
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

Automated Emerging Cyber Threat Identification And Profiling Based On Natural Language Processing

Dr.M.Vinaya Babu Associate professor [1], Raparthi Santhosha[2]

[1]PG Scholars, Department of CSE, Teegala Krishna Reddy EngineeringCollege (Autonomous Institution),Medbowli, Meerpet, Saroornagar,Hyderabad
[2]Professor, Department of CSE, Teegala Krishna Reddy Engineering College (Autonomous Institution), Medbowli, Meerpet, Saroornagar, Hyderabad

ABSTRACT

The time window between the disclosure of a new cyber vulnerability and its use by Cybercriminals have been getting smaller and smaller over time. Recent episodes, such as Log4j vulnerability, exemplify this well. Within hours after the exploit being released, attackers started scanning the internet looking for vulnerable hosts to deploy threats like crypto currency miners and ransom ware on vulnerable systems. Thus, it becomes imperative for the cyber security defense strategy to detect threats and their capabilities as early as possible to maximize the success of prevention actions. Although crucial, discovering new threats is a challenging activity for security analysts due to the immense volume of data and information sources to be analyzed for signs that a threat is emerging. In this sense, we present a framework for automatic identification and profiling of emerging threats using Twitter messages as a source of events and MITRE ATT&CK as a source of knowledge for threat characterization. The framework comprises three main parts: identification of cyber threats and their names; profiling the identified threat in terms of its intentions or goals by employing two machine learning layers to filter and classify tweets; and alarm generation based on the threat's risk. The main contribution of our work is the approach to characterize or profile the identified threats in terms of their intentions or goals, providing additional context on the threat and avenues for mitigation. In our experiments, the profiling stage reached an F1 score of 77% in correctly profiling discovered threats.

Keywords: cyber vulnerability, crypto currency, ransom ware, MITRE ATT&CK

I. INTRODUCTION

Recently there has been increasing reliance on the Internet for business, government, and social interactions because of a trend of hyper- different motivations and intentions. Preventing organizations from cyber exploits needs timely intelligence about cyber vulnerabilities and attacks, referred to as threats .

Threat intelligence is defined as ‘‘evidence-based knowledge, including context, mechanisms, indicators, implications, and actionable advice, about an existing or emerging menace or hazard

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

to assets that can be used to inform decisions regarding the subject's response to that menace or hazard' . Threat intelligence in cyber security domain, or cyber threat intelligence, provides timely and relevant information, such as signatures of the attacks, that can help reduce the uncertainty in identifying potential security vulnerabilities and attacks.

Cyber threat intelligence can generally be extracted from informal or formal sources, which officially release threat information in structured data format. Structured threat intelligence adheres to a well-defined data model, with a common format and structure. Structured cyber threat intelligence, therefore, can be easily parsed by security tools to analyze and respond to security threats accordingly. Examples of formal sources of cyber threat intelligence include the Common Vulnerabilities and Exposures (CVE) databaseand the National Vulnerability Database (NVD).

Cyber threat intelligence is also available on informal sources, such as public blogs, dark webs, forums, and social media platforms. Informal sources allow any person or entity on the Internet to publish, in real-time, the threat information in natural language, or unstructured data format. The unstructured and publicly available threat intelligence is also called Open-Source Intelligence (OSINT) . Cyber security-related OSINT are early warning sources for cyber security events such as security vulnerability exploits .

To conduct a cyber-attack, malicious actors typically must 1) identify vulnerabilities, 2) acquire the necessary tools and tradecraft to successfully exploit them, 3) choose a target and recruit participants, 4) create or purchase the infrastructure needed, and 5) plan and execute the attack. Other actors— system administrators, security analysts, and even victims— may discuss vulnerabilities or coordinate a response to attacks [. These activities are often conducted online through social media, (open and dark) Web forums, and professional blogs, leaving digital traces behind.

II.RELATED WORK

Literature Survey on Automated Cyber Threat Identification

Introduction

Automated cyber threat identification has become a critical area of research due to the increasing number of cyber-attacks and the vast amount of data generated on social media platforms and other sources. This survey reviews key methodologies and contributions in this field, focusing on the use of machine learning, natural language processing (NLP), and data mining techniques.

Gathering Cyber Threat Intelligence from Social MediaTwitter as a Source of Cyber Threat Intelligence

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

B. D. Le, G. Wang, M. Nasim, and A. Babar (2019) proposed a novel approach for gathering cyber threat intelligence from Twitter using novelty classification. Their method involves:

Collecting tweets related to cyber security.

Applying NLP techniques to preprocess the text data.

Using machine learning models to identify novel cyber threats based on the collected tweets. Their findings suggest that Twitter can be an effective source of real-time cyber threat intelligence, significantly enhancing the capability to detect emerging threats promptly (Le et al., 2019).

Machine Learning Techniques in Cyber Threat Identification

Supervised learning algorithms such as Support Vector Machines (SVM), Random Forest, and Neural Networks have been widely used in the identification of cyber threats. These algorithms require labelled data to train models that can predict threats based on historical attack patterns.

Unsupervised learning methods, including clustering and anomaly detection, are used to identify unusual patterns in network traffic or user behaviour that may indicate a cyber threat. These techniques do not require labelled data, making them suitable for detecting unknown or emerging threats.

Natural Language Processing in Cyber security

NLP techniques are extensively used to extract relevant information from unstructured text data, such as tweets, forums, and blogs. Text mining involves tasks such as tokenization, stemming, and named entity recognition (NER) to identify key terms and entities related to cyber threats.

Sentiment analysis can help in understanding the context and potential impact of identified cyber threats. By analysing the sentiment of posts related to cybersecurity, researchers can prioritize threats based on the perceived severity and urgency.

Data Mining for Cyber Threat Intelligence

Association rule mining helps in discovering relationships between different types of cyber threats and attack patterns. This technique can reveal common sequences of events leading to a cyber-attack, providing insights into preventive measures.

Classification and Regression Trees (CART)

CART models are used to classify data into different threat categories and predict the likelihood of future attacks. These models help in automating the decision-making process in threat identification and response.

III.IMPLEMENTATION

Service Provider

In this module, the Service Provider has to login by using valid username and password. After login successfully he can do some operations such as Browse Datasets and Train & Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Cyber Threat Identification Type,

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

View Cyber Threat Identification Type Ratio, Download Predicted Data Sets, View Cyber Threat Identification Type Ratio Results, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who are all registered. In this, the admin can view the user's details such as, username, email, address and admin authorize the users.

IV.ALGORITHMS

Decision Tree Classifiers

Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision-making knowledge from the supplied data. Decision tree can be generated from training sets. The procedure for such generation based on the set of objects (S), each belonging to one of the classes C1, C2, Ck is as follows:

Step 1. If all the objects in S belong to the same class, for example Ci, the decision tree for S consists of a leaf labelled with this class

Step 2. Otherwise, let T be some test with possible outcomes O1, O2…, On. Each object in S has one outcome for T so the test partitions S into subsets S1, S2… Sn where each object in Si has outcome Oi for T. T becomes the root of the decision tree and for each outcome Oi we build a subsidiary decision tree by invoking the same procedure recursively on the set Si.

Gradient boosting

Gradient boosting is a machine learning technique used in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees. When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms random forest. A gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

- Learning based on instances, and thus also works lazily because instance close to the input vector for test or prediction may take time to occur in the training dataset

Yet, despite this, it appears robust and efficient. Its performance is comparable to other supervised learning techniques. Various reasons have been advanced in the literature. In this tutorial, we highlight an explanation based on the representation bias. The naive bayes classifier is a linear classifier, as well as linear discriminate analysis, logistic regression or linear SVM (support vector machine). The difference lies on the method of estimating the parameters of the classifier (the learning bias).

While the Naive Bayes classifier is widely used in the research world, it is not widespread among practitioners which want to obtain usable results.

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

On the one hand, the researchers found especially it is very easy to program and implement it, its parameters are easy to estimate, learning is very fast even on very large databases, its accuracy is reasonably good in comparison to the other approaches. On the other hand, the final users do not obtain a model easy to interpret and deploy, they does not understand the interest of such a technique.

This largely explains the good performance of the method in comparison to others. In the second part, we use various tools on the same dataset (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b and Rapid Miner 4.6.0). We try above all to understand the obtained results.

An extension of the algorithm was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark in 2006 (as of 2019, owned by Minitab, Inc.).The extension combines Breiman's "bagging" idea and random selection of features, introduced first by Ho[1] and later independently by Amit and Geman[13] in order to construct a collection of decision trees with controlled variance.

Random forests are frequently used as "black box" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration.

SVM

In classification tasks a discriminate machine learning technique aims at finding, based on an independent and identically distributed (iid) training dataset, a discriminate function that can correctly predict labels fornewly acquired instances. Unlike generative machine learning approaches, which require computations of conditional probability distributions, a discriminate classification function takes a data point x and assigns it to one of the different classes that are a part of the classification task. Less powerful than generative approaches, which are mostly used when prediction involves outlier detection, discriminate approaches require fewer computational resources and less training data, especially for a multidimensional feature space and when only posterior probabilities are needed. From a geometric perspective, learning a classifier is equivalent to finding the equation for a multi dimensional surface that best separates the different classes in the feature space.

SVM is a discriminate technique, and, because it solves the convex optimization problem analytically, it always returns the same optimal hyper plane parameter—in contrast to genetic algorithms(*GAs*) or perceptrons, both of which are widely used for classification in machine learning. For perceptrons, solutions are highly dependent on the initialization and termination criteria. For a specific kernel that transforms the data from the input space to the feature space, training returns uniquely defined SVM model parameters for a given training set, whereas the perceptron and GA

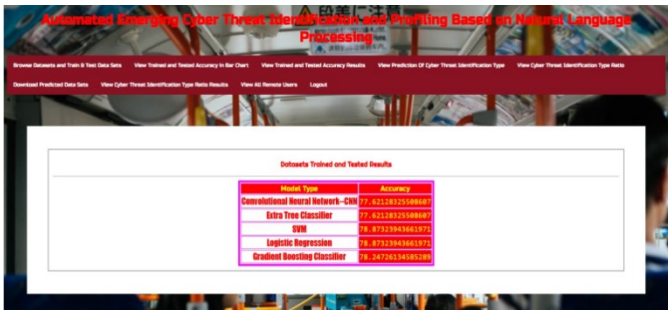
ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

classifier models are different each time training is initialized. The aim of GAs and perceptrons is only to minimize error during training, which will translate into several hyper planes' meeting this requirement.

V.RESULTS

Fig:1 , Home Page

Fig:2 Trained and Tested Result



Fig,3 View all Remote Users

Fig.4 View Trained and Tested Accuracy Bar

Fig.5 Cyber Theft Identification Type Details

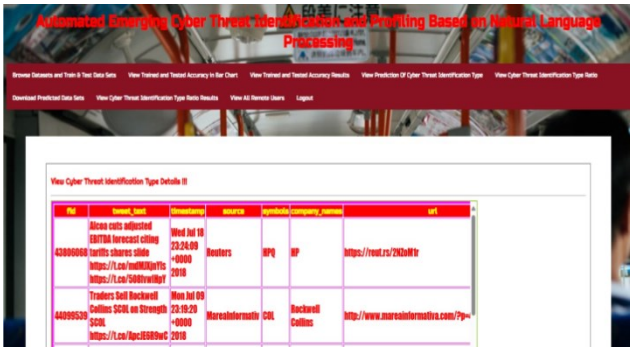


Fig:6 View Trained and Tested Accuracy

Fig:7.View all Remote Users

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

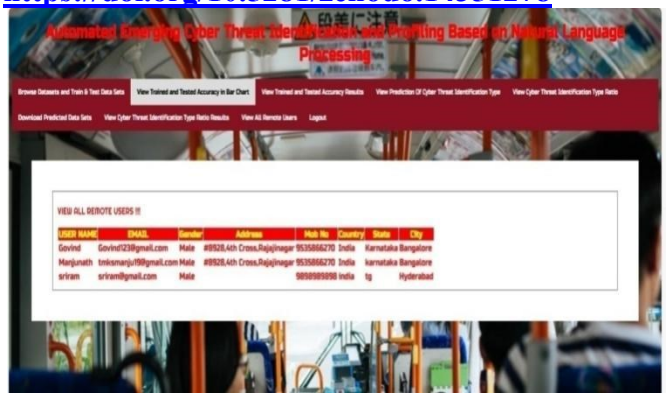


Fig:8 Cyber Theft Identification Type Ratio

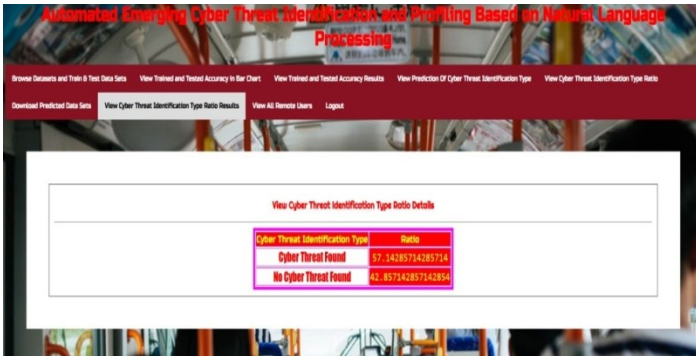


Fig:9 Prediction of Cyber Theft Type

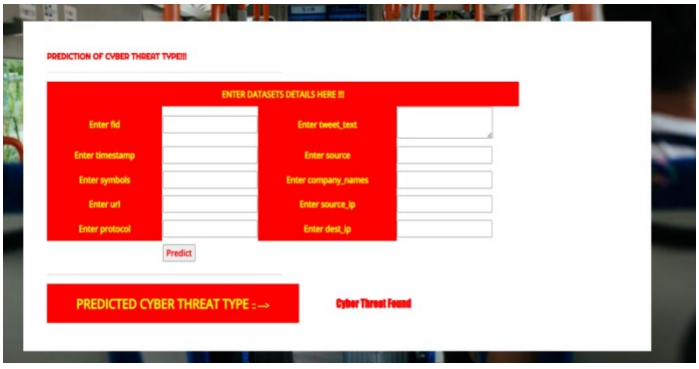


Fig:10 Prediction of Cyber Theft Type Result

VI.CONCLUSION

Given the dynamism of the cyber security field, with new vulnerabilities and threats appearing at any time, keeping up to date on them is a challenging but important task for analysts. Even following the best practices and applying the best controls, a new threat may bring an unusual way to subvert the defences requiring a quick response. This way, timely information about emerging cyber threats becomes paramount to a complete cyber security system.

This research proposes automated cyber threat identification and profiling based on the natural language processing of Twitter messages. The objective is exactly to cooperate with the hard work of following the rich source of information that is Twitter to extract valuable information about emerging threats in a timely manner.

This work differentiates itself from others by going a step beyond identifying the threat. It seeks to identify the goals of the threat by mapping the text from tweets to the procedures conducted by real threats described in MITRE ATT&CK knowledge base. Taking advantage of this evolving and collaborative knowledge base to train machine learning algorithms is a way to leverage the efforts of cyber security community to automatically profile identified cyber threats in terms of their intents.

To put in test our approach, in addition to the research experiment, we implemented the proposed pipeline and run it for 70 days generating online alerts for the Threat Intelligence Team of a big financial institution in Brazil. During this period, at least three threats made the team take preventive actions, such as the Petit Potam case, described in section V. Our system alerted the team making them aware of Petit- Potam 17 days

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

before the official patch was published by Microsoft. Within this period, the defense team was able to implement mitigations avoiding potential exploits and, consequently, incidents.

Our experiments showed that the profiling stage reached an F1 score of 77% in correctly profiling discovered threats among 14 different tactics and the percentage of false alerts of 15%. In future work, we consider it important to advance in tweets selection stages (Unknown Words and One-class), to improve the false positives rate and in the profiling stage, to reach higher accuracy in determining the technique associated with the identified threat. We are working on this way by experimenting with a different NLP approach using the part of speech (POS) algorithm implementation from Spacy29 Python library. The object is to identify the root verb, the subject, and the object of the phrases to select tweets where the action described (the root verb) is referencing the unknown word (the subject).

REFERENCES

[1] B. D. Le, G. Wang, M. Nasim, and A. Babar, ‘‘Gathering cyberthreat intelligence from Twitter using novelty classification,’’ 2019,arXiv:1907.01755.

[2] Definition: Threat Intelligence, Gartner Research, Stamford, CO, USA,2013.

[3] R. D. Steele, ‘‘Open source intelligence: What is it? why is it important tothe military,’’ Journal, vol. 17, no. 1, pp. 35–41, 1996.

[4] C. Sabottke, O. Suciu, and T. Dumitras, ‘‘Vulnerability disclosure in the ageof social media: Exploiting Twitter for predicting real-world exploits,’’ inProc. 24th USENIX Secur. Symp. (USENIX Secur.), 2015, pp. 1041–1056.

[5] A. Sapienza, A. Bessi, S. Damodaran, P. Shakarian, K. Lerman, andE. Ferrara, ‘‘Early warnings of cyber threats in online discussions,’’ inProc. IEEE Int. Conf. Data Mining Workshops (ICDMW), Nov. 2017,pp. 667–674.

[6] E. Nunes, A. Diab, A. Gunn, E. Marin, V. Mishra, V. Paliath, J. Robertson,J. Shakarian, A. Thart, and P. Shakarian, ‘‘Darknet and deepnet mining for proactive cybersecurity threat intelligence,’’ in Proc. IEEE Conf. Intell.Secur. Informat. (ISI), Sep. 2016, pp. 7–12.

[7] S. Mittal, P. K. Das, V. Mulwad, A. Joshi, and T. Finin, ‘‘CyberTwitter:Using Twitter to generate alerts for cybersecurity threats and vulnerabilities,’’in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining(ASONAM), Aug. 2016, pp. 860–867.

[8] A. Attarwala, S. Dimitrov, and A. Obeidi, ‘‘How efficient is Twitter:Predicting 2012 U.S. presidential elections using support vector machine via Twitter and comparing against Iowa electronic markets,’’ in Proc. Intell.Syst. Conf. (IntelliSys), Sep. 2017, pp. 646–652.

[9] N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, ‘‘Towards end-to-end cyberthreat detection from Twitter using multi-task learning,’’ in Proc. Int.Joint Conf. Neural Netw. (IJCNN), Jul. 2020, pp. 1–8. [10] O. Oh, M. Agrawal, and H. R. Rao, ‘‘Information control and terrorism:Tracking the Mumbai terrorist attack through Twitter,’’ Inf. Syst. Frontiers,vol. 13, no. 1, pp. 33–43, Mar. 2011.

[11] T. Sakaki, M. Okazaki, and Y. Matsuo, ‘‘Earthquake shakes Twitter users:Real-time event detection by social sensors,’’ in Proc. 19th Int. Conf.WorldWide Web, Apr. 2010, pp. 851–860.

[12] B. De Longueville, R. S. Smith, and G. Luraschi, ‘‘‘OMG, from here, ican see the flames!’: A use case of

ISSN 2321-2152
www.ijmece.com
Vol 12, Issue 4, 2024
https://doi.org/10.5281/zenodo.14351278

mining location based social networksto acquire spatio-temporal data on forest fires,'' in Proc. Int. Workshop Location Based Social Netw., Nov. 2009, pp. 73–80.

[13] A. Sapienza, S. K. Ernala, A. Bessi, K. Lerman, and E. Ferrara, ''DISCOVER:Mining online chatter for emerging cyber threats,'' in Proc.Companion Web Conf. Web Conf. (WWW), 2018, pp. 983–990.

[14] R. P. Khandpur, T. Ji, S. Jan, G. Wang, C.-T. Lu, and N. Ramakrishnan,''Crowdsourcing cybersecurity: Cyber attack detection using social media,'' in Proc. ACM Conf. Inf. Knowl. Manage., Nov. 2017,pp. 1049–1057.

[15] Q. Le Sceller, E. B. Karbab, M. Debbabi, and F. Iqbal, ''SONAR: Automatic detection of cyber security events over the Twitter stream,'' in Proc.12th Int. Conf. Availability, Rel. Secur., Aug. 2017, pp. 1–11.

[16] K.-C. Lee, C.-H. Hsieh, L.-J.Wei, C.-H. Mao, J.-H. Dai, and Y.-T. Kuang,''Sec-buzzer: Cyber security emerging topic mining with open threat intelligence retrieval and timeline event annotation,'' Soft Comput., vol. 21,no. 11, pp. 2883–2896, Jun. 2017.

[17] A. Ritter, E. Wright, W. Casey, and T. Mitchell, ''Weakly supervised extraction of computer security events from Twitter,'' in Proc. 24th Int.Conf. World Wide Web, May 2015, pp. 896–905.

[18] A. Queiroz, B. Keegan, and F. Mtenzi, ''Predicting software vulnerabilityusing security discussion in social media,'' in Proc. Eur. Conf. Cyber Warfare Secur., 2017, pp. 628–634.

[19] A. Bose, V. Behzadan, C. Aguirre, and W. H. Hsu, ''A novel approach fordetection and ranking of trendy and emerging cyber threat events in Twitter streams,'' in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining(ASONAM), Aug. 2019, pp. 871–878.

[20] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, ''Mitre ATT&CK: Design and philosophy,'' MITRE Corp., McLean, VA, USA, Tech. Rep. 19-01075-28, 2018.

[21] B.-J. Koops, J.-H. Hoepman, and R. Leenes, ''Open-source intelligence
and privacy by design,'' Comput. Law Secur. Rev., vol. 29, no. 6,
pp. 676–688, Dec. 2013.

[22] R. Campiolo, L. A. F. Santos, D. M. Batista, and M. A. Gerosa, ''Evaluating the utilization of Twitter messages as a source of security alerts,'' in Proc.28th Annu. ACM Symp. Appl. Comput., Mar. 2013, pp. 942–943.

[23] N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, ''Cyberthreat detection from Twitter using deep neural networks,'' in Proc. Int. Joint Conf.
Neural Netw. (IJCNN), Jul. 2019, pp. 1–8.

[24] A. Niakanlahiji, J.Wei, and B. Chu, ''A natural language processing basedtrend analysis of advanced persistent threat techniques,'' in Proc. IEEE Int.Conf. Big Data (Big Data), Dec. 2018, pp. 2995–3000.

[25] G. Ayoade, S. Chandra, L. Khan, K. Hamlen, and B. Thuraisingham,''Automated threat report classification over multi-source data,'' in Proc.IEEE 4th Int. Conf. Collaboration Internet Comput. (CIC), Oct. 2018,pp. 236–245.