Love Bytes Back: Cybercrime Following Relationship Breakdown

Quincy Taylor
Computer Laboratory
University of Cambridge
Cambridge, United Kingdom
qct20@cantab.ac.uk

Anna Talas
Computer Laboratory
University of Cambridge
Cambridge, United Kingdom
anna.talas@cl.cam.ac.uk

Alice Hutchings
Computer Laboratory
University of Cambridge
Cambridge, United Kingdom
alice.hutchings@cl.cam.ac.uk

***Abstract*—Relationship breakdown and dissolution can be a period of high stress and difficulty. This research uses a subset of underground cybercrime forum data to understand the motivations and discussions in response to a relationship breakdown. Underground cybercrime forums provide insights into cybercrime from the perspective of the perpetrator. To understand the connection between cybercrime and relationship breakdown, we analyze a large, long-running English-language-based underground forum. To analyze data on a scale, we developed three machine learning classifiers that identify the relationship represented, motivation, and moderation of each of the posts. Additionally, we use topic-modeling techniques to surface prominent themes within the dataset. We find forum posts related to relationship breakdown are frequently posted by very active users who use the community as a support system while they vent. The cybercrime types discussed are mostly motivated by an intention to spy, typically through gaining system access. Expanding prior research, we systematically address the role of de-escalators on the platform and suggest improvements to dissuade illegal behaviors.**

I. Introduction

Developing and maintaining close relationships can be a significant source of satisfaction and joy [1]. However, their erosion and eventual breakdown can have a profound and consuming impact [2], [3]. Recent research has attempted to better understand the role of technology as these relationships erode and dissolve [4], [5]. Despite major examples of harassment and the prevalence of Intimate Partner Surveillance (IPS) technologies in discussions with survivors of domestic abuse [6], [7], [8], few studies have investigated the role of online communities where individuals discuss relationship breakdown. Tseng et al. [9] analyzed on-line infidelity forums to understand the patterns associated with IPS which uniquely leveraged data from the perspective of the perpetrators rather than from the recollections of the victims. As these behaviors are highly localized to close relationships, they are hard to measure and frequently result in these incidents being ignored or understudied. Despite studies of IPS, there is limited work exploring how relationship breakdown influences cybercrime as reflected in discussion on online forums. Thus, this study endeavors to address some of these limitations and surface new questions.

Previous work has demonstrated that underground forums provide insights into the solicitation and exchange of cyber-criminal services [10], [11]. However, to the best of our knowledge, there is no prior research specifically into discussions of intimate relationship breakdown within these cybercrime platforms. This research analyzes Hack Forums, an English-language cybercrime forum on the surface web with an active community, to identify if and how relationship breakdown acts as a motivator for cybercrime. To address the gap in research, this study aims to better understand the types of relationships, stated motivations, and trends surrounding these posts.

We address the following questions:

- What information do users seek as their relationships break down? What cybercrime activity is discussed? Do these partners aim to avoid the breakdown of the relationship? Are they seeking to inflict physical or psychological harm? Are they trying to spy on or harass their former partner?
- How do users interact with each other when presented with vulnerability?
- Are there attempts at moderation, such as de-escalation? What is the efficacy of de-escalation?
- Do users join the forum to seek information at the time of the breakdown or are they existing members?

This research uses a data-driven approach utilizing a subset of the CrimeBB dataset collected and maintained by the Cambridge Cybercrime Centre [12]. The specific data used in this study is scraped from Hack Forums which offers many different boards that are frequently contributed to and maintained. Many of the boards allow users to anonymously get support, assist others, and develop skills related to hacking. As we are focusing on relationship breakdown, keywords were used to identify and filter posts that explicitly utilize labels that indicate a past or eroding relationship.

As the corpus of posts is large, manual information extraction is infeasible. Therefore, automatic classification models were trained to label data points. We developed three information classes to extract for each post to answer our research questions. The models tested to automate the classification task included Logistic Regression, Random Forest, Support Vector Machines (SVM), and XGBoost. The purpose of this study is to understand these communities by identifying the technical trends and moderation patterns displayed in the

forums over thousands of posts. This study aims to look at the tension between the highly emotional vulnerability of romantic relationships and the underground messaging boards that facilitate significant online communities.

In addressing the outlined questions, we made the following contributions:

- We develop three classifiers that automate the labeling of underground forum posts given their context and connection to relationship breakdown. The first classifier identifies relationship context and, the second classifier identifies the motivation of the post given the mention of a relationship breakdown. The third classifier identifies instances of moderation, namely de-escalation (and also escalation), within a thread of posts.
- Previous research documents indicators of IPS escalation and de-escalation in infidelity forums [9]. We hypothesised that de-escalation posts within a forum will not have the same level of impact due to the nature of discussion on Hack Forums. The results of the moderation classifier supported this hypothesis, with only limited presence of de-escalators, perhaps due to the audience primarily having a positive view of hacking as evidenced by their presence on the forum. We also see few attempts at escalation, perhaps due to the unclear line between resounding encouragement, casual abetment, and apathetic complicity in forum posts.
- Topic analysis, using Latent Dirichlet Allocation (LDA), of the relationship breakdown dataset surfaced unique themes. Many of the topics have a strong connection to storytelling and reminiscing, such as the use of action terminology or an emphasis on communicating a sense of time. Naturally, clearly defined relationship labels such as "long distance relationship" or "parents divorced" reoccur. The analysis also surfaced religious themes which connects to the popularity of the "Science, Religion, Philosophy, and Politics" board within the relationship breakdown dataset.
- The relationship classifier labels posts based on the connection to the author. We find that relationships in which the author is a subject are the most common type of relationships represented. There are very few instances of individuals posting about their partner's ex-partner, although within that group there is a high rate of cybercrime mentioned and frequent discussion of stalking.

We contextualize the connection between underground forums and criminology theory as well as the current understanding of the role of technology in relationship breakdown in §II. In §III we outline our research methods, including ethical considerations and the machine learning (ML) models we used for classification and topic modeling. We evaluate the classifiers in §IV, and outline our contextualized results of the topic modeling, classifiers, and post analysis. We discuss our findings in §V, followed by a summary of the study and future research directions.

II. Background and Related Work

We surveyed the prior research to gain a better understanding of the significance and vulnerability inherent in the breakdown of a relationship. We establish the connection between online forums and cybercrime by surveying previous research that analyzes cybercrime forums and cybercrime pathways. Next we address the research surrounding relationship breakdown, particularly studies that focus on the role of technology. We also outline research that address the online discussion and prevalence of IPS tools and techniques.

A. Underground Forums and Criminological Theory

With the proliferation of technologies that bring more people online, criminal activity has shifted and become cybercrime [13]. According to Sutherland's [14] differential association theory, the technical skills and mindset associated with criminal activity are learned and shared through frequent association with intimate personal groups. Underground forums are an easily accessible entry point into cybercrime, providing technical information and assistance that can be leveraged to commit illegal activities [15]. Samtani et al. [16] argue that offenders associate with each other in both physical and online spaces, including forums. Pastrana et al. [10] theorize that forums provide the necessary opportunities to cultivate specific techniques and mindsets related to cybercrime. Therefore, the posts presented on underground hacking forums can provide insights into the culture surrounding these groups. We suggest that these online groups can provide insights into cybercriminal activities where the victims are a past partner or they in other ways motivated by relationship breakdown.

Sutherland [14] theorized that individuals become delinquent due to an "excess of definitions favorable to the violation of law". Within the realm of underground forums, anonymized users cultivate the definitions that can persuade or dissuade criminal activities. We seek to identify how individuals in these forums could be setting a tone favorable to violations of the law. In this research, the moderation classifier was used to identify instances of escalation and de-escalation within these interactions. Posts identified by the classifier were reviewed to understand the patterns of participation.

A challenge inherent in forum datasets is the large volume of posts that necessitate automatic tools for extracting intelligence about the content. Previous work addresses this issue by producing automatic tools that leverage natural language processing (NLP) techniques to extract information from underground forums such as those introduced by Portnoff et al. [17]. Their results demonstrate that large-scale automated exploration can provide holistic insights into forum datasets. They suggest this information can be used to better understand the cybercrime ecosystem and bolster defenses against specific strategies and tools. Similar to the methodology of this study, Portnoff et al. [17] leverage a classifier to label different types of posts, some of which are collected from Hack Forums.

The CrimeBB dataset is a convenient source of underground forum data [12]. There is a body of research using this dataset to better understand underground forums. Inspired by the post

type classifier from Portnoff et al. [17], Caines et al. [18] applied labels for post type, author intent, and addressee to the CrimeBB data. Later, Atondo Siu et al. [19] applied labels to understand the connection between currency exchange and discussions about cybercriminal activities. Notably, Atondo Siu et al.'s [19] study contributed the crime type labels for the Hack Forums posts that we use to understand the connection between relationship breakdown and cybercrime. Man et al. [20] also leverage CrimeBB data to understand autism disclosure within these online communities. Talas and Hutchings [21] analyze music shared on Hack Forums to assess if they glamorize crime. Similar to Man et al. [20] and Talas and Hutchings [21], we are looking at a subset of the Hack Forums community.

B. Technology and Relationship Breakdown

Relationship breakdown is difficult to define, but refers to the dissolution of civil unions or non-formalized relationships as well as divorce (the erosion of a formalized marriage). Relationship breakdown is believed to be a process rather than a discrete event [22]. Duck [22] describes the phases of the process, which starts with estrangement and concludes with individuals forming identities separate from their relationship. These phases are primarily internal to an individual or private between the parties, therefore, apart from insights gained through interviews, observations about these phases tend to be limited. The anonymity of online forums may provide insights into these phases as people may feel more willing to be vulnerable before the dissolution is socialized in the subsequent phases.

Duck's social phase includes going public and seeking support. As people's lives bridge the physical and online worlds, the socialization phase is partially migrated online [23]. If online forum communities can act like individual groups in differential association theory, perhaps these communities also form a support network that can shape experience. Amato [24] constructed the Divorce-Stress-Adjustment perspective which builds on Duck's breakdown process [22] by identifying stressors that can further strain the relationship and increase the risk of negative outcomes for those involved. Protective factors absorb negative impacts to reduce the stress experienced during the divorce process. Although explicitly addressing instances of divorce, the Divorce-Stress-Adjustment framework emphasizes the role of moderators as protective factors that can be extended to include instances of less-formalized relationship dissolution. As moderators can positively or negatively impact stressors in the physical world, perhaps individuals on forums play a similar role online as willing listeners and supporters (de-escalators) or as escalators that can encourage vulnerable individuals toward cybercriminal acts.

Relationship breakdown has social, psychological and technical factors that are unique to a technology saturated society [5]. Through a qualitative interview study, Moncur et al. [5] identify challenges associated with relationship breakdown and technology which causes a sense of impression management where individuals feel pressure to act appropriately throughout the process. The effect of deliberate impression management may be eroded in an online forum where there is a perceived sense of anonymity. This could provide a sense of liberation and freedom in their willingness to share personal stories. We posit that the anonymity of Hack Forums as well as the distinct reputation as a hobbyist community could provide a space for venting. People could be more willing to disclose internal feelings of dissatisfaction and confrontation to a community that may not have a relationship with the user in real life. Naturally, relationships lead to entanglement with peer groups that could complicate the social phase, thus forums provide an opportunity for perceived impartiality from the community and the opportunity to share exclusively one side of the experience of the breakdown.

C. Technology and Intimate Partner Surveillance

IPS involves the use of technical and non-technical means to surveil an intimate partner. Qualitative studies have researched the approaches to surveillance through interviews with survivors [6], [7]. These authors report on the threat model and suggest defensive mechanisms to guard against these activities, but exclusively highlight the perspective of the victim. Chatterjee et al. [25] and Tseng et al. [9] take a different approach to understanding instances of IPS by seeking the perspective of the (ab)user. Chatterjee et al. [25] study of the IPS ecosystem by analyzing mobile applications that can monitor and track partners. Notable contributions from this work include a compiled list of apps with the potential to track other users and an assessment of the user reviews associated with each app. They find significant limitations with commercial anti-spyware software in detecting and addressing many of the dual-use applications.

Inspired by Chatterjee et al.'s review of online app marketplaces, Tseng et al. [9] analyze online infidelity forums to identify potential surveillance tactics. They taxonomize strategies used by attackers constructed through the human review of posts from multiple infidelity forums. They also identify patterns of escalation and de-escalation within forum discussions. A limitation of this research was the use of human reviewers to perform a thematic analysis, which limits the number of posts and associated threads that can be reviewed.

This research differs from its predecessors as it utilizes data from a publicly available cybercrime forum rather than infidelity forums and suggests an automated approach to post review. The focus on a cybercrime forum is important as the user-base is either already utilizing Hack Forums which indicates an interest in hacking as a technical skill or the user intentionally chose to seek guidance on Hack Forums indicating an interest in developing the skill or obtaining the service. Inspired by Tseng et al.'s work on escalation and de-escalation, we trained a classifier to identify instances of moderation within a thread that encouraged and discouraged users from utilizing technical tools in response to their relationship breakdown [9]. Unlike Tseng et al., we are able to utilize the users' history to identify if they were new to the forum or

long-time members. Additionally, this study will address if history on the platform impacts the response to their requests.

III. Design and Implementation

This section provides an overview of the ethical considerations that guided our research. We then outline the process for extracting the relationship breakdown dataset, manual annotations, and the background and process for developing each of the classification models.

A. Ethics

This research was approved by the *blinded for review* research ethics committee. The longevity and scale of Hack Forums makes it impossible to retrieve informed consent from all users. Based on British Society of Criminology's Ethics Statement, the CrimeBB dataset does not require informed consent due to the public nature of the community [26]. Throughout the work, we took precautions to ensure privacy for users including anonymizing and paraphrasing text references in this written work. Further, we made no attempt to identify users based on publicly available information such as their posts or usernames. Additionally, this research aimed to identify collective behaviors and trends rather than individual activity.

B. Dataset Extraction

This study utilizes a subset of posts extracted from the CrimeBB maintained by the Cambridge Cybercrime Centre.[1] The CrimeBB data is collected over time using web scrapers [12]. The subset used in this study was extracted from the Hack Forums, a long-running and large English-language forum. Hack Forums has a reputation for cultivating a community around hobby hackers, but provides many discussion boards dedicated to different interests. We selected this forum for analysis due to its extensive history and size [12]. The Hack Forums subset of CrimeBB we used spans 2007 to January 2024. It includes more than 640,000 users with varying levels of engagement on the platform who have collectively authored more than 42 million posts.

To filter the dataset, we identified terms associated with relationship dissolution (shown in Table I). Chatterjee et al. [25] and Tseng et al. [9] also use keywords to identify relevant posts within forum data, but their terms are limited to word associated with IPS rather than the nature of the relationship. After a preliminary review of Hack Forums posts and inspired by the approach taken by Man et al. [20], the terms were discussed and selected by committee to get a wide range of keywords that focus on dissolution. We exclude terms such as "cheating" as we are analyzing posts relating to relationship breakdown rather than infidelity. Nevertheless, themes of infidelity were observed within the breakdown-related posts. To identify posts that could relate to relationship breakdown, we used `regex` patterns to query the CrimeBB database for the associated terms taking into account differences in tense and punctuation. The terms used for filtering includes both subjects and actions, as shown in Table I. These terms were included in a SQL query to extract posts that had a match within the text content of the post.

Category	Keywords
Subjects	my ex, their ex, your ex, her ex, his ex, ex wife, ex husband, ex partner, ex girlfriend, ex boyfriend, ex gf, ex bf, ex spouse, former spouse
Actions	divorce, breakup, break up, broke up, left me

TABLE I: Keywords used to extract the relationship breakdown dataset

Our first attempt to filter relevant posts resulted in a subset of 50,975 posts, but during the first round of annotations the results were found to include many false positives. Many of the false identifications utilized "ex" as shorthand for example, experiment, or exception rather than referencing a former partner. Therefore, we honed the `regex` query to explicitly identify instances of "ex" as a subject. This was accomplished by making the terms possessive, "my ex", "their ex", "your ex", "her ex", and "his ex". This additional filtering resulted in the identification of 29,666 relevant posts. The resulting subset includes posts from 165 boards (78% of the total boards available on the Hack Forums platform). The earliest post in the relationship breakdown dataset is from September of 2007 and the latest post dates to October 2022. The median post date is from the year 2013, which, based on the forum's trends, was the period of peak activity on the forum [11].

C. Pre-Processing

The text content of extracted posts was pre-processed using standard NLP techniques such as removing punctuation, links, and stop words. In addition to the stop words included in the standard `gensim` library, we added terms that appear frequently in posts, but do not provide contextual insight in forming meaningful topics. We also included common internet terms, such as 'lol' and 'bro', because of their limited communicated meaning within a forum context. We used the `gensim` utilities to tokenize and process the content.[2] Once processed, the terms were added into a dictionary and token2id was applied to identify the frequency of individual terms. To create the dataset corpus, we used two forms of vectorization: bag of words (BoW) and Term Frequency Inverse Document Frequency (TF-IDF). Next, we identified and analyzed the frequency of individual terms, bigrams, and trigrams which are sets of word combinations that appear within the dataset.

D. Topic Modeling

Given the scale of the filtered dataset, we used NLP tools to classify posts and surface frequent topics addressed in the posts. LDA is an unsupervised ML technique to identify patterns and repeating topics in the content of the posts. LDA uses a Bayesian network for modeling extracted topics using a text-based corpus. LDA uses terms within a text as an indicator of the topics within the document and can be

[1] https://www.cambridgecybercrime.uk/process.html

[2] Gensim library for pre-processing data https://radimrehurek.com/gensim_3.8.3/parsing/preprocessing.html

Category	Description	Anonymized Example
No Relationship	The post does not directly refer to a relationship.	"This series was actually good. Based on the trailer, we were skeptical. But the cinematography was beautiful, the music catchy, and many of the episodes engaging and hyped me up."
Subject Relationship	The original poster is a member of the relationship or ex-relationship in question. Subsequent posts also hold this classification.	"It's my ex wife, she stole my money. I went to the police but the case was dismissed due to lack of evidence. I'm sure that I would find good evidence on her messenger that would hold up in court. Should i just give up?"
Relationship Proxy	The poster is referencing a relationship that is personal to them (parent, extended family member, friend), but they are not a member of the relationship.	"After my parent's divorced, my mom started dating my future step-dad. I think she was dating him before the divorce. Every sentence that comes out of his mouth is crass and he does it in front of my friends and partner. i need to vent here and calm down."
Partner's Past Relationship	The poster is referring to a former relationship of their current partner.	"Me and my fiance have a stalker. Her ex bf. He's a total psychopath. He has followed us to work, broke into my house, threatened to kidnap her. He uses texting apps to harass her. Any advice?"
External Relationship	References a relationship external to the poster, like a public celebrity relationship. This also includes abstract references to relationships, divorce, or break ups without a specific relationship stated.	"He seems to be that celebrities ex boyfriend. she is an American songwriter. From my search, this seems to be the only reason why he's relevant. He also seems to have a lot of followers online."

TABLE II: Annotation guide for relationship type classifier

effectively applied to online data [21], [27]. The LDA analysis was completed using the `gensim` library which has been used in similar studies [21], [28]. When running the LDA model we limited the number of words in each topic to eight.

We ran the LDA analysis with two different approaches to vectorization first BoW directly and then applying TF-IDF normalization.[3] BoW is the most simple approach to vectorization for text-based data because each word corresponds to a column of a vector with the listed value being the frequency of the word in the document. Next, the frequency of the word in the document can be normalized using the inverse of document frequency (TF-IDF) which adds weighting to terms concentrated in a few documents over those that are perhaps more universal across the entire corpus [29]. This approach allows for less overt topics to surface.

The dataset includes 29,666 posts related to relationship breakdown based on the filtering using keywords as previously described. We initially identified 20 topics from the dataset, but, in an attempt to extract more meaningful topics from the content, the number of topics was limited to five. By reducing the number of topics, the analysis identified topics that were more distinct compared to the original attempt, which included nested and overlapping topics.

E. Manual Annotation and Classifier Models

This study also utilized supervised ML models to train classifiers for labeling thousands of posts, thus it is necessary to provide a sample that has been manually annotated to train the model and evaluate it based on a established standard of truth. To mitigate potential bias, three annotators manually annotated randomly selected posts based on an established annotation guide. The level of agreement between the annotators was quantified using Fleiss' Kappa for each of the annotation sets that form the training dataset for the classification models. Any differences in the classification of the sample data were addressed and mediated during in-person moderation meetings. The annotation guidelines for each of the classifiers are included in Tables II, III, and IV with associated descriptions and anonymized examples.

1) Inter-rater Agreement: Fleiss' κ is a metric to demonstrate the degree of agreement between categorical labels assigned by the annotators [30]. For each iteration of the annotation process, the Fleiss' κ metric is listed in Table V. The Landis and Koch interpretation of the Fleiss' κ is also included to contextualize the value [31]. Landis and Koch interpret a $\kappa <= 0.20$ as 'Poor' and 'Slight,' 'Fair' is assigned to $0.20 < \kappa <= 0.40$, 'Moderate' is assigned to $0.40 < \kappa <= 0.60$, 'Substantial' is assigned to $0.60 < \kappa <= 0.80$, and 'Almost Perfect' is when $\kappa > 0.80$.

The first round of annotations was completed on a sample of 100 posts from the relationship breakdown dataset. Each sample of posts annotated were selected using the random sample function from `Scilearn`. The agreement between the different categories of labels was 'Almost Perfect' for the relationship type classification and 'Substantial' for the motivation and moderator. On completion of this sample, the annotators recognized many false positives associated with the inclusion of the "ex" keyword when identifying posts. Further, the original annotation guide did not include the 'venting' motivation, but it was believed necessary after the initial review of a sample of the posts. To gauge the types of conversation included in full threads, it was necessary to identify initial posts from the sample and extract five posts from the thread to be included in the annotation sample. Each of these observation were applied to improve the sample for the next round of annotations.

Next the first five posts from the threads were collected to form a dataset of 837 posts. They were annotated and the inter-rater agreement was found to be 'Almost Perfect' for the relationship type and 'Substantial' for motivation and

[3]The LDA analysis was completed using the `gensim Ldamodel` tool. https://radimrehurek.com/gensim/models/ldamodel.html

Category	Description	Anonymized Example
None	No overt motivation.	"I'm waiting for it to jump. Some person will jump it at some point then I'm going to dump this coin harder than my ex girlfriend."
Venting	The post is soliciting advice or emotional support, but does not show indications of action in response to breakdown, e.g. "what do you think?"	"Despite being young, I believe I had experienced love, but i don't think she felt the same. I cant stop thinking about her. We broke up a while ago and it was so sudden and hurt. Help me stop obsessing. Help me MOVE ON!"
Avoid Breakdown	The post relates to avoiding a relationship breakdown or resuming a previous relationship.	"She's a beautiful girl, but like i said i don't feel the sparks i did a couple months ago. I don't want to break up with her, i just want to get back to how i felt a couple months ago. I don't know guys/:"
Spy	The post relates to gathering information or spying on their ex-partner or another, e.g. spyware.	"i have heard of programs that you can purchase called "Mobile Spy" and stuff like that but i dont want to buy anything is there a way to see calls and text messages on someone elses phone without that software?"
Revenge	The post relates to getting revenge on an ex- or current partner, e.g. financial, image-based.	"I was in a relationship with a narcissist. For example, he threw me out of the car. I was homeless. I would like to access his email/social accts, gps track to make him pay. I know most of his passwords, but cant get past two Factor."
Harass	The post relates to harassing or annoying an ex- or current partner, e.g. doxing.	"I need access to Skype messenger but i prefer to not create a fake page, i want to keylog her to mentally punish her like she has me. She changed her myspace email and i just wanna delete her acc or toy with her mind."
Physical harm	The post relates to an attempt to inflict physical harm on their ex-partner, another, or themselves.	"If I send you my ex, will you cut her up and scatter her in random places? I will pay shipping fee for the weight."
Mental harm	The post relates to an attempt to inflict psychological harm on their ex-partner, another, or themselves.	"i would rather keylogged or use a flooder to mentally punish her. She changed her myspace email and i dont have a clue to what it is i just wanna delete her acc or toy with her mind either or. Thanks in advance Hack Forums."
Offer	The post is attempting to sell a product potentially by appealing to a client base in a compromised relationship position.	"If you need to hack into any database, delete record, monitor your spouse's divorce papers, improve credit score, spy on your partners whatsapp, text, phone, emails, bank account and many more. I give a trial to convince you."

TABLE III: Annotation guide for motivation classifier

Category	Description	Anonymized Example
Neutral	The post does not address the actions or motivations of the post; nor makes a value judgement on the addressee.	"If you're not that close to them, I wouldn't worry about it. Especially if he's been with other people, I think you're good."
Deescalation	The post intends to dissuade the addressee from enacting harm.	"I bet you want your ex or someone that did you wrong hacked. You will not get access to the accounts. Instead move on. It is easier. You will probably be scammed."
Escalation	The post intends to persuade the addressee towards enacting harm.	"Make sure to use proxies to stay anonymous your efforts. Good luck."

TABLE IV: Annotation guide for escalation/moderator classifier

moderator. The threads provided greater context for assessing the moderation labels of escalation and de-escalation. The third annotation round also included the first five posts from a thread but focused on posts where the classification was a criminal crime type based on the Atondo Siu's [19] classification. This sample included 790 posts and the inter-rater agreement was 'Substantial' for each of the categories. The κ for relationship type in this sample is just shy of the 'Almost perfect' threshold. The fourth round was a much smaller sample of 150 initial posts rather than thread responses with a criminal label to further bolster the training dataset. This set was only annotated with relationship type and motivation, because moderation labels only apply to thread responses. The final round of annotations was completed on a sample that included additional initial posts. The set used to train the classification models was a conglomeration of each round of annotated samples.

As shown in Table V, the relationship type labels had a higher level of agreement between annotators. Despite a few instances of complex relationship descriptions, relationship type is generally objective or directly stated using well-established terminology. This level of clarity within the posts allows for a higher level of agreement compared to motivation which was more convoluted or based in interpretation. The moderator labels were the least similar for each iteration of annotations. When discussing instances of disagreement in the post responses, it was difficult to discern between persuasion towards an action and generally being complicit in the behavior.

Data classification is a task addressed in the field of NLP research. Each ML model employed in this study, namely Random Forest, Logistic Regression, SVM, and XGBoost, has a unique approach to classification that affects the performance of the model. The tested models were selected based on similar research that assesses cybercrime forums and reflect

Iteration	N	Rel Type κ	Rel Type Agreement	Motivation κ	Motivation Agreement	Moderator κ	Moderator Agreement
1) All extracted	100	0.860	Almost Perfect	0.738	Substantial	0.656	Substantial
2) Sample with Threads	837	0.826	Almost Perfect	0.722	Substantial	0.712	Substantial
3) Crimetype Filtered	790	0.795	Substantial	0.693	Substantial	0.659	Substantial
4) Initial Posts	150	0.858	Almost Perfect	0.710	Substantial	-	-
5) Aggregated Final	790	0.794	Substantial	0.680	Substantial	0.639	Substantial

TABLE V: Fleiss' κ between different annotated datasets

a sample of popular supervised ML models used in NLP research [17], [18], [19], [20]. In order to combat imbalanced data, we employed Synthetic Minority Oversampling TEchnique (SMOTE) [32] through the `scikit-learn` library. As detailed in future sections, the models are tested and evaluated to identify the best performing model to be employed as a classifier on the relationship breakdown dataset.

IV. Results

We begin our analysis by outlining the results from the topic modeling to gauge the general topics frequently addressed by posts within the relationship breakdown dataset. Next, we present the results of the relationship type and motivation classifiers, analyzing the relationship between these categories. We then investigate the escalation and de-escalation patterns as identified by the moderation classifier. Finally, we identify patterns associated with the top users and new users that appear in the relationship breakdown dataset.

A. Topic Modeling

1) Bigrams and Trigrams: The first step in identifying recurring themes from the relationship breakdown dataset was systematically identifying collocations, specifically bigrams and trigrams. The top 20 documented bigrams and trigrams and their frequency within the relationship breakdown dataset are charted in Table VI.

Bigram	Freq.
(feel like)	1116
(best friend)	850
(year old)	701
(parents divorced)	631
(years old)	580
(high school)	575
(years ago)	558
(yahoo com)	544
(good luck)	521
(long time)	475
(months ago)	458
(dont know)	439
(let know)	436
(need help)	404
(felt like)	403
(hotmail com)	375
(sounds like)	366
(little bit)	357
(things like)	348
(look like)	342

(a) Bigram analysis

Trigram	Freq.
(long story short)	198
(parents got divorced)	158
(mail address verified)	110
(address verified says)	105
(want join empire)	91
(long distance relationship)	84
(tell little bit)	82
(paragraphs tell little)	78
(memorable happened life)	78
(feel like sh*t)	74
(fbi gov mail)	67
(gov mail address)	67
(long time ago)	64
(gender male religion)	62
(blah blah blah)	59
(feel accepted empire)	59
(wish best luck)	58
(age gender male)	55
(year old girl)	53
(little bit memorable)	52

(b) Trigram analysis

TABLE VI: Bigrams and trigrams sorted by frequency

The bigrams that dominate the relationship breakdown dataset can be divided into categories: interpreting emotion, relationship explanation, references to online resources, and measurements of time. First, the highest frequency bigram is "feel like" with "let know", "need help", " felt like", "sounds like", "things like", "looks like", and "sounds like" also within the top twenty. These results show the use of simile, which could be attributed to the need to explain intangible emotion to the Hack Forums' audience. Second, the bigrams surface two critically important relationships in the dataset: "best friend" and "parents divorced". These labels represent relationships types that are categorized using the relationship type classifier and show that the discussion on the forum is not limited to first-person intimate relationships. Third, the most frequent bigram patterns identified in the top twenty are categorized as measurements of time. This category includes instances of "year old", "years old", "long time", and "months ago", statements about time contextualize post contents as real events that have a time and place. Internet terms, such as "yahoo com" and "hotmail com", appear in the most frequent bigrams, presumably due to the sharing of email addresses.

The results of the trigram analysis shown in Table VIb support many of the same themes highlighted by the bigrams. The original trigram analysis identified usernames from copied transcripts of private messages into public posts. These posts were excluded as the outlier phrases do not represent the trends of the full dataset and to avoid re-identification of users. Similar to the bigram results, the trigrams surface connections to time and storytelling such as the phrase "long story short" and "long time ago." This further supports the idea that users use the platform to share personal stories on the platform. Again similar to the relationship labels in the bigrams, the phrase "parents got divorced" is repeated frequently as is "long distance relationship." The trigrams also surface references to "The Empire". Further investigations revealed this user group started in 2012 to represent the most active participants, with members responsible for providing support to other users. Many of these posts ended up in the relationship breakdown dataset due to users explaining their motivations for applying to join the group. Popular trigrams such as "gender male religion" and "age gender male" are also likely to come from these application posts, which included demographic fields in the template.

2) Latent Dirichlet Allocation (LDA): Collocations provide insight into adjacent terms, but identifying connected topics requires a different approach. LDA topic modeling surfaced themes across the full dataset of 29,666 posts and produced

Topic (% of Tokens)	Related Words
1) Action (35.1%)	like, got, know, left, time, want, said, break, going
2) Business (22.3%)	people, time, help, need, money, want, like, work
3) Family (21.3%)	family, years, parents, old, mother, year, empire, divorce
4) Online forum (12.3%)	account, user, server, use, email, file, paypal, click
5) Religious (9%)	women, god, woman, muslims, man, religion, men, islam

TABLE VII: BoW topics

Topic (% of Tokens)	Related Words
1) Breakdown (53.9%)	left, divorce, got, people, break, like, good, time
2) Emotional (36.3%)	girlfriend, account, breakup, feel, divorced, said, sh*t, phone
3) Harm (6.1%)	rep, f*ck, job, girlfriends, number, file, reputation, check
4) Custody (3.3%)	point, mother, mean, partner, stay, kids, pictures, boy

TABLE VIII: TF-IDF topics

theme clusters indicated by associated word tokens. Based on the result of the LDA model, we defined associated word clusters to topic titles for further analysis. The resulting topics based in BoW vectorization is shown in Table VII and the normalized version using TF-IDF is shown in Table VIII.[4]

BoW Topic 1 is dominated by verbs including "like", "got", "know", "left", and "said". The use of verbs highlight the storytelling cadence of the posts. By contrast, BoW Topic 2 integrates primarily nouns, such as "money" and "work", indicating a connection to transactions. Solicitations of and offers for hacking services frequently appear. Other instances of these terms arise as part of a narrative. BoW Topic 3 is characterised by familial relationship labels, such as "parents", "mother", "family", and "divorce". These words are paired with "years" and "old" which indicate measures of time or age. The emphasis on time and the inclusion of "Empire" in this topic likely connects to "The Empire" connection observed earlier. BoW Topic 4 reveals overt terms related to forums, online services, and technical lingo, such as "account" and "user". The final topic uses religious terminology such as "god", "muslims", "islam", and "religion".

Over half of the tokens identified in the normalized vectorization are captured in the TF-IDF topic 1, which references actions such as "divorce", "left", and "break". Similarly, the second topic includes relationship dissolution terms such as "breakup" and "divorced" paired with the label "girlfriend". As we do not see the term "boyfriend" prominent in the topic analysis, subjects are primarily speaking about female identifying partners. The inclusion of "account" and "phone" in this cluster also highlights the interest in online social media and messaging accounts. There is also a emphasis on emotion with the inclusion of "feel" as a term. The next topic has a characteristic focus on "reputation". The relationship label highlighted in the cluster is "girlfriends", specifically underscoring an effect on the reputation of female partners. Similarly, the last topic projects more of a familial connection with terms like "mother", "partner", "kids", and "boy". Many of these terms are found in posts requesting hacking services in an attempt to get custody of children.

[4]TF-IDF results for the fifth topic included the personal usernames of users likely due to posts that include a transcript of online discussions. To preserve the privacy of these posts, the fifth topic was excluded from the written work.

Classifier	Precision	Recall	F1	Accuracy
Random Forest	0.925	0.921	0.921	0.920
Logistic Regression	0.936	0.936	0.935	0.935
SVM	**0.959**	**0.960**	**0.959**	**0.958**
XGBoost	0.923	0.912	0.910	0.921

TABLE IX: Performance metrics for relationship type classifier

Classifier	Precision	Recall	F1	Accuracy
Random Forest	0.938	0.933	0.934	0.935
Logistic Regression	0.949	0.948	0.948	0.948
SVM	**0.965**	**0.964**	**0.964**	**0.964**
XGBoost	0.928	0.925	0.918	0.922

TABLE X: Performance metrics for post motivation classifier

B. Relationship Type and Post Motivation

We compared the performance of four classification approaches outlined in §III-E. The relationship type and motivation classifiers used the same annotated data to train each of the models. Performance metrics are shown in Tables IX and X. Based on the performance, the SVM classifier was selected and applied to the full relationship breakdown dataset.

The results of the motivation classifier are used to address the question, "what information do individual users seek when facing a relationship breakdown on Hack Forums?" The primary proportion of the posts were labeled with no motivation

Label	Post Count	Percentage
Subject Relationship	16293	55%
External Relationship	5138	17%
No Relationship	7160	24%
Proxy Relationship	1067	3.5%
Partner's Ex	8	0.5%
Venting	6375	72%
Spy	1948	22%
Revenge	362	4%
Offer	19	0.2%
Avoid Breakdown	148	1.7%
Harassment	11	0.1%

TABLE XI: Distribution of the relationship type and motivation classifier results

Label	Post Count	Percentage
access to systems/sql injection	459	31%
bots/malware	403	27.2%
trading credentials	269	18.2%
DDoS/booting/stress testing	163	11%
spam related/sharing email addresses/marketing	82	5.5%
VPN/proxy/hosting	61	4.1%
currency exchange	40	2.7%
identity theft/identity fraud/credit card fraud	1	0.3%

TABLE XII: Crime type distribution in the full dataset

potentially because they were not explicitly 'venting' or mentioning a malicious intention. Table XI shows the distribution of the remaining motive categories within the dataset. Besides the 'no motivation' label, the Table XI shows that 'venting' is the most common motivation along with 'spy,' 'revenge,' and 'avoid breakdown' respectively. To further understand the types of posts, the relationship type and crime type labels provide insight into the structure of their relationships and the potential attacks.

During the manual annotation we noted the often blurred line between 'revenge,' 'harassment,' and harms both 'physical' and 'psychological' within the content of the posts. Additionally, users will include ample details about the relationship in question, potentially in anticipation of future questions or justification for their action. Posters frequently seek justice and/or avengement, while others solicit personal advice or solidarity while sharing their side of a complex relationship story. For example, one user posted *"I am currently in a custody battle for my kid. I think I can prove she is an unfit mother using her social media messages. Would a RAT be the best option?"* This user uses a custody battle as a justification in their appeal for technical advice. Many will tell stories for context when seeking advice from the community. A post labeled as 'venting' began, *"I need some advice or someone to talk to. I don't usually show my personal feelings or emotions on the internet, but i need some advice..."*

The distribution of labels indicate the majority of posts are made from the subject perspective. The 'no relationship' label includes false positives picked up by the keywords or posts that could not be easily defined in any of the other categories. Those who post about their own relationship provide greater detail about the inner working of their thoughts and feelings compared to posts that discuss third person relationships. For example, a user describes a recent breakup, *"My ex gf is gone, because we broke up. I regret the choices i made up to this point in our relationship. I am so sad we will never see each other again. My stomach is sinking and I feel like i am suffocating. I didn't appreciate what I had an now it is gone."* In contrast, those who discuss external or proxy relationships often speak in general terms.

Atondo Siu et al. [19] developed a crime type classifier. The majority of posts within the dataset were labeled 'not criminal', aligning with the finding that most posts do not have a malicious motivation. The distribution of crime types excluding 'not criminal' posts are shown in Table XII. The top three crime type labels are 'access to systems/sql injection', 'bots/malware', and 'trading credentials'. Many users seeking assistance with hacking or requesting hacking services want access to another person's online accounts like social media or email. The emphasis on account services is shown in the requests for hacking services where most posts resemble, *"I need access to my ex's email or Facebook account ASAP."*

We explore the relationship between motivation and crime type, using Atondo Siu et al.'s [19] labels. The distribution violates the assumption of a Chi-square test, which requires a minimum proportion of expected frequencies, therefore we do not test the relationship statistically. Figure 1 shows each label within each motivation category in reference to the associated cybercrime type. We find system access was most frequently discussed in relation to posts that had expressed a motivation for spying.

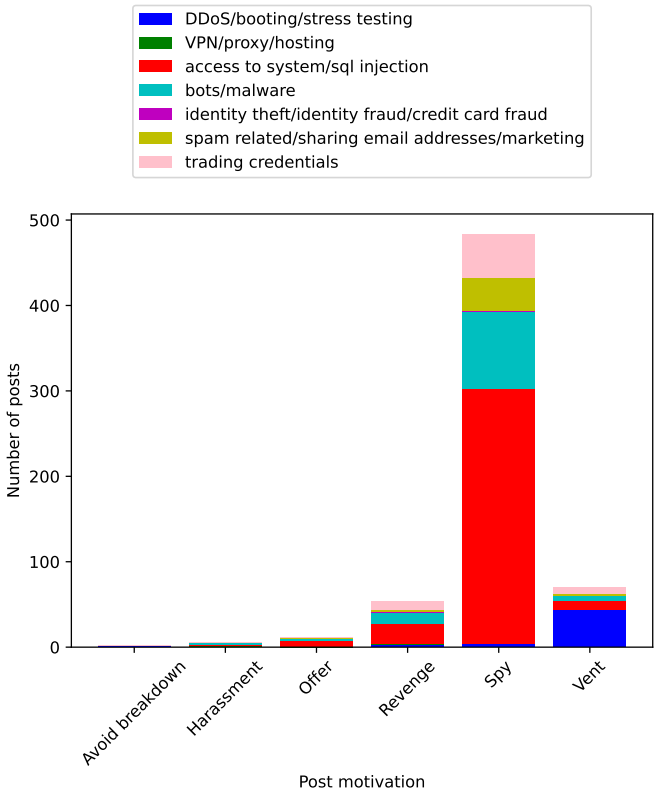

Fig. 1: Cybercrime type and motivation breakdown

A Chi-square test found a significant relationship between motivation and relationship type ($\chi^2(24, N = 8,731) = 7,616.64, p < 0.001$). The contingency plot in Table XIII shows that subjects are more likely to discuss spying in relation to their own relationships, and more frequently vent about proxy relationships. This also shows that offers are significantly more likely to be made in relation to posts external to the author, and harassment is significantly more likely to relate to external relationships and partners' ex-partners than expected. The hue of each tile within the contingency plot shows the general range of the standardized residual value which is the deviation between the observed and the expected value.

C. Patterns of Moderation

The moderation classifier labels the thread of initial posts that are included in the relationship breakdown dataset, excluding those with non-criminal crime type labels [19]. Most threads have fewer than five posts. Replies reflect the reaction and response of the community to the initial post. The best performing model for the moderation classifier was the Random Forest model, as shown in Table XIV. The model was

Motivation	Relationship Type				Total
	Subject	Proxy	External	Partner Ex	
Spy	1827 (1731.5) Std. Res 8.3	65 (121.4) Std. Res -6.0	24 (65.3) Std. Res -5.9	4 (1.8) Std. Res 1.9	1920
Harassment	4 (9.9) Std. Res -6.0	1 (0.7) Std. Res 0.4	5 (0.3) Std. Res 7.7	1 (0.01) Std. Res 9.9	11
Revenge	310 (317.4) Std. Res -1.4	13 (22.2) Std. Res -2.1	27 (11.9) Std. Res 4.5	2 (0.3) Std. Res 3.0	352
Vent	5591 (5665.4) Std. Res -6.0	473 (397.2) Std. Res 7.4	217 (213.7) Std. Res 0.4	1 (5.8) Std. Res -3.7	6282
Avoid Breakdown	139 (133.4) Std. Res 1.5	0 (9.4) Std. Res -3.2	9 (5.0) Std. Res 1.8	0 (0.1) Std. Res -0.4	148
Offer	3 (16.2) Std. Res -10.5	0 (1.1) Std. Res -1.10	15 (0.6) Std. Res 18.7	0 (0.0) Std. Res -0.1	18
Total	7874	552	297	8	8731

TABLE XIII: Contingency Table for Cybercrime and Motivation (Expected Frequencies are shown in the Parenthesis)

Classifier	Precision	Recall	F1	Accuracy
Random Forest	**0.955**	**0.953**	**0.953**	**0.952**
Logistic Regression	0.947	0.947	0.947	0.946
SVM	0.952	0.952	0.951	0.950
XGBoost	0.951	0.949	0.951	0.949

TABLE XIV: Performance metrics for escalation classifier

Label	Post Count	Percentage
De-escalation	43	0.5%
Escalation	491	5.5%
Neutral	8799	94%

TABLE XV: Escalation distribution on full dataset

applied to a dataset of 9,333 posts which were all responses to initial posts from the relationship breakdown dataset.

The classifier results in Table XV show the majority of responses are neutral towards the post rather than attempting to escalation or de-escalate the specific behavior. Escalators are those that encourage the behavior, which is communicated in the form of encouragement or solidarity. In response to a user attempting to access an account, a user responds, *"You could send a keylogger to the boyfriend. maybe you will get the info you want after a while. Good luck."* Hack Forums has an engaged community that frequently provides technical support, so advising on approaches, such as recommending the use of a keylogger, can be viewed as an endorsement.

The role of the de-escalator is to dissuade the behavior presented in the original post. Reviewing results, we find de-escalation frequently comes in the form of informing the original poster of the infeasibility or illegality of the behavior. One example simply states, *"Hack her social media? Is this possible? I think not."* Another user outlines questions for the original poster to consider before attempting to access another individual's system, *"Before attempting a hackback you should consider the following: Are you hurting a third party with this activity? Is it even possible to access the target machine?"* Others dissuade the behavior by encouraging the original poster to move on, *"The best way to resolve the issue is to stop investing time and effort into stalking your ex girlfriend"*. In reviewing the 43 threads with de-escalation present, the effectiveness of these attempts appears limited. The subsequent messages tend to disregard the attempts to dissuade, with some original post authors continuing conversations with others in the thread that play a more encouraging role.

D. User Trends

The relationship breakdown dataset includes 13,586 users with varying levels of engagement and account longevity. There are users whose first post is related to relationship breakdown. Some of these first time users are joining Hack Forums for the express reason of posting about a relationship. By contrast, there is a subset of users who have a high number of posts on the platform addressing relationship breakdown. This section will analyze trends from the accounts whose lifespans fall in these two extremes.

1) First Time Posters: Given the vulnerability and limited support for individuals during a relationship breakdown, we expected to observe a large proportion of posters joining the forum to receive support and find a community. We identified 377 first time posts that create accounts to discuss relationship concerns. As the total number of posts within the dataset is 29,666, first posts comprise only 1.2% of the total dataset, and first time posters represent just 2.8% of users in the relationship breakdown dataset. These results indicate most of the conversation about relationship breakdown occurs between established forum members.

Although first time users are a small proportion of the dataset, many of these accounts are short lived. Note that an individual can have multiple accounts registered on Hack Forums, therefore we refer to account longevity rather than users. 23% of these posts are the only post made for the registered account and others post only a few times. For example, one user posted, *"i need help getting access to my ex-gf's facebook. i have been trying, but i haven't been able to get access. i need someone to help me hack a facebook account,"* twice within a few minutes on both the "Beginner Hacking" and "Hacking Tutorials" boards with no further posts.

The reputation of users is built over time as they engage with the community. As the post count is displayed on the accounts in Hack Forums, it contributes to the users' reputation. Some users requesting relationship advice anticipate

negative reactions. A user prefaces the substance of their post with *"Get ready for a long story. I know I have a low post count, but please I just need some advice...On a side note, please take this thread seriously. I do not post often, and would appreciate if the trolls stayed away."* Not only does the poster anticipate little attention due to the "low post count," the user expects "trolls" which could highlight the hazing that happens to new members. The culture of hazing and trolling new members could dissuade newer users from revealing vulnerability central to relationship breakdown.

2) Top Posters: In the other extreme, there are 12 users with more than 50 posts in the relationship breakdown dataset which we have assessed as the top users. All of the top posters frequent "The Lounge", the default board on the platform. Other popular boards amongst at least 37% top posters include "Science, Religion, Philosophy, and Politics", "Vices", and "Beginner Hacking". In surveying the posts of the top users, most are labeled as venting by the motivation classifier and are labeled non-criminal in nature. Additionally, based on the results of the relationship classifier, their posts are primarily addressing relationships in which they are the subject. Additionally, the proportion of posts that align with the relationship breakdown dataset compared to their total number of posts is very small. Within our sub-dataset, these frequent posters have a mix of initial posts and responses to other people's posts. This varied posting pattern indicates that their engagement in the forum is varied and active.

3) Popular Boards: The most popular board within the relationship breakdown dataset is "The Lounge" with over 18,000 posts. Posts with no overt motivation are the majority of activity on the board from the relationship breakdown dataset. Like the overall dataset, venting and spying were next in concentration. "The Lounge" is a repository of posts that are aimed at general viewers of Hack Forums, therefore appeals for personal and technical advice collect.

Additional boards on Hack Forums beyond "The Lounge" are sorted and explored based on the users interests. The distribution of motivation for each of the top boards is shown in Figure 2. Within the relationship breakdown dataset, the second most popular board is "Science, Religion, Philosophy, and Politics." The breakdown of this board shows that the majority of activity has been classified as venting. Posts sampled from this board show an array of discussions about varied topics such as same-sex marriage, sexual assault, divorce, and religious affiliation. The range of topics provides ample opportunity to intersect with relationship breakdown with many discussing these concepts without much overt personal connection.

The next most popular boards are "Beginner Hacking" and "Requests for Hacking" which demonstrate the prominent intention of utilizing hacking within the relationship breakdown posts. In comparison to the other boards which demonstrate a greater connection to venting, these two boards are dominated by attempts to spy. Revenge is also a clear presence on these boards as the intention is usually to utilize hacking as an act of retaliation against another individual. The board called "Vices"

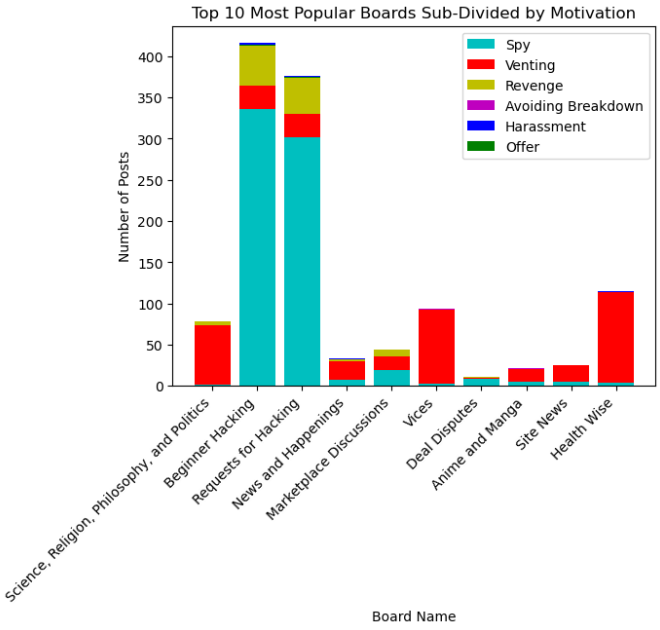

Fig. 2: Most popular boards motivation breakdown

attracts many instances of venting, which would indicate that the board is partly dedicated to storytelling.

As they are established around interest and intention, the boards provide a unique lens to view the posts. The most nefarious motivations, such as spying or revenge, are concentrated on boards with overt connections to hacking. Alternatively, boards that encourage debate ("Science, Religion, Philosophy, and Politics") and confessional style storytelling ("Vices") are dominated by less malicious forum activity like venting. Although the stratification by board can be an effective view to understand the locations of malicious discussion, this strategy is challenged by "The Lounge" which attracts a variety of different motivations because of its broadcast approach and massive audience.

V. Conclusion and Further Work

We concur with Tseng et al. that forums provide a source of threat intelligence that should inform the development of interventions that dissuade abuse [9]. The aim of this study is to present a methodology that can more effectively extract intelligence from underground forums compared to a human review approach previously utilized [9], [7]. Based on our findings, we suggest that the efficacy and techniques for de-escalation be further studied and linked to the overarching culture of the platform. Further, analyzing the concentration of posts that demonstrate malicious intentions can point moderators to particular boards for targeted intervention.

A significant limitation is that Hack Forums is specifically an English language forum, which excludes many different populations that could demonstrate different trends and characteristics. Even within English language forums, this study only includes posts from Hack Forums, thus observations could be influenced by the platform itself. Further work could

expand the number of forums analyzed including those in other languages. Further, the findings from this study may not generalize to private forums or social media groups.

The moderator classifier, which identifies instances of escalation and de-escalation within threads, could be expanded beyond threads related to relationship breakdown to understand the exchanges about cybercrime activity. The results could inform suggestions to better moderate and even dissuade cybercrime within underground cybercrime forums. Special attention could also be taken to study the effect of de-escalators on a platform where they are more integrated into the culture of the forums. This approach would provide a better perspective by which to gauge their impact.

Another limitation is the use of keywords as a way to identify related posts to generate the relationship breakdown dataset as it could exclude related posts that were not overtly using the selected keywords. A different collection of terms would generate a different set of posts for the dataset. Additionally, there could be specific slang on the forums that also connects to relationship breakdowns, but is unknown to the researchers. Another design limitation is the lack of a tuning process for the classifiers. For example, throughout this process, there were misclassifications identified for crime type labels [19].

Expanding on Tseng et al. [9], further work could analyze the technical approaches taken by individuals in the aftermath of a relationship breakdown. We found many requests for approach that required limited, if any, contact with a device. A next step for this research could train a classifier to identify the particular attack phases explicitly discussed. Incident response frameworks could inform categories for phases of access for future post classification [33].

Underground hacking forums provide insights into the motivations and tools associated with committing cybercrime. Building off previous work that leverages such forums, this work analyzed how relationship breakdown motivates and impacts cybercrime based on the discussions recorded on the Hack Forums platform. First, we used topic modeling techniques to surface themes within the text using both bi-gram/tri-gram analysis and LDA modeling. Next, to understand the types of relationships represented, the motivations of posters, and the moderation patterns of the community, we developed three classifiers to identify different characteristics of each post.

We found that Hack Forums provided a platform for users to get support, vent, and request advice about their relationships. Despite being a community that provides a support system to potentially vulnerable people, it also provides resources to those who intend to spy and harass others. We found that the platform also attracts different types of posts with cybercriminal motivations such as spying and harassment concentrated on specific boards. Through topic analysis we discovered that many of the posts focus on storytelling, demonstrating the support role this platform fills for those experiencing a relationship breakdown. We also found that those addressing relationship breakdown are primarily already members of the forum with very few joining the platform specifically to post about their relationship breakdown. We also labeled escalation and de-escalation attempts within the threads of the dataset and found little evidence of effective de-escalation. Based on our findings, we suggest more targeted de-escalation attempts to combat those that attempting to solicit and use hacking services. We believe that this research provides threat intelligence and methods that can be utilized to better understand and defend against relationship breakdown motivated cybercrime.

ACKNOWLEDGMENT

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/T517847/1 (for AT and AH) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 949127) (for AH).

REFERENCES

[1] E. Berscheid and H. T. Reis, "Attraction and close relationships." in *The Handbook of Social Psychology, Vols. 1-2, 4th ed.* New York, NY, US: McGraw-Hill, 1998, pp. 193–281.

[2] T. Tashiro and P. Frazier, ""I'll never be in a relationship like that again": Personal growth following romantic relationship breakups," *Personal Relationships*, vol. 10, no. 1, pp. 113–128, 2003. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6811.00039

[3] G. K. Rhoades, C. M. Kamp Dush, D. C. Atkins, S. M. Stanley, and H. J. Markman, "Breaking up is hard to do: The impact of unmarried relationship dissolution on mental health and life satisfaction." *Journal of Family Psychology*, vol. 25, no. 3, pp. 366–374, 2011.

[4] C. Sas and S. Whittaker, "Design for forgetting: disposing of digital possessions after a breakup," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1823–1832.

[5] W. Moncur, L. Gibson, and D. Herron, "The role of digital technologies during relationship breakdowns," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 2016, pp. 371–382.

[6] J. P. Dimond, C. Fiesler, and A. S. Bruckman, "Domestic violence and information communication technologies," *Interacting with Computers*, vol. 23, no. 5, pp. 413–421, May 2011. [Online]. Available: https://doi.org/10.1016/j.intcom.2011.04.006

[7] D. Freed, J. Palmer, D. Minchala, K. Levy, T. Ristenpart, and N. Dell, ""A Stalker's Paradise": How Intimate Partner Abusers Exploit Technology," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–13. [Online]. Available: https://doi.org/10.1145/3173574.3174241

[8] T. Matthews, K. O'Leary, A. Turner, M. Sleeper, J. P. Woelfer, M. Shelton, C. Manthorne, E. F. Churchill, and S. Consolvo, "Stories from Survivors: Privacy & Security Practices when Coping with Intimate Partner Abuse," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 2189–2201. [Online]. Available: https://doi.org/10.1145/3025453.3025875

[9] E. Tseng, R. Bellini, N. McDonald, M. Danos, R. Greenstadt, D. McCoy, N. Dell, and T. Ristenpart, "The tools and tactics used in intimate partner surveillance: An analysis of online infidelity forums," in *29th USENIX Security Symposium (USENIX Security 20)*, 2020, pp. 1893–1909.

[10] S. Pastrana, A. Hutchings, A. Caines, and P. Buttery, "Characterizing Eve: Analysing cybercrime actors in a large underground forum," in *Research in Attacks, Intrusions, and Defenses: 21st International Symposium, RAID 2018, Heraklion, Crete, Greece, September 10-12, 2018, Proceedings 21*. Springer, 2018, pp. 207–227.

[11] J. Hughes and A. Hutchings, "Digital drift and the evolution of a large cybercrime forum," in *2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 2023, pp. 183–193.

[12] S. Pastrana, D. R. Thomas, A. Hutchings, and R. Clayton, "CrimeBB: Enabling cybercrime research on underground forums at scale," in *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 1845–1854.

[13] P. N. Grabosky, "Virtual criminality: Old wine in new bottles?" in *Cyberspace Crime*. Routledge, 2017, pp. 75–81.

[14] E. H. Sutherland, *The Theory of Differential Association*. New York Chichester, West Sussex: Columbia University Press, 1972, pp. 365–371. [Online]. Available: https://doi.org/10.7312/dres92534-039

[15] A. Hutchings, "Cybercrime trajectories: An integrated theory of initiation, maintenance, and desistance," in *Crime Online: Correlates, Causes, and Context*, T. J. Holt, Ed. Carolina Academic Press Durham, NC, 2016.

[16] S. Samtani, R. Chinn, and H. Chen, "Exploring hacker assets in underground forums," in *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*. IEEE, 2015, pp. 31–36.

[17] R. S. Portnoff, S. Afroz, G. Durrett, J. K. Kummerfeld, T. Berg-Kirkpatrick, D. McCoy, K. Levchenko, and V. Paxson, "Tools for automated analysis of cybercriminal markets," in *Proceedings of the 26th International Conference on World Wide Web*, 2017, pp. 657–666.

[18] A. Caines, S. Pastrana, A. Hutchings, and P. J. Buttery, "Automatically identifying the function and intent of posts in underground forums," *Crime Science*, vol. 7, no. 1, p. 19, 2018.

[19] G. Atondo Siu, B. Collier, and A. Hutchings, "Follow the money: The relationship between currency exchange and illicit behaviour in an underground forum," in *2021 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 2021, pp. 191–201.

[20] J. Man, G. Atondo Siu, and A. Hutchings, "Autism Disclosures and Cybercrime Discourse on a Large Underground Forum," in *2023 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, 2023, pp. 1–14.

[21] A. Talas and A. Hutchings, "Hacker's Paradise: Analysing Music in a Cybercrime Forum," in *2023 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, 2023, pp. 1–14.

[22] S. Duck, *Human Relationships Fourth Edition*. SAGE Publications, 2007.

[23] G. Aeby and J. van Hooff, "Who gets custody of the friends? Online narratives of changes in friendship networks following relationship breakdown," *Families, Relationships and Societies*, vol. 8, no. 3, pp. 411 – 426, 2019. [Online]. Available: https://bristoluniversitypressdigital.com/view/journals/frs/8/3/article-p411.xml

[24] P. R. Amato, "The consequences of divorce for adults and children," *Journal of Marriage and Family*, vol. 62, no. 4, pp. 1269–1287, 2000.

[25] R. Chatterjee, P. Doerfler, H. Orgad, S. Havron, J. Palmer, D. Freed, K. Levy, N. Dell, D. McCoy, and T. Ristenpart, "The spyware used in intimate partner violence," in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 441–458.

[26] British Society of Criminology, "British Society of Criminology Statement of Ethics," https://www.britsoccrim.org/ethics/, 2015.

[27] M. Hoffman, F. Bach, and D. Blei, "Online Learning for Latent Dirichlet Allocation," in *Advances in Neural Information Processing Systems*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds., vol. 23. Curran Associates, Inc., 2010. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2010/file/71f6278d140af599e06ad9bf1ba03cb0-Paper.pdf

[28] S. Bellaouar, M. M. Bellaouar, and I. E. Ghada, "Topic Modeling: Comparison of LSA and LDA on Scientific Publications," in *Proceedings of the 2021 4th International Conference on Data Storage and Data Engineering*, ser. DSDE '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 59–64. [Online]. Available: https://doi.org/10.1145/3456146.3456156

[29] K. Spärck Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of Documentation*, vol. 28, no. 1, pp. 11–21, Jan. 1972, publisher: MCB UP Ltd. [Online]. Available: https://doi.org/10.1108/eb026526

[30] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological Bulletin*, vol. 76, no. 5, pp. 378–382, 1971.

[31] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977. [Online]. Available: http://www.jstor.org/stable/2529310

[32] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[33] N. Naik, P. Jenkins, P. Grace, and J. Song, "Comparing Attack Models for IT Systems: Lockheed Martin's Cyber Kill Chain, MITRE ATT&CK Framework and Diamond Model," in *2022 IEEE International Symposium on Systems Engineering (ISSE)*, 2022, pp. 1–7.