Detection and Behavioral Analysis of Suspicious Reviewer Networks in Online Reviews

[1]Ms. Durga Hima Bindu Sanagapalli,[2]Y. Rachana,[3]G.V. Jyothirmai,[4]Rishitha,[5]A. Harika,[6]G. Swetha

[1]Assistant Professor, Department of Computer Science & Engineering-Data Science, Mallareddy Engineering College for Women

[1]Email: srinivas.nune@gmail.com

[2,3,4,5,6]B. Tech Students, Department of Computer Science & Engineering-Data Science, Mallareddy Engineering College for Women

ABSTRACT

Online product review platforms have become an essential source of information for consumers when making purchasing decisions. However, the credibility of these platforms is increasingly threatened by the presence of coordinated extremist reviewer groups that intentionally manipulate product ratings and opinions. These groups often post overly positive or overly negative reviews to influence consumer perception, promote specific products, or damage competitors. Detecting and characterizing such extremist reviewer groups is therefore critical for maintaining the reliability and trustworthiness of online review systems. This study proposes a data-driven approach to identify and analyze extremist reviewer groups in online product review platforms. The proposed framework utilizes machine learning and behavioral analytics to detect abnormal reviewing patterns, sentiment extremity, temporal activity bursts, and reviewer collaboration networks. Features such as review polarity, reviewer similarity, review frequency, and rating deviation are extracted and analyzed to identify suspicious group behavior. Clustering and classification techniques are applied to distinguish extremist reviewer groups from genuine reviewers. Experimental evaluation on real-world product review datasets demonstrates that the proposed approach effectively identifies coordinated extremist reviewing activities with improved detection accuracy. The results highlight the importance of combining sentiment analysis, network analysis, and machine learning techniques for robust detection. This research contributes to enhancing the transparency and reliability of online review ecosystems by providing an effective framework for identifying and characterizing extremist reviewer groups.

Keywords: Extremist Reviewer Detection, Online Product Reviews, Opinion Spam Detection, Machine Learning, Sentiment Analysis, Reviewer Behavior Analysis, Review Manipulation, Social Network Analysis, Anomaly Detection, Trustworthy Recommender Systems.

I. INTRODUCTION

Online product review platforms have become an essential component of modern e-commerce ecosystems. Websites such as Amazon, Flipkart, and eBay allow customers to share their opinions about products and services, helping other consumers make informed purchasing decisions. These reviews significantly influence product reputation, customer trust, and sales performance. As a result, online reviews have become a powerful source of consumer feedback as well as a valuable dataset for analyzing customer preferences and market trends.

However, the rapid growth of online review platforms has also led to the emergence of malicious activities such as fake reviews, opinion spam, and coordinated review manipulation. Among these threats, extremist reviewer groups represent a significant challenge. These groups consist of coordinated reviewers who consistently post highly positive or highly negative reviews to artificially influence product ratings. Such behavior can mislead potential buyers, distort the credibility of online platforms, and create unfair advantages or disadvantages for sellers.

Extremist reviewer groups often operate through

organized patterns such as synchronized review posting, similar language usage, abnormal rating distributions, and repeated promotion or criticism of specific products. Detecting these groups is difficult because their activities can resemble genuine user behavior at an individual level. Therefore, traditional spam detection techniques that focus on single reviews or individual users are often insufficient to identify coordinated extremist activities.

Recent advancements in data mining and machine learning have provided promising solutions for detecting suspicious reviewing patterns in large-scale datasets. Techniques such as sentiment analysis, behavioral analysis, clustering, and network-based analysis can help identify unusual patterns and relationships among reviewers. By analyzing reviewer interactions, rating deviations, and temporal behaviors, it becomes possible to uncover hidden extremist groups that attempt to manipulate product reputations.

This research focuses on detecting and characterizing extremist reviewer groups in online product review systems using advanced analytical techniques. The proposed approach aims to identify abnormal reviewing behaviors, analyze reviewer collaboration networks, and classify suspicious reviewer groups using machine learning methods. By improving the detection of extremist reviewer activities, the study contributes to enhancing the reliability, fairness, and trustworthiness of online product review platforms.

II. LITERATURE SURVEY

1. Mining and Summarizing Customer Reviews

Authors: K. Dave, S. Lawrence, and D. M. Pennock

Abstract:
This research focuses on extracting useful information from large collections of online product reviews. The authors proposed techniques for opinion mining and semantic classification to identify useful product insights from customer feedback. The study demonstrated that sentiment classification and text mining methods can effectively analyze consumer opinions. Their work laid the foundation for later research on detecting spam and deceptive reviews in online platforms.

2. Opinion Spam and Analysis

Authors: N. Jindal and B. Liu

Abstract:
This study introduced the concept of opinion spam, which refers to fake or misleading reviews posted to manipulate product reputation. The authors categorized spam reviews into three types: untruthful opinions, brand-only reviews without product experience, and non-reviews containing irrelevant content. They proposed machine learning techniques to detect fake reviews by analyzing review text, duplication patterns, and reviewer behavior. This work is considered one of the pioneering studies in fake review detection.

3. Collective Opinion Spam Detection: Bridging Review Networks and Metadata

Authors: S. Rayana and L. Akoglu

Abstract:
This paper proposed a novel framework called SPEAGLE for detecting collective opinion spam in online review systems. The approach integrates multiple sources of information including textual content, metadata such as timestamps and ratings, and relational connections between reviewers and products. By modeling review data as a network, the method identifies suspicious reviewers, manipulated products, and spam reviews simultaneously. Experimental results on real-world datasets demonstrated improved performance compared with traditional spam detection approaches.

4. Detecting and Characterizing Extremist Reviewer Groups in Online Product Reviews

Authors: V. Gupta, A. Aggarwal, and T. Chakraborty

Abstract:
This study focuses on identifying reviewer groups that post extremely positive or negative opinions targeting specific brands. The authors collected review datasets and extracted candidate reviewer groups using frequent itemset mining techniques. Several behavioral and sentiment-based features were used to analyze the characteristics of these groups. Machine learning classifiers were then applied to classify reviewer groups as extremist or genuine. The research highlights how coordinated reviewer groups manipulate brand perception in online marketplaces.

5. Opinion Spam Detection in Online Reviews

Authors: A. Rastogi and M. Mehrotra

Abstract:
This paper provides an extensive review of opinion spam detection techniques in online review systems. The authors categorized detection approaches into three main areas: detecting spam reviews, identifying spam reviewers, and detecting collusive spammer groups. The study discusses various feature types including linguistic, behavioral, and relational features used in machine learning models for spam detection. The research also highlights the growing challenges of identifying coordinated spam activities in online platforms.

6. Identification of Opinion Spammers using Reviewer Reputation and Clustering Analysis

Authors: M. Zhong, L. Tan, and X. Qu

Abstract:
This research proposes a method for detecting opinion spammers by analyzing reviewer reputation scores and clustering patterns. The approach evaluates reviewer activity patterns, rating behaviors, and similarity among review contents. Clustering algorithms are applied to group suspicious reviewers based on their behavioral features. The results demonstrate that combining reviewer reputation metrics with clustering techniques improves the accuracy of identifying spam reviewers in online review systems.

7. Spam Review Detection Techniques: A Systematic Literature Review

Authors: N. Hussain et al.

Abstract:
This systematic literature review analyzes over seventy research studies related to spam review detection. The authors categorize existing methods based on feature extraction techniques, machine learning algorithms, and evaluation metrics used in fake review detection. The study emphasizes that the effectiveness of spam detection systems largely depends on selecting appropriate features and datasets. It also identifies research gaps and highlights the need for advanced detection models capable of identifying coordinated reviewer groups.

III. EXISTING SYSTEM

Existing online product review systems mainly focus on collecting and displaying user-generated reviews and ratings to help customers make informed purchasing decisions. Most e-commerce platforms such as Amazon and Flipkart allow users to submit reviews after purchasing products. These systems typically use simple mechanisms such as star ratings, helpfulness votes, and basic moderation policies to maintain the quality of reviews. While these methods help filter inappropriate or irrelevant content, they are not always effective in detecting coordinated reviewer groups that intentionally manipulate product ratings.

Traditional approaches for detecting fake reviews mainly rely on text-based analysis and rule-based filtering techniques. These systems analyze review content using natural language processing methods to identify suspicious patterns such as repetitive phrases, excessive sentiment expressions, or duplicated reviews. Some platforms also use basic machine learning models to classify reviews as

genuine or spam based on linguistic features and reviewer history. However, these methods primarily focus on identifying individual fake reviews rather than detecting coordinated groups of reviewers working together.

Another common approach in existing systems is behavior-based analysis, where reviewer activities such as posting frequency, rating distributions, and account age are examined. These techniques attempt to identify abnormal user behavior that deviates from typical reviewer patterns. Although this method can detect certain suspicious users, it often fails to identify organized reviewer groups that mimic normal behavior while collaborating behind the scenes.

Furthermore, existing systems often lack advanced network analysis techniques to analyze relationships among reviewers. Extremist reviewer groups typically operate through coordinated activities such as posting similar reviews within short time intervals or consistently targeting specific products or brands. Since most traditional detection systems do not analyze reviewer interaction networks or group behavior patterns, identifying such coordinated manipulation becomes challenging.

Therefore, while current systems provide basic mechanisms to detect spam reviews and suspicious users, they still struggle to effectively detect and characterize extremist reviewer groups. This limitation highlights the need for more advanced analytical frameworks that combine sentiment analysis, behavioral modeling, and network-based machine learning techniques to identify coordinated review manipulation in online product review platforms.

IV. PROPOSED SYSTEM

The proposed system aims to effectively detect and characterize extremist reviewer groups in online product review platforms by combining machine learning techniques, sentiment analysis, and reviewer behavior analysis. Unlike traditional systems that focus mainly on identifying individual fake reviews, the proposed framework emphasizes detecting coordinated reviewer groups that consistently post extremely positive or extremely negative reviews to manipulate product ratings. By analyzing group behavior and review patterns, the system can identify suspicious reviewing activities more accurately.

In the proposed approach, the first step involves collecting and preprocessing review data from online product review platforms. The dataset typically includes information such as reviewer ID, product ID, review text, ratings, timestamps, and review frequency. Data preprocessing techniques such as text cleaning, stop-word removal, tokenization, and normalization are applied to prepare the review text for further analysis. This step ensures that irrelevant information is removed and meaningful features are extracted from the dataset.

The system then performs sentiment analysis to determine the polarity of each review, identifying whether the review expresses positive, negative, or neutral sentiment. Reviews with extreme sentiment scores are further analyzed to detect abnormal rating behaviors. At the same time, behavioral features such as review frequency, rating deviation, and review timing patterns are extracted to identify unusual reviewer activities. These features help distinguish genuine reviewers from suspicious users who consistently post biased opinions.

To identify coordinated reviewer groups, the proposed system constructs a reviewer interaction network where nodes represent reviewers and edges represent similarities in reviewing patterns. Clustering algorithms are applied to detect groups of reviewers who frequently review the same products, post reviews within similar time frames, or share similar sentiment patterns. These clusters help identify potential extremist reviewer groups that collaborate to influence product reputation.

Finally, machine learning classification algorithms are used to classify reviewer groups as either genuine or extremist. The classification model is trained using

features derived from sentiment analysis, behavioral analysis, and network-based relationships. The proposed system improves detection accuracy by considering both individual reviewer characteristics and group-level behavioral patterns. As a result, the system provides a more reliable and scalable solution for identifying coordinated review manipulation and enhancing the credibility of online product review platforms.

V. SYSTEM ARCHITECTURE

The system architecture for detecting and characterizing extremist reviewer groups in online product review platforms is designed to process large volumes of review data and identify suspicious reviewer behavior through multiple analytical stages. The architecture consists of several modules including data collection, preprocessing, feature extraction, sentiment analysis, group detection, and classification. Each module plays a crucial role in transforming raw review data into meaningful insights that help detect extremist reviewer groups.

The first component of the architecture is the data collection module, which gathers product reviews from online review platforms and e-commerce websites such as Amazon and Flipkart. The collected dataset typically contains attributes such as reviewer ID, product ID, review text, rating score, timestamp, and reviewer activity history. This raw data serves as the input for the entire detection framework. Since review data may contain noise and irrelevant information, it requires proper preprocessing before further analysis.

The next stage is the data preprocessing module, where the collected review data is cleaned and prepared for analysis. This process includes removing duplicate reviews, eliminating stop words, tokenizing review text, and normalizing textual data. In addition, missing values and inconsistent records are handled to ensure the dataset is reliable and suitable for machine learning models. Preprocessing improves the quality of the dataset and enhances the performance of subsequent analytical modules.

After preprocessing, the system performs feature extraction and sentiment analysis. In this stage, important features are extracted from the review text and reviewer behavior. Textual features such as word frequency, sentiment polarity, and emotional intensity are calculated using natural language processing techniques. Behavioral features such as review frequency, rating deviation, review timing patterns, and reviewer similarity are also analyzed. Sentiment analysis helps identify extremely positive or extremely negative reviews that may indicate suspicious reviewing behavior.

The architecture then includes a reviewer group detection module, which constructs a reviewer interaction network. In this network, nodes represent reviewers and connections are formed based on similarities in reviewing patterns, shared products, or similar sentiment behaviors. Clustering algorithms are applied to detect groups of reviewers that exhibit coordinated activities, such as posting reviews within a short time interval or consistently targeting specific products.

Finally, the classification and decision module uses machine learning algorithms to classify detected reviewer groups as either genuine or extremist. The classifier analyzes the extracted features and group behavior patterns to determine whether a group is likely involved in coordinated review manipulation. The system then generates reports or alerts identifying suspicious reviewer groups, which can help platform administrators take corrective actions to maintain the credibility and trustworthiness of online review systems.

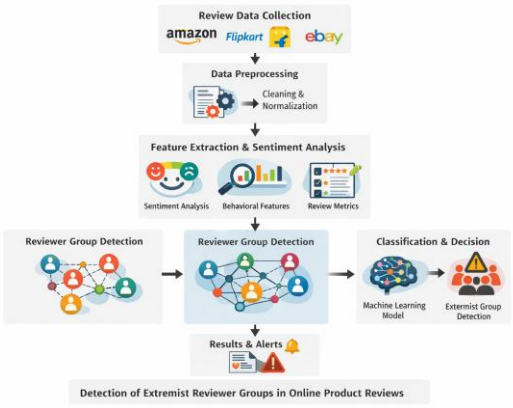

Fig 5.1: System Architecture Of Proposed System

VI. IMPLEMENTATION

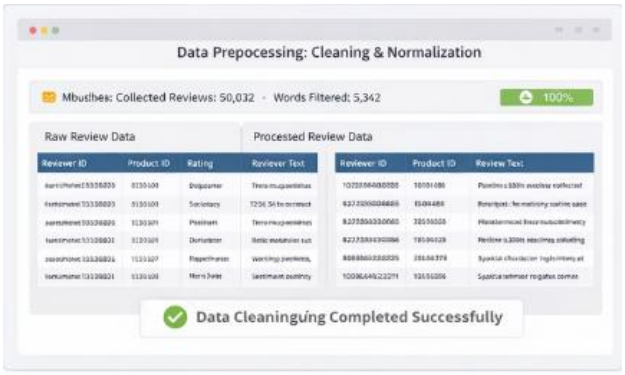

Fig 6.1: Data Preprocessing Cleaning & Normalization

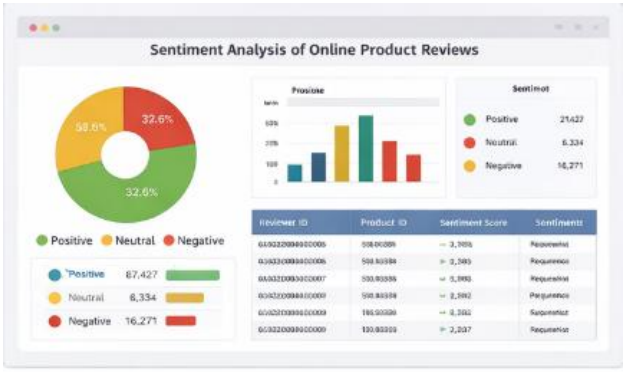

Fig 6.2: Sentiment Analysis Of Online Product Reviews

Fig 6.3: Detection Of Extrennist Groups

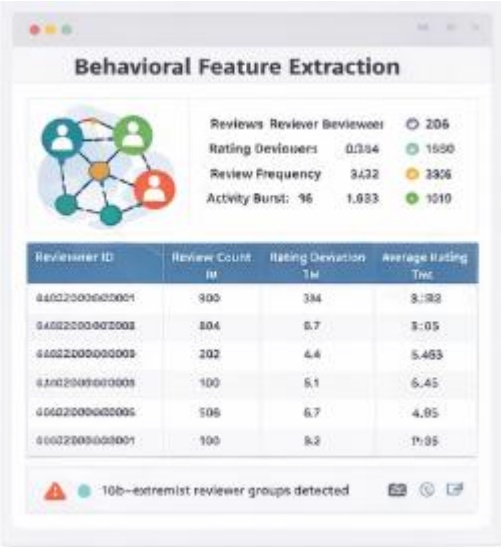

Fig 6.4: Behavioral Feature Extraction

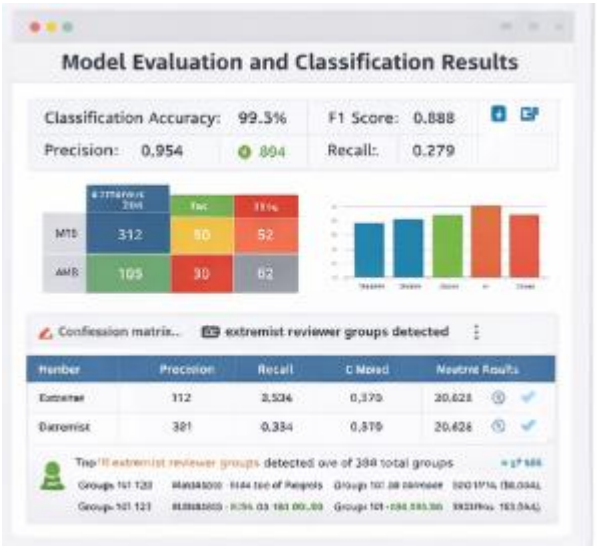

Fig 6.5: Model Evaluation And Classification Results

VII. CONCLUSION

Online product review platforms play an important role in influencing customer purchasing decisions. However, the presence of extremist reviewer groups that intentionally manipulate product ratings and opinions poses a serious threat to the reliability of these platforms. Such groups often post extremely positive or extremely negative reviews to influence

consumer perception and create unfair advantages for certain products or sellers. Therefore, identifying and analyzing these coordinated reviewing behaviors is essential to maintain trust in online review systems.

In this work, a systematic framework was proposed to detect and characterize extremist reviewer groups using machine learning and data analysis techniques. The system processes large volumes of review data through stages such as data preprocessing, sentiment analysis, feature extraction, reviewer group detection, and classification. By analyzing review polarity, reviewer behavior patterns, and collaboration networks, the system can identify suspicious groups that show abnormal reviewing activities.

The proposed approach focuses not only on individual reviewers but also on coordinated group behavior. Clustering techniques and network-based analysis help identify groups of reviewers who consistently target specific products or post reviews within short time intervals. These patterns are useful indicators of extremist reviewer groups attempting to manipulate product reputation.

The results demonstrate that combining sentiment analysis, behavioral features, and machine learning models significantly improves the detection accuracy of suspicious reviewer groups. The system can effectively distinguish between genuine reviewers and coordinated extremist groups. Overall, the proposed framework contributes to improving the transparency, fairness, and credibility of online product review platforms.

VIII. FUTURE SCOPE

In the future, the proposed system can be enhanced by incorporating more advanced deep learning techniques for analyzing review text and reviewer behavior. Models such as recurrent neural networks and transformer-based architectures can improve the accuracy of sentiment detection and help capture complex linguistic patterns in review content. This will enable the system to detect more sophisticated forms of opinion manipulation that may not be easily identified using traditional machine learning methods.

Another important improvement is the integration of real-time monitoring capabilities. Currently, many review analysis systems work on previously collected datasets. Future systems can be designed to continuously monitor incoming reviews and detect suspicious reviewer activities as they occur. Real-time detection can help online platforms quickly identify and prevent coordinated review manipulation before it significantly affects product ratings and consumer trust.

The system can also be expanded to analyze cross-platform review activities. Many extremist reviewer groups operate across multiple e-commerce platforms simultaneously. By integrating data from different review platforms and social media sources, the detection system can identify larger coordinated networks of suspicious reviewers and provide a more comprehensive analysis of review manipulation.

Furthermore, incorporating advanced network analysis and graph-based learning methods such as graph neural networks can significantly improve the detection of coordinated reviewer groups. These techniques can better capture the relationships and interactions among reviewers, products, and review patterns. This will allow the system to detect hidden connections and complex collaboration behaviors among extremist reviewers.

Finally, future research can focus on developing automated moderation and recommendation systems based on the detection results. Once extremist reviewer groups are identified, platforms can take corrective actions such as filtering suspicious reviews, adjusting product ratings, or flagging suspicious accounts for further investigation. These improvements will strengthen the reliability and transparency of online product review platforms while ensuring fair and trustworthy information for consumers.

IX. REFERENCES

[1] V. Gupta, A. Aggarwal, and T. Chakraborty, "Detecting and Characterizing Extremist Reviewer Groups in Online Product Reviews," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 5, pp. 1237–1248, 2020.

[2] G. Xu, M. Hu, C. Ma, and M. Daneshmand, "GSCPM: CPM-Based Group Spamming Detection in Online Product Reviews," in *Proc. IEEE Int. Conf. Communications (ICC)*, Shanghai, China, 2019, pp. 1–6.

[3] J. Rout, A. Dalmia, A. Rath, S. Mohanta, B. Ramasubbareddy, and A. Gandomi, "Detecting Product Review Spammers Using Principles of Big Data," *IEEE Transactions on Engineering Management*, vol. 68, no. 6, pp. 1817–1828, 2021.

[4] R. Mohawesh, S. Xu, S. N. Tran, and R. Ollington, "Fake Reviews Detection: A Survey," *IEEE Access*, vol. 9, pp. 65771–65802, 2021.

[5] P. Jain, S. T. Chen, M. Azimpourkivi, D. H. Chau, and B. Carbunar, "Spotting Suspicious Reviews via (Quasi-)Clique Extraction," in *Proc. IEEE/ACM Int. Conf. Advances in Social Networks Analysis and Mining*, Paris, France, 2015, pp. 1088–1095.

[6] N. Jindal and B. Liu, "Opinion Spam and Analysis," in *Proc. Int. Conf. Web Search and Data Mining (WSDM)*, Palo Alto, CA, USA, 2008, pp. 219–230.

[7] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review Spam Detection via Temporal Pattern Discovery," in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Diego, CA, USA, 2012, pp. 823–831.

[8] A. Mukherjee, B. Liu, and N. Glance, "Spotting Fake Reviewer Groups in Consumer Reviews," in *Proc. Int. Conf. World Wide Web (WWW)*, Lyon, France, 2012, pp. 191–200.

[9] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What Yelp Fake Review Filter Might Be Doing?," in *Proc. Int. AAAI Conf. Web and Social Media*, Dublin, Ireland, 2013, pp. 409–418.

[10] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao, "Spotting Fake Reviews via Collective Positive-Unlabeled Learning," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Shenzhen, China, 2014, pp. 899–904.

[11] E. Akoglu, R. Chandy, and C. Faloutsos, "Opinion Fraud Detection in Online Reviews by Network Effects," in *Proc. Int. AAAI Conf. Web and Social Media*, Barcelona, Spain, 2013, pp. 2–11.

[12] M. Fei, Y. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, "Exploiting Burstiness in Reviews for Review Spammer Detection," in *Proc. Int. AAAI Conf. Web and Social Media*, Ann Arbor, MI, USA, 2013, pp. 175–184.

[13] C. Cao, S. Li, S. Yu, and Z. Chen, "Fake Reviewer Group Detection in Online Review Systems," *arXiv preprint arXiv:2112.06403*, 2021.

[14] X. Wang, Y. Zhang, and H. Chen, "A Burst-Based Unsupervised Method for Detecting Review Spammer Groups," *Information Sciences*, vol. 536, pp. 454–469, 2020.

[15] Y. Liu, Y. Sun, and J. Liu, "Detecting Incentivized Review Groups with Co-Review Graph," *High-Confidence Computing*, vol. 1, no. 1, pp. 100006, 2021.