Offers.txt

Concordance Hits 0

Search Term ☑ Words ☐ Case ☐ Regex: friend*

Search Window Size: 50

Show Every Nth Row: 1

Total No. 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Corpus Files: Scam Job Offers.txt

Concordance Hits 7

Hit	KWIC	File
1	REPRESENTATIVE NEEDED Dear Friend, I am Mr.[SPAM_S	Scam Job Offers.txt
2	REPRESENTATIVE NEEDED Dear Friend, I am Mr. [SPAM_S	Scam Job Offers.txt
3	PHONE_NUMBER] Hello.. Dear friend, I represent [CC	Scam Job Offers.txt
4	REPRESENTATIVE NEEDED, Dear friend, I represent [COMF	Scam Job Offers.txt
5	HONESTY IS ALL I WANT Dear Friend, I represent [COMI	Scam Job Offers.txt
6	r). HI, [COMPANY_NAME] Dear Friend, Top of the day to	Scam Job Offers.txt
7	OUR REPRESENTATIVE!!! Dear Friend, Top of the day to	Scam Job Offers.txt

Search Term ☑ Words ☐ Case ☐ Regex: friend*

Search Window Size: 55

Show Every Nth Row: 1

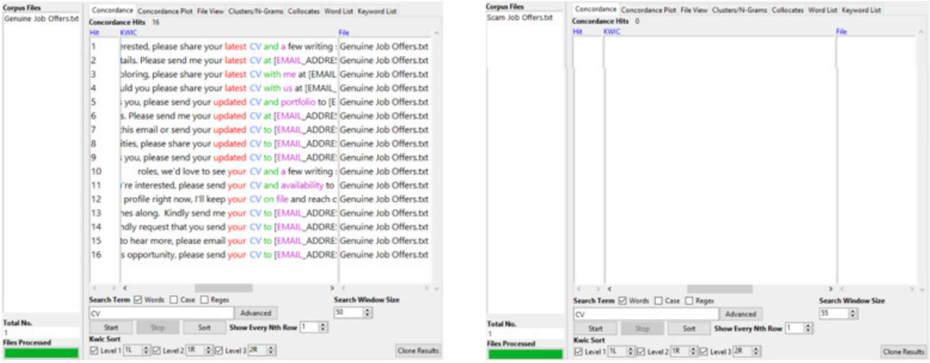
Total No. 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Similarly, in the case of the search term ‘**CV**’, the obtained results asserted the lack of professionalism in scam job offers. As shown in **Table 2**, the search term ‘**CV**’ returned 0 concordance hits in the corpus of scam job offers, while 16 hits were obtained at the level of authentic offers. The fact that real recruitment offers require the receivers to share their Curriculum Vitae (CV) to assess their eligibility based on their career summary serves as a pivotal indicator of the flaws of scam offers. Therefore, the absence of the term ‘**CV**’ together with the presence of informal language such as ‘**friend***’ function as a lucid indicator of the inauthenticity and deception in scam job offers.

Table 2

Concordance hits of 'CV' in genuine and scam job offers



Corpus Files: Genuine Job Offers.txt

Concordance Hits 16

Hit	KWIC	File
1	erested, please share your latest CV and a few writing :	Genuine Job Offers.txt
2	tails. Please send me your latest CV at [EMAIL_ADDRE!	Genuine Job Offers.txt
3	ploring, please share your latest CV with me at [EMAIL	Genuine Job Offers.txt
4	uld you please share your latest CV with us at [EMAIL_	Genuine Job Offers.txt
5	; you, please send your updated CV and portfolio to [E	Genuine Job Offers.txt
6	s. Please send me your updated CV at [EMAIL_ADDRE!	Genuine Job Offers.txt
7	this email or send your updated CV to [EMAIL_ADDRE:	Genuine Job Offers.txt
8	ities, please share your updated CV to [EMAIL_ADDRE:	Genuine Job Offers.txt
9	; you, please send your updated CV to [EMAIL_ADDRE:	Genuine Job Offers.txt
10	roles, we'd love to see your CV and a few writing :	Genuine Job Offers.txt
11	're interested, please send your CV and availability to	Genuine Job Offers.txt
12	profile right now, I'll keep your CV on file and reach c	Genuine Job Offers.txt
13	nes along. Kindly send me your CV to [EMAIL_ADDRE:	Genuine Job Offers.txt
14	ndly request that you send your CV to [EMAIL_ADDRE:	Genuine Job Offers.txt
15	to hear more, please email your CV to [EMAIL_ADDRE:	Genuine Job Offers.txt
16	s opportunity, please send your CV to [EMAIL_ADDRE:	Genuine Job Offers.txt

Search Term ☑ Words □ Case □ Regex: CV

Search Window Size: 50

Show Every Nth Row: 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No.: 1

Corpus Files: Scam Job Offers.txt

Concordance Hits 0

Hit	KWIC	File

Search Term ☑ Words □ Case □ Regex: CV

Search Window Size: 55

Show Every Nth Row: 1

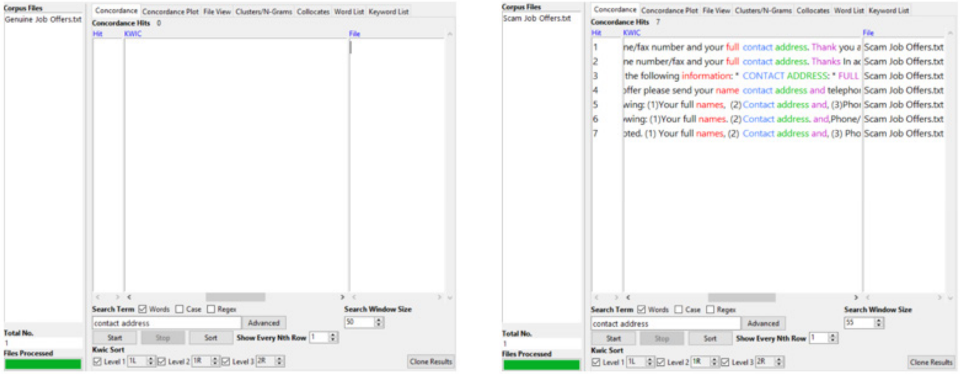
Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No.: 1

Moreover, in terms of respecting the receiver's privacy, authentic job offers are rarely expected to require the receiver to share personal information. Conversely, scammers tend to hunt on sensitive data, therefore, their offers are expected to drive the receiver to fall into their trap by sharing personal data. This is further asserted through investigating the search term '**contact address**' in both corpora. **Table 3** shows that 0 hits were obtained in the context of genuine offers while 7 hits appeared in the corpus of scam recruitment offers. Therefore, inquiring about personal information is another indicator of deception characterizing scam job offers as a critical form of e-fraud.

Table 3

Concordance hits of ‘contact address’ in genuine and scam job offers



Corpus Files: Genuine Job Offers.txt

Concordance Hits 0

Hit	KWIC	File

Search Term ☑ Words ☐ Case ☐ Regex: contact address

Search Window Size: 50

Show Every Nth Row: 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No. 1

Files Processed

Corpus Files: Scam Job Offers.txt

Concordance Hits 7

Hit	KWIC	File
1	ne/fax number and your full contact address. Thank you a	Scam Job Offers.txt
2	ne number/fax and your full contact address. Thanks In ac	Scam Job Offers.txt
3	the following information: * CONTACT ADDRESS: * FULL	Scam Job Offers.txt
4	offer please send your name contact address and telephoı	Scam Job Offers.txt
5	wing: (1)Your full names, (2)Contact address and, (3)Phoı	Scam Job Offers.txt
6	wing: (1)Your full names. (2)Contact address. and,Phone/	Scam Job Offers.txt
7	oted. (1) Your full names, (2) Contact address and, (3) Pho	Scam Job Offers.txt

Search Term ☑ Words ☐ Case ☐ Regex: contact address

Search Window Size: 55

Show Every Nth Row: 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

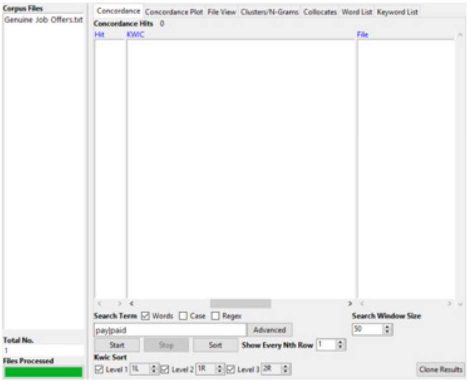
Total No. 1

Files Processed

In a similar vein, unlike professional recruiters, scammers usually rely on financial promises to persuade their victims and influence their behavior. This feature is asserted by the search carried out on the terms ‘**pay|paid**’ in AntConc. The findings revealed that this search term yielded 0 hits in the context of genuine offers, while 15 occurrences were reported at the level of scam job offers. This is displayed in **Table 4**. Similarly, searching for ‘**money**’ as a part of the semantic field of the previous search term asserts the previously obtained results. As shown in **Table 5**, the search term ‘**money**’ returned 0 hits in the corpus of authentic offers while 8 hits were obtained in scams. Thus, this serves as another critical indicator differentiating genuine offers from scams.

Table 4

Concordance hits of ‘pay|paid’ in genuine and scam job offers



Corpus Files: Genuine Job Offers.txt

Concordance Hits 0

Search Term ☑ Words □ Case □ Regex: pay|paid

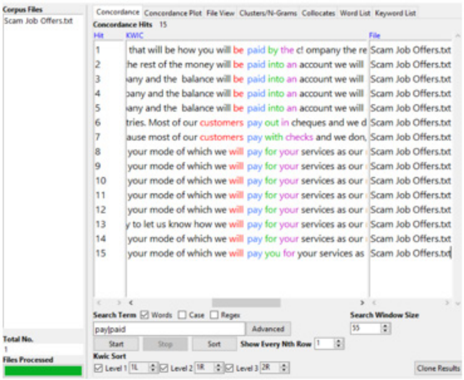
Search Window Size: 50

Show Every Nth Row: 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No. 1

Files Processed



Corpus Files: Scam Job Offers.txt

Concordance Hits 15

Hit	KWIC	File
1	that will be how you will be paid by the c! ompany the re	Scam Job Offers.txt
2	he rest of the money will be paid into an account we will	Scam Job Offers.txt
3	any and the balance will be paid into an account we will	Scam Job Offers.txt
4	pany and the balance will be paid into an account we will	Scam Job Offers.txt
5	any and the balance will be paid into an account we will	Scam Job Offers.txt
6	tries. Most of our customers pay out in cheques and we d	Scam Job Offers.txt
7	ause most of our customers pay with checks and we don,	Scam Job Offers.txt
8	your mode of which we will pay for your services as our	Scam Job Offers.txt
9	your mode of which we will pay for your services as our	Scam Job Offers.txt
10	your mode of which we will pay for your services as our	Scam Job Offers.txt
11	your mode of which we will pay for your services as our	Scam Job Offers.txt
12	your mode of which we will pay for your services as our	Scam Job Offers.txt
13	y to let us know how we will pay for your services as our	Scam Job Offers.txt
14	your mode of which we will pay for your services as our	Scam Job Offers.txt
15	your mode of which we will pay you for your services as	Scam Job Offers.txt

Search Term ☑ Words □ Case □ Regex: pay|paid

Search Window Size: 55

Show Every Nth Row: 1

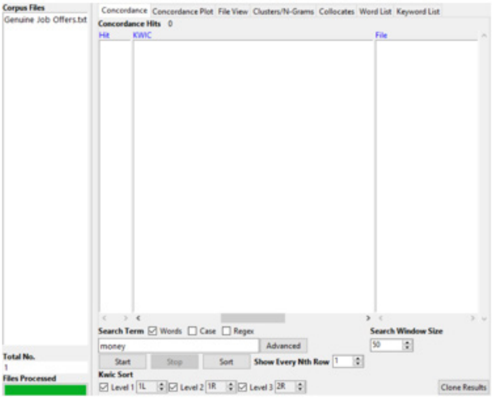
Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No. 1

Files Processed

Table 5

Concordance hits of ‘money’ in genuine and scam job offers



Corpus Files: Genuine Job Offers.txt

Concordance Hits 0

Search Term ☑ Words □ Case □ Regex: money

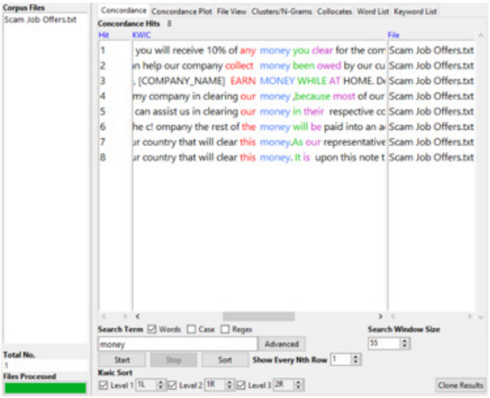
Search Window Size: 50

Show Every Nth Row: 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No. 1

Files Processed



Corpus Files: Scam Job Offers.txt

Concordance Hits 8

Hit	KWIC	File
1	you will receive 10% of any money you clear for the com	Scam Job Offers.txt
2	n help our company collect money been owed by our cu	Scam Job Offers.txt
3	. [COMPANY_NAME] EARN MONEY WHILE AT HOME. D	Scam Job Offers.txt
4	my company in clearing our money ,because most of our	Scam Job Offers.txt
5	can assist us in clearing our money in their respective cc	Scam Job Offers.txt
6	he c! ompany the rest of the money will be paid into an a	Scam Job Offers.txt
7	ur country that will clear this money.As our representative	Scam Job Offers.txt
8	ur country that will clear this money. It is upon this note t	Scam Job Offers.txt

Search Term ☑ Words □ Case □ Regex: money

Search Window Size: 55

Show Every Nth Row: 1

Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

Total No. 1

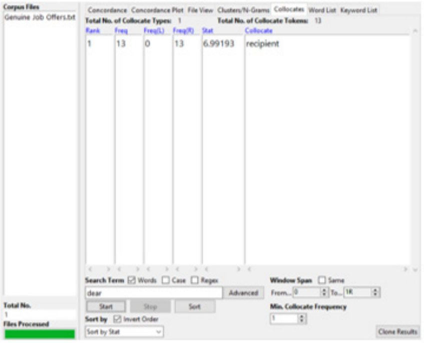
Files Processed

In line with the illustrated concordances, investigating collocates also provided deep insights into deception indicators characterizing scam job offers. At this level, the researcher explored the company of common terms characterizing job offers and identified their contextual use. For example, the term ‘**dear**’ is a common term in the greeting of emails. However,

Table 6 shows that despite its presence in both scam and genuine mails, its collocation varied widely. The results showed that the term ‘**dear**’ existed 13 times in each corpus, yet it collocated with professional terms in authentic offers, mainly ‘**recipient**’, while it had 7 collocations with ‘**friend**’ in scam offers and 6 times with ‘**Sir**’. This emphasizes the significance of considering both hits and context in exploring the use of any term. In this case, although the term ‘**dear**’ yielded equal number of hits in both corpora, the company it had in authentic offers broadly differed from that in scam offers since the former had a professional collocation while the latter involved instances of informality. Thus, this serves as a critical indicator of deception characterizing scam job offers.

Table 6

The collocates of ‘dear’ in genuine job offers compared to scams

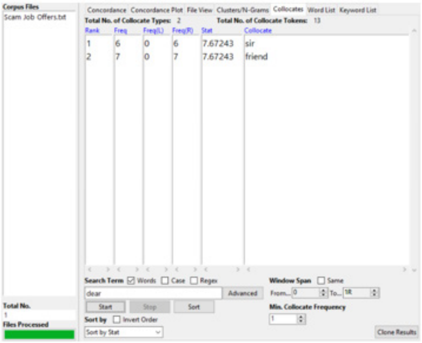


Corpus Files
Genuine Job Offers.txt

Total No. of Collocate Types: 1 Total No. of Collocate Tokens: 13

Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	13	0	13	6.99193	recipient

Search Term ☑ Words □ Case □ Regex
dear
Window Span □ Same
From... 0 To... 1R
Min. Collocate Frequency
1
Sort by ☑ Invert Order
Sort by Stat
Total No.
1
Files Processed



Corpus Files
Scam Job Offers.txt

Total No. of Collocate Types: 2 Total No. of Collocate Tokens: 13

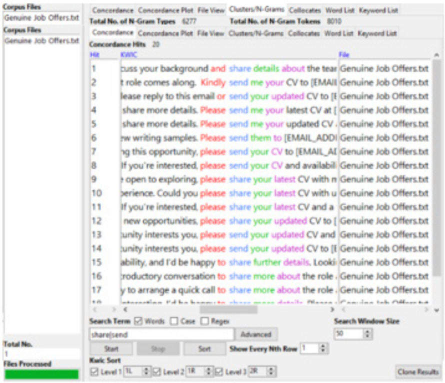
Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	6	0	6	7.67243	sir
2	7	0	7	7.67243	friend

Search Term ☑ Words □ Case □ Regex
dear
Window Span □ Same
From... 0 To... 1R
Min. Collocate Frequency
1
Sort by □ Invert Order
Sort by Stat
Total No.
1
Files Processed

Proceeding with the scope of collocation, exploring the collocation of ‘share|send’ in both corpora yielded significant results. **Table 7** shows that ‘share|send’ existed 20 times in authentic offers and mainly collocated with ‘CV, further details, etc.”. Besides, the recipient was kindly and professionally invited to share their CV as reflected by the words and phrases such as ‘please, kindly, I’d be happy to, etc.”. Conversely, the search term/s ‘share|send’ returned only 4 hits in scam job offers and did not collocate with professional terms. Instead, recipients were required to share personal information such as ‘contact address’, and the sender seemed to lack professional communicative skills as there are instances of authoritarian commands such as ‘do send me’ in requesting any information from the recipients. Therefore, this presents a lucid demonstration of the professional, communicative, and linguistic flaws characterizing scam job offers and serving as deception indicators within this context.

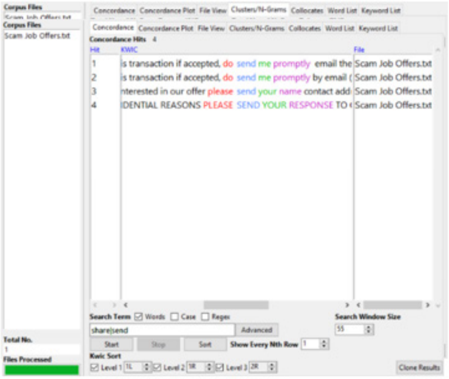
Table 7

Concordance lines contextually showing the collocates of share|send’ in genuine and scam offers



Hit	KWIC	File
1	cuss your background and share details about the tear	Genuine Job Offers.txt
2	t role comes along. Kindly send me your CV to [EMAII	Genuine Job Offers.txt
3	lease reply to this email or send your updated CV to [E	Genuine Job Offers.txt
4	share more details. Please send me your latest CV at [	Genuine Job Offers.txt
5	share more details. Please send me your updated CV .	Genuine Job Offers.txt
6	ew writing samples. Please send them to [EMAIL_ADDI	Genuine Job Offers.txt
7	ng this opportunity, please send your CV to [EMAIL_AI	Genuine Job Offers.txt
8	If you're interested, please send your CV and availabil	Genuine Job Offers.txt
9	e open to exploring, please share your latest CV with m	Genuine Job Offers.txt
10	perience. Could you please share your latest CV with u	Genuine Job Offers.txt
11	If you're interested, please share your latest CV and a	Genuine Job Offers.txt
12	new opportunities, please share your updated CV to [	Genuine Job Offers.txt
13	tunity interests you, please send your updated CV and	Genuine Job Offers.txt
14	tunity interests you, please send your updated CV to [E	Genuine Job Offers.txt
15	ability, and I'd be happy to share further details. Looki	Genuine Job Offers.txt
16	troductory conversation to share more about the role .	Genuine Job Offers.txt
17	y to arrange a quick call to share more about the role .	Genuine Job Offers.txt

Search Term ☑ Words ☐ Case ☐ Regex: share|send
Search Window Size: 50
Show Every Nth Row: 1
Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R



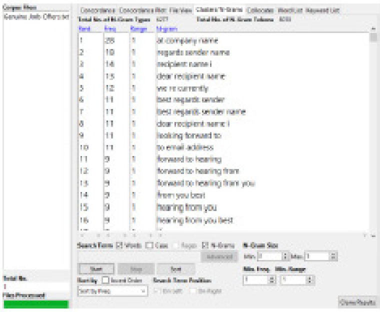
Hit	KWIC	File
1	is transaction if accepted, do send me promptly email the	Scam Job Offers.txt
2	is transaction if accepted, do send me promptly by email (	Scam Job Offers.txt
3	nterested in our offer please send your name contact add	Scam Job Offers.txt
4	IDENTIAL REASONS PLEASE SEND YOUR RESPONSE TO (	Scam Job Offers.txt

Search Term ☑ Words ☐ Case ☐ Regex: share|send
Search Window Size: 55
Show Every Nth Row: 1
Kwic Sort: ☑ Level 1 1L ☑ Level 2 1R ☑ Level 3 2R

In line with the investigated concordances and collocates, n-grams/ clusters also provide a fertile window into uncovering deception through text mining. To explore frequent lexical characterizing each of the studied corpora, the researcher looked up 3-grams, 4-grams, and 5-grams as a sample of the recurrent linguistic patterns occurring in scam and genuine job offers. **Table 8** displays the reported findings. As shown in **Table 8**, the top 10 frequencies of the studied n-grams in genuine job offers mainly include identifying information, salutations, and email closing such as ‘**company name, dear recipient name, forward to hearing from you, etc.”**. This type of information is usual in professional recruitment offers.

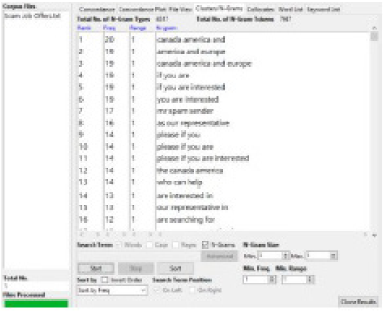
Table 8

3-grams, 4-grams, and 5-grams in genuine and scam job offers



Total No. of N-Gram Types [illegible] Total No. of N-Gram Tokens [illegible]

Rank	Freq	Range	N-gram
1	28	1	at company name
2	18	1	regards sender name
3	16	1	recipient name i
4	15	1	dear recipient name
5	12	1	we re currently
6	11	1	best regards sender
7	11	1	best regards sender name
8	11	1	dear recipient name i
9	11	1	looking forward to
10	11	1	to email address
11	9	1	forward to hearing
12	9	1	forward to hearing from
13	9	1	forward to hearing from you
14	9	1	from you best
15	9	1	hearing from you
16	9	1	hearing from you best



Total No. of N-Gram Types [illegible] Total No. of N-Gram Tokens [illegible]

Rank	Freq	Range	N-gram
1	20	1	canada america and
2	19	1	america and europe
3	19	1	canada america and europe
4	19	1	if you are
5	19	1	if you are interested
6	19	1	you are interested
7	17	1	mr spam sender
8	16	1	as our representative
9	14	1	please if you
10	14	1	please if you are
11	14	1	please if you are interested
12	14	1	the canada america
13	14	1	who can help
14	13	1	are interested in
15	13	1	our representative in
16	12	1	are searching for

Nevertheless, investigating the first 10 frequencies in scam job offers revealed uncommon patterns for the studied n-grams. These patterns included '**canada america and, america and europe, who can help, etc.**'. The obtained results in terms of 3-grams, 4-grams, and 5-grams served both a quantitative and qualitative function. The reported frequencies and the displayed rank of recurrent patterns defined word sequences and occurrences in each corpus. Besides, the carried-out comparison illustrated the linguistic content and collocations differentiating scam offers from genuine ones. Therefore, through this analysis of n-grams, the frequent linguistic patterns and collocational clusters of scam job offers were differentiated from authentic ones, serving as a lucid indicator of deception.

Drawing on the findings brought forth by the conducted comparative text mining analysis, it is revealed that words' hits, frequencies, collocations, and clustering patterns provide detailed insights into identifying deception indicators. Compared to genuine job offers, scam offers lack professionalism and often violate professional ethical standards mainly by disrespecting privacy concerns. Besides, the unrealistic financial promises also serve as clear indicators of deception and e-fraud. At the level of collocates, the formal context of authentic offers compared to the informality displayed in the context of scams also functions as a critical feature differentiating real offers from scams. Moreover, the analysis of selected n-grams revealed the lexical patterns and clusters characterizing scam recruitment offers. Together, these features function as critical indicators of deception.

A Forensic Stylometric Analysis of Scam Versus Authentic Recruitment Offers

In line with the conducted text mining, the researcher also delved into stylometric analysis as a critical investigative procedure characterizing forensic linguistic analysis. Grounded in corpus linguistics, stylometry is a computational tool through which e-fraud is effectively investigated within the realm of forensic linguistics. Through the conducted stylometric analysis, the researcher explored the style of scam job offers and compared it to that of genuine ones. This provided insights into the diverse style markers distinguishing authentic offers from scams in terms of the sentence length and punctuation patterns. As illustrated before, only the email body was considered at this level to avoid any algorithmic errors in identifying real linguistic sentences due to line breaks.

In terms of the sentence length, **Figure 1** shows the variation between genuine and scam job offers. While the y-axis denotes the sentence count, the x-axis represents the number of words contained in each sentence. This is briefly clarified below:

X-axis= the number of words in each sentence

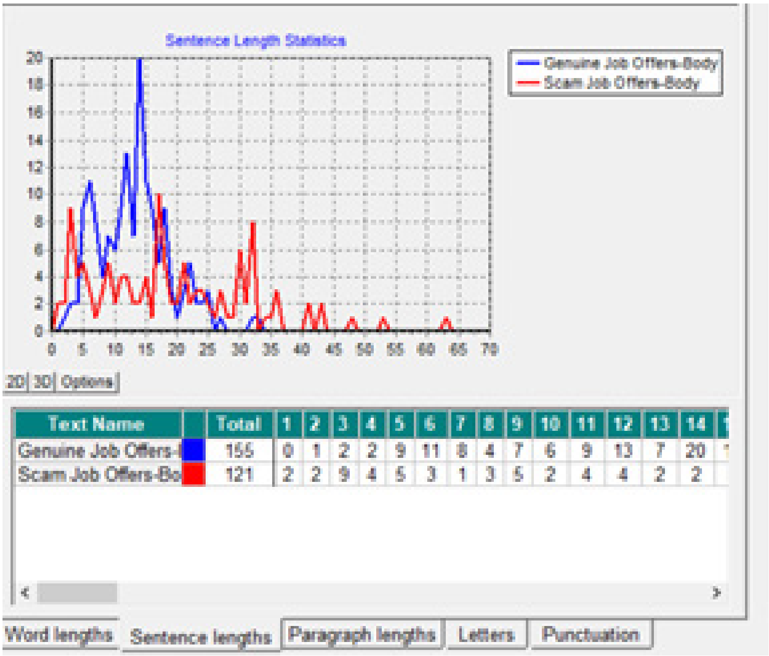
Y-axis= the number of sentences corresponding to the word counts displayed on the X-axis

Signature Stylometric System (1.0) measures sentence length according to the number of words making up a sentence. Besides, for a string of words to be counted as a sentence, it should end with a period, a question mark, or an exclamation mark. Signature Stylometric System (1.0) follows this punctuation convention in detecting sentences. Accordingly, **Figure**

1 shows that the selected corpus of genuine job offers consists of 155 sentences while that of scam job offers includes 121 sentences. At the level of genuine job offers, the graph portrays that the majority of sentences (20 sentences) often include 14 words/sentence, which is a linguistically appropriate length of a sentence in the email context. In addition, the graph reveals that the sentence length in this dataset mainly ranges between 2 to 33 words/sentence, showing a usual and logical distribution in natural written discourses, mainly in the email context. Moreover, this reflects the regularity and homogeneity among sentences, offering a smooth flow of meaning and enhancing the communicative purpose of the shared message.

Figure 1

A graph showing stylometric variation between genuine job offers and scams in terms of sentence length



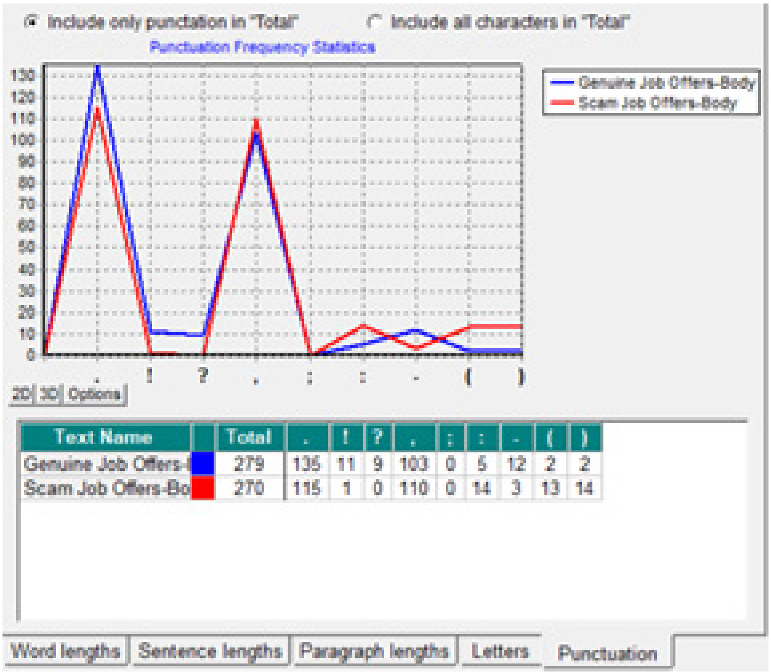
Text Name		Total	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1
Genuine Job Offers-I	🟦	155	0	1	2	2	9	11	8	4	7	6	9	13	7	20	1
Scam Job Offers-Bo	🟥	121	2	2	9	4	5	3	1	3	5	2	4	4	2	2	

On the other hand, scam job offers involve an atypical word count making up sentences. As reflected in **Figure 1**, sentences' length in the context of scam job offers oscillates between two divergent points, particularly, 1 to 63 words/sentence. This vast span mirrors the lack of consistency and homogeneity among sentences, showing their poor linguistic structure. Besides, having 2 occurrences of 1-word sentences reflects issues either in structure or in punctuation which are serious flaws characterizing scam offers. Moreover, the presence of 48-word, 53-word, and 63-word sentences together with 1-word and 2-word sentences in the same dataset mirrors poor regularity and lack of professionality as two critical features characterizing the style of scams. Therefore, unlike genuine job offers emails which adhere to proper linguistic standards and involve a homogenous distribution of sentences, scams lack the proper linguistic style needed in constructing a balanced, consistent, and communicative context due to the irregular variation and poor consistency in sentence length.

In line with the sentence length, punctuation patterns also serve as pivotal style markers distinguishing genuine job offers from scams. As displayed in **Figure 2**, the studied corpus of genuine job offers includes 279 punctuation marks, with a variety of types, particularly, 135 periods, 11 exclamation marks, 9 question marks, 103 commas, 0 semicolons, 5 colons, 12 dashes, and 2 pairs of brackets. Linked to the aforementioned analysis of sentence numbers and length, the distribution of commas shows harmony with the average sentence length. Also, the sum of periods, exclamation marks, and question marks denoting the end of sentences is: 135+11+9=155. This sum matches the total number of sentences displayed in **Figure 1**, which asserts that no flaws or errors exist in the studied dataset.

Figure 2

Punctuation frequencies in genuine and scam job offers



◉ Include only punctuation in "Total" ○ Include all characters in "Total"

Text Name		Total	.	!	?	,	;	:	-	(	)
Genuine Job Offers-I	🟦	279	135	11	9	103	0	5	12	2	2
Scam Job Offers-Bo	🟥	270	115	1	0	110	0	14	3	13	14

However, although the corpus of scam job offers includes a variety of punctuation marks, their distribution involves some errors and the sum of periods, exclamation marks, and question marks showing sentence endings does not match the total number of sentences displayed in **Figure 1**. In specifying the displayed punctuation marks, **Figure 2** shows that the corpus of scam job offers includes 115 periods, 1 exclamation mark, 0 question marks, 110 commas, 0 semicolons, 14 colons, 3 dashes, and an odd number of brackets, i.e., 13 left brackets and 14 right brackets. Starting with sum of sentence ending punctuation marks, the total number of periods, exclamation marks, and question marks denoting sentence endings is: 115+1+0=116. This number (116) does not match the total number of sentences in scam job offers displayed in **Figure 1**, which is 121. This means that some sentences were left unpunctuated although they were followed by a line break denoting the shift into a new sentence. Additionally, the presence of an odd number of brackets reveals the lack of correctness and accuracy in the investigated dataset. Hence, these are serious errors characterizing scam job offers and acting as style markers differentiating scams from authentic job offers.

The Deceptive Strategies of Scam Job Offers as Perceived through the Aristotelian Triad

In addition to the conducted digital and computational analysis, the researcher also carried out theoretical linguistic analysis grounded in **Aristotle's Rhetorical Triangle** as a theoretical framework. This approach assisted in uncovering deception indicators through a qualitative interpretive lens structured on the use of rhetoric as a persuasive strategy in scam job offers. According to the **Aristotelian Triad**, rhetoric is constructed through ethos, pathos, and logos. At the level of ethos as a central element of trustworthiness, scammers lack the authorized identity needed in establishing reliability and enhancing credibility. Thus, they often invent stories to enforce the power of ethos in establishing persuasion and manipulation. For example, scammers pretend to be representatives of companies, and they often impersonate international and reputable companies to enforce the power of ethos as a critical rhetorical element fostering persuasion. This is displayed in saying "we are a company who deal on mechanical equipment…, I am Mr. [SPAM_SENDER], managinig director of [COMPANY_NAME]…, I represent [COMPANY_NAME] import and export company…, My name is [SPAM_SENDER], CEO of [COMPANY_NAME] [CITY_NAME]…". Providing this type of data functions as a significant tool in deceiving the recipients through appearing legitimate and establishing trust with the recipient.

In line with fabricated ethos, scammers mainly benefit from emotional appeal as a powerful means of persuasion. At this level, pathos is activated in terms of exaggerated promises and amplified care and interest in the recipient's skills. In other words, scammers usually rely on fake financial promises to convince their victims of certain operation needed in executing their fraudulent activity. Some instances of these promises include “you will receive 10% of the total amount, earn 10% of every payment made through you to us, you will receive 5% of whatever amount you clear for the company, etc.”. As these assurances appeal to the victim’s emotions, they are expected to respond to the scammer’s requests, and eventually, fall as victims of e-fraud through sharing personal information or even paying advance fees. Besides, some other instances of pathos are displayed in scammers’ salutations and closings such as “Dear friend, yours truly, etc.”. This is a strategic trick used as an icebreaker to implicitly instigate positive emotions in the recipient.

At the level of logos, this is poorly displayed in scam job offers compared to other rhetorical elements. Some instances of logos overlap with ethos, manifested in inventing stories and fake numbers to influence the recipient’s behavior. This strategy enhances the credibility and trustworthiness of the scammers and simultaneously provides a logical and reasonable appeal for the recipient. This is displayed in saying “It is upon this note that we seek your assistance to stand as our representative in your country…, We are one of the fastest growing software company based in…, your basic task will consist of checking emails, processing orders and receiving payments on behalf of [COMPANY_NAME], etc.”. Providing this type of information aligns with the natural procedures and responsibilities of real companies. This serves as a reasonable factor driving the recipients to reply to the scammers’ messages in which they fall into the trap.

Therefore, through the lens of the **Aristotelian Triad**, it was illustrated how scam job offers deploy rhetorical components as persuasive strategies enforcing their fraudulent activities. Since e-fraud in this context is practiced via language, it is thus made clear how language is not only a means of communication, but it is also a critical tool through which cybersecurity is either supported or hindered. In deploying ethos, pathos, and logos, scam job offers victimize the recipients and invisibly manipulate their behavior.

In light of the conducted digital and theoretical analysis, it is concluded that informality, poor professionalism, privacy disrespect, and exaggerated promises are lucid indicators of deception. In addition, the lack of consistency and poor regularity and homogeneity in sentence length, coupled with the unbalanced and incorrect punctuation patterns, are common flaws distinguishing scams from genuine offers. Besides, the deployment of ethical, emotional, and logical appeal is a common feature characterizing the persuasive strategy of scams.

Conclusion

Throughout the conducted study, the interplay between language, cybersecurity, and e-fraud is demonstrated. The performed text mining processes uncovered different linguistic components serving as deception indicators in scam job offers. Besides, the executed forensic stylometric analysis provided evidence-based insights into the diverse style markers distinguishing scams from genuine offers in terms of their sentences length and punctuation patterns. Identifying these indicators through the lens of forensic linguistics contributes to fostering cybersecurity and limiting the spectrum of language crimes manifested in the realm of e-fraud. In addition, the adoption of the **Aristotelian Triad** supported the conducted digital analysis showing that scammers often deploy ethos, pathos, and logos as critical rhetoric elements contributing to manipulating the victims' behavior and fostering persuasion. Building on the unveiled deception indicators, this study responded to a serious issue impeding digital users' experience in the cyberworld. This offered a crystallized demonstration of the power of language and the effectiveness of forensic linguistics in resolving real-life issues and assisting the justice through detecting cyber fraud. Hence, this study extends beyond the traditional scope of language, touching on a serious language crime and offering clear insights into the different indicators distinguishing genuine job offers from scams. This contributed to supporting cybersecurity as a substantial matter indispensable in today's cyberspace.

References

Ahmed, H. (2021). The role of forensic linguistics in crime investigation: Uses in legal proceedings. **Anglisticum Journal (IJLLIS), 10**(2), 23-31. https://doi.org/10.5281/zenodo.4609333

Ali, J. (2020). Forensic linguistics: A study in criminal speech acts. **BSU International Journal of Humanities and Social Science, 2**(1), 39-65.

Amossy, R. (2000). L'Argumentation dans le discours. Paris: Nathan Univertsty.

Anthony, L. (2004). **AntConc: A Learner and Classroom Friendly, Multi-Platform Corpus Analysis Toolkit**. IWLeL 2004: An Interactive Workshop on Language e-Learning, pp. 7- 13.

Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software/antconc/releases/AntConc359/AntConc_64 bit.exe

Ariani, M., Sajedi, F., & Sajedi, M. (2014). Forensic linguistics: A brief overview of the key elements. **Procedia- Social and Behavioral Sciences, 158**, 222-225.

Creswell, J. W. (2014). **Research design: Qualitative, quantitative, and mixed methods approaches** (4th ed.). Sage Publications.

Dzomira, S. (2014). Electronic fraud (cyber fraud) risk in the banking industry, Zimbabwe. **Risk Governance & Control: Financial Markets & Institutions, 4**(2), 16-26.

Dusane, S. (2021). Forensic linguistics - An emerging area in law and justice. **International Journal of Law Management & Humanities, 4**(3), 5170-5179. https://doij.org/10.10000/IJLMH.111118

Fromkin, V., Rodman, R., & Hyams, N. (2014). **An introduction to language**. Boston: Wadsworth.

Galinec, D., Možnik, D., & Guberina, B. (2017). Cybersecurity and cyber defence: national level strategic approach. **Automatika, 58**, 273 - 286. https://doi.org/10.1080/00051144.2017.1407022.

Geeta, D. V. (2011). Online identity theft-an Indian perspective. **Journal of Financial Crime, 18**(3), 235-246. Emerald Group Publishing Ltd.

Guillen-Nieto, V., Vargas-Sierra, C., Pardino-Juan, M., Martinez-Barco, P., & Suarez-Cueto, A. (2008). Exploring state-of-the-art software for forensic authorship identification. **IJES, 8**(1), 1-28.

Ham, J. (2021). Toward a better understanding of "cybersecurity". **Digital Threats: Research and Practice, 2**(3), 1-3. https://doi.org/10.1145/3442445.

Kolupuri, S., Paul, a., Bhowmick, R., & Ganguli, I. (2025). **S**cams and frauds in the digital age: ML-based detection and prevention strategies. In **Proceedings of the 26th International Conference on Distributed Computing and Networking** (1-6). ACM. http://dx.doi.org/10.1145/3700838.3703672

Kumar, R. (2011). **Research methodology: A step-by-step guide for beginners** (3rd ed.). Sage.

Kuzior, A., Tiutiunyk, I., Zielińska, A., & Kelemen, R. (2024). Cybersecurity and cybercrime: Current trends and threats. **Journal of International Studies, 17**(2), 220 239. doi:10.14254/2071-8330.2024/17-2/12

MacLeod, N., & Wright, D. (2020). Forensic linguistics. In S. Adolphs & D. Knight (Eds)., **Routledge handbook of English Language and digital humanities** (pp. 360-377). Routledge.

McMenamin, G. (2002). **Forensic linguistics: Advances in forensic stylistics**. USA: CRC Press.

Mijwil, M., Unogwu, O., Filali, Y., Bala, I., & Al-Shahwani, H. (2023). Exploring the top five evolving threats in cybersecurity: An in-depth overview. **Mesopotamian Journal of Cyber Security**, 57-63. https://doi.org/10.58496/mjcs/2023/010.

Momeni, N. (2012). ‘Fraud in judicial system’ as a language crime: Forensic linguistics approach. **Theory and Practice in Language Studies**, **2**(6), 1263-1269. doi:10.4304/tpls.2.6.1263-1269

Mshvenieradze, T. (2013). Logos, ethos and pathos in political discourse. **Theory and Practice in Language Studies**, **3**(11), 1939-1945.

Murthy, M. L., & Ghosal, M. (2014). A study on Aristotle’s rhetoric. **Research Journal of English Language and Literature**, **2**(4), 249-255.

National Institute of Standards and Technology. (2018). **Framework for improving critical infrastructure cybersecurity** (Version 1.1) [Technical report]. https://doi.org/10.6028/NIST.CSWP.04162018

Nurrosyidah, H. Y. (2016). **Persuasive strategies in Joko Widodo’s political speeches** [Bachelor’s Thesis, Maulana Malik Ibrahim State Islamic University]. https://core.ac.uk/download/pdf/44743443.pdf

Obeyd, S. (2021). Research methods in linguistics: An Overview. **Studies in Linguistics,**

Culture, and FLT, **9**(1), 54-82. 10.46687/SILC.2021.v09i01.004

Olsson, J. (2004). **Forensic linguistics: An introduction to language, crime, and the law**. New York: Continuum.

Perloff, R. M. (2003). **The dynamics of persuasion. Communication and attitudes in the 21st century**. Mahwah, N.J.: Lawrence Erlbaum Associates.

Ramezani, F., Khosousi, A., & Moghadam, K. (2016). Forensic linguistics in the light of crime investigation. **Pertanika Journal**, **24**(1), 375-384.

Sakakini, A. (2020). Forensic linguistics: An applied theory. **BAU Journal - Society, Culture and Human Behavior**, **1**(2). https://doi.org/10.54729/2789-8296.1022

Shuy, R.W. (2006). **Linguistics in the courtroom: A practical guide**. New York: Oxford University Press.

Sousa-Silva, R. (2024). Fighting cyber-malice: A forensic linguistics approach to detecting ai-generated malicious texts. In **Proceedings of the 1st International Conference on NLP & AI for Cyber Security** (pp. 164-174). https://www.researchgate.net/publication/385312304_Fighting_Cyber-malice_A_Forensic_Linguistics_Approach_to_Detecting_AI-generated_Malicious_Texts

Syam, S. (2018). Aspects of forensic linguistics in policing. **Language in India**, **18**(12), 100-111.

Verizon Business. (2024). **2024 Data Breach Investigations Report**. https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf?utm_source=chatgpt.com

Williams, M. (2001). The language of cybercrime. In D. S. Wall (Ed.), **Crime and the Internet** (pp. 152–166). Routledge.