Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

Naive Bayes and Entropy based Analysis and Classification of Humans and Chat Bots

Dr. S. Smys,
Professor,
Department of CSE,
RVS Technical Campus,
Coimbatore, India.
Email id: smys375@gmail.com

Dr. Haoxiang Wang,
Director and Lead Executive Faculty Member,
GoPerception Laboratory,
Cornell University, Ithaca, USA.
Email: wanghaoxiang1102@hotmail.com

Abstract: Internet users are largely threatened by abuse and manipulation of several automated chat service programs called as chat bots. Malware and spam is distributed by the popular chat networks using chat bots. The commercial chat network is surveyed in this paper with a series of measurements. A series of 15 advanced to simple chatbots are used for this purpose. When compared to the bot behavior, the complexity of human behavior is high. A classification system is proposed for accurate distinguishing between human user and chatbots based on the measurements obtained from the study. Naive Bayes Classifier and entropy classifier are used for the purpose of classification. Chat bot detection is performed with improved efficiency and accuracy using these classifiers. The speed of Naive Bayes Classifier and accuracy of entropy classifier compliments each other in the process of detection of chat bots. The improved efficiency of the proposed system is proved by testing and comparison with the existing schemes.

Keywords: Naive Bayes Classifier, Entropy classifier, chatbots, malware, spam, classification

1. Introduction

Communication in the form of text on real time is enabled by internet chat applications. Internet chat is used by millions of people across the globe for online information exchange over a wide range of subjects [1]. Low bandwidth consumption and enabling interaction

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

between humans are the unique features of the internet chat application. However, malicious exploitation often occurs in internet chat due to its open nature and large user base. The online users are threatened by chatbots and automated programs leading to abuse of the chat services [2]. Open chat networks like Jabber, IRC and huge commercial chat networks like MSN, Yahoo!, AOL and so on contains chatbots on several chat systems. Even in World of Warcraft and other online games, which are basically non-chat applications with chat features, bots have been reported [3]. Through mounting phishing attacks, spreading malware and sending spam, the online systems are exploited by chat bots.

Human interactive proofs and filtering based on key words are the major approaches used so far to combat the chat bots [4]. However, the keyword lists published are often evaded by the chat bots by frequent updates by the bot makers leading to high false negative rates when third-party chat clients use knowledge-based message filters [5]. The chat bots are assisted by the bot operators for logging into the chat rooms by taking up the passing tests leading to inefficiency of the human interactive proofs like CAPTCHAs. Huge number of bots enter the chat rooms despite the implementation of CAPTCHA by Yahoo! In the year 2007 for blocking the bots from entering the chat rooms [6]. Several online petitions have been raised requesting the chat service providers like Yahoo! And AOL to address the growing issues of chat bots. The chat bots are not investigated in a systematic manner by the online systems despite being plagued with the chat bots [7]. There is a great demand to implement an effective chat bot detection system. However, an all efficient system does not still exist.

Yahoo! chat, a huge commercial chat network is analyzed and a series of measurements are taken for studying the online chat behavior of humans and chat bots. 15 diverse chat bots are considered for capturing the required measurements in this work [8]. Different text obfuscation schemes and triggering mechanisms are used by different chat bot varieties. The message content and message timing are determined based on these factors. The chat bot classification may be performed by measuring complexity by means of entropy rate in order to reveal the complexity of human behavior over bot behavior. For accurate distinguishing between human and chat bots, a classification system is proposed based on the measurement analysis [9]. Naive Bayes Classifier and entropy classifier are the major components of the classification model. The chart flow complexity is measured by the entropy classifier with the help of message size and time characteristics for classifying them as humans or bots. However, the detection is

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

largely dependent on the message content in case of Naive Bayes Classifier. In the chat bot detection process, both these classifiers complement each other. In detection of unknown chat bots, the accuracy is high using entropy classifier despite the being a slow process as more messages are required for detection [10]. The Naive Bayes Classifier is trained by the entropy classifier. The unknown bots cannot be detected accurately by the machine learning classifiers despite their speed of classification. The speed and accuracy of the system is improved by combining the Naive Bayes Classifier and entropy classifier. Efficiency of the classification system is validated by experimental tests.

2. Related Works

The chat traffic and their statistical properties are measured by systematic analysis of Web and IRC. When compared to the web chat sessions, the duration of total IRC sessions is longer leading to longer chat sessions [11]. An exponential distribution is seen in the interarrival time between chat sessions and a non-exponential interarrival time exists between message distribution. Small packets dominate the chat sessions in a large amount when considering the message size. 10 times the data transmitted is received by the user typically over a complete session. Conversely, more data is transmitted than received by extremely active web chat users who use IRC based automated scripts. In terms of user base and protocol, there is a significant overlap between the instant messaging (IM) and chat systems [12]. Systems supporting channels and chat rooms are generally referred to as chat whereas direct presence and messaging are supported by the IM systems. For example, IRC is a chat system and AIM is an IM system. The protocol design and IM system are largely impacted by the rise of IM systems as predated due to the wide use of chat systems like IRC. The chat systems back-port certain user-friendly features in the IM systems as a response [13].

File transfer and presence are some of the recent features similar to IM implemented in the classic IRC chat system [14]. IM as well as chat can access the clients and end users in certain message service providers like Yahoo. Related literature based on IM systems are outlined here. The spim and spam IM are filtered and detected by means of server-side and client-side schemes [15]. As the spim data is insufficient, spam messages are evaluated based on short e-mail corpuses in this evaluation. The IM contact list is used for studying the IM systems and spread of automated malware through IM worms. The honeypot concept is used

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

for detection and suppression of IM malware by leveraging the IM malware spreading characteristics [16]. The spim content characteristics and network analysis is performed with honeypots despite the fact that they have no direct relationship with instant messaging or chat. Socially interactive malware issues are also discussed by certain researchers.

3. Proposed Work

The chat bot classification model is designed and developed based on the operation of two classifiers namely the Naive Bayes Classifier and entropy classifier. Figure 1 represents the architecture of the proposed chat bot classification technique. The input data is processed for making classification decisions in a concurrent manner by both the Naive Bayes as well as the entropy classifier. The bot corpus is built by the Naive Bayes classifier with the information obtained from the entropy classifier. The chat users are scored based on the corrected conditional entropy and entropy by the entropy classifier for classifying them as humans or bots. New chat bots are captured by the entropy classifier and included to the corpus of chat bots. Classification is performed based on manual log or the clean chat logs database to deduce human corpus. The patterns of humans and bots can be learnt from the human and bot corpora by the Naive Bayes classifier [14]. This pattern aids in quick classification of chat bots.

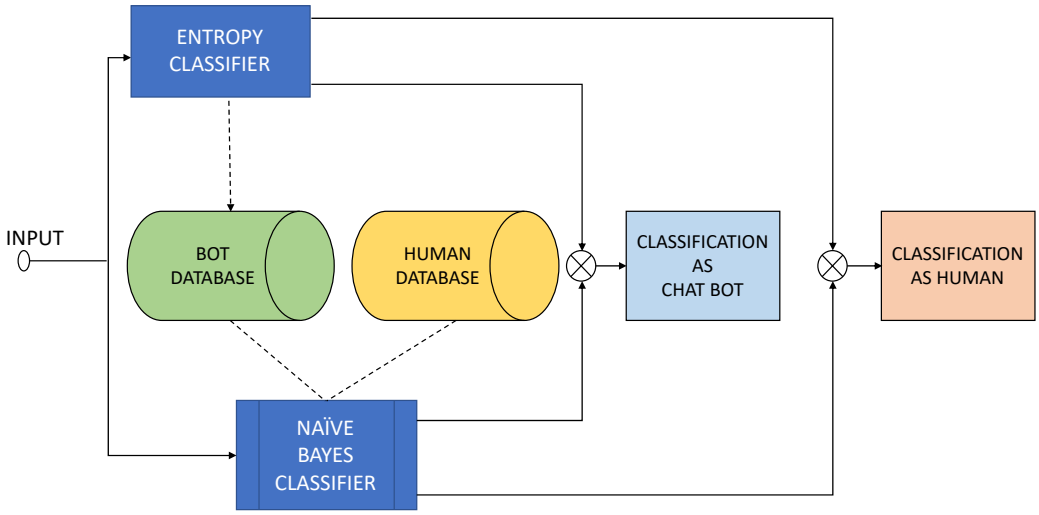

Figure 1. Architecture of the proposed chat bot classification model

The chat user inter-message delays and message sizes providing the measure of entropy rate an entropy are used for making classification decision by the entropy classifier. Chat bots

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

are identified based on the predictable and regular behavior observed from the characteristics like low entropy rate or entropy. Unpredictable and irregular behavior is observed when high entropy rate and entropy is detected indicating possible human behavior. For each measure of entropy, a cutoff is set for the purpose of classification. Classification of chat user as a human is done if the test score is equal to or larger than the cut off score. Classification of chat user as a bot is done if the test score is lesser than the cut off score. For an entropy classifier, in order to determine the true positive and false positive rates, cut off score is a significant factor [12]. Misclassification of humans as bots may occur with a huge entropy score, while misclassification of bots as humans may occur with small entropy score. The cut off scores are selected to attain the desired false positive rate based on the entropy score of humans. This is because, attaining low false positive rate is significant.

The Naive Bayes Classifier identifies chat bits with the help of the chat message content. Naive Bayes text classification may be performed for identifying the chat bots with text and emoticons from the chat messages. Here, a set of predefined classes, text for classification and the classifier are used for formalizing the text classification issue within the paradigm. Other techniques like decision trees, Bayesian classifier and support vector machines may also be used for classification of text. However, in e-mail spam detection and other text classification applications, the Naive Bayes Classifier is more efficient. For this purpose, it has been employed for the chat bot detection in this paper.

4. Results and Discussion

The results of detection by the Naive Bayes Classifier as well as the entropy classifier are presented in this section. Chat bots of various types like replay responder, advanced responder, replay, responder, random and periodic levels are categorized based on their detection difficulty for classifying the tests and their results. Human results are presented following the bot results. Figure 2 represents the message size and human inter-message delay distribution probability. The probability mass function (PMF) curves are derived due to the persistent human behavior over a duration of 30 days. A heavily tailed distribution has been observed on the inter-message delay distribution in internet chat systems in prior studies. The observation is confirmed by this study. On the other hand, messages are posted by periodic bots at regular intervals of time. Several seconds of delay and variation may be used by the

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

periodic bots. Chat server delay, network traffic congestion or transmission delay may contribute to this delay period variation. Utilization of periodic timers by chat bots form a substantial portion as message distribution is performed in an efficient and simple manner with periodic message posting.

Messages can be posted at random time intervals using random bots. They use continuous, discrete or random distributions for generating delay between messages. When compared to periodic bots, human-like appearance may be observed in random bots due to the use of random timers. However, when compared to humans, the distribution of inter-message delay is quite different by random bots. For random bots, the message size and inter-message delay distribution is analyzed. The chatroom message content acts as the input to responder bots. A vague response may be transmitted by the responder bot when a messages with question marks are posted. This tricks human users to believe that the responder is a human and lead them to click the link further.

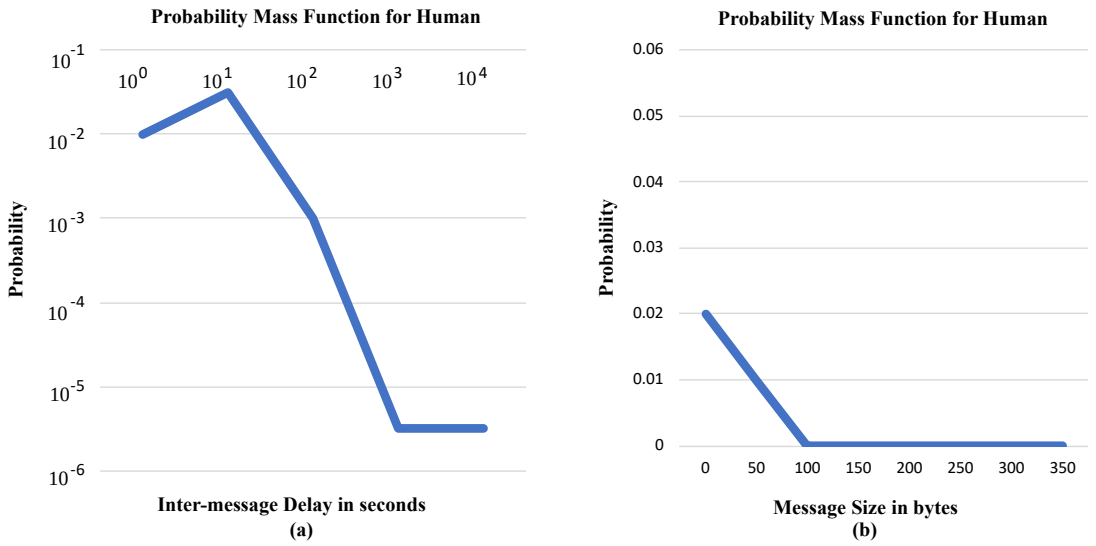

Figure 2. Probability Mass Function for humans in terms of (a) inter-message delay and (b) message size

The replay bots appear more like a human user by transmitting messages from other users while sending its own messages. However, the replayed phrases may not occur in the chatroom and are related to the same topic. They may be previously saved in a database or may be taken from conversations on similar topics in other chatrooms. Integration of responder and

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

replay bots enable the bot to provide response to user messages based on keywords while providing inter-message delay in a human-like manner. The advanced responder bot has a highly detailed configuration. It consists of a huge database of responses and keywords. Figure 3 represents the probability mass function of periodic bots in terms of message size as well as inter-message delay. The duration of analysis is divided into periods of 15 days.

Human training set is essential for the entropy classifier for determining the cutoff scores. Entropy based and fully supervised training are used for testing Naive Bayes Classifier. Samples of 50 user messages are used for entropy classifier while as few as 10 user message samples are sufficient for the Naive Bayes Classifier. Leading internet chat service providers are studied over a period of 30 days and the chat logs are collected from 10 different chat rooms. 15 varieties of chat bots are identified from these chat logs and are classified into advanced responder bots, replay-responder bots, replay bots, responder bots, random bots and periodic bots. When compared to the human users, the behavior of chat bots is significantly different based on the statistical analysis obtained from message size and inter-message delay for humans and chat bots. In terms of message size and inter-message delay, certain regularities are specifically exhibited by chat bots. The overall sophistication found in humans is missing in the replay and responder bots despite their advanced human-like aspects.

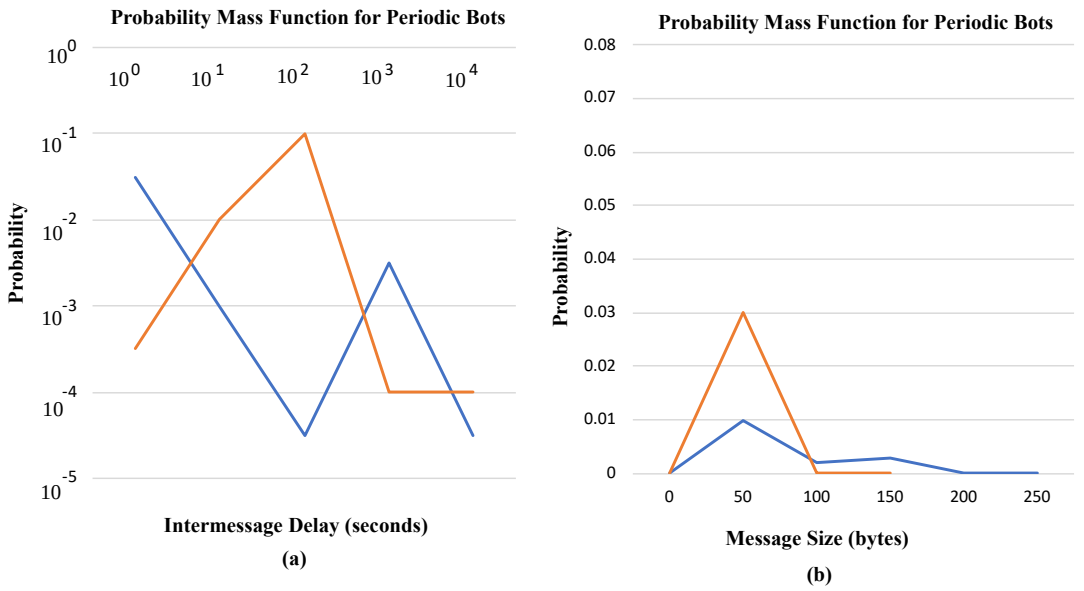

Figure 3. Probability Mass Function for Periodic Bots in terms of (a) Inter-message delay and (b) Message size

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

5. Conclusion

A large scale analysis and study on internet chat is presented in this paper. Naive Bayes Classifier and entropy classifier are used for accurate detection and classification of chat bots by analyzing various measurements. The message size and inter-message delay in chat bots that are low entropy characteristics are exploited by the entropy classifier while the difference between the message content among chat bots and human users are leveraged in the Naive Bayes Classifier. The replay and responder bots that transmits messages in a human-like manner are also detected along with other unknown bots with the help of the entropy classifier. However, the time taken for detection is relatively long due to the need for a large number of messages. Naive Bayes Classifier operates in a faster manner when compared to the entropy classifier as the number of messages required is relatively small. A bot corpus is built and maintained by the entropy classifier along with bot detection. This bot corpus is used by the Naive Bayes Classifier for training purpose leading to accurate and quick detection of chat bots. Based on the experimental results, it is evident that this hybrid classification system is quick and accurate in detection and identification of bots that were previously unknown. Future work is directed towards introducing exponential aging and micro grooming schemes for managing the bot corpus.

References

[1] Parimala, M., Swarna Priya, R. M., Praveen Kumar Reddy, M., Lal Chowdhary, C., Kumar Poluru, R., & Khan, S. (2021). Spatiotemporal-based sentiment analysis on tweets for risk assessment of event using deep learning approach. Software: Practice and Experience, 51(3), 550-570.

[2] Tyagi, P., & Tripathi, R. C. (2019, February). A review towards the sentiment analysis techniques for the analysis of twitter data. In Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE).

[3] Karanja, E. M., Masupe, S., & Jeffrey, M. G. (2020). Analysis of internet of things malware using image texture features and machine learning techniques. Internet of Things, 9, 100153.

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

[4] Bird, J. J., Ekárt, A., Buckingham, C. D., & Faria, D. R. (2019, July). High resolution sentiment analysis by ensemble classification. In Intelligent Computing-Proceedings of the Computing Conference (pp. 593-606). Springer, Cham.

[5] Pophale, S., Gandhi, H., & Gupta, A. K. (2021). Emotion Recognition Using Chatbot System. In Proceedings of International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications (pp. 579-587). Springer, Singapore.

[6] Ahuja, R., Chug, A., Gupta, S., Ahuja, P., & Kohli, S. (2020). Classification and clustering algorithms of machine learning with their applications. In Nature-Inspired Computation in Data Mining and Machine Learning (pp. 225-248). Springer, Cham.

[7] Ismail, Z., Jantan, A., Yusoff, M. N., & Kiru, M. U. (2021). The effects of feature selection on the classification of encrypted botnet. Journal of Computer Virology and Hacking Techniques, 17(1), 61-74.

[8] Bird, J. J., Ekárt, A., Buckingham, C. D., & Faria, D. R. Ensemble Classification in Multi-level Sentiment Analysis for Cross-Domain Application.

[9] Victor, D. B., Kawsher, J., Labib, M. S., & Latif, S. (2020, November). Machine Learning Techniques for Depression Analysis on Social Media-Case Study on Bengali Community. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1118-1126). IEEE.

[10] Tamizharasi, B., Livingston, L. J., & Rajkumar, S. (2020, December). Building a Medical Chatbot using Support Vector Machine Learning Algorithm. In Journal of Physics: Conference Series (Vol. 1716, No. 1, p. 012059). IOP Publishing.

[11] Leonova, V. (2020, June). Review of Non-English Corpora Annotated for Emotion Classification in Text. In International Baltic Conference on Databases and Information Systems (pp. 96-108). Springer, Cham.

[12] Vijayasekaran, G., & Rosi, S. (2018). Spam and email detection in big data platform using naives bayesian classifier. International Journal of Computer Science and Mobile Computing, 7(4), 53-58.

[13] Jacob, I. J. (2020). Performance evaluation of caps-net based multitask learning architecture for text classification. Journal of Artificial Intelligence, 2(01), 1-10.

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021

Journal of ISMAC (2021)
Vol.03/ No.01
Pages: 40-49
http://irojournals.com/iroismac/
DOI: https://doi.org/10.36548/jismac.2021.1.004

[14] Joseph, S. I. T., & Thanakumar, I. (2019). Survey of data mining algorithm's for intelligent computing system. Journal of trends in Computer Science and Smart technology (TCSST), 1(01), 14-24.

[15] Manoharan, S. (2020). Geospatial and social media analytics for emotion analysis of theme park visitors using text mining and gis. Journal of Information Technology, 2(02), 100-107.

[16] Sungheetha, A., & Sharma, R. (2020). Transcapsule model for sentiment classification. Journal of Artificial Intelligence, 2(03), 163-169.

ISSN: 2582-1369 (online)
Submitted:19.01.2021
Revised: 28.02.2021
Accepted: 21.03.2021
Published: 07.04.2021