EasyChair Preprint

№ 7151

Retrieval of Redundant Hyperlinks After Attack Based on Hyperbolic Geometry of Web Complex Networks

Mahdi Moshiri and Farshad Safaei

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

December 4, 2021

Retrieval of redundant hyperlinks after attack based on hyperbolic geometry of web Complex Networks

Mahdi Moshiri, Farshad Safaei,
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, IRAN
{m_moshiri, f_safaei}@sbu.ac.ir

ABSTRACT

The Internet and the Web can be described as huge networks of connected computers, connected web pages, or connected users. Analyzing link retrieval methods on the Internet and the Web as examples of complex networks is of particular importance. The recovery of complex networks is an important issue that has been extensively used in various fields. Much work has been done to measure and improve the stability of complex networks during attacks. Recently, many studies have focused on the network recovery strategies after the attack. Predicting the appropriate redundant links in a way that the network can be recovered at the lowest cost and fastest time after attacks or interruptions will be critical in a disaster. In addition, real-world networks such as the World Wide Web are no exception, and many attacks are made on hyperlinks between web pages, and the issue of predicting redundant hyperlinks on this World Wide Web is also very important.

In this paper, different kinds of attack strategies are provided and some retrieval strategies based on link prediction methods are proposed to recover the hyperlinks after failure or attack. Besides that, a new link prediction method based on the hyperbolic geometry of the complex network is proposed to retrieve redundant hyperlinks and the numerical simulation reveals its superiority that the state-of-the-art algorithms in recovering the attacked hyperlinks especially in the case of attacks based on edge betweenness strategy.

1. INTRODUCTION

For the past decades, many studies have focused on network recovery in various contexts and many different kinds of recovery strategies were developed considering failure characteristics and network types. Matisziw et al. [1] suggested a multi-objective recovery. Chaoqi et al. [2] proposed a repair model to study the impact of network structure and analyzed the energy of the network during recovery. Hu et al. [3] suggested an optimal recovery strategy for geographical networks after attacks. Yu and Yang [4] selected appropriate components of the failed network to repair it and reach a stable condition with limited recovery resources. Besides that, there are many studies in the area of network resilience.Resilience is the power of the network to return to a stable state as soon as possible [5]. The successful implementation of recovery strategies has a direct impact on the resilience of complex networks. Based on this, many researchers have focused on studying the concept of resilience in a complex networks. Majdandzic et al. [6] developed a model to demonstrate the automatic recovery in a networked system. Also, Afrin et al. [7] have recently given a thorough survey about the attack and recovery strategies and future approaches to this problem. The recovery strategies mentioned above are quite diverse and can be applied to different kinds of networks including web networks under various circumstances and based on this reason, it is necessary to study the retrieval methods and develop a more general retrieval strategy that can confront more types of networks.

In recent years, the link prediction problem [8] has achieved great progress in the area of network science. Link prediction methods try to predict missing, spurious , and future links in different kinds of complex networks [9-13]. This problem is generally solved based on unsupervised similarity measures, supervised methods (machine-learning based methods) [12], maximum likelihood methods [14], stochastic block model [15, 16], and probabilistic models [17, 18]. Recently many new applications have been defined for link prediction methods such as community detection, network reconstruction, and recommendation systems [19-21]. Recently robustness of link prediction methods under attack was examined by Wang et al. [22]. Different tools have been applied to enhance the performance of link prediction methods. One of these tools is the hyperbolic geometry of the complex networks. Krioukov et al. [23] and Papadopoulos et al. [24] introduced the networks' mapping to their underlying hyperbolic space and studied the impact of the two factors of popularity and similarity in networks' growth and suggested that new connections are made between nodes with popularity and similarity trade-off. After that Papadopoulos et al. [25] proposed the HyperMap method to map a network to its hyperbolic geometry and used the hyperbolic distance to solve the link prediction problem. Different from these works. Recently, Muscoloni et al. [26, 27] introduced a model (N-PSO) to predict the missing links based on the community structure of the networks. Samei and Jalili [28] proposed an enhanced method based on the hyperbolic distance to solve missing and spurious link prediction in the multiplex networks.

The objective of this article is to propose a method to recover the attacked or failed hyperlinks in the web network after random or targeted attacks instantly via unsupervised similarity measures used in solving the link prediction problem. The similarity measures are defined based on some existing measures such as Common Neighbors (CN), Resource Allocation (RA), Hyperbolic distance, Cannistraci-Resource-Allocation (CRA), CH2-L2, and a proposed measure based on underlying hyperbolic geometry of the complex networks.

This paper is organized as follows. In section 2, the problem is defined clearly as well as the attack and retrieval strategies and the evaluation metrics are presented. In section 3 the datasets are proposed and in section 4 the experimental results and analysis are provided. Finally, Section 5 is the conclusion.

2. METHODS AND EVALUATION METRICS

2.1. Problem definition

Complex web networks are always threatened with malfunctions or disruptions due to technical or intentional failure or attacks, which can eventually lead to total system failure. A graph that has the nodes of those web pages and the edges pointed to the link between those pages can be used to study behavior for several reasons. One of these is the disruption or attack on hyperlinks between web pages that will cause web pages to crash.

For a given simple web network represented as $G = (V, E), V$ denotes the set of all nodes (web pages) and E denotes the set of all hyperlinks. After attack or failure, a percentage of hyperlinks E called E^A is removed from the web network and the goal is to add a set $E^R$of redundant hyperlinks to the network that restores the web efficiency to its original state before attack.

In this section, attack strategies, retrieval strategies, and evaluation metrics are proposed.

2.2. Attack Strategies

Real web networks suffer from various kinds of failures or attacks. Here we consider representative attack strategies [22]. The attack strategies to complex networks are generally proposed based on network characteristics such as degree centrality, betweenness centrality, eigenvector, closeness centrality, entropy, and so on [29-33]. Here both random hyperlink failure and some targeted hyperlink attacks strategies are considered which are defined as below and for each network, 15% of the hyperlinks are removed based on these attack strategies.

(1) Random Attack (RA): This is an ordinary attack strategy and hyperlinks are accidentally removed from the web network. This strategy is often used as a baseline for other strategies.

(2) Edge Betweenness Attack (EBA): This attack strategy evaluates the importance of hyperlinks based on the edge betweenness centrality. The higher the edge betweenness of a link the more important it is. The edge betweenness is defined as below:

$$BC_e = \sum_{i,j \in V} \frac{\sigma(i,j|e)}{\sigma(i,j)} \tag{1}$$

Where $\sigma(i,j)$ is the number of shortest paths between nodes(web pages) i and j, and $\sigma(i,j|e)$ is the number of those paths that also pass through hyperlink e.

(3) Preferential Attachment (PA): This attack strategy is defined based on the node (web page) degrees. If i and j are the end webpages of the hyperlink e, the weight of hyperlink e is defined as below:

$$PA = k_i . k_j \tag{2}$$

Where k_i and k_j represent the degree of nodes (web pages) i and j, respectively.

(4) Similarity-based Attack (SA): This attack strategy is based on the similarity measures that are mainly defined in link prediction methods. Here, the Common Neighbors is used as the similarity measure. The higher the number of common neighbors of a node (web page) pairs the more important the hyperlink between them is.

2.3. Retrieval strategies

To develop a resilient characteristic for the network against random and targeted attacks, many different strategies have been suggested and applied to complex networks structures [7]. Variable retrieval strategies can be categorized to retrieve based on resilience consideration, failures characteristics, networks properties, and retrieval priorities. Some retrieval strategies can be members of more than one category group. The method proposed in this article can be categorized as a retrieval strategy based on retrieval priorities which is an important factor in instant retrieval of different networks such as web and blog networks. In fact in this category, retrieval strategies follow rules based on priority to repair failed or attacked hyperlinks or elements. The selection rules consider different factors such as betweenness, distance, load, or crucial attacked elements in the network. In link prediction methods, the goal is identifying the hyperlinks that can be replaced with removed hyperlinks (via attack or failure) instantly, so that the network performance with different criteria can be restored to a pre-attack state as soon as possible.

2.4. Redundant hyperlinks identification

After a targeted attack or accidental failure, the ordinary way to recover the network is by choosing random links to replace the removed ones. The most effective way to recover damaged links instantly is to check network efficiency at every step of adding redundant links and check the performance of the recovered network. In this paper, the main approach is based on network retrieval and predicting redundant links. This goal is done via proposed methods which are based on link prediction similarity measures to improve network efficiency during the recovery state.

Assuming the structure of the network after an intentional attack or random failure, the proposed link prediction-based methods are supposed to predict the appropriate redundant links to be replaced with the original ones in the situation of the real attack. So a list of reserved links is considered that can be added to the network after failure or attack based on the strategies defined above. In another word, first, it is assumed that the network has been under attack based on the attack strategies and then try to find the appropriate links that can increase the network efficiency to the same condition before an attack. The existing link prediction methods and a proposed method based on the local information of removed links are used to choose the appropriate list of reserved links:

- *Common Neighbors(CN)*

$$s_{ij}^{CN} = \left\| \Gamma_i \cap \Gamma_j \right\| \tag{3}$$

where Γ_i is the set of neighbors of node i and $\| . \|$ indicates the number of nodes in a set.

- *Resource Allocation(RA)*

$$s_{ij}^{RA} = \sum\nolimits_{k \in \Gamma_i \cap \Gamma_j} \frac{1}{\|\Gamma_k\|} \tag{4}$$

- *Preferential Attachment(PA)*

$$s_{ij}^{PA} = \|\Gamma_i\| \times \|\Gamma_j\| \tag{5}$$

- *Cannistraci-Resource-Allocation (CRA)*

$$s_{ij}^{CRA} = \sum\nolimits_{k \in \Gamma_i \cap \Gamma_j} \frac{\|iLCL(k)\|}{\|\Gamma_k\|} \tag{6}$$

where $iLCL(k)$ is the internal local community links defined in[34]

- *CH2-L2*

$$s_{ij}^{CH2-L2} = \sum\nolimits_{k \in L_2} \frac{1 + di_k}{1 + de_k} \tag{7}$$

Where k is the intermediate node on the path of length two (L_2) between i and j and di_k is the respective internal node degree and de_k is the respective external node degree. More information about this measure is provided in [34]

- *Hyperbolic distance (HP):* This measure uses the hyperbolic distance of links based on the HyperMap method which is based on Maximum Likelihood Estimation and finds the radial and angular coordinates for all nodes and maximizes the likelihood:

$$L = \prod_{1 \le i < j \le N} p(x_{ij})^{\alpha_{ij}} \left[1 - p\big(x_{ij}\big)\right]^{1-\alpha_{ij}} \tag{8}$$

where x_{ij} is defined as the hyperbolic distance between pair i, j:

$$x_{ij} = \operatorname{arccosh}\big(\cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos\Delta\theta_{ij}\big) \tag{9}$$

$$\approx r_i + r_j + 2ln \sin\big(\Delta\theta_{ij}/2\big) \approx r_i + r_j + 2ln\, (\Delta\theta_{ij}/2)$$

where

$$\Delta\theta_{ij} = \pi - |\pi - |\theta_i - \theta_j|| \tag{10}$$

and $p(x_{ij})$ is the Fermi-Dirac connection probability:

$$p(x_{ij}) = \frac{1}{1 + e^{\frac{1}{2T}(x_{ij}-R)}} \tag{11}$$

where $R \sim \ln N$ [35].

The pseudo-code of the proposed method is shown in Figure 1.

After the attack to hyperlinks or failure:

1. Identify the hyperlinks between the web pages, convert the web pages and hyperlinks into the graph.
2. Approximate the hyperbolic coordinates of each node (web page or blog).
3. Compute the matrix H of the hyperbolic distance of the existing hyperlinks in the web network.
4. E^A is the list of removed hyperlinks via failure or attack and E^R is the list of reserved hyperlinks for retrieval.
5. For each $(i,j) \in E^A$:
   a. Γ_i is the list of neighbors of node (web page or blog) i and Γ_j is the list of neighbors of node (web page or blog) j
   b. for each $k \in (\Gamma_i \cup \Gamma_j) - (\Gamma_i \cap \Gamma_j)$ (k is a neighbor of i or a neighbor of j and not both)

   $$(k,k') = \arg\max \frac{degree(k)}{H(i,k)+H(k,j)} \tag{12}$$

   $where: k' = i \quad if\ k \in \Gamma_j\ and\ k' = j \quad if\ k \in \Gamma_i$

   $$\rightarrow (k,k') \in E^R$$
6. Add a hyperlink (k, k') to the web network (instead of attacked or failed hyperlink (i,j)).
7. Repeat step 5 for all members of E^A.

Figure 1: The pseudo-code of the proposed method

Figure 2 shows an example of the proposed method. When the hyperlink between the red nodes(web pages 3,6) is damaged, all neighbors(other web pages) of the red nodes are the candidate for making a new hyperlink with one of the red nodes and the one with the highest computed score is the winner.

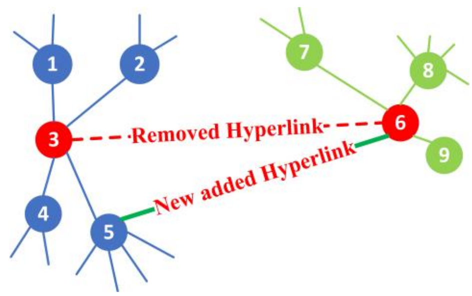


Figure 2: Example of the proposed method. Nodes 1 to 9 are web pages or blogs that are connected via hyperlinks.

2.5. Evaluation Metrics

To evaluate the proposed methods two metrics are used that are defined as below:

- ***Efficiency:***

$$E = \frac{EF^R}{EF^A} \tag{13}$$

Where E^A is the efficiency of the web network before the attack and E^R is the efficiency of the web network after retrieval. The efficiency is defined as the efficiency of the web network as below:

$$EF = \frac{1}{|N|(|N|-1)} \sum_{i,j \in N} \frac{1}{d_{ij}} \tag{14}$$

The greater the F is the better the method performance is.

- ***Retrieval power(RP):***

$$RP = \frac{|E^{Recovery}|}{|E^{Attacked}|} \tag{15}$$

where $|E^{Attacked}|$ is the number of removed hyperlinks after the attack and $|E^{Recovery}|$ is the number of redundant hyperlinks that should be added to the web network to retrieve the network to its original efficiency before attack or failure. The smaller the RP is the better the method performance is. As it can be seen in the Result section, in some methods $|E^{Recovery}|$ is much less than $|E^{Attacked}|$ that shows their good performance. But in some other methods $|E^{Recovery}|$ is much more than $|E^{Attacked}|$ that shows their poor performance. For the cases that $RP > 1$ the performance of the method is not acceptable and in this case, this measure is not computed and is shown with "**-**".In cases where $RP = 0$ retrieval is not achieved by the relevant algorithm.

3. DATASETS

To evaluate the proposed method and compare them with other existing methods, the information of four real web networks is used. To approximate the underlying hyperbolic of the complex networks via the HyperMap method two parameters γ and T are required. Parameter γ is the power-law degree distribution exponent which is approximated via the method introduced by Clauset et al.[36], and T is the temperature that is estimated using the Nonuniform Popularity × Similarity Optimization N-PSO model [27]. In the following table 1, we provide an explanation of these web or blog networks.

Table 1: Summary of the datasets

Name	#Nodes	Node Type	#Edges	Edge Type	Ref.	Description
AIDS blogs (2005)	146	Blog	187	Hyperlink	[37]	A network of hyperlinks among blogs related to AIDS, patients, and their support networks, collected by Gopal over a three-day period in August 2005.
moreno_blogs	1224	Blog	19025	Hyperlink	[38, 39]	This directed network contains front-page hyperlinks between blogs in the context of the 2004 US election.
web-polblogs	643	Blog	2300	Hyperlink	[40]	The graph data sets are donated by a number of different authors and organizations and in many cases have provided the citation information that should be used upon request.
Amazon pages(2012)	2880	Web page	5037	Hyperlink	[41]	A small sample of web pages from Amazon.com and its sister companies. The manner in which this network was sampled is unclear, and the direction of the hyperlink has been discarded.

4. EXPERIMENTAL RESULTS AND ANALYSIS

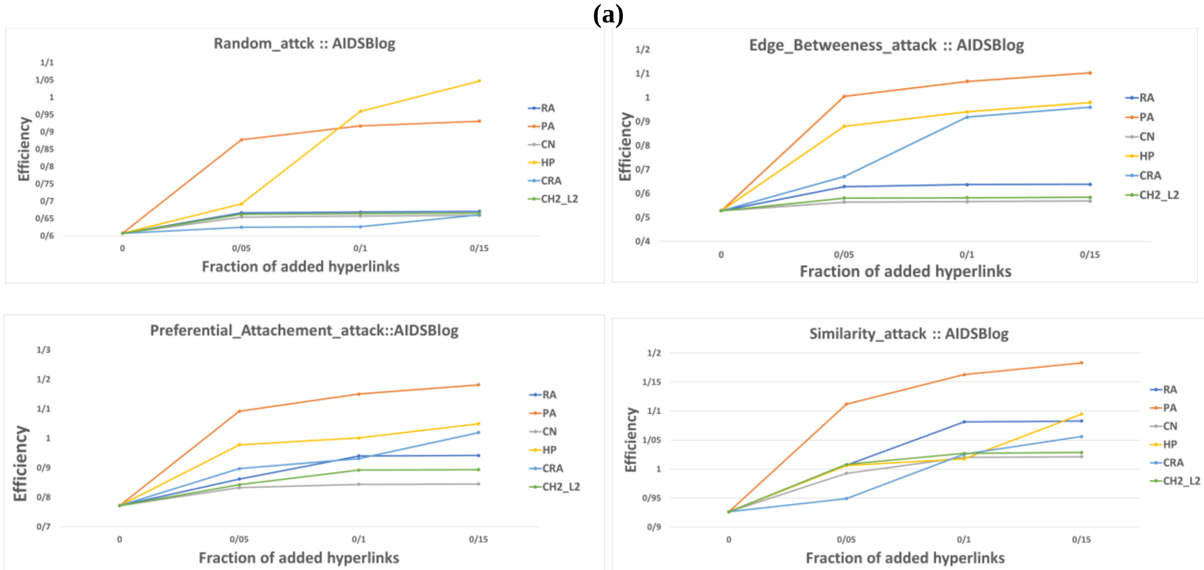
The experimental results of the proposed methods on four real web networks are presented in this section. For each network, at first 15% of hyperlinks are removed via different attack strategies explained before, and then the predicted set of redundant hyperlinks are added gradually and the evaluation metrics are used to investigate the performance of the retrieved network.

4.1. Comparison of retrieval methods

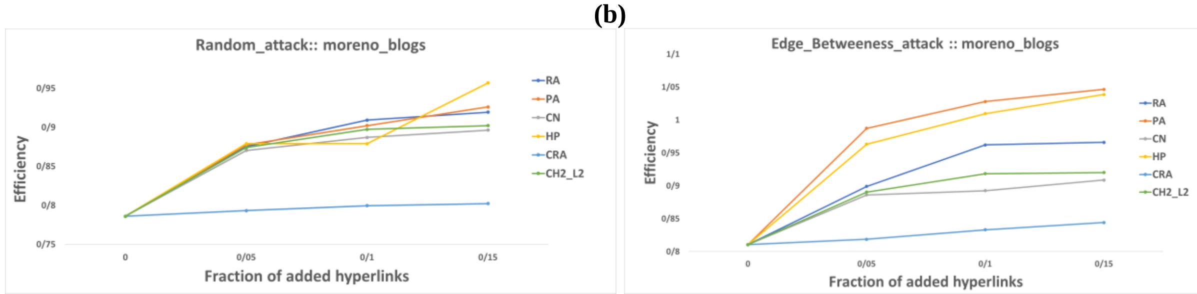
Figure 3 shows the efficiency of different web networks via the percentage of added hyperlinks to the attacked network in the range of [0, 0.15]. As it is shown, the proposed retrieval methods result in increasing the efficiency of all networks and as it can be distinguished, the efficiency of the proposed retrieval method (HP) performs significantly better in the case of attack based on Edge Betweenness. When the attack is considering the edge efficiency betweenness as the important factor of a hyperlink, removing those hyperlinks results in crucial damages in the network efficiency and it may cause the network to become disconnected. So adding redundant hyperlinks that are locally near to the damaged hyperlink helps the efficiency of the attacked network significantly. So the proposed method could have an impressive impact on restoring the web network structure as soon as possible.

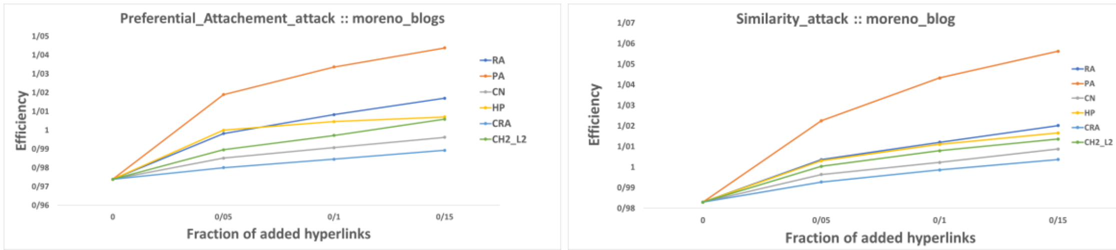
Table 2 shows the Retrieval power of different recovery methods while restoring the attacked network. The number in each cell shows the percentages of hyperlinks that should be added to the attacked network to restore the efficiency of the network to the state before the attack. In most cases except the proposed method (HP), adding 15% of redundant hyperlinks to the damaged network can not achieve the original efficiency of the network. But in the case of (HP) adding less or equal percentage of redundant hyperlinks to the damaged network result in the original efficiency of the attacked network. Since in the attacking situation the instant and cheap recovery strategy is of high importance, so the proposed method profits of a very important advantage.

(a)

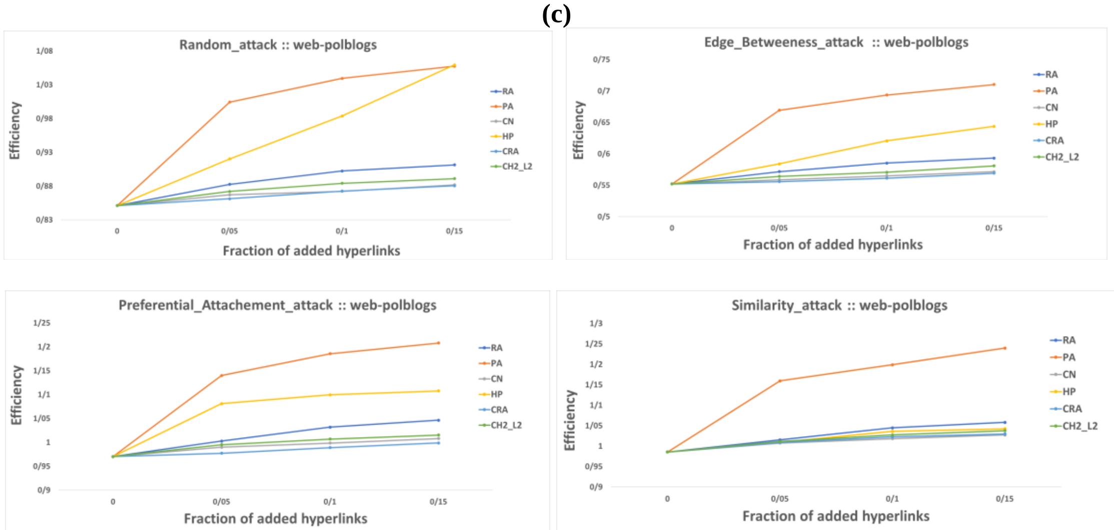


(b)





(c)



(d)

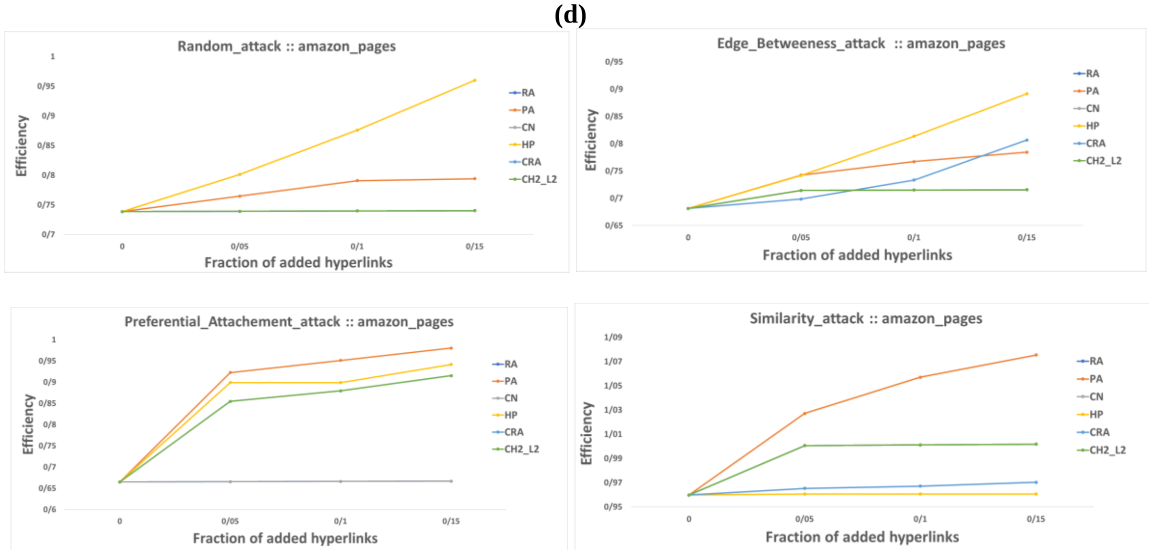


Figure 3: Efficiency of proposed methods under different kinds of attacks in four real web/blog networks (a) AIDS blogs (2005) (b) moreno_blogs and (c) web-polblogs (d) Amazon_pages 2012

Table 2: Retrieval Power of proposed retrieval strategy under different kinds of attack

Network	Attack Strategy	HP	RA	PA	CN	CRA	CH2_L2
AIDS blogs (2005)	*Random Failure*	1.0	0	-	0	-	-
	Edge Betweenness	1.0	0	0.3	0	-	-
	Preferential Attachment	1.0	0	0.26	-	0.81	-
	Similarity	0.96	0.26	0.04	0.52	0.56	0.3
moreno_blogs	*Random Failure*	0.97	0	-	0	-	0
	Edge Betweenness	1.0	-	0.37	0	-	-
	Preferential Attachment	1.0	0.39	0.1	-	-	0.72
	Similarity	1.0	0.23	0.09	0.52	0.75	0.31
web-polblogs	*Random Failure*	1.0	-	0.31	0	0	-
	Edge Betweenness	0.96	0	0	0	0	0
	Preferential Attachment	1.0	0.26	0.01	0.7	1.01	0.45
	Similarity	1.0	0.07	0.01	0.13	0.13	0.11
Amazon_pages	*Random Failure*	1.0	0	0	0	0	0
	Edge Betweenness	1.0	0	0	0	-	0
	Preferential Attachment	1.0	0	-	0	-	-
	Similarity	0.3	0	0	0	-	0

4.2. Comparison of the network structure under different retrieval strategy

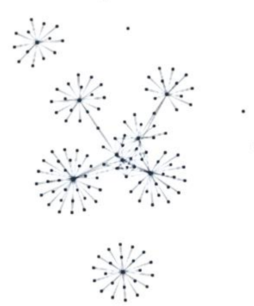
To evaluate the efficiency of the proposed retrieval method for different sizes of attacked hyperlinks, the fraction of damaged hyperlinks is changed in the range of [0.1,0.5]. Figure 4 shows the network structure of AIDS blogs (2005) for example before and after the EBA attack as well as after blog network retrieval using a different retrieval strategy. As can be seen, the proposed HP retrieval method preserves the network structure before the attack and has a structure closer to the original blog network before the attack. In other methods, node rupture is observed after network retrieval.

(a)The initial Network (Before attack)

(b) Network after EBA attack

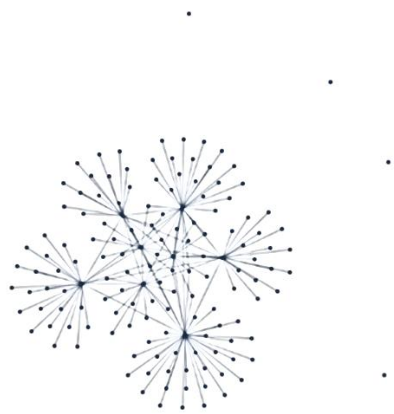
(c) Network after HP Retrieval Strategy

(c) Network after CH2_L2 Retrieval Strategy

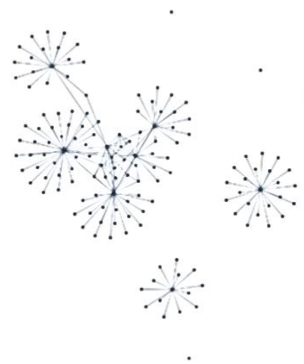


(d) Network after CN Retrieval Strategy

(e) Network after CRA Retrieval Strategy



(f) Network after PA Retrieval Strategy



(g) Network after RA Retrieval Strategy

Figure 4: A comparison of the structure of AIDS blogs (2005) (N=146) when hyperlinks have been removed by EBA attack in different retrieval strategy (Visualization generated using python pyvis lib[42]).

5. CONCLUSION

Link prediction methods have been used widely in solving different kinds of problems. In this paper, some existing and proposed methods are suggested to retrieve the web network structure after random or targeted attacks. Firstly, some attack strategies are defined, and based on them some retrieval strategies based on the link prediction similarity measures are proposed. The main proposed method is based on the hyperbolic geometry and local information of the attacked network. The underlying hyperbolic geometry of complex networks uses two important parameters of similarity and popularity of nodes (web pages or blogs) that both play important role in solving the link prediction problem. On the other side using the local information of the damaged hyperlinks helps the method to choose appropriate hyperlinks to be replaced with the removed ones as soon as possible. The results show that the proposed method outperforms the other existing link prediction methods in most cases and can achieve the efficiency of the network before the attack even with adding less number of hyperlinks than the removed ones. Also, the network structure after the attacks in the proposed retrieval method based on hyperbolic geometry is largely preserved and will be closer to the main network before the attacks.

REFERENCES

1. Matisziw, T.C., A.T. Murray, and T.H. Grubesic, *Strategic network restoration.* Networks and Spatial Economics, 2010. **10**(3): p. 345-361.
2. Chaoqi, F., et al., *Complex networks under dynamic repair model.* Physica A: Statistical Mechanics and its Applications, 2018. **490**: p. 323-330.
3. Hu, F., et al., *Recovery of infrastructure networks after localised attacks.* Scientific reports, 2016. **6**(1): p. 1-10.
4. Yu, H. and C. Yang, *Partial network recovery to maximize traffic demand.* IEEE communications letters, 2011. **15**(12): p. 1388-1390.
5. Yodo, N. and P. Wang, *Engineering resilience quantification and system design implications: A literature survey.* Journal of Mechanical Design, 2016. **138**(11).
6. Majdandzic, A., et al., *Spontaneous recovery in dynamical networks.* Nature Physics, 2014. **10**(1): p. 34-38.
7. Afrin, T. and N. Yodo, *A concise survey of advancements in recovery strategies for resilient complex networks.* Journal of Complex Networks, 2019. **7**(3): p. 393-420.

8. Liben-Nowell, D. and J. Kleinberg, *The link-prediction problem for social networks.* Journal of the American society for information science and technology, 2007. **58**(7): p. 1019-1031.
9. Clauset, A., C. Moore, and M.E. Newman, *Hierarchical structure and the prediction of missing links in networks.* Nature, 2008. **453**(7191): p. 98-101.
10. Fu, C., et al., *Link weight prediction using supervised learning methods and its application to yelp layered network.* IEEE Transactions on Knowledge and Data Engineering, 2018. **30**(8): p. 1507-1518.
11. Lü, L., et al., *Toward link predictability of complex networks.* Proceedings of the National Academy of Sciences, 2015. **112**(8): p. 2325-2330.
12. Lü, L. and T. Zhou, *Link prediction in complex networks: A survey.* Physica A: statistical mechanics and its applications, 2011. **390**(6): p. 1150-1170.
13. Samei, Z. and M. Jalili, *Discovering spurious links in multiplex networks based on interlayer relevance.* Journal of Complex Networks, 2019. **7**(5): p. 641-658.
14. Sales-Pardo, M., et al., *Extracting the hierarchical organization of complex systems.* Proceedings of the National Academy of Sciences, 2007. **104**(39): p. 15224-15229.
15. Airoldi, E.M., et al., *Mixed membership stochastic blockmodels.* Journal of machine learning research, 2008. **9**(Sep): p. 1981-2014.
16. Holland, P.W., K.B. Laskey, and S. Leinhardt, *Stochastic blockmodels: First steps.* Social networks, 1983. **5**(2): p. 109-137.
17. Heckerman, D., C. Meek, and D. Koller, *Probabilistic entity-relationship models, PRMs, and plate models.* Introduction to statistical relational learning, 2007: p. 201-238.
18. Neville, J., *Statistical models and analysis techniques for learning in relational data.* 2006.
19. Herrgård, M.J., et al., *A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology.* Nature biotechnology, 2008. **26**(10): p. 1155-1160.
20. Linden, G., B. Smith, and J.Y.A. Com. *Industry report: Amazon. com recommendations: Item-to-item collaborative filtering*. in *IEEE Distributed Systems Online*. 2003. Citeseer.
21. Radicchi, F., et al., *Defining and identifying communities in networks.* Proceedings of the national academy of sciences, 2004. **101**(9): p. 2658-2663.
22. Wang, K., L. Li, and C. Pu, *Robustness of link prediction under network attacks.* arXiv preprint arXiv:1811.04528, 2018.
23. Krioukov, D., et al., *Hyperbolic geometry of complex networks.* Physical Review E, 2010. **82**(3): p. 036106.
24. Papadopoulos, F., et al., *Popularity versus similarity in growing networks.* Nature, 2012. **489**(7417): p. 537-540.
25. Papadopoulos, F., C. Psomas, and D. Krioukov, *Network mapping by replaying hyperbolic growth.* IEEE/ACM Transactions on Networking, 2014. **23**(1): p. 198-211.
26. Alessandro, M. and C.C. Vittorio, *Leveraging the nonuniform PSO network model as a benchmark for performance evaluation in community detection and link prediction.* New Journal of Physics, 2018. **20**(6): p. 063022.
27. Muscoloni, A. and C.V. Cannistraci, *A nonuniform popularity-similarity optimization (nPSO) model to efficiently generate realistic complex networks with communities.* New Journal of Physics, 2018. **20**(5): p. 052002.
28. Samei, Z. and M. Jalili, *Application of hyperbolic geometry in link prediction of multiplex networks.* Scientific reports, 2019. **9**(1): p. 1-11.
29. Albert, R., H. Jeong, and A.-L. Barabási, *Error and attack tolerance of complex networks.* nature, 2000. **406**(6794): p. 378-382.
30. Cohen, R., et al., *Breakdown of the internet under intentional attack.* Physical review letters, 2001. **86**(16): p. 3682.
31. Crucitti, P., et al., *Error and attack tolerance of complex networks.* Physica A: Statistical mechanics and its applications, 2004. **340**(1-3): p. 388-394.
32. Allesina, S. and M. Pascual, *Googling food webs: can an eigenvector measure species' importance for coextinctions?* PLoS computational biology, 2009. **5**(9).
33. Iyer, S., et al., *Attack robustness and centrality of complex networks.* PloS one, 2013. **8**(4).
34. Muscoloni, A., I. Abdelhamid, and C.V. Cannistraci, *Local-community network automata modelling based on length-three-paths for prediction of complex network structures in protein interactomes, food webs and more.* bioRxiv, 2018: p. 346916.

35. Kleineberg, K.-K., et al., *Hidden geometric correlations in real multiplex networks.* Nature Physics, 2016. **12**(11): p. 1076-1081.
36. Clauset, A., C.R. Shalizi, and M.E. Newman, *Power-law distributions in empirical data.* SIAM review, 2009. **51**(4): p. 661-703.
37. Gopal, S., *The evolving social geography of blogs*, in *Societies and cities in the age of instant access*. 2007, Springer. p. 275-293.
38. Kunegis, J. *Konect: the koblenz network collection*. in *Proceedings of the 22nd international conference on World Wide Web*. 2013.
39. Adamic, L.A. and N. Glance. *The political blogosphere and the 2004 US election: divided they blog*. in *Proceedings of the 3rd international workshop on Link discovery*. 2005.
40. https://networkrepository.com/web-polblogs.php.
41. Šubelj, L. and M. Bajec, *Ubiquitousness of link-density and link-pattern communities in real-world networks.* The European Physical Journal B, 2012. **85**(1): p. 1-11.
42. https://pyvis.readthedocs.io/en/latest/.