IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 24, 2021, accepted December 2, 2021, date of publication December 8, 2021,
date of current version December 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3134076

Financial Cybercrime: A Comprehensive Survey of Deep Learning Approaches to Tackle the Evolving Financial Crime Landscape

JACK NICHOLLS, ADITYA KUPPA, AND NHIEN-AN LE-KHAC, (Member, IEEE)
School of Computer Science, University College Dublin, Dublin 4, D04 V1W8 Ireland

Corresponding authors: Jack Nicholls (jack.nicholls@ucdconnect.ie), Aditya Kuppa (aditya.kuppa@ucdconnect.ie), and Nhien-An Le-Khac (an.lekhac@ucd.ie)

This work was supported by the Science Foundation Ireland under Grant 18/CRT/6183.

ABSTRACT Machine Learning and Deep Learning methods are widely adopted across financial domains to support trading activities, mobile banking, payments, and making customer credit decisions. These methods also play a vital role in combating financial crime, fraud, and cyberattacks. Financial crime is increasingly being committed over cyberspace, and cybercriminals are using a combination of hacking and social engineering techniques which are bypassing current financial and corporate institution security. With this comes a new umbrella term to capture the evolving landscape which is financial cybercrime. It is a combination of financial crime, hacking, and social engineering committed over cyberspace for the sole purpose of illegal economic gain. Identifying financial cybercrime-related activities is a hard problem, for example, a highly restrictive algorithm may block all suspicious activity obstructing genuine customer business. Navigating and identifying legitimate illicit transactions is not the only issue faced by financial institutions, there is a growing demand of transparency, fairness, and privacy from customers and regulators, which imposes unique constraints on the application of artificial intelligence methods to detect fraud-related activities. Traditionally, rule based systems and shallow anomaly detection methods have been applied to detect financial crime and fraud, but recent developments have seen graph based techniques and neural network models being used to tackle financial cybercrime. There is still a lack of a holistic understanding of the financial cybercrime ecosystem, relevant methods, and their drawbacks and new emerging open problems in this domain in spite of their popularity. In this survey, we aim to bridge the gap by studying the financial cybercrime ecosystem based on four axes: (a) different fraud methods adopted by criminals; (b) relevant systems, algorithms, drawbacks, constraints, and metrics used to combat each fraud type; (c) the relevant personas and stakeholders involved; (d) open and emerging problems in the financial cybercrime domain.

INDEX TERMS Anomaly detection, artificial intelligence, cybersecurity, cryptocurrency analysis, SIM-swap analysis, deep learning, financial crime, hacking, social engineering.

I. INTRODUCTION

Financial crime, or economic crime, is defined by Europol [1] as *"illegal acts committed by an individual or a group of individuals to obtain a financial or professional advantage. The principal motive in such crimes is economic gain"*. This includes money laundering, tax evasion, investment fraud, mass-marketing fraud, and many more. In financial crime, the World Economic Forum considered it to be a trillion-dollar industry [2]. Financial crime is being committed over cyberspace with the use of hacking tools and social engineering techniques which are bypassing financial and corporate institutions security [3]. This leads researchers and corporations to view financial crime in a different light, in that the distinction between financial crime, hacking, and social engineering for economic gain has faded. This is where the authors introduce the term financial cybercrime, which encapsulates the combination of financial crime, hacking, and social engineering committed over cyberspace for the sole purpose of illegal economic gain.

The associate editor coordinating the review of this manuscript and approving it for publication was Mostafa M. Fouda.

As the technical skills and advancement of technology are more available to criminals, their tactics for committing criminal offenses become more difficult to combat. McKinsey & Company [3] described to the industry how current methods and operating models of tackling financial crime and fraud are approaching the holistic view with cybersecurity embedded deeply in the business risk architecture. This is in line with the expanding digitization of finance with the explosion in popularity of cryptocurrency, including central banks developing and releasing their digital currency. China has already paved the way in this regard by releasing the digital yuan [4].

This symbiosis of financial crime and cybersecurity is leading financial institutions to use their in-house developed methods to protect their assets using tools like real-time analytics and interdiction to prevent financial loss [3]. However, as the models are showing signs of lacking the ability to prevent and address these attacks [5], new methods must be developed and deployed across organizations to prevent further loss to their business, customer data, and their own reputation. The new methods being deployed in the research community and industry are machine learning and deep learning models. There are pros and cons attributed to each technique which are discussed in this paper, and the sub-fields of financial cybercrime (i.e. money laundering) pose specific challenges regarding the underlying data and anomaly detection.

To combat financial cybercrime, anomaly detection (AD) is one method of identifying bad actors in a financial network or preventing illicit transactions from occurring. With the increasing technical ability of cybercriminals, and the evolving tools available for masking identity, the task of protecting public and private assets has become more difficult. Group anomaly detection (GAD) is the next phase of anomaly detection in combating complex networks of purposely disguised criminals who are potentially working alone with multiple malicious accounts, or in collaboration with an organized criminal ring.

To the best of our knowledge, there is no survey paper that has analyzed deep learning and machine learning AD research published with a specific focus on combating financial cybercrime. This could be due to the evolving characteristics of what financial cybercrime is, the methods of how criminals are committing financial crime and fraud, or how anomaly detection is branching into specifically analyzing groups of outliers rather than points. Group anomaly detection has garnered more attention from researchers with newly published research in group anomaly detection [6] and cybersecurity [7].

Our contribution to the research community is as follows:

- Providing a holistic view of financial cybercrime and its definition.
- Discuss the various fraud methods adopted by criminals.
- Assessing the performance, trends, drawbacks, and constraints of state-of-the-art anomaly detection techniques being applied to financial cybercrime.
- Describing threat models and relevant persona's involved in financial cybercrime.
- Discussing the future of ML and DL in financial cybercrime including the open and emerging problems across the domain.

In section II, a background is given on AD and GAD with supplementary definitions of the categories and supervision types. Also included is a table listing the popular algorithms used in state-of-the-art research. Section III defines financial cybercrime and lists the various methods. Section IV evaluates performance and defines metrics used in the methods, defines personas, and lists drawbacks within reviewed literature. Section V discusses the challenges, future research and contribution, and the associated difficulty in industry applying state-of-the-art research to their platforms. Table 1 is a glossary for all abbreviations used throughout this survey.

TABLE 1. Glossary.

ABBREVIATION	PHRASE
AD	Anomaly Detection
AI	Artificial Intelligence
AML	Anti-Money Laundering
AUC	Area Under Curve
CPS	Crown Prosecution Service
DATE	Dual Attentive Tree Aware Embedding
DBN	Deep Belief Network
DEFI	Decentralized Finance
DL	Deep Learning
EU	European Union
FBI	Federal Bureau of Investigation
FINRA	Financial Industry Regulatory Authority
FPR	False Positive Rate
GAD	Group Anomaly Detection
GCN	Graph Convolutional Network
GDPR	General Data Protection Regulation
GNN	Graph Neural Network
GPT-2	Generative Pre-trained Transformer 2
H-GCN	Hierarchical Graph Convolutional Network
IP	Internet Protocol
KYC	Know Your Customer
LSTM	Long Short Term Memory
ML	Machine Learning
MLP	Multilayer Perceptron
NLP	Natural Language Processing
PCA	Principal Components Analysis
PEDA	Pre-Encryption Detection Algorithm
POC	Point of Compromise
RF	Random Forest
RNN	Recurrent Neural Network
SEC	Securities Exchange Commission
SIM	Subscriber Identity Module
SMS	Short Message Service
STGNN	Spatio-Temporal Graph Neural Network
SVM	Support Vector Machine
T-EGAT	Temporal-Edge Enhanced Graph Attention Network
T-GCN	Temporal-Graph Convolutional Network
TAN	Tax Audit Network
USA	United States of America
VAT	Value Added Tax
VPN	Virtual Private Network
VPT	Virtual Property Theft
XAI	Explainable Artificial Intelligence

II. BACKGROUND AND RELATED WORK

A. BACKGROUND

1) ANOMALY DETECTION (AD)

Anomaly Detection (AD) involves the use of various computational and mathematical techniques to detect points of abnormality in a dataset. In related literature, anomaly detection has different names, such as *outlier detection*, *novelty detection*, *noise detection*, and *deviation detection*. Anomaly detection is the process of analyzing the dataset to identify deviant cases. It involves either one or both of the tasks: (a) *identification* of the abnormal data, e.g., noise, deviations or outliers from the original dataset and (b) *discovery* of novel data instances based on the knowledge learned according to the original dataset. Anomaly detection can be used for various goals, [8] such as fraud detection, data quality analysis, security scanning, process and system monitoring, image/video surveillance, spam detection, malicious insider attack detection, data cleansing prior to training statistical models, human behavior analytics [9], and sensor-fault detection [10].

2) GROUP ANOMALY DETECTION (GAD)

Group Anomaly Detection (GAD) is the technique of identifying collections or clusters of data points that are abnormal or inconsistent with ordinary group patterns [11]. Similar to traditional anomaly detection, GAD refers to a problem of finding patterns in groups of data that do not conform to expected behaviors. Anomalous groups may consist of individually anomalous points, which are relatively easy to detect. However, anomalous groups of relatively normal points, whose behavior as a group is unusual, is much more difficult to detect. Reference [11] extends the idea of GAD by classifying both dynamic and static situations. Static GAD identifies groups which go against the normal group behavior, while dynamic GAD examines the differences in the state of a group over a period of time.

3) NETWORK-BASED METHODS

Networks or graphs play an important role in GAD and particularly within financial cybercrime. The state-of-the-art algorithms being published and researched include either a preprocessing stage or direct analysis of the graph or network structures for identification of micro-clusters or sub-communities. Reference [8] due to the variety and mix of the different kinds of networks and graphs available in real domains, it is crucial to use application-specific properties to define anomalies in networks or graphs. Outliers in networks can be defined as the nodes, edges, subgraphs or sub-communities, and spatio-temporal graphs have the same factors except with an evolving and dynamic nature added to the difficulty of identifying outliers. Reference [8] network-based methods allow for some of the most powerful and meaningful forms of data representation, and allow for the expression of an array of entities like social networks, banking networks, chemical compounds, knowledge graphs such as a citation network or bibliography, and many more.

4) DEEP LEARNING (DL)

A sub-field of Machine Learning which uses artificial neural networks to learn representations or features of an input dataset. Similar to graphs, DL plays a crucial role in the future of AD and GAD. Advantages to using deep learning models is that the neural networks can learn their own connections to certain data points through the backpropagation method. This weights particular input differently throughout the input dataset. This allows bypassing of manual feature engineering [12]. However, the backpropagation method typically uses gradient descent methods as a form of loss function, and this can result in the loss minimization not reaching an optimal point due to the surface area of a loss function having multiple local minima, and the global minimum may not be discovered. Types of DL models are autoencoders, recurrent neural networks (RNNs), and graph neural networks (GNNs). Hybrid models exist with a blend of DL and ML, introduced by [13] who created a hybrid model using an unsupervised Deep Belief Network (DBN) trained to extract generic underlying features, and then training a one-class SVM on the features learned by the DBN model. Details on the different algorithms and their formulas can be seen in Table 2.

The AD supervision types focus on the ground truth and ability of models to correctly classify anomalies with the information received. They are:

- *Supervised anomaly detection*: [8] models require the availability of labels for the definitions of normality and abnormality.
- *Semi-supervised anomaly detection*: [8] has only normal data samples or only abnormal ones as the inputs. It endeavors to model a single concept and achieves anomaly detection according to the fitness of the data in the concept.
- *Unsupervised anomaly detection*: [8] is typically employed in the situation where no prior knowledge of the dataset or label information is known.
- *Human-in-the-loop*: [8] or active learning corresponds to the setup where the learning algorithm can selectively query a human analyst for labels of input instances to improve its prediction accuracy.

Recent literature in AD has steered towards using graph models and deep learning algorithms. State-of-the-art algorithms now being used are a combination of the two coined Graph Neural Networks. Some examples of these models would be Recurrent Graph Neural Networks, Graph Convolutional Networks, and Spatio-Temporal Graph Neural Networks. There exists surveys with rich content on the background of graph neural networks which can be read in [14], [15].

Table 2 below displays popular deep learning algorithms used currently in anomaly detection research and throughout the financial cybercrime methods section. The algorithms discussed in Table 2 include graphs, and deep learning algorithms like autoencoders, GNNs and their variations.

TABLE 2. Background algorithms.

Algorithms	Function	Description
Graphs	$G = (V, E)$	G represents the graph dataset, V is a finite non-empty set also called the vertex or nodes set, and E are the pairs of distinct vertices or edges - E may be empty.
Autoencoder	$L_r(G_m, \hat{G}_m) = \|\|G_m - \hat{G}_m\|\|^2$	The autoencoder consists of encoder f_ϕ to embed the input to latent or hidden representations and decoder g_ψ which reconstructs the input from hidden representation. The reconstruction loss of an autoencoder is defined as the squared error between input G_m and output $\hat{G}_m$.
Variational Autoencoder (VAE)	$L(G_m, \hat{G}_m) = L_r(G_m\hat{G}_m + KL(f_\phi(z\|x)\|\|g_\psi(z))$	Variational Autoencoders [16] are generative analogues to the standard deterministic autoencoder. VAE impose constraint while inferring latent variable z. The hidden latent codes produced by encoder f_ϕ is constrained to follow prior data distribution $P(G_m)$. The core idea of VAE is to infer P(z) from $P(z\|G_m)$ using Variational Inference (VI) technique. KL is the Kullback-Leibler divergence.
Adversarial Autoencoder (AAE)	$L_G = \frac{1}{M'}\sum_{m=1}^{M'} \log D(z_m)$ and $L_D = -\frac{1}{M'}\sum_{m=1}^{M'} [\log D(z'_m) + \log(1 - D(z_m))]$	To address the VAE's limitation which is the lack of closed form solution, AAE [17] avoids using the KL divergence. Using adversarial learning to learn broader sets of distributions as priors for the latent code. The loss function for the autoencoder (or generator) L_G is composed of the reconstruction error along with the loss for discriminator L_D. Where M' is the minibatch size while z represents the latent code generated by encoder and z' is a sample from the true prior $P(z)$.
Recurrent Graph Neural Network (RGNN)	$H_v^{(t)} = \sum_{u \in N(v)} f(x_v, x^e, x^e{}_{(v,u)}, x_u, H_v^{(t-1)})$	RGNNs were first introduced by [18] by extending the recursive neural network, and Markov chain approach. [15] In order to extract the node representations the same set of parameters are applied over the nodes across the graph recurrently. Where $f(.)$ is a parametric function and H_v^0 is initialized randomly. The sum operator enables the RGNN to be applicable to all nodes.
Graph Convolutional Network (GCN)	$Z = \tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}X\Theta$	GCNs (fully derived in [19]) are a first-order approximation of spectral graph convolutions. [14] This convolution operation is defined in the Fourier domain by computing the eigendecomposition of the graph Laplacian. Where the Laplacian matrix provides a matrix representation when applied to a graph. The equation shown on the left is the combination of a multi-layer GCN with a layer-wise propagation rule and a first-order approximation of localized spectral filters on graphs. The derivation circumvents the prohibition of computational cost for large graphs through Chebyshev polynomials and combining the convolution of a signal with a filter. The final equation is a generalized representation where $\Theta \in \mathbb{R}^{C \times F}$ is a matrix of filter parameters, $Z \in \mathbb{R}^{N \times F}$ is the convolved matrix signal.
Spatio-Temporal Graph Neural Network (STGNN)	$\tilde{A_{adp}} = SoftMax(ReLU(E_1 E_2^T))$	[15] Key goal of the STGNN is to capture the spatial and temporal dependencies of a graph simultaneously. An example of an STGNN is shown in the adjacent equation column. This STGNN is titled Graph WaveNet and it proposes a self-adaptive adjacency matrix to perform graph convolutions. The full derivation of the formula is available in [20]. E_1 is the source node embedding and E_2 is the target node embedding. The spatial dependency weights between the nodes are derived through multiplication of both embeddings. The ReLU activation function removes weak connections. The SoftMax function is used to normalize the self-adaptive adjacency matrix.

B. RELATED WORK

We examine the financial cybercrime ecosystem based on four axes. They are:

- Different fraud methods adopted by criminals.
- Relevant systems, algorithms, drawbacks, constraints and metrics used to combat each fraud type.
- Relevant personas and stakeholders involved.
- Open and emerging problems in the financial cybercrime domain.

To the best of our knowledge, there does not exist any survey which categorizes financial cybercrime and the methods used to combat it. There is however a wealth of published research including surveys which cover specific topics within financial cybercrime which we acknowledge and reference.

We explored the related work on each axis individually. This exploration included popularly cited and newly published research probing the future of financial fraud and cybercrime.

Addressing the first axis, as we examine multiple fraud cases spanning from money laundering to ransomware, there is information available describing and detailing the fraud methods used by criminals. Some examples include, research on malware categories [21], exploring inner workings of a romance fraud [22], various phishing attack techniques [23], how electronic payment systems are misused [24], and large scale insider trading analysis [25].

For related work on the second axis, we identified surveys that looked at the systems, algorithms, and their challenges combating financial cybercrime fraud types. A survey by [26] examined the financial fraud detection methods applying data mining techniques from 2009 to 2019. This survey did not focus on deep learning. Their findings, reviewing 75 articles, saw roughly 23% of financial fraud detection techniques use SVM, followed by Naive Bayes and Random Forest. Of the 75 papers reviewed, [26] found only 3 papers researching cryptocurrency fraud. Reference [27] review 45 papers for financial accounting fraud detection based on data mining techniques spanning from neural networks to linear regression. Graph anomaly detection with deep learning is examined by [28]. They provide a taxonomy that follows a task-driven strategy and review existing work to categorize the anomalous graph objects that they can detect. Reference [29] provide an overview on how ML has been used for malware analysis in the Windows environment. A survey by [30] provided a state-of-the-art snapshot of the current DL models being applied in finance, highlighting the major advancements. In their work they did not encounter any noteworthy, published research on the applications of Generative Adversarial Networks, or Deep Gaussian Processes in the financial landscape. Research by [31] is a strong beginning point to understanding the fundamentals of cryptocurrency network analysis providing an overview of popular cryptocurrencies and their nuances with regards to graph construction. Reference [15] performed a comprehensive survey on GNNs. They provide a taxonomy for GNNs and also discuss the applications and future direction. A survey on the Group Deviation Detection Methods by [11] provided a thorough review of static and dynamic situations involving group deviation detection. Proactive fraud detection strategies are analyzed by [32] by evaluating Fourier transform and Wavelet transform experiments. These experiments do not involve previous fraudulent transactions with an attempt to move away from the retrospective fraud case analyses. The results show it performs closely to current state-of-the-art algorithms (i.e., random forests) but only exploits a single class of data. The researchers propose a hybrid approach to blend their proactive approach with non-proactive approaches to achieve higher performance of fraud detection. Credit card fraud detection is a popularly researched area with many recent publications surveying the current methods being used for detecting and preventing it. Reference [33] published a survey of credit card fraud detection in machine learning. They reviewed some of the latest techniques being used to detect fraud across supervised and unsupervised methods. They also provided information on the datasets being used by researchers. Another survey on credit card fraud detection techniques by [34] examine popular machine learning techniques like random forests, ANNs, SVMs, k-nearest neighbor, and finally propose their own genetic algorithm which reacts to fraudulent transactions.

The third axis of our work studies the relevant personas and stakeholders involved. Related work includes surveys which look at the victims of cybercrime [35], direct insights into victims of cybercrime [36], and a comprehensive study on cybercrime and cybercriminals by [37]. White papers published by industry capture potential victims of cybercrime, and the capabilities of cybercriminals [3], [5] are some examples.

There is a wealth of information available on the background of anomaly detection including popularly cited literature such as [8] who published an important book in anomaly detection which describes in detail the categories of AD and their respective challenges and applications, the same author produced an earlier work on outlier detection for high dimensional data [38]. A survey by [39] discusses the 'curse of dimensionality' and survey specialized algorithms for unsupervised outlier detection. Reference [12] performed a survey on deep learning from anomaly detection which provides a comprehensive overview the techniques including their challenges. Reference [40] performed a survey of network anomaly detection techniques which look to confront malicious cyber activities performed by criminals and network intruders. Another popularly cited survey on AD by [41] groups existing AD techniques into different categories based on their underlying approach.

III. FINANCIAL CYBERCRIME

The term financial cybercrime is a new term capturing the umbrella of crime committed over cyberspace for the sole purpose of illegal economic gain. In 2019, the FBI released an Internet Crime Report for 2018 [42]. This report detailed the number of complaints and victims, the wide variety of crime types used, and the reported amounts of money both stolen by criminals and then recovered by the FBI Recovery Assets Team (RAT). Table 4 displays the number of victims for each crime type reported in the United States for 2018, and Table 5 shows the amounts stolen. The highest value of criminal activity is reportedly from BEC/EAC (Business Email Compromise/Email Account Compromise) totaling nearly $1.3billion. This report highlights the huge amount of money that can be defrauded from victims of financial cybercrime.

Financial cybercrime perpetrators are difficult to identify. They purposely mask their activities to blend their actions with the normal behavior of any other customer or user of a website or financial service, however when grouped together the activity is more obvious of its abnormality. Group

TABLE 3. Related work taxonomy.

Axis	Paper	Description	Financial Cybercrime Category or Associated Topic
1) Different fraud methods adopted by criminals	[21]	Malware categorization and discussion.	Ransomware
	[22]	Romance fraud analysis.	Romance Fraud
	[23]	Analysis of phishing techniques.	Phishing Attacks
	[24]	Romance fraud analysis.	Money Laundering
	[25]	Romance fraud analysis.	Insider Trading
2) Relevant systems, algorithms, drawbacks, constraints and metrics used to combat each fraud type	[26]	Financial fraud detection survey and associated models.	Fraud
	[27]	Analysis of detecting accounting fraud.	Fraud
	[28]	Review of GAD and DL techniques.	Applicable Algorithms
	[29]	ML techniques to identify Windows malware.	Applicable Algorithms
	[30]	Analysis of DL applications in finance.	Applicable Algorithms
	[31]	Analysis of popular cryptocurrencies.	Cryptocurrency graph preprocessing
	[15]	GNN survey and taxonomy.	Applicable Algorithms
	[11]	Survey on GAD.	Applicable Algorithms
	[32]	Proactive Fraud Strategies.	Fraud
	[33]	Survey on ML in credit card fraud.	Fraud
	[34]	Survey on ML in credit card fraud including genetic algorithm.	Fraud
	[8]	Book on AD techniques.	Applicable Algorithms
	[38]	AD techniques in high dimensional data.	Applicable Algorithms
	[39]	Survey on unsupervised AD techniques.	Applicable Algorithms
	[12]	Survey on DL techniques in AD.	Applicable Algorithms
	[40]	Network AD analysis for malicious cyber activity.	Applicable Algorithm
	[41]	Survey on AD.	Applicable Algorithms
3) Relevant personas and stakeholders involved.	[35]	Behavioral analysis of cybercrime victims.	Fraud, social engineering, hacking
	[36]	Behavioral analysis of cybercrime victims.	Fraud
	[37]	Behavioral analysis of cybercriminals.	Hacking, social engineering, fraud, extortion
	[3]	McKinsey: cybersecurity analysis for financial fraud prevention.	Fraud
	[5]	IBM: Preventing financial crime with AI.	Fraud

TABLE 4. Number of reported victims of selected financial cybercrimes—2018 FBI internet crime report [42].

Crime Type	Number of Victims
Non-Payment/Non-Delivery	65,116
Phishing and variants	26,379
BEC/EAC	20,373
Romance Fraud	18,493
Identity Theft	16,128
Credit Card Fraud	15,210
Lottery/Sweepstakes	7,146
Investment	3,693

TABLE 5. US victims monetary loss for selected financial cybercrimes—2018 FBI internet crime report [42].

Crime Type	Victim $ Loss
Non-Payment/Non-Delivery	$183,826,809
Phishing and variants	$48,241,748
BEC/EAC	$1,297,803,489
Romance Fraud	$362,500,761
Identity Theft	$100,429,691
Credit Card Fraud	$88,991,436
Lottery/Sweepstakes	$60,241,814
Investment	$252,955,320

anomaly detection is a method which can identify the irregular patterns of this behavior, and in instances, more accurately than point anomaly detection in tackling financial cybercrime. The following techniques discussed use a combination of graphs, deep learning, anomaly detection techniques, and feature engineering to identify malicious behavior, and protect the assets of customers and other users.

This section of the paper discusses the evolving intertwining behavior of cybercriminal activity and financial crime. There are numerous methods being deployed by criminals to attack financial institutions, corporations, public agencies, and individuals of the public. This subsection will describe both the financial cybercriminal methodology and the current research methods being implemented by researchers and industry to prevent and tackle financial cybercrime.

1) STOCKS AND SECURITIES INVESTMENT FRAUD

The stock market and financial securities allow people to invest their money with the ambition of making a positive return based on either performing research, or just a hunch. However, it is known that a proportion of the market participants cheat, and by doing so make huge profits at the cost of institutional and retail investors.

Catching these fraudulent actors is not easy, and typically requires a large workforce to gather evidence over a long period of time, particularly in cases of insider trading. However, recent developments in machine learning applications and techniques are aiding in identification of bad actors in a more efficient and faster way. Some of the methods used by people committing fraudulent investing activity include market manipulation, insider trading, money laundering, terrorist

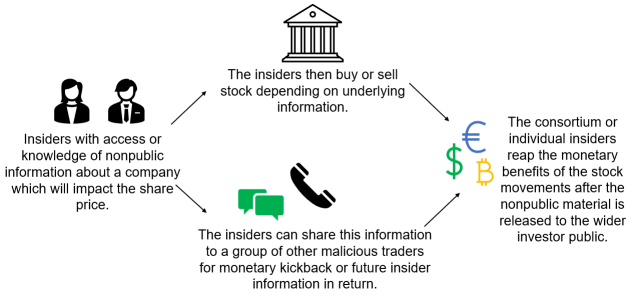

FIGURE 1. Diagram of the illegal insider trading process. Phase 1: Insiders with access or knowledge of nonpublic information about a company which will impact the share price. Phase 2 (Option 1): The insiders then buy or sell stock depending on underlying information. Phase 2 (Option 2): The insiders can share this information to a group of other malicious traders for monetary kickback or future insider information in return. Phase 3: The consortium or individual insiders reap the monetary benefits of the stock movements after nonpublic material is released to the wider investor public.

financing and many more. Market manipulation is considered an act of selling or buying a financial security with the objective of purposefully manipulating the price of the underlying asset or security. Illegal insider trading is when 'insiders', or people who are privy to private and non-public company material use that information ahead of its public dissemination to benefit monetarily. This includes not only the act of trading on securities, but also the leaking of non-public information to third parties. The process of this criminal methodology can be seen in Figure 1. Table 7 outlines the methods and characteristics within the Stocks and Securities Investment Fraud topic.

There has been progress made in Machine Learning and Deep Learning with respect to tackling these two areas. In particular, graph outlier detection methods can aid in the detection of groups of insiders or co-conspirators with regards to insider trading. It is also capable of identifying potential market manipulation by analyzing the order book of stocks to discover connections of traders or brokers acting in unusual methods which would be against the benefit of their clients or the wider market.

Wang *et al.* [43] believe they are the first to successfully produce research that incorporates deep learning algorithms and ensemble learning with techniques for stock price manipulation detection. They propose a novel RNN-based ensemble learning framework to detect stock price manipulation. The group handcrafted a dataset from cases extracted from the China Securities Regulatory Commission with accompanying trading data. The group proposed further possible work implementing more suitable methods of analyzing the stock trading data time series using methods like LSTMs. They specifically mention that integrating more social relationships of executives of the listed company and announcement content can improve the manipulation detection system with possibilities of identifying insider trading. This is an important caveat as [44] produced a paper "Mining Illegal Insider Trading of Stocks: A Proactive Approach" who used various mixed unstructured and structured data to detect anomalous trading patterns for particular companies in an attempt to discover illegal insider trading. The researchers created their dataset using SEC Litigation and Press Release Archive information combined with Yahoo Finance trade data to build their algorithm. A combination of these techniques could prove to be very successful in connecting manipulation of the market and insider trading.

Another paper explores the use of machine learning to detect illegal insider trading [45]. By using insider trade filing made publicly available by the SEC through the EDGAR system and combining that with historical price information from Google Finance, [45] were able to construct networks capable of capturing the relationship between trading behaviors of insiders with the ambition of identifying indicators of potential anomalies. The researchers identified possibly anomalous graphs which could represent cliques of traders. These graphs with higher anomaly ranking score would be passed on to investigators to develop a human-in-the-loop method.

Reference [46] published research on anomaly detection on big data in financial markets. This was not an analysis into fraud, but an attempt to detect rare anomalies present in the previous five years of daily trading information on the Australian Security Exchange. As this was an experiment in big data AD, the researchers attempted to identify anomalies denoted by zero values across five key variables extracted from the exchange including price and volume information. These zero values represent erroneous transaction information as it represents incorrect information surrounding a true transaction. They discovered Local Outlier Factor and Clustering-based Multivariate Gaussian Outlier Score techniques to be the best performing methods.

"Pump and Dump" schemes are used by fraudulent investors to manipulate the price increase of a security in the market and then sell at the apex of the price to gain large profits, shown in Figure 2. This leaves vulnerable and sometimes gullible investors with significant losses as the price decreases after a considerable sell off. This can be done in collusion with numerous investors in a cartel like manner. Particularly stock tickers or cryptocurrency coins are targeted by groups of fraudulent investors to increase the price of a stock through social media. Reference [47] produced a method of detecting pump and dump schemes through analyzing multiple social media platforms and predicting whether the scheme would be successful. They use both SVM and random tree as the methods of classification. Reference [47] did not use deep learning methods in their research.

2) FRAUD DETECTION AND ANTI-MONEY LAUNDERING (AML)

Money laundering is a method used by criminals and people in possession of 'dirty' or illegally obtained funds through criminal activity to transform the money to a 'clean' or legitimate state in the eyes of the law and governments.

TABLE 6. Stocks and securities investment fraud methods.

Method	Characteristics
Market Manipulation	Typically a network of traders and/or brokers working in collusion executing abnormal trades directly affecting the price of an asset of financial security for illegal economic benefit. Analyzing the order book over a period of time can result in identifying outliers, but requires vast amounts of trading information.
Insider Trading	Can be a network of malicious actors or individual using nonpublic material information for illegal economic gain. Criminal cases take an extremely long time to build involving multiple agencies and numerous resources. The trade data is sometimes not enough on its own to identify someone as an insider trader. The communication between traders can also be difficult to identify. The available data for researchers to work on this poses a difficult problem, particularly ground truth trades indicating confirmed insider dealings.
Pump and Dump Schemes	This method requires multiple malicious actors working in synchronization to execute buy or sell orders. The manipulated price allows the group of criminals a quick profit if exited at the optimal time. The scope of information is vast requiring messaging platform data, and corroborating trade data to identify the criminals.

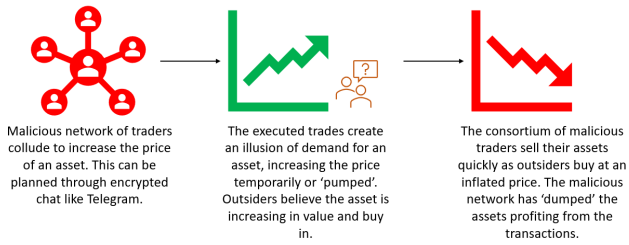

FIGURE 2. Diagram of pump and dump process. Phase 1: Malicious networks of traders collude to increase the price of an asset. This can be planned through encrypted chat like telegram. Phase 2: The executed trades create an illusion of demand for an asset, increasing the price temporarily or 'pumped.' Outsiders believe the asset is increasing in value and buy in. Phase 3: The consortium of malicious traders sell their assets quickly as outsiders buy at an inflated price. The malicious network has 'dumped' the assets thereby profiting from the transactions.

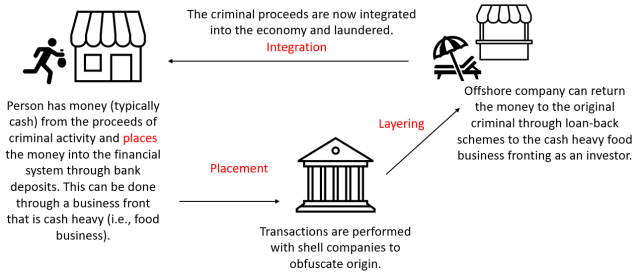

FIGURE 3. Diagram of money laundering example. Phase 1: Person has money (typically cash) from the proceeds of criminal activities and places the money into the financial system through bank deposits. This can be done through a business front that is cash heavy (i.e., food business), Phase 2: *Placement*—Transactions are performed with shell companies to obfuscate origin. Phase 3: *Layering*—Offshore company can return the money to the original criminal through loan-back schemes to the cash heavy food business fronting as an investor. Phase 4: *Integration*—The criminal proceeds are now integrated into the economy and laundered.

The Crown Prosecution Service (CPS) of the United Kingdom [48] define the money laundering scheme to typically involve three stages. The first is *placement* which is the process of depositing criminal money into the financial system. The second is *layering* which is moving the money within the financial system through complex webs of transactions with the goal of obfuscation. Layering is typically performed through offshore companies. Finally, *integration* is the criminal money being absorbed or blended into the real economy, through investments like real estate, stock purchases, and luxury items. An example of money laundering using placement, layering, and integration can be seen in Figure 3.

Machine learning and deep learning have become more popular in their application to tackling money laundering and attempting to identify illegitimate transactions. Zhou *et al.* [49] researched methods to analyze and detect money laundering accounts in online social networks. A large social networking site in China called QQ owned by Tencent has a digital currency within their network, allowing users to perform transactions with other users to make purchases, and also the ability to gift the digital currency to others in the network. One issue within QQ is the laundering of its digital currency. The researchers first collected approximately 500,000 accounts and meticulously labeled these as benign or malicious by hand by following advertisements of cheap virtual currency on major e-commerce sites and associating IP address logins to further identify malicious account activity. Features were engineered to identify account behavior such as account activity like uploading pictures, or site engagement outside of finance, methods of topping up the digital currency, withdrawal, spending and gifting. Sequences of financial activity were modeled using a discrete-time Markov chain model. The sequences captured are used as features in the model. These features are then loaded into a graph and is now used as a global overview of the currency transfer behaviors between the accounts. Using the "Fast Unfolding" [50] method, which finds community structures in large networks, subgraphs are identified which map the malicious-to-malicious accounts, benign-to-benign, and malicious-to-benign. Statistical classifiers are used to identify malicious accounts in the created features. SVM, random forest, and logistic regression classifiers are used and produced extremely high accuracy results of 94.2% with very low false positive rates of 0.97% when using all the features produced. An information gain metric was extracted from the features to identify the significance to the model and the top features included *Percentage of number of expenditures as gifts in the community* and *Normalization of the number of destination accounts in the community*. The top five features, all with information gain > 0.5, were created from the graph

model. Without using group anomaly detection methods such as ''Fast Unfolding'' to detect sub-groups in the overall social network community, the accuracy of the model would have been affected as the top information gain features would not have been included.

Further anti-money laundering papers have been published focusing on the use of social network analysis with deep learning to identify malicious transactions. Reference [51] analyzed Italian factoring businesses which purchase cash receivables, such as invoices, from a company at a discounted price. Factor businesses are used as VAT evasion tools and method of laundering money. The key stage of the process in factor businesses is the money transfer made from the debtor to the factor. This operation is the laundering stage of transforming criminal proceeds into clean funds. Relational graphs were mapped to analyze operational risks associated to the economic sector of activities, consider the risk associated to geographical areas, study the number of transactions, and finally to identify potentially malicious links between different companies sharing the same owner or agents. Through the binned values of transactions, visual assessment of the networks, and using social network analysis like centrality to assess the graphical structures, the researchers were able to identify the risk profile of companies involved in the factoring business. This study used graphing techniques to accomplish their objectives, however, no machine learning was implemented and is an opportunity for advancement of performance for the risk profiling of factoring business customers.

Reference [52] criminals create organizational structures with the goal of obfuscation. In order to combat it, it is essential to identify entire networks to understand and identify the network member's roles. By integrating social network analysis algorithms [52] has harnessed data from bank statements and a national court register to construct and analyze the social networks during an investigation into AML. The researchers were successful in identifying key components of money laundering rings. They were able to reveal true leaders of the network and expose vulnerabilities. They were also able to detect which accounts are held by the same person. They implemented clustering techniques which allowed them to find the roles of persons in the network. These roles are discussed in the paper but include organizers, insulators, protectors, etc. It is quite evident from the published work that the combination of machine learning techniques with social network analysis can be a powerful tool in establishing networks of criminality and tackling money laundering. These tools in combination with humans-in-the-loop, such as law enforcement or AML specialists, could yield highly accurate and promising results.

A paper by Chen *et al.* [53] discussed a graph-based fraud detection system for e-commerce insurance called InfDetect. The experiment was performed on private insurance company data including security deposit insurance and return-freight insurance. They demonstrate that using a combination of graphs and raw features as the input data, and feature construction through supervised and unsupervised graph learning model called DeepWalk. A denoising autoencoder and feature processing is also implemented. They then apply a parameter server based gradient boosted decision tree called PSMART to output a fraud probability. The researchers have claimed to help save millions of dollars per year for these e-commerce companies. One key feature engineered by the researchers was the use of assigning bin scores for transaction amounts which contributed to improved performance.

Araujo *et al.* [54] have recently developed ''BreachRadar,'' a distributed alternating algorithm that assigns a probability of a banking card being compromised to the different possible locations of card use. These Points-of-Compromise (POC) are the point of initiation for fraud in bank transactions. An example would be a data breach resulting in the banking customers information to be obtained, including their credit card information which is then laundered online via DarkNet sites or used personally by criminals. Another example would be card skimming, where the details of a customer's card are copied or cloned using a device purposely altered to illegally obtain this information. The purpose for detecting a POC is to prevent fraudulent use of customer data. The researchers boast a high performing model with 90% precision and recall against a dataset with 10% compromised customer credit cards. The technique includes forming a bipartite graph model to represent all the cards and their locations. With implementation of an in-memory algorithm allowing the updated POC probabilities, the researchers have produced the first distributed procedure capable of automatic detection of POCs.

A newly published vision-guided algorithm called EagleMine [55] has been developed to recognize and summarize node groups in a feature correlation histogram plot. It is capable of identifying anomalous micro-clusters which have underlying nodes of similarity and suspicious behavior that deviate from the normal behavior of the dataset population. EagleMine is capable of detecting a micro-cluster containing hundreds of bots of real-world micro-blog data from Weibo in China. One of the key features identified in the bots was their unusual login-name prefixes, and exhibit similar behavior in the feature spaces. These bots are spamming sales links for cheap technology, and possibly promoting fraudulent activity. EagleMine's key strengths are automated summarization, effectiveness, anomaly detection, and scalability.

Reference [56] a joint research collaboration between MIT and IBM has seen the use of Graph Convolutional Networks (GCNs) as a promising method to combat money laundering. In particular, FastGCN [57] was demonstrated on benchmark datasets to perform at a magnitude of two orders higher than its peers. The researchers noted the key challenge of scalability in AML when working with the large transaction datasets. The AML graphs in practice would always be increasing due to the new transactions occurring over time, and the history of accounts begins to grow. To aid in the further research of AML, the researchers have released a synthetic dataset generator called AMLSim [58].

AMLSim attempts to construct money laundering graph datasets through simulation.

Cryptocurrency: According to CipherTrace, [59] as of April 2021, $432million is the total figure due to major crypto thefts, hacks, and frauds. [59] As of 2019, DeFi (Decentralized Finance) hacks were a rarity, while now DeFi related hacks are responsible for 60% of the total hack and theft volume. Laundering and terrorist financing is possible through cryptocurrency. [59] Noted that 72,000 unique IP addresses were directly linked to Iran, a heavily economically sanctioned country. Many examples of the various hacks, frauds, and illegal transmission of funds through laundering methods are documented by CipherTrace [59]. There is increased research being published in the field of machine learning and deep learning in an attempt to identify the wallets of the associated owners.

Reference [60] the same research group as [56] apply GCNs to cryptocurrency in an attempt to identify money laundering transactions. The researchers have caveated a possible research opportunity of combining the promising results of the GCN and random forest techniques. This paper also provides an accessible Bitcoin transaction dataset provided by Elliptic containing a time series graph of over 200,000 Bitcoin transactions. Challenges faced in trying to apply deep learning or machine learning techniques to the Bitcoin blockchain is the volume of data available. As the blockchain is publicly available there is an abundance of information. The Elliptic dataset is just a snapshot of the blockchain, therefore for practical deployment in order to investigate a single wallet's transactions the full blockchain would need to be accessed. Following the release of this dataset and paper, numerous researchers have also attempted to deploy GCN variations to tackle money laundering problems [61], [62].

Cryptocurrency mixing/tumbling is a method used by cyber criminals to launder cryptocurrency through different wallets to conceal origin of the funds. This is done by using a trusted third party to receive cryptocurrency from an original address, and using an alternative address to send the original funds to a newly set up address by the user [63]. This is also performed through multiple addresses to create a difficult trail to map back to an original address, which could in turn identify a person. This process is shown in Figure 4. Reference [63] published a paper on the use of a deep autoencoder on identifying bitcoin mixing. The experiment was performed in four stages. The first was a transaction graph construction stage which used 10,000 transactions from a public bitcoin ledger. The users were assigned to the graph nodes, and undirected edges created when two users make transactions. The next stage applies a deep autoencoder to the previously built transaction graph to perform graph node embedding. The third step uses a clustering classification model to detect communities within the overall graph, similar to Fast Unfolding. Finally, an outlier detection stage allows the researchers to identify a node belonging to mixing transactions. Altogether, [63] demonstrated the ability to create community structures in bitcoin transaction graphs, detect bitcoin mixing services

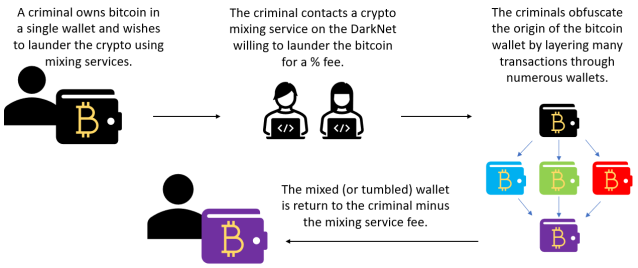

FIGURE 4. Diagram of cryptocurrency mixing process. Phase 1: A criminal owns Bitcoin in a single wallet and wishes to launder the crypto using mixing services. Phase 2: The criminal contacts a crypto mixing service on the DarkNet willing to launder the Bitcoin for a % fee. Phase 3: The criminals obfuscate the origin of the Bitcoin wallet by layering many transactions through numerous wallets. Phase 4: The mixed (or tumbled) wallet is returned to the criminals minus the mixing service fee.

through deep learning methods and also outperform previously implemented heuristic methods attempting to recognize the bitcoin communities. Industrial leaders such as Elliptic have shown methods of tracking wallets through the DarkNet and identifying final wallets after cryptocurrency mixing has been performed.

Deanonymizing Cryptocurrency Blockchains: In order to combat financial cybercrime taking place over cyberspace, it is imperative to be able to identify the controllers of cryptocurrency accounts. As the use of cryptocurrency begins to grow, regulators have put into law the necessity of certain cryptocurrency exchanges to practice AML. This includes implementing Know Your Customer (KYC) being required across the userbase. Reference [64] describes an approach to deanonymizing the Bitcoin blockchain through use of a GCN. Reference [64] also details the challenges faced in tackling large networks such as the Bitcoin blockchain, this includes:

- Large and extremely skewed graphs
- Dynamically increasing graphs
- Semantic graphs

The semantic graphs describe the different activities the blockchain can be used for including smart contracts. Reference [64] show the friendly properties that allow the GCN to favor, this includes:

- Inactive/zero balance address
- Publicly available address labels

The GCN is capable of reducing the size of the graph through identifying inactive/zero balance addresses, with the result of a reduced graph for improved computational embedding speeds. Mentioned above, due to certain exchanges or websites which can require KYC, some wallets identities are already known.

3) CUSTOMS FRAUD

Customs fraud is the evasion of payment for the importing of goods into a country. This is not a crime committed over cyberspace exclusively, however, it is a financial crime that is being tackled using machine learning. The EU anti-fraud office [65] have stated that customs fraud is financially

TABLE 7. Money laundering methods.

METHOD	CHARACTERISTICS
SOCIAL MEDIA DIGITAL MONEY LAUNDERING	With the increasing use of digital currencies created by social media websites, QQ for example, users will begin to use this in lieu of traditional FIAT currency due to in-site benefits. With this increases the likelihood of criminal behavior taking place. This will include the buying and selling of stolen currency and advertise them 'cheap' or discounted, with the criminals exiting the site for FIAT currency usable outside of the social media
FRAUDULENT SOCIAL MEDIA ACCOUNTS	A challenge for social media sites is combating the increasing amount of fraudulent, scam, or disingenuous accounts being created. Some of these accounts are scripted to comment automatically under popular posts with copy-and-pasted content luring genuine users on to their page, inviting them to click on a malicious link, or a link inviting them to a fraudulent transaction or business invitation. This creates a hostile environment for users unaware of account that look legitimate. Tools are being created to examine the usernames, activity, content posted, and profile pictures in order to establish if the accounts are genuine or not.
CRYPTOCURRENCY MIXING	Cryptocurrency mixing is a popular method of laundering cryptocurrency by criminals. DarkNet sites offer services to users to mix or tumble cryptocurrency funds by layering multiple transactions through dozens of wallets to obfuscate the origin. Due to the fact that the entire blockchain is available for the public to view, once the original wallet address is known, it is possible to map transactions in a graph-like manner. Elliptic company have produced multiple blog posts examining ransomware wallet end destination after cryptocurrency mixing. This has spawned new ultra private cryptocurrencies which are more difficult to track such as Monero. Key difficulties in combating cryptocurrency mixing is identifying mixing accounts, and establishing datasets for the use of machine learning models to tackle them.

damaging to legitimate industry and the EU taxpayer. In 2012 and 2013 a combined €110million was recovered from the EU anti-fraud office in customs fraud investigations. Due to the massive transaction volume and manual inspection of goods required in customs fraud investigations, it is not possible without a huge pool of resources to screen every good being imported into a country. There are strides being made in the AI community and their research that is attempting to aid customs fraud officers in the detection of import goods to be inspected. A paper by Kim *et al.* [66] explored a human-in-the-loop customs selection scenario. The data fed into the model are import declaration forms required for customs. The selection model will highlight goods it believes should be inspected by a customs officer, and the officer will then give feedback to the selection model by accepting or rejecting the model's identification of a possible suspect declarations form.

In the case of tax evasion, or specifically customs evasion, the matter of revenue gained by the inspection and identification of evasion is an objective that is taken into account in modeling. Reference [66] the main model used in Kim *et al.* research was a *DATE* model. This is a tree-enhanced dual-attentive model which allows the optimization of dual objectives. The dual objectives in this case were both the illicit transaction classification and the revenue return prediction. The researchers used an exploration strategy and exploitation strategy to search for the best model in the selection process. The exploration strategy is defined as an approach to select uncertain items at the risk of an instant revenue loss, with the potential to detect more novel fraud patterns in the future. The exploitation approach tries to select the most likely fraudulent and highly profitable items to secure the short-term revenue for customs administration. The dataset assembled for this experiment used three African countries transaction level import declarations information. The labels were very accurate including the amount of tariffs charged due to the near 100% inspection rate of imported goods. Numerous hybrid techniques were implemented but the main model used in the selection process was DATE, displaying the highest evaluation performance. Further details on the different models used can be found in the paper.

Another example of deep learning being applied to anti-money laundering and exports was performed Paula *et al.* [67]. The paper applied autoencoders to support fraud investigation in Brazilian exports and anti-money laundering. The use of foreign trade as a method of laundering money takes advantage of the fact that countries have a separation of operations and information, so converting 'dirty' money to 'clean' is an easier task. The results of their work found that the time to perform dimension reduction was twenty times faster using an autoencoder when compared to linear PCA. However, they found the evaluation of the unsupervised technique to be difficult as it required third party experts, in this case tax auditors, which could be a possible barrier for many domains applying unsupervised anomaly detection techniques without the specific expertise for evaluation.

4) TAX EVASION

Tax evasion is the unlawful act of taxpayers purposefully neglecting to pay their tax liabilities to the appropriate revenue commissioner authorities. There is a substantial amount

of research that has taken place in the tax evasion domain with the application of deep learning and machine learning. Including but not exclusive to:

- Missing Trader(MTIC)/VAT Carousel [68]
- Taxpayer Evasion/Tax Auditor Assistant [69], [71]
- Social Media Platform Tax Evasion [70]

Reference [68] investigated Bulgarian taxpayers or traders who were non-compliant in paying their VAT obligations. Experimenting with a dataset made of Bulgarian taxpayers and traders who amount to a total of 312,726 with an average of 75% conducting a single transaction per month, the researchers identified a high % of non-compliant taxpayers/traders within the Bulgarian VAT returns and ledgers. Reference [68] implemented a blended use of graph based theory and machine learning regression model, and identified more high risk subjects who were not originally labeled by the Bulgarian National Revenue Agency.

Both [69], [71] used deep learning graph methods to identify suspicious transactions in taxpayer data. Reference [69] presented a novel tax evasion detection method, T-EGAT or temporal edge enhanced graph attention network. This combines EGAT with recurrent weighted average units (RWA). They showed through a broad comparative analysis of similar deep learning graph methods that T-EGAT performed better at detecting tax evaders than current methods. Outside of the EGAT models tested, the Dynamic GCN, and T-GCN (Temporal-GCN) performed well on the same dataset. Some of the key takeaways from [69] was the ability of the EGAT models to make use of the multi-dimensional edge features between the nodes, a limitation experienced for basic GCNs and GATs. Reference [71] presented TaxAA, a tax auditor assistant helping tax auditors explore and analyze suspicious transactions. The researchers used a dataset consisting of 28,373 taxpayers with a label rate of only 0.071. Two experiments were performed, one using the transaction layer only, and the other used a multi-view TAN (Tax Audit Network). By using extended algorithms on the GCN, a Hierarchical-GCN (H-GCN) was selected through comparative results. Due to some limitations of the GCN such as adding more convolutional layers resulting in an 'over-smoothing' output, the H-GCN is more capable of capturing the global information of the graph due to expansion on the receptive field of each node. To allow this to be a practical tool used by a tax auditor, the researchers created an accompanying visual analytical system to customize suspicious indicators.

5) SIM-SWAPPING, PHISHING, AND SOCIAL ENGINEERING

SIM-Swapping is an attack which allows a cybercriminal to gain unauthorized control of a wireless customer's mobile phone number. This gives an attacker access to the SMS-based text messages which enable resetting of account passwords on websites that rely on the security of a mobile phone number [72]. A successful SIM-Swap attack requires a malicious actor to have the target's phone number, and depending on what account they wish to access, their email

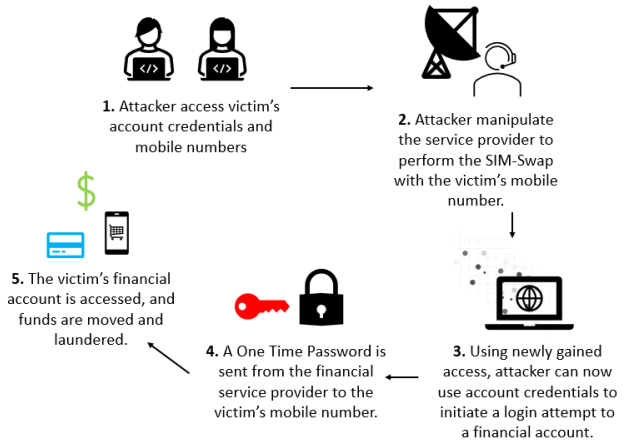

FIGURE 5. Diagram of the SIM-Swap process. Phase 1: Attacker accesses victim's account credentials and mobile numbers. Phase 2: Attacker manipulates the service provider to perform the SIM-Swap with the victim's mobile number. Phase 3: Using newly gained access, attacker can now use account credentials to initiate a login attempt to a financial account. Phase 4: A One Time Password is sent from the financial service provider to the victim's mobile number. Phase 5: The victim's financial account is accessed, and funds are moved and laundered.

as well. The attackers will either contact a victim's service provider and imitate the victim in order to transfer the phone number to a new SIM card, or the attackers have cooperating employees of a service provider which will allow them to an easier route of access. Once the attacker has access to the victim's phone number on their own SIM, they can extract SMS messages, including One Time Passwords sent by financial services such as Coinbase. Figure 5 is an example of how an attacker could gain access to a victim's account and extract funds through a SIM-Swap attack.

There are multiple aspects including phishing, social engineering, and cybersecurity which surround SIM-Swapping, but the main motivation for committing the act has been for the financial gain of the attacker. Highly publicized trials have been covered by the media detailing the theft of millions of euro worth of cryptocurrency through this SIM-Swapping attack. For example, Ortiz [73], and Freeman [74] represent two separate incidents in which their victims have had their Coinbase account hacked through this SIM-Swap attack. Ortiz, of the United States, is reported to have stolen approximately $5million across 40 different SIM-Swap attacks. He targeted high-profile cryptocurrency investors at the Consensus conference. Freeman, of Ireland, was reported to have stolen over $2million from multiple SIM-Swapping attacks. He was also part of a larger online group which worked together in numerous SIM-Swapping attacks.

There has not been definitive papers published in the area of preventing the specific SIM-Swapping attacks performed by cybercriminals on telecommunications service providers. Research has been performed looking at the ease of obtaining mobile phone numbers in the US through various messaging services such as WhatsApp, Signal, and Telegram which is an absolute requirement in the SIM-Swapping scam.

Hagen *et al.* [75] published a paper detailing the use of the contact discovery method to abuse mobile messengers applications and extract networks of mobile phone users and their private data. Mobile messenger applications are able to determine contacts in a user's address book that are registered with the messaging service through the procedure contact discovery. This is used widely across social media platforms. Through a combination of crawling and hash reversal attacks [75] over a fixed amount of time and limited resources, the researchers were capable of obtaining 100% of US mobile phone numbers for Signal, 10% for WhatsApp, and reveal the Telegram API weaknesses revealing a wide range of sensitive information. These types of security and privacy holes in messaging applications are allowing potential attackers access and identification of victims for the SIM-Swapping schemes.

Phishing is considered a social engineering technique with the interest of luring victims to unwillingly hand over their personal information including passwords, email addresses, phone numbers, addresses, usernames and financial information [23]. Examining Table 5, phishing and its variants are reported to have stolen over $48million from victims of the United States in 2018 alone. Quite clearly, this activity is a lucrative market for cybercriminals.

Recent research has been published by Lansley *et al.* [76] developed a tool called ''SEADer++'', a social engineering attack detection in online environments using machine learning. The system attempts to detect social engineering attacks based on NLP and artificial neural networks. The researchers have based their concept on attempting to detect a social engineering attack in an online chat environment. The proposed process has three steps which are data preprocessing, feature extraction, and aggregation of results. The aggregation of results includes a classification technique to identify the social engineering attempts. The researchers have reported high accuracy results in classification using decision tree, random forest, and an MLP. Based on the AUC results, the proposed soft-voting Ensemble Learning method is decided as the best solution for industrial application.

Understanding the psychology behind social engineering and its ability to manipulate people can help in the tackling of this field. As seen in the ''SEADer++'' tool [76], a feature used was principles of persuasion. Montañez *et al.* [77] produced a paper titled ''Human Cognition Through the Lens of Social Engineering Cyberattacks''. They advocate treating social engineering cyberattacks as a psychological attack and wish to propose an extension of the standard framework of human cognition to recognize and accommodate social engineering cyberattacks. This framework by the researchers has led to a quantitative representation for characterizing persuasion mathematically. This work demonstrates the range of science and research required to effectively combat social engineering cyberattacks on industry and the civilian population.

More research is to be performed on analyzing and preventing social engineering attacks. It can be accepted that financial cybercrime requires an element of social engineering when attempting to gain access to private information of customers on financial platforms. SIM-Swapping is just one example this paper has looked at. In time, it can be expected that machine learning will have a greater impact on the prevention of social engineering and boosting the security of customers account across financial and telecommunications platforms.

6) ROMANCE FRAUD

Romance fraud as defined by the FBI [78] as a scam that occurs when a criminal adopts a fake online identity to gain a victim's affection and trust. The scammers will use that trust to build up an illusion of romance or close relationship and manipulate victims with the ambition of illegal financial gain. An example of the romance fraud process can be seen in Figure 6.

This fraud has seen a rise in popularity with scammers particularly through the global lockdown due to COVID-19 with reports of up to 20% increase in bank transfer fraud linked to romance scams in 2020 when compared with 2019 [79]. When examining Table 5, romance fraud has reportedly been responsible for the theft of over $362 million US dollars alone in 2018. Not only are victims scammed from their own money but can be used as money laundering mules unassumingly by being asked to transfer received money from the criminal to various accounts the criminal will instruct.

There is little publication on the tackling of romance fraud using machine learning techniques. Reference [80] published research attempting to automatically dismantle online dating fraud. The researchers deployed an ensemble classifier method combining the predictions of multiple SVM or RF classifier outputs in order to identify whether a profile was fraudulent or not on a dating website. Reference [80] produced a system which can accurately detect online dating fraud or scam profiles with high precision, but the results included a number of false negative classifications due to real profiles having very similar traits as scam ones.

7) RANSOMWARE

Reference [21] ransomware is a form of malware that has the ability to encrypt victim's computer systems and digital information, prohibiting access to it until a ransom is paid to the attackers. Malware is *malicious software*, it is created with an intent for criminality to gain access undetected into the computer systems of its victims. There are various forms of malware including Trojan horses, rootkits, and viruses [21]. Typical payment demanded by the criminals is in the form of cryptocurrency due to the anonymity surrounding the owner of wallets.

Ransomware is a more sophisticated method used by financial cybercriminals. It is growing in popularity as seen in recent attacks in 2021 with the Colonial Pipeline in the US [81], the Irish Health Service Executive [82], and the highest ever ransomware attack on Acer [83] with a demand of $50million dollars from the cybercriminals. It is reported

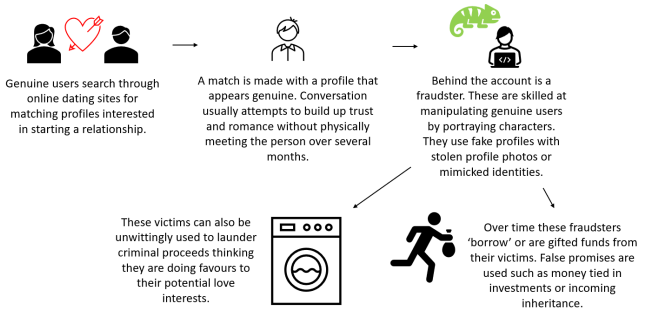

FIGURE 6. Diagram of romance fraud example. Phase 1: Genuine users search through online dating sites for matching profiles interested in starting a relationship. Phase 2: A match is made with a profile that appears genuine. Conversation usually attempts to build up trust and romance without physically meeting the person over several months. Phase 3: Behind the account is a fraudster. These are skilled at manipulating genuine users by portraying characters. They use fake profiles with stole profile photos or mimicked identities. Phase 4 (Option 1): Victims can be used unwittingly as launderers, cleaning criminals proceeds thinking they are doing favors to their potential love interests by performing illegal transactions. Phase 4 (Option 2): Over time the fraudsters 'borrow' or are gifted funds from their victims. False promises are used such as money tied in investments or incoming inheritance to which they will repay their victims.

in the United States that approximately 50% to 75% of ransomware victims are small businesses [84].

There is extensive literature on the use of machine learning in combating malware and analyzing it. One such example by [85] proposed an ML method to detecting ransomware. Their machine learning algorithm, PEDA (Pre-Encryption Detection Algorithm), is possible to detect ransomware within Windows environments before it is activated and locks the victims out of their systems. Other industrial leaders in cryptocurrency anti-money laundering and compliance like Elliptic [86] have posted blog entries covering ransomware groups like REvil [87]. Their forensics teams have shown through graph analysis the end destination of the ransomware payments by following the blockchain.

Ransomware shows a combination of the previously discussed methods in financial cybercrime. It is a blend of cryptocurrency money laundering and social engineering in order to hack into and access corporations, businesses and institutions. Preventative measures for combating ransomware includes improved cybersecurity for potential victims, and reactive measures like those seen in the Elliptic blog articles where the money can be followed and identify these criminal organizations.

8) DEEPFAKES AND GPT-2

With the development of ML models to counteract financial cybercrime, so has the advancement of the attackers. Deepfakes and advanced chat bots like GPT-2 are capable of spoofing and manipulating staff at all levels of an organization. In March 2019, the CEO of a UK Based energy company believed he was speaking to his senior, the CEO of the German based parent company. This was in fact a sophisticated deepfake model deployed to socially engineer the UK-based CEO in transferring approximately €250,000 to criminals [88]. Deepfakes are not only audio manipulation but also visual. Deepfake programs are capable of creating completely fictitious identities of individuals. Websites such as [89] uses a Generative Adversarial Network to create a 'person' or even generate modified images of a person without their consent, i.e., deepfake celebrity pornography [90]. These images can also be used in online profiles which can spoof genuine users of websites such as dating sites or social media vendor sites. Research by [91] takes aim at non-authorized deepfakes. By using adapted adversarial attacks, they disrupt the conditional image translation facial manipulation networks. With regards to financial cybercrime and deepfakes, research published by [92] examined the camouflaging of accounting journal entries using deepfakes in a direct attack against the current ML models deployed by auditors. This form of research is regarded as adversarial attacks, these attacks are deliberately designed to exploit vulnerabilities in ML models and cause them to make a mistake such as misclassification [92]. They successfully spoofed existing auditing ML models using an adversarial autoencoder to unwind the entries underlying the latent generative factors. This shows the possibility of adversarial ML models bypassing what regulators and auditors believe to be state-of-the-art auditing models, possibly covering fraud and money laundering offenses.

GPT-2, an open-AI chat bot which is trained to predict the next word in a sentence and has shown it can produce human-like passages of text such as news articles [93]. GPT-2 has been used to create false reviews for vendor websites such as Amazon. False reviews have the ability of fooling genuine customers into transacting with either illegitimate suppliers or low-quality goods manufacturers or damaging a rival business's total review score and reputation. Manually doing this across vendor websites is a method known as ''crowdturfing'' and is considered an attack on online review systems. A developed AI method by [94] implemented the GPT-2 system to create a bundle of false reviews which were not distinguishable against genuine reviews. The researchers also showed through subjective and automated fake review evaluation that the fake reviews were just as influential as the genuine.

IV. PERFORMANCE, DRAWBACKS, AND PERSONAS

A. METRICS AND PERFORMANCE

In this section, we examine the algorithm performance of the papers surveyed in the literature, personas involved in financial cybercrime, and associated drawbacks. This section has been split into three subsections which cover the following:

- **Metrics and Performance**: We examine the performance and output of the models presented in the survey. We highlight the best performing models and methodology in the published research.
- **Personas and stakeholders**: This subsection discusses the users and data sources involved in financial

cybercrime. This includes both the malicious and genuine actors in the network. Here we discuss the security, privacy, explainability, and threat models which have an impact on the financial cybercrime community.
- **Drawbacks**: We comment on the drawbacks experienced by researchers in the field of financial cybercrime, taking examples from the published material discussed.

One of the main key insights is the importance of preprocessing techniques and additional contextual information that may result in high performance gains. For example, InfDetect's [49] Information Gain analysis performed on their features showed the high impact of the graph output. This reinforces the use of group anomaly detection techniques and approaching certain financial cybercrime as attempting to identify micro-clusters of bad actors using ML and DL techniques like Fast Unfolding [50]. Similarly, researchers used autoencoders to denoise their data for preprocessing and also dimension reduction techniques. Both [49] and [67] applied autoencoders in their preprocessing stage, with [67] citing that the autoencoder was 20 times faster than PCA for time complexity. The use of Social Network Analysis capturing node information and importance has shown promise in anti-fraud research with both [51] and [52] produces results of identifying anomalous accounts or transactions without the use of any machine learning or deep learning implementation.

The various forms of evaluation include accuracy, precision, recall, True Positive Rate (TPR), False Positive Rate (FPR), Area Under the Curve (AUC), and node centrality. We explain the evaluation methods used and provide the mathematical equations.

Accuracy is a method of evaluating the performance of a binary classification model distinguishing between True or False. Accuracy is determined by summing the correctly classified outputs (True Positive and True Negative) and dividing by the total number of classifications which is True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Equation 1 below represents the accuracy calculation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Recall is considered a percentage of ground-truth anomalies, which have been classified as anomalies at a given threshold [8]. Equation 2 calculates the recall metric.

$$Recall = \frac{TP}{TP + FP} \tag{2}$$

Precision is considered to be the percentage of reported anomalies, which are actually true anomalies [8]. Equation 3 calculates the precision metric.

$$Precision = \frac{TP}{TP + FN} \tag{3}$$

The True Positive Rate (TPR) is defined in the same way as Recall. The False Positive Rate (FPR) is the percentage of falsely classified positives from the true negatives. It is represented in Equation 4.

$$FPR = \frac{FP}{FP + TN} \tag{4}$$

The Area Under the Curve (AUC) or Area under ROC curve, where ROC is the receiver operating characteristic, represents a graphical plot of the TPR against the FPR [95]. An AUC with a value of 0.5 means the model is no better at classifying a model than randomly guessing. As the AUC approaches a higher percentage such as 90% or 0.90, this means the model is correctly classifying the test data. Therefore, the AUC is a metric which represents the model's performance at distinguishing between two classes [96].

The advantage to using AUC over other evaluation metrics like accuracy is because the AUC decouples classifier performance from class skew and error costs [95].

In networks and graphs, common evaluation methods includes node and degree centrality. These are subsets of Social Network Analysis (SNA) [97]. [98] Centrality analysis is used to identify key actors or nodes in a graph, metrics like closeness, degree, and betweenness are some examples. [52] Centrality can be used to measure who is the most important person in a network. Below we present the equations to calculate betweenness centrality, closeness centrality, and degree centrality covered in [99].

Betweenness centrality measure is a way to examine actors positional advantage or influence or power, where an actor falls on the shortest pathway between other pairs of actors. It is represented in Equation 5.

$$C_B(v) = \sum_{s \notin v \notin t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{5}$$

where σ_{st} represents the number of shortest paths, and $\sigma_{st}(v)$ is the number of shortest paths from s to t containing the vertex v.

Closeness centrality allows us to examine how close a node is to other nodes, and allows closer examination into how quickly they can spread information, or how influential they are. It is represented in Equation 6.

$$C_C(v) = \frac{1}{\sum_{u \notin v \in V} g_{vu}} \tag{6}$$

where g_{vu} is the length of the shortest path from v to u.

Degree centrality is simply the number of edges connected to that node. It is represented in Equation 7.

$$C_D(v) = deg_{(v)} \tag{7}$$

Table 8 summarises different algorithms and their performance metrics in financial crime domain.

B. DRAWBACKS

A key problem experienced across several papers is the access to labeled data to train and evaluate their model performance. This is noticeable in cases of cryptocurrency money laundering, the Elliptic bitcoin dataset seen in [60] has labeled only 2% of their nodes as illicit, and 21% as licit. Work

TABLE 8. Summary of algorithms and their performance in financial crime domain.

Financial Cybercrime Method	Paper	Method	Performance or Method of Evaluation
Stocks & Securities Investment Fraud - Price Manipulation	[43]	RNN-Ensemble	AUC = 0.907, F1 = 0.47, Precision = 0.317, Recall = 0.902
Stocks & Securities Investment Fraud - Insider Trading	[44]	LSTM-RNN & ANOMALOUS	The researchers used Normalized Cross Correlation (NCC) to compare the discrete signals produced from the ANOMALOUS algorithm.
Stocks & Securities Investment Fraud - Pump & Dump	[45]	Hypergraphs	The researchers used the signed normalized dollar amount. This is used to compare the reported price of a transaction with the market closing price of the company's stock on the same day as the transaction. This value lies between -1 and 1.
Stocks & Securities Investment Fraud - Pump & Dump	[47]	SVM/Random Forest	Average AUC of 1) predicting pump attempts = 0.75 2) predicting pump success 0.71
Fraud Detection & AML - Social Network Digital Currency	[49]	SVM	AUC = 0.97, FPR = 0.0097, Detection Rate = 0.94
Fraud Detection & AML - Account Network Analysis	[51]	Social Network Analysis (SNA)	Correlation coefficients used to infer risk profiles of companies along with node centrality.
Fraud Detection & AML - Account Network Analysis	[52]	SNA, Clustering, Frequent Pattern Mining	Node centrality key metric used for inference.
Fraud Detection & AML - Insurance Fraud	[53]	InfDetect	AUC = 0.97
Fraud Detection & AML - Credit Card Fraud	[54]	BreachRadar	>90% Precision and Recall when 10% of stolen cards have been used in fraud
Fraud Detection & AML - Spam Bot Detection	[55]	EagleMine	Sina Weibo Dataset: AUC for examining fraudulent messages = 0.83, AUC for fraudulent user = 0.68
Fraud Detection & AML - Cryptocurrency AML	[56]	GCNs	Not Provided
Fraud Detection & AML - Cryptocurrency AML	[60]	GCN & Skip-GCN	GCN F1 Score = 0.628, Precision = 0.812, Recall = 0.512; Skip-GCN F1 Score = 0.705, Precision = 0.812, Recall = 0.623
Fraud Detection & AML - Cryptocurrency AML	[62]	GCN & MLP	F1 Score = 0.773, Precision = 0.899, Recall = 0.678. Accuracy = 0.974
Fraud Detection & AML - Cryptocurrency Mixing	[63]	Autoencoder	F1 Score = 0.208
Customs Fraud	[66]	DATE	Two metrics used to evaluate the dual objective performance. Pre@n% which explains how many transactions are illicit, and Rev@n% which explains how much customs duties can be generated from the top n% of transactions than the revenue generated by inspecting all transactions. [66] show in their work that the hybrid strategy yielded the best performance.
Customs Fraud	[67]	Autoencoder	Researchers faced difficulty in evaluating results; with expert help confirming at least twenty fraudulent exporters.
Tax Evasion - Chinese Taxpayers	[69]	T-EGAT	Accuracy = 0.939, Precision = 0.893, Recall = 0.922, F1 = 0.907, AUC = 0.915
Tax Evasion - Audit Assistant	[70]	TaxAA (Hierarchical-GCN)	Transaction layer experiment: Accuracy = 0.85, F1 = 0.82, AUC = 0.915; Multi-View TAN: Accuracy = 0.94, F1 = 0.87
Tax Evasion - VAT Evasion	[68]	Network Analysis and GBT	Low level information of results; False Negative Rate of 10% and True Positive Rate of 90%
Romance Fraud - Dating Site Profile Analysis	[80]	Ensemble Approach	Accuracy = 0.97, F1 = 0.945
Social Engineering - Attack Detection	[76]	SEADer++	AUC = 0.722, Accuracy = 0.924, F1 = 0.759

performed on the same dataset by [62] noted that the dataset has withheld proprietary information regarding the features, and the real-time associated with time stamps which would be a beneficial addition for future researchers. The requirement of subject matter experts to review data leads to a necessity for wider collaboration with industry and law enforcement specialists. [67] identified potential customs cases that without ground truths were impossible to confirm whether they were truly outliers and fraudulent or just misclassified by their models. Similarly, [68] identified potentially fraudulent taxpayers not labeled by the National Revenue Agency of Bulgaria, the researchers had to consult subject matter experts with the access and ability to assess the taxpayers appropriately. Analyzing QQ for malicious accounts also required manual labeling of thousands of accounts [49].

Sub-graphs are key methods discussed in money laundering identification and promising tools for group anomaly detection. However, a limitation to dense sub-graph-based techniques is that they focus on single-step transfers [100]. This means the sub-graph based methods require specific adjustments for AML modeling.

The scope of researchers' tasks can encompass a large area of examination in order to identify criminality. This includes analyzing social media platforms performed by [47] and comparing that against the correct asset prices. The researchers required cryptocurrency market price data, Telegram messaging information, and Twitter data. The securities and stock investment fraud research is difficult to tackle with the lack of publicly available datasets for readily available testing on ML and DL models. Researchers investigating price manipulation [43], and insider trading [44] require numerous sources of unstructured and unlabeled information from governmental agencies, causing a large and arduous task for preprocessing.

Interpretability and explainability of the model and its outputs can pose as a challenge for researchers and industry partners. Key aspects of anomaly detection include the ability to interpret the results of a model and understand why something is considered an outlier. This comprehension can lead to constructing rules within an organization for example, particularly in tax that if a taxpayer is seen to be flagged breaking a particular logical rule, the auditors can increase the risk rating for fraudulent activity on the suspected individual's profile. Using DL and ML models can hinder this as seen in deep learning model outputs where unsupervised models such as an autoencoder simply classifies an account as non-compliant. An example from [67] classifies an account as potentially fraudulent but without context around that classification, such as information gain on features, future work on improving subject matter experts understanding of the fraud is lost.

As GNNs are an area being explored to tackle financial cybercrime, we examine several drawbacks to using GNNs. These drawbacks are discussed thoroughly in the GNN survey performed by Wu *et al.* [15] which include:

- **Model Depth**: An issue with the GCN model is the restrictive architecture in the layers. GCNs are incapable of having many hidden layers as each layer passthrough performs aggregation on the node information and converges the nodes to a single point. This has been circumvented with GCN variants such as the Hierarchical-GCN shown in [71]. Also, cryptocurrency money laundering research following the work by [60] focused more on the feature extract concatenation from the GCN model to then pass this through an MLP [62]. The output of their research shows promise in implementing GCN architecture in conjunction with other existing methods.
- **Dynamicity**: Graphs tend to evolve over time. Particularly in the case of financial cybercrime, nodes which could represent accounts for example, can be closed or become inactive. Temporal changes will cause the graph to contract and expand, Spatial-Temporal Graph Neural Networks (STGNNs) have been created to address this by factoring space and time in the modeling architecture. Such advancements are seen in [98] who use STGNNs to model criminal networks over time.
- **Heterogeneity**: Many GNN methods depend on the graph being homogeneous. A homogeneous graph is one where the nodes and edges represent the same information across the graph. A heterogeneous graph is one where the nodes and edges can represent different things. For example, nodes in a heterogeneous graph can be represented as buildings or people, and the edges could be money transferred or traveled to. A knowledge graph would be another example of a heterogeneous graph. Relational-GCNs have been created to tackle this issue such as one developed by [101] coined the Knowledge Embedding Based Graph Convolutional Network.

C. PERSONAS

The personas of financial cybercrime take a look into the various actors and data sources which appear across the sector. They are:

- *Fraud Analysts/Investigators*: The end users for many models are the financial, governmental, or corporate analysts. Taking the research performed by [66], the human-in-the-loop method involves a subject matter expert verifying the results of the customs fraud ML model. The ideal end result for many ML practitioners is to aid analysts and the overall business model in their day-to-day activities in financial cybercrime prevention and identification. To enhance the user experience and education for analysts it should be factored in for researchers' ML models to include explainability of their model outputs. XAI (Explainable Artificial Intelligence) is becoming an increasingly more important field of research, particularly when it comes to complying to the regulatory standards of the required transparency around processing customer data. Another factor to be considered is the time benefit or drawback. Human-in-the-loop algorithms are created with a goal to improve the productivity of an analyst but a complicated and encumbering procedure for updating data systems or encountering issues which need to pull in other teams can damage an analyst's valuable time.
- *Fraudulent/Malicious Users*: The spectrum of technical ability for financial cybercriminals varies widely as seen across the range of methods used to illegally gain economically. A comprehensive study on the various identities of cybercriminals and their practices including their categories and classes is well described in [37]. The researchers also discuss some of the motivations of cybercriminals, although it is difficult to classify each individually. Some of the motives include curiosity, manipulation, political, ego gratification, notoriety, and more obviously, financial benefit.
- *Genuine Users*: Customers, employees, employers, and any authentic user of a service provided by a corporation, government or financial institution represent the body of genuine users. This body of people are also the victims of financial cybercrime. A study by [36] analyzed and surveyed victim's perceptions of Virtual Property Theft (VPT). They concluded that despite users' knowledge and awareness of potential VPT, 23% still become victims. This shows genuine users are still susceptible to financial cybercrime as the cybercriminals have found loopholes in the security of the systems used, and their ability to manipulate their victims.
- *Auditors, Regulators, and Governing Bodies*: Law enforcement agencies, auditors, and regulators govern much of the financial landscape where the analysts, genuine and malicious users operate. The regulators in the financial cyberspace include central banks, national data protection commissioners, and national or international

governing bodies like the European Commission. Auditors are positioned within companies both internally and externally to give a degree of confidence for investors and regulators regarding the published financial accounts. As noted in [92], auditors are in a position to be targeted using adversarial DL models to force misclassification of fraudulent accounting entries by attacking the very models designed for the auditors to use.

- *Data Sources*: Data sources and datasets across the reviewed literature vary between public datasets, which allow comparative experiments to be performed by fellow researchers, and private industry held data which only allow for replication of the algorithm if it is provided appropriately. Industrial companies do release datasets although not all the proprietary information is there, hindering full research capabilities from the community. For example, [60] Elliptic's Bitcoin dataset on Kaggle has allowed for research to be published from experiments ran on its dataset, but the columns are not labeled. The same group [56] published information on their ALMSim tool which will create synthetic datasets allowing money laundering research to be completed and compared across literature. The disadvantage for all researchers comparing their algorithms against public datasets is whether the datasets capture the latest techniques being deployed by financial cybercriminals and whether the techniques are truly effective. Thus, industrial datasets hold much more weight especially with monetary value being published to back up the effectiveness of the ML or DL models as seen in [53]. Table 9 displays the datasets suitable for AD with financial cybercrime, and some publicly available datasets that were presented in the reviewed literature.

V. CHALLENGES AND FUTURE RESEARCH DIRECTION

Tackling financial cybercrime is not an easy task. Financial cybercriminals are sophisticated in their methods of attack, and cunning in their social engineering ability. Although the research is advancing as seen in this survey, it is an ever-evolving battle of being in a defensive landscape for cybersecurity professionals. Below we highlight the challenges and potential future research areas to explore.

- *Latency Sensitive Applications*: Creating practical and effective machine learning or deep learning techniques requires not only accuracy in the predictions but also speed. Particularly in the finance domain customers expect transactions to be delivered seamlessly to their respective payees. The ambition of deploying a real-world application within the industry requires an ability to update the underlying datasets with response to new transactions in real-time. Reference [56] sparse dynamic recomputation can improve the performance of algorithms, and could be seen incorporated into future research in financial cybercrime.
- *Anonymity*: The advent of cryptocurrency in combination with the use of the DarkNet and associated IP masking tools such as VPNs make the task of identifying financial cybercriminals difficult. As discussed above, [63] have modeled methods to detect cryptocurrency mixing accounts which are key in the obfuscation stage of laundering cryptocurrency. Key area of investigation for laundering cryptocurrency is the exit strategy used by criminals. To be able to convert the cryptocurrency into a more favorable FIAT currency a number of methods exist. These methods include using DarkNet marketplace vendors who accept cryptocurrency transactions in exchange for FIAT currency which will be transported physically through mail systems. Other methods include the more commonly used exchanges such as Coinbase and exchange marketplaces on the DarkNet such as Hydra for gift cards and vouchers to be used to purchase tangible goods. Reference [104] Hydra, a Russian-based DarkNet marketplace, received over $1.4 billion dollars worth of bitcoin in 2020. In exchange for bitcoin, a user will receive prepaid debit cards, or FIAT currency deposited into a nominated bank account. The commonality across the methods is the movement of cryptocurrency from one wallet to another, thereby logging a traceable transaction in the blockchain. Identifying these accounts/transactions are a step in the right direction for combating the methods in un-masking fraud activity. More specifically, the use of Group/Graph-based Anomaly Detection is fundamentally necessary to tackle community structures within networks.
- *Robustness and Adversarial Attacks*: Creating algorithms designed to combat criminals who purposely obfuscate their actions requires robustness. This robustness of a model can be identified and measured through its weaknesses. A challenge of designing a method of preventing criminal activity is also maintaining it to update along with the criminal activities and sophisticated techniques. To address this, future research will include adversarial attacks. These attacks will try and penetrate the model in an attempt to elude any detection by the algorithms as seen in the research performed by [92]. Through these tests, weaknesses can be identified and then addressed through further model experimentation or re-evaluation of the algorithm in its entirety. Included in the robustness of a model is its ability to adapt to changes in both data and feature drifts. These are common in the financial domain due to changes in economic trends, the spending habits of customers, and the introduction of modern technologies.
- *Graph Construction*: [31] Due to the heterogenic properties of cryptocurrencies, it is challenging to model them in a graph form. Ethereum and Bitcoin have different blockchain structures, where Ethereum is capable of incorporating contracts into their transactions, while Bitcoin is a simpler transaction method but there are

TABLE 9. Financial cybercrime datasets.

Paper or Source	Details
[47]	Researchers interested in identifying and analyzing cryptocurrency manipulation in social media can access the dataset that includes Telegram, Twitter, and cryptocurrency market price data.
[58]	AMLSim was created by IBM researchers to allow the research community to assess the performance of their models on synthetically generated data.
[60]	One of the leading providers of cryptocurrency compliance solutions, Elliptic, has provided a Bitcoin transaction dataset. This dataset allows for the classification of already labeled licit and illicit nodes in a graph. The graph is made of 203,769 nodes and 234,355 edges with 21% of the graph labeled licit and 2% labeled illicit.
[80]	Researchers interested in tackling romance fraud can access scripts which scrape data from a public scam list and dating site. The task is identifying online dating fraud profiles using their accessible profile elements.
[102]	ODDS: Outlier Detection DataSets provide a large collection of anomaly detection datasets with ground truth labels (if present). They have branches of AD datasets including multi-dimensional point datasets, time series graph datasets for event detection, time series point datasets, and adversarial/attack scenario and security datasets.
[103]	For extended assessment on the performance of a newly created GNN, a benchmarking framework has been created for researchers to evaluate the performance of their GNNs on medium-scale datasets.

still multiple ways to represent the nodes and edges for the graph construction stage. These various forms of representation require the user to build a specific graph to capture the transaction information necessary for an end task such as money laundering or perhaps liquidity analysis of outstanding contracts on a network.

- *Uncertain and Incomplete data at decision time*: [31] Financial fraud detection methods have to make decisions based on incomplete and uncertain data. As ground truth labels are sought through manual review by analysts, an updated fraud model could be obsolete due to the delay in the real-time availability of rich incident data. There is a challenge in having ML systems adapt and make decisions on incomplete data. For example, in a typical AML use case, not every observation and latent relationships are available at decision time. This makes the algorithm design and evaluation a challenge. Self-supervision methods have potentially massive impacts on the future of research in tackling financial cybercrime. Self-supervised graph learning approaches can allow us to understand and forecast events with no priors. There is a potential use case by adapting DeepMind's traffic prediction that uses advanced graph neural networks and transforms it for financial networks to identify malicious users [105].
- *Human In the Loop Evaluation*: Without ground truth labels, certain evaluation requires expertise to correctly identify whether a given transaction is fraud or not. As seen in [66], a human-in-the-loop method is required to absolutely determine the outcome of the prediction. Especially when an accusation of theft or deception has been cast at a user or customer.
- *Regulations*: FINRA (Financial Industry Regulatory Authority) released a white paper [106] discussing the potential regulatory considerations towards the security market participants who wish to develop and deploy AI models. Reference [106] FINRA discuss four key challenges around model explainability, data integrity, customer privacy, and model risk management. With regards to model explainability, FINRA Rule 3110 (Supervision) states a requirement for "*firms to establish and maintain a system to supervise the activities of its associated persons that is reasonably designed to achieve compliance with the applicable securities laws and regulations and FINRA Rules*". This means firms who are deploying AI models within financial intuitions will seek to understand how outputs are derived, particularly with deep learning or "black box" models so they are in line with legal and compliance regulations. Similarly, data bias is a key discussion point. FINRA Rule 2010 requires firms to "*observe high standards of commercial honor and just and equitable principles of trade*". Data bias can result in incorrect outputs due to a lack of relevant data, or skewed subsets of information. Customer privacy is protected by SEC and GDPR regulations which requires data controllers and processors to take appropriate steps to ensure the security of the data is correct and that customer consent is received for processing particular tasks. With these various regulatory requirements comes a phase of building algorithms with "privacy by design" as an architectural foundation. Explainability is the next step of algorithm architecture by understanding which features are key to a model input, and why certain outputs are calculated. A survey by [107] examined XAI (explainable artificial intelligence) and its role in the future of ML and DL.

VI. CONCLUSION

Examining the state-of-the-art algorithms, models, and techniques used to tackle the various facets of financial cybercrime, it can be accepted that it is not a trivial task. The manner of obfuscation, manipulation, and masking of behavior creates a difficult job for researchers and industry to identify, prevent, and detect malicious illegal activity. As seen in the above literature, group anomaly detection, deep learning, and graph theory are being combined to identify networks of malicious actors within overall groups of users and customers. With large sums of money being extracted

from the financial system, there is a penalty being paid by the public through increased fees and mistrust of their private information being held by companies. It can be argued that financial cybercrime has an impact on our society at a sociological level, particularly in the area of money laundering and tax evasion. The lack of consequence or retribution for criminal activity can upset and cause discohesion within a society's view of its policing and tax agenda.

The models presented in this review compound the necessity for use of detection techniques based on Graph/Group based Anomaly Detection to combat financial crime. The authors accept the challenge particularly in obtaining labeled datasets, and the expertise required in labeling ground truths where one is not already available. With the advancement of cryptocurrency and its deepening entrenchment into the financial ether, it is not surprising that anti-money laundering in cryptocurrency research has been initiated. The authors suspect a closer examination of cryptocurrency and its integration into the public domain by the respective Revenue Commissioners and law enforcement authorities of the varying countries worldwide, resulting in increased output of research, particularly in the Group/Graph based Anomaly Detection domain.

ACKNOWLEDGMENT

This publication emanated from research conducted with the financial support of Science Foundation Ireland under grant number 18/CRT/6183. For the purpose of open access, the author has applied a cc by public copyright licence to any author accepted manuscript version arising from this submission.

REFERENCES

[1] *Economic Crime*. Accessed: Jul. 19, 2021. [Online]. Available: https://www.europol.europa.eu/crime-areas-and-trends/crime-areas/economic-crime

[2] (2017). *LexisNexis Risk Solutions 2017 True Cost of Fraud*. [Online]. Available: https://risk.lexisnexis.com/-/media/files/financial%20services/research/2018-true-cost-of-fraud-overall-rep%20pdf.pdf?la=en-us

[3] S. Hasham, S. Joshi, and D. Mikkelsen, *Financial Crime and Fraud in the Age of Cybersecurity*. Shanghai, China: McKinsey & Company, 2019, pp. 1–11.

[4] *Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case*. [Online]. Available: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

[5] (2019). *Fighting Financial Crime With AI*. [Online]. Available: https://www.ibm.com/downloads/cas/WKLQKD3W

[6] A. Feroze, A. Daud, T. Amjad, and M. K. Hayat, "Group anomaly detection: Past notions, present insights, and future prospects," *Social Netw. Comput. Sci.*, vol. 2, no. 3, May 2021, Art. no. 219, doi: 10.1007/s42979-021-00603-x.

[7] A. Kuppa, S. Grzonkowski, M. R. Asghar, and N.-A. Le-Khac, "Finding rats in cats: Detecting stealthy attacks using group anomaly detection," in *Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./13th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2019, pp. 442–449.

[8] C. C. Aggarwal, *Outlier Analysis*. Cham, Switzerland: Springer, 2017.

[9] S. Choi, C. Kim, Y.-S. Kang, and S. Youm, "Human behavioral pattern analysis-based anomaly detection system in residential space," *J. Supercomput.*, vol. 77, no. 8, pp. 9248–9265, Aug. 2021, doi: 10.1007/s11227-021-03641-7.

[10] H. Darvishi, D. Ciuonzo, E. R. Eide, and P. S. Rossi, "Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture," *IEEE Sensors J.*, vol. 21, no. 4, pp. 4827–4838, Feb. 2021.

[11] E. Toth and S. Chawla, "Group deviation detection methods: A survey," *ACM Comput. Surv.*, vol. 51, no. 4, pp. 1–38, Sep. 2018.

[12] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," pp. 1–50, 2019, *arXiv:1901.03407*.

[13] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognit.*, vol. 58, pp. 121–134, Oct. 2016, doi: 10.1016/j.patcog.2016.03.028.

[14] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," pp. 1–22, 2018, *arXiv:1812.08434*.

[15] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.

[16] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. 2nd Int. Conf. Learn. Represent. (ICLR)*, 2014, pp. 1–14.

[17] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," 2015, *arXiv:1511.05644*.

[18] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Dec. 2009.

[19] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. 5th Int. Conf. Learn. Represent. (ICLR)*, 2017, pp. 1–14.

[20] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph WaveNet for deep spatial-temporal graph modeling," 2019, *arXiv:1906.00121*.

[21] W. Z. A. Zakaria, M. F. Abdollah, O. Mohd, and A. F. M. Ariffin, "The rise of ransomware," in *Proc. ACM Int. Conf.*, 2017, pp. 66–70.

[22] E. Carter, "Distort, extort, deceive and exploit: Exploring the inner workings of a romance fraud," *Brit. J. Criminol.*, vol. 61, no. 2, pp. 283–302, Feb. 2021.

[23] R. Alabdan, "Phishing attacks survey: Types, vectors, and technical approaches," *Future Internet*, vol. 12, pp. 1–39, Oct. 2020.

[24] V. Todorovic and M. Jaksic, *Misuses of Electronic Payment Systems*, vol. 4. Kragujevac, Siberia, 2016.

[25] A. Tamersoy, E. Khalil, B. Xie, S. L. Lenkey, B. R. Routledge, D. H. Chau, and S. B. Navathe, "Large-scale insider trading analysis: Patterns and discoveries," *Social Netw. Anal. Mining*, vol. 4, no. 1, pp. 1–17, Dec. 2014.

[26] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Comput. Sci. Rev.*, vol. 40, May 2021, Art. no. 100402.

[27] A. Sharma and P. K. Panigrahi, "A review of financial accounting fraud detection based on data mining techniques," *Int. J. Comput. Appl.*, vol. 39, no. 1, pp. 37–47, Feb. 2012.

[28] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q. Z. Sheng, H. Xiong, and L. Akoglu, "A comprehensive survey on graph anomaly detection with deep learning," Aug. 2021, *arXiv:2106.07178*.

[29] D. Ucci, L. Aniello, and R. Baldoni, "Survey of machine learning techniques for malware analysis," *Comput. Secur.*, vol. 81, pp. 123–147, Mar. 2019, doi: 10.1016/j.cose.2018.11.001.

[30] A. M. Ozbayoglu, M. U. Gudelek, and O. B. Sezer, "Deep learning for financial applications : A survey," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106384, doi: 10.1016/j.asoc.2020.106384.

[31] J. Wu, J. Liu, Y. Zhao, and Z. Zheng, "Analysis of cryptocurrency transactions from a network perspective: An overview," *J. Netw. Comput. Appl.*, vol. 190, Sep. 2021, Art. no. 103139.

[32] R. Saia and S. Carta, "Evaluating the benefits of using proactive transformed-domain-based techniques in fraud detection tasks," *Future Gener. Comput. Syst.*, vol. 93, pp. 18–32, Apr. 2019, doi: 10.1016/j.future.2018.10.016.

[33] A. H. Alhazmi and N. Aljehane, "A survey of credit card fraud detection use machine learning," in *Proc. Int. Conf. Comput. Inf. Technol. (ICCIT)*, Sep. 2020, pp. 10–15.

[34] N. Shirodkar, P. Mandrekar, R. S. Mandrekar, R. Sakhalkar, K. M. C. Kumar, and S. Aswale, "Credit card fraud detection techniques—A survey," in *Proc. Int. Conf. Emerg. Trends Inf. Technol. Eng. (ic-ETITE)*, 2020, pp. 1–7.

[35] C. M. M. R. van den Bergh and M. Junger, "Victims of cybercrime in Europe: A review of victim surveys," *Crime Sci.*, vol. 7, no. 1, pp. 1–15, Dec. 2018, doi: 10.1186/s40163-018-0079-3.

[36] N. Patterson, M. Hobbs, and D. Palmer, "A direct insight into victims of cybercrime," in *Proc. 12th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Jul. 2013, pp. 603–610.

[37] R. Sabillon, J. Cano, V. Cavaller, and J. Serra, "Cybercrime and cybercriminals: A comprehensive study," *Int. J. Comput. Netw. Commun. Secur.*, vol. 4, no. 6, pp. 165–176, 2016. [Online]. Available: http://www.ijcncs.org

[38] C. C. Aggarwal and P. S. Yu, "Outlier detection for high dimensional data," *ACM SIGMOD Rec.*, vol. 30, no. 2, pp. 37–46, 2001.

[39] A. Zimek, E. Schubert, and H.-P. Kriegel, "A survey on unsupervised outlier detection in high-dimensional numerical data," *Stat. Anal. Data Mining*, vol. 5, no. 5, pp. 363–387, Oct. 2012.

[40] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *J. Netw. Comput. Appl.*, vol. 60, pp. 19–31, Jan. 2016, doi: 10.1016/j.jnca.2015.11.016.

[41] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–72, 2009.

[42] *FBI—2018 Internet Crime Report*. Accessed: Sep. 9, 2021. [Online]. Available: https://www.ic3.gov/Media/PDF/AnnualReport/2018_IC3Report.pdf

[43] Q. Wang, W. Xu, X. Huang, and K. Yang, "Enhancing intraday stock price manipulation detection by leveraging recurrent neural networks with ensemble learning," *Neurocomputing*, vol. 347, pp. 46–58, Jun. 2019, doi: 10.1016/j.neucom.2019.03.006.

[44] S. R. Islam, S. Khaled Ghafoor, and W. Eberle, "Mining illegal insider trading of stocks: A proactive approach," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 1397–1406.

[45] A. Kulkarni, P. Mani, and C. Domeniconi, "Network-based anomaly detection for insider trading," 2017, *arXiv:1702.05809*.

[46] M. Ahmed, N. Choudhury, and S. Uddin, "Anomaly detection on big data in financial markets," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Jul. 2017, pp. 998–1001.

[47] M. Mirtaheri, S. Abu-El-Haija, F. Morstatter, G. V. Steeg, and A. Galstyan, "Identifying and analyzing cryptocurrency manipulations in social media," *IEEE Trans. Comput. Social Syst.*, vol. 8, no. 3, pp. 607–617, Jun. 2021.

[48] *Money Laundering Offences*. Accessed: Jul. 25, 2021. [Online]. Available: https://www.cps.gov.U.K./legal-guidance/proceeds-crime-act-2002-part-7-money-laundering-offences

[49] Y. Zhou, X. Wang, J. Zhang, P. Zhang, L. Liu, H. Jin, and H. Jin, "Analyzing and detecting money-laundering accounts in online social networks," *IEEE Netw.*, vol. 32, no. 3, pp. 115–121, May 2018.

[50] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech., Theory Exp.*, vol. 2008, pp. 1–12, Oct. 2008.

[51] A. F. Colladon and E. Remondi, "Using social network analysis to prevent money laundering," *Expert Syst. Appl.*, vol. 67, pp. 49–58, Jan. 2017, doi: 10.1016/j.eswa.2016.09.029.

[52] R. Dreżewski, J. Sepielak, and W. Filipkowski, "The application of social network analysis algorithms in a system supporting money laundering detection," *Inf. Sci.*, vol. 295, pp. 18–32, Feb. 2015.

[53] C. Chen, C. Liang, J. Lin, L. Wang, Z. Liu, X. Yang, J. Zhou, Y. Shuang, and Y. Qi, "InfDetect: A large scale graph-based fraud detection system for E-commerce insurance," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 1765–1773.

[54] M. Araujo, M. Almeida, J. Ferreira, L. Silva, and P. Bizarro, "BreachRadar: Automatic detection of points-of-compromise," 2020, *arXiv:2009.11751*.

[55] W. Feng, S. Liu, C. Faloutsos, B. Hooi, H. Shen, and X. Cheng, "EagleMine: Vision-guided micro-clusters recognition and collective anomaly detection," *Future Gener. Comput. Syst.*, vol. 115, pp. 236–250, Feb. 2021, doi: 10.1016/j.future.2020.08.033.

[56] M. Weber, J. Chen, T. Suzumura, A. Pareja, T. Ma, H. Kanezashi, T. Kaler, C. E. Leiserson, and T. B. Schardl, "Scalable graph learning for anti-money laundering: A first look," 2018, *arXiv:1812.00076*.

[57] J. Chen, T. Ma, and C. Xiao, "FastGCN: Fast learning with graph convolutional networks via importance sampling," pp. 1–15, 2018, *arXiv:1801.10247*.

[58] T. Suzumura and H. Kanezashi. (2021). *Anti-Money Laundering Datasets: InPlusLab Anti-Money Laundering Data Datasets*. [Online]. Available: http://github.com/IBM/AMLSim/

[59] *Cryptocurrency Crime and Anti-Money Laundering Report Ciphertrace Cryptocurrency Intelligence*, CipherTrace, Menlo Park, CA, USA, 2021.

[60] M. Weber, G. Domeniconi, J. Chen, D. Karl I. Weidele, C. Bellei, T. Robinson, and C. E. Leiserson, "Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics," 2019, *arXiv:1908.02591*.

[61] D. Vassallo, V. Vella, and J. Ellul, "Application of gradient boosting algorithms for anti-money laundering in cryptocurrencies," *Social Netw. Comput. Sci.*, vol. 2, no. 3, May 2021, Art. no. 143, doi: 10.1007/s42979-021-00558-z.

[62] I. Alarab, S. Prakoonwit, and M. I. Nacer, "Competence of graph convolutional networks for anti-money laundering in bitcoin blockchain," in *Proc. ACM Int. Conf.*, Jun. 2020, pp. 23–27.

[63] L. Nan and D. Tao, "Bitcoin mixing detection using deep autoencoder," in *Proc. IEEE 3rd Int. Conf. Data Sci. Cyberspace (DSC)*, Jun. 2018, pp. 280–287.

[64] A. Gaihre, S. Pandey, and H. Liu, "Deanonymizing cryptocurrency with graph learning: The promises and challenges," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Jun. 2019, pp. 2019–2021.

[65] *Customs Fraud—EU Anti-Fraud Office*. [Online]. Available: https://ec.europa.eu/anti-fraud/investigations/eu-revenue/trade_customs_fraud_en

[66] S. Kim, T.-D. Mai, T. N. D. Khanh, S. Han, S. Park, K. Singh, and M. Cha, "Take a chance: Managing the exploitation-exploration dilemma in customs fraud detection via online active learning," 2020, *arXiv:2010.14282*.

[67] E. L. Paula, M. Ladeira, R. N. Carvalho, and T. Marzagão, "Deep learning anomaly detection as support fraud investigation in Brazilian exports and anti-money laundering," in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2016, pp. 954–960.

[68] A. Alexopoulos, P. Dellaportas, S. Gyoshev, C. Kotsogiannis, and T. Pavkov. (2020). *Detecting Network Anomalies in the Value Added Taxes (VAT) System*. [Online]. Available: https://tarc.exeter.ac.U.K./media/universityofexeter/businessschool/documents/centres/tarc/publications/reports/Detecting_Network_Anomalies_in_the_VAT_system.pdf

[69] Y. Wang, Q. Zheng, J. Ruan, Y. Gao, Y. Chen, X. Li, and B. Dong, "T-EGAT: A temporal edge enhanced graph attention network for tax evasion detection," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2020, pp. 1410–1415.

[70] L. Zhang, X. Nan, E. Huang, and S. Liu, "Detecting transaction-based tax evasion activities on social media platforms using multi-modal deep neural networks," 2020, *arXiv:2007.13525*.

[71] Z. Zha, "TaxAA: A reliable tax auditor assistant for exploring suspicious transactions," in *Proc. Web Conf., Companion World Wide Web Conf. (WWW)*, Apr. 2020, pp. 240–244.

[72] N. Andrews, "'Can I get your digits?': Illegal acquisition of wireless phone numbers for SIM-swap attacks and wireless provider liability," *Northwestern J. Technol. Intellectual Property*, vol. 16, pp. 79–106, Nov. 2018.

[73] *How a Hacker Allegedly Stole Millions by Hijacking Phone Numbers*. Accessed: Feb. 26, 2021. [Online]. Available: https://www.vice.com/en/article/a3q7mz/hacker-allegedly-stole-millions-bitcoin-sim-swapping

[74] *Convicted SIM Swapper Gets 3 Years in Jail*. Accessed: Feb. 26, 2021. [Online]. Available: https://krebsonsecurity.com/2020/11/convicted-sim-swapper-gets-3-years-in-jail/

[75] C. Hagen, C. Weinert, C. Sendner, A. Dmitrienko, and T. Schneider, "All the numbers are U.S.: Large-scale abuse of contact discovery in mobile messengers," in *Proc. Netw. Distrib. Syst. Secur. Symp. (NDSS)*, 2021, pp. 1–17. [Online]. Available: https://encrypto.de/papers/HWSDS21.pdf

[76] M. Lansley, F. Mouton, S. Kapetanakis, and N. Polatidis, "SEADer++: Social engineering attack detection in online environments using machine learning," *J. Inf. Telecommun.*, vol. 4, no. 3, pp. 346–362, Jul. 2020, doi: 10.1080/24751839.2020.1747001.

[77] R. Montañez, E. Golob, and S. Xu, "Human cognition through the lens of social engineering cyberattacks," *Frontiers Psychol.*, vol. 11, pp. 1–18, Sep. 2020.

[78] *FBI Romance Fraud*. Accessed: Apr. 29, 2021. [Online]. Available: https://www.fbi.gov/scams-and-safety/common-scams-and-crimes/romance-scams

[79] *Romance Fraud on Rise in Coronavirus Lockdown*. Accessed: Apr. 29, 2021. [Online]. Available: https://www.bbc.com/news/technology-55997611

[80] G. Suarez-Tangil, M. Edwards, C. Peersman, G. Stringhini, A. Rashid, and M. Whitty, "Automatically dismantling online dating fraud," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1128–1137, 2019.

[81] *Hackers Breached Colonial Pipeline Using Compromised Password*. Accessed: Aug. 12, 2021. [Online]. Available: https://www.bloomberg.com/news/articles/2021-06-04/hackers-breached-colonial-pipeline-using-compromised-password
[82] *HSE—Cyber Attack Response*. Accessed: Aug. 12, 2021. [Online]. Available: https://www2.hse.ie/services/cyber-attack/how-it-may-affect-you.html
[83] *$50m Ransomware Demand on Acer is Highest Ever*. Accessed: Aug. 12, 2021. [Online]. Available: https://www.computerweekly.com/news/252498227/50m-ransomware-demand-on-Acer-is-highest-ever
[84] *DHS Secretary Warns Ransomware Attacks on the Rise, Targets Include Small Businesses*. Accessed: Aug. 13, 2021. [Online]. Available: https://abcnews.go.com/Politics/dhs-secretary-warns-ransomware-attacks-rise-targets-include/story?id=77512872
[85] S. H. Kok, A. Azween, and N. Jhanjhi, "Evaluation metric for crypto-ransomware detection using machine learning," *J. Inf. Secur. Appl.*, vol. 55, Dec. 2020, Art. no. 102646, doi: 10.1016/j.jisa.2020.102646.
[86] *Elliptic Website*. Accessed: Aug. 13, 2021. [Online]. Available: https://www.elliptic.co/
[87] *Revil Revealed—Tracking a Ransomware Negotiation and Payment*. Accessed: Aug. 13, 2021. [Online]. Available: https://www.elliptic.co/blog/revil-revealed-tracking-ransomware-negotiation-and-payment
[88] *Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case*. Accessed: Aug. 26, 2021. [Online]. Available: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402
[89] *Person Does Not Exist*. Accessed: Aug. 26, 2021. [Online]. Available: https://thispersondoesnotexist.com
[90] *Deepfakes Porn has Serious Consequences*. Accessed: Aug. 26, 2021. [Online]. Available: https://www.bbc.com/news/technology-42912529
[91] N. Ruiz, S. A. Bargal, and S. Sclaroff, "Disrupting deepfakes: Adversarial attacks against conditional image translation networks and facial manipulation systems," in *Proc. Eur. Conf. Comput. Vis.*, in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 12538, 2020, pp. 236–251.
[92] M. Schreyer, T. Sattarov, B. Reimer, and D. Borth, "Adversarial learning of deepfakes in accounting," Oct. 2019, *arXiv:1910.03810*.
[93] *Better Language Models and Their Implications*. Accessed: Aug. 26, 2021. [Online]. Available: https://openai.com/blog/better-language-models/
[94] D. I. Adelani, H. Mai, F. Fang, H. H. Nguyen, J. Yamagishi, and I. Echizen, "Generating sentiment-preserving fake online reviews using neural language models and their human- and machine-based detection," Jul. 2019, *arXiv:1907.09177*.
[95] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2005.
[96] M. A. H. Farquad and I. Bose, "Preprocessing unbalanced data using support vector machine," *Decis. Support Syst.*, vol. 53, no. 1, pp. 226–233, Apr. 2012, doi: 10.1016/j.dss.2012.01.016.
[97] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, vol. 8. Cambridge, U.K.: Cambridge Univ. Press, 1994.
[98] S. Min, Z. Gao, J. Peng, L. Wang, K. Qin, and B. Fang, "STGSN—A spatial–temporal graph neural network framework for time-evolving social networks," *Knowl.-Based Syst.*, vol. 214, Feb. 2021, Art. no. 106746, doi: 10.1016/j.knosys.2021.106746.
[99] R. Hanneman and M. Riddle, *Introduction to Social Network Methods*. Berkeley, CA, USA: Univ. California Press, 2005. [Online]. Available: https://wiki.gonzaga.edu/dpls707/images/6/6e/Introduction_to_Social_Network_Methods.pdf
[100] E. Kurshan, H. Shen, and H. Yu, "Financial crime & fraud detection using graph computing: Application considerations & outlook," in *Proc. 2nd Int. Conf. Transdisciplinary AI (TransAI)*, Sep. 2020, pp. 125–130.
[101] D. Yu, Y. Yang, R. Zhang, and Y. Wu, *Knowledge Embedding Based Graph Convolutional Network*, vol. 1. New York, NY, USA: Association for Computing Machinery, 2021.
[102] S. Rayana. (2016). *Odds Library*. [Online]. Available: http://odds.cs.stonybrook.edu
[103] V. Prakash Dwivedi, C. K. Joshi, T. Laurent, Y. Bengio, and X. Bresson, "Benchmarking graph neural networks," 2020, *arXiv:2003.00982*.
[104] *Buried Treasure—How Tightening Regulation is Forcing Criminals to go to Extreme Lengths to Cash-Out Their Cryptoassets*. Accessed: Aug. 28, 2021. [Online]. Available: https://www.elliptic.co/blog/buried-treasure-criminals-to-go-to-extreme-lengths-to-cash-out-crypto
[105] *Traffic Prediction With Advanced Graph Neural Networks*. Accessed: Aug. 12, 2021. [Online]. Available: https://deepmind.com/blog/article/traffic-prediction-with-advanced-graph-neural-networks
[106] FINRA. (Jun. 2020). *Artificial Intelligence (AI) in the Securities Industry*. [Online]. Available: https://www.finra.org/sites/default/files/2020-06/ai-report-061020.pdf
[107] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.

JACK NICHOLLS received the B.Sc. degree in physics with medical physics and bioengineering from Technological University Dublin and the M.Sc. degree in business analytics from the UCD Michael Smurfit Graduate Business School. He is currently pursuing the Ph.D. degree with the School of Computer Science (CS), University College Dublin (UCD). He has several years of experience working in financial risk as a Consultant and as an Analyst in an Irish bank.

ADITYA KUPPA is currently pursuing the Ph.D. degree with the School of Computer Science (CS), University College Dublin (UCD). He has worked in cyber security industry for last 17 years at various roles. He has published scientific papers in peer-reviewed conferences in related research fields. He is also active as a reviewer for many key conferences and journals in related disciplines.

NHIEN-AN LE-KHAC (Member, IEEE) received the Ph.D. degree in computer science from the Instituto Politécnico Nacional de Grenoble (INPG), France, in 2006. He is a Lecturer at the School of Computer Science (CS), University College Dublin (UCD). He is currently the Program Director of M.Sc. Program in Forensic Computing and Cybercrime Investigation. He is also the Co-Founder of UCD-GNECB Postgraduate Certificate in fraud and e-crime investigation. Since 2008, he has been a Research Fellow at Citibank, Ireland (Citi). Since 2013, he has collaborated on many international and national research projects as a principal/co-PI/funded investigator. His research interests include cybersecurity and digital forensics, machine learning for security, fraud and criminal detection, cloud security and privacy, grid and high performance computing, and knowledge engineering. He has published more than 200 scientific papers in peer-reviewed journals and conferences in related research fields.