

Viki-mühitdə gizli sosial şəbəkələrin aşkarlanması metodu

Ramiz Alıquliyev¹, İradə Ələkbərova²

^{1,2}AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan

¹a.ramiz@science.az, ²airada.09@gmail.com

Xülasə– Məqalədə viki-mühitdə fəaliyyət göstərən sosial şəbəkələrin analizi ilə bağlı bəzi yanaşmalar şərh olunur. İnformasiya qarşılıqlı əlaqəsində iştirak edən gizli sosial şəbəkələrin aşkarlanması üçün metod təklif edilir.

Açar sözlər - viki-texnologiyalar, sosial şəbəkələr, informasiya qarşılıqlı əlaqəsi, viki-səhifə, loq-fayl

I. GİRİŞ

Viki-texnologiyaları ilə idarə olunan Vikipediya virtual ensiklopediyasının dinamikliyi, viki-səhifələrin sayının və həcmnin son illər həddən artıq çoxalması, ensiklopedik məqalələrin yaradılması prosesinə milyonlarla İnternet istifadəçisinin cəlb olunması viki-mühitdə informasiya təhlükəsizliyi və viki-səhifələrin keyfiyyəti ilə bağlı bir sıra problemlərin meydana çıxmasına səbəb olmuşdur [1, 2, 3].

Uord Kenninqem və həmmüəllifi Bo Leyf "The Wiki Way: Quick Collaboration on the Web" kitabında "Viki" konsepsiyasının mahiyyətini izah edərkən yazırlar [4]:

- viki bütün istifadəçilərə yalnız brauzerdən istifadə etməklə virtual ensiklopediyanın yaradılması işində iştirak etməyi təklif edir.
- viki müxtəlif səhifələr arasında əlaqəni təmin edir və bu məqalələrin mövcud olub-olmaması mühüm rol oynamır.
- viki məqalənin yaradılması və dəyişdirilməsində kollektiv biliyə üstünlük verir.

Nəzərə almaq lazımdır ki, "kollektiv biliyə üstünlük verən" viki-layihələrin istifadəçiləri müxtəlif ideologiyaya, dinə, dilə, təhsilə, irqə malik, dünyanın müxtəlif yerlərində yaşayan insanlardır. Bu səbəbdən onların hər hansı hadisəyə münasibəti yaratdıqları və ya dəyişdirdikləri informasiyanın məzmununda və keyfiyyətində əks olunur [5].

Viki-mühitdə informasiya qarşılıqlı əlaqələrinin kəskinləşməsi və uzun müddət davam etməsi Vikipediya və onun layihələrinin (Vikilüğət, Vikisitə, Vikikitəb, Vikimənbə, Vikixəbər, Vikiversitet və s.) inkişafına, ensiklopedik məqalələrin keyfiyyətinə, Vikipediya fəlsəfəsinin əsasını təşkil edən istifadəçilər arasında qarşılıqlı hörmət və tərəfsizlik prinsiplərinə ciddi zərbə vurmaqdadır. İstifadəçilər arasında konfliktlərin aradan qaldırılması, ensiklopedik məqalələrin vandalizm hallarından, əks təbliğət, dezinformasiya tipli məlumatlardan qorunması üçün viki-mühitin aktiv istifadəçiləri də cəlb olunmaqla administratorlar tərəfindən müxtəlif qaydalar işlənmişdir [6].

Vikipediya layihələrinə daxil edilən istənilən növ kontent (mətn, şəkil, audio- və video-fayllar) avtomatik Vikipediyanın verilənlər bazasına daxil edilir. Bu baxımdan istifadəçilərin viki-layihələrdəki davranışlarına nəzarət etmək: kimlərlə və hansı mövzularda müzakirələr

aparmaları, hansı məqalələrdə və hansı dəyişikliklər etmələri haqqında məlumat əldə etmək mümkündür. Mövcud ensiklopedik məqalələrin keyfiyyətini təyin etmək və informasiya qarşılıqlı əlaqəsində iştirak edən gizli sosial qrupların aşkarlanması mühüm məsələdir. Vikimetrik tədqiqatlarla məşğul olan bir çox mütəxəssislər etiraf edirlər ki, cəmiyyətdə baş verən hadisələr, problemlər və münasibətlər öz əksini viki-mühitdə tapmaqdadır. Viki-istifadəçilərin davranışları onların məqsədini, maraq dairəsini və fəaliyyətlərində informasiya müharibəsi texnologiyalarından nə dərəcədə istifadə edib-etmədiklərini üzə çıxarır [7, 8].

II. VIKI-MÜHİTDƏ FƏALİYYƏT GÖSTƏRƏN SOSIAL ŞƏBƏKƏLƏRİN ANALİZİ

Viki-mühitdə fəaliyyət göstərən sosial şəbəkələrin və istifadəçilərin nüfuzunun (*user reputation*) analizi, həmçinin, istifadəçilər arasında əməkdaşlıq və münasibətlərin kontentə təsiri müxtəlif metod və alqoritmlərlə tədqiq edilmişdir [9, 10]. Kaliforniya Universitetinin professoru Sara Cavanmardinin viki-istifadəçilər arasında sosial münasibətləri və hansının daha çox nüfuz malik olmasını təyin etmək üçün təklif etdiyi model təsnifatlandırma metodları əsasında viki-mühitdə vandalları və təcrübəsiz istifadəçiləri təyin etməyə imkan verir [11].

Viki-mühitdə sosial münasibətləri, mübahisələri və əməkdaşlığı ölçmək üçün digər yanaşma İsrail Universitetinin tədqiqatçıları Hagit Meşa-Tal və Edna Tal-Əlhəsid tərəfindən təklif edilmişdir [12]. Bu metodda 3 əlamətdən istifadə edilir:

1. Viki-istifadəçilərin sayı;
2. İnteraktivlik, yəni müəyyən vaxt intervalında istifadəçilərin viki-səhifələrdə etdikləri redaktələrin sayı;
3. İntensivlik, yəni redaktələr nəticəsində viki-səhifədə olan dəyişikliklərin sayı.

Redaktə dedikdə, viki-səhifələrdə edilən dəyişikliklər – informasiyanın silinməsi, digər informasiya ilə əvəz olunması və ya yeni məlumatların əlavə edilməsi nəzərdə tutulur.

Meşa-Tal və Tal-Əlhəsid tərəfindən təklif olunan metod əsasında "redaktələr müharibəsi" (*Edit warring*) [13], viki-istifadəçilərin aktivliyi və onların viki-səhifələrdə etdikləri redaktələrin keyfiyyəti təyin edilir. Redaktələr müharibəsi zamanı viki-səhifələrdə edilən dəyişikliklər digər istifadəçi qrupu tərəfindən qəbul edilmir və nəticədə iki və daha artıq qruplar arasında qarşılıqlı yaranır. Viki-səhifədə hər hansı bir qrup tərəfindən daxil edilən informasiya digər qrup tərəfindən silinir və bu proses müəyyən müddət davam edir.

Nəticədə belə bir vəziyyət yaranır: məqalə üzərində işləyən istifadəçilərin sayı çoxdur və çox sayda redaktələr edilmişdir, lakin son nəticədə ensiklopedik məqalənin həcmi və keyfiyyəti az dəyişmişdir.

Bəzi vikimetrik tədqiqatlarda viki-istifadəçilərin nüfuzu və qruplar əsasən onların davranışları ilə təyin edilir. Nüfuzun ölçülməsi vasitəsilə vandalizmin təyini və məqalələrin keyfiyyətinin analizi ideyasını ilk dəfə Kaliforniyanın Santa-Kruz Universitetinin professorları Luka de Alfaro, Yan Pay və Tomas Adler irəli sürmüşlər. Onlar viki-mühitdə ensiklopedik məqalələrin keyfiyyətinin və istifadəçilərin nüfuzunun analizində “nüfuz əsaslanan sistem” (*a reputation-based system, Wiki-Trust*) təklif etmişlər [14, 15].

Təklif olunan yanaşma ilk baxışdan çox sadədir və müəyyən vəziyyətlərdə Sara Cavanmardi tərəfindən təklif edilən alqoritmə bənzəyir: istifadəçilərin nüfuzu onların etdiyi dəyişikliklərin nə dərəcədə qəbul olunmasından asılıdır [11]. Wiki-Trust həm də vandalizmə məruz qalan ensiklopedik məqalələri təyin etmək üçün nəzərdə tutulmuşdur. Burada vandalizm hallarının təyini anonimlərin (qeydiyyatdan keçmədən fəaliyyət göstərən istifadəçilərin) və qeydiyyatdan keçmiş yeni istifadəçilərin təsnifatlandırılması yolu ilə həyata keçirilir [16, 17].

Lakin, yuxarıda göstərilən yanaşmalardan heç biri informasiya qarşılıqlaşdırmasında iştirak edən gizli sosial şəbəkənin aşkarlanmasında effektiv deyildir. Belə ki, təklif olunan metodlardan yalnız informasiya qarşılıqlaşdırmasının və vandalizm hallarının yeni istifadəçilər və anonimlər tərəfindən həyata keçirildiyi halda istifadə edilməsi nəzərdə tutulur. Lakin araşdırmalar göstərir ki, informasiya qarşılıqlaşdırmasında yalnız anonimlər və yeni istifadəçilər iştirak etmirlər [15, 18].

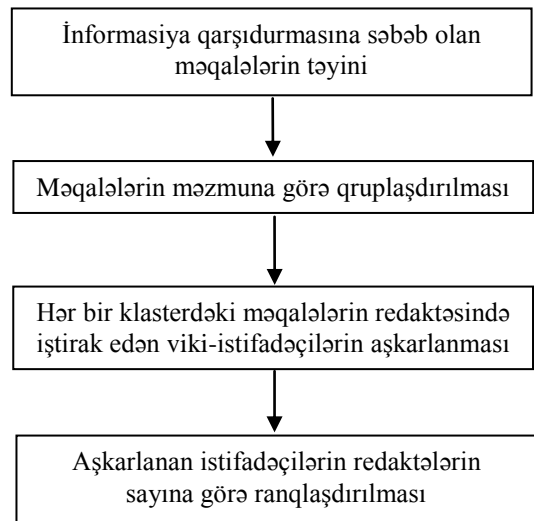
III. GİZLİ SOSIAL ŞƏBƏKƏNİN AŞKARLANMASI MODELİ

Mübahisə və qarşılıqlaşdırma yaradan viki-səhifələr fəal və təcrübəli istifadəçilərin də maraqlı obyektinə çevrilirlər. Bunları nəzərə alaraq, məqalədə viki-mühitdəki gizli sosial qrupları aşkarlamaq üçün yeni metod təklif edilmişdir.

Viki-mühitdə gizli sosial şəbəkələri aşkarlamaq üçün məsələnin mərhələlərlə həll olunması təklif olunur. Məsələnin həlli informasiya qarşılıqlaşdırmasına səbəb olan viki-səhifələrin aşkarlanması, konfliktli viki-səhifələrin məzmununa görə qruplaşdırılması, hər bir qrupdakı səhifələrin redaktəsində iştirak edən istifadəçilərin aşkarlanması və fəaliyyətlərinin analizini həyata keçirməklə mümkündür (şəkil 1).

Tədqiqatda istifadə olunan göstəricilər:

- ensiklopedik məqalələrin həcmi (baytlarla);
- məqalələrin müzakirə səhifələrinin həcmi (baytlarla);
- məqalənin predmetini təyin edən ilk abzasdakı sözlər;



Şəkil 1. Viki-mühitdə gizli sosial şəbəkələrin aşkarlanması modeli

- məqalənin yaradılmasında və redaktəsində iştirak edən istifadəçilərin sayı;
- məqalədə edilən redaktələrin sayı;
- məqalədə "silib-bərpa etmək" (*Reverting*) əməliyyatlarının sayı.

“Bərpa etmək” viki-səhifələrdə müəyyən müddət öncə edilən redaktələrin ləğv edilməsi deməkdir [18].

IV. İNFORMASIYA QARŞILIQLAŞDIRMASINA SƏBƏB OLAN MƏQALƏLƏRİN TƏYİNİ

Viki-mühitdə informasiya qarşılıqlaşdırmasına səbəb olan məqalələrin aşkarlanması üçün hər bir məqalədəki "silib-bərpa etmək" əməliyyatları, məqalənin və müzakirə səhifəsinin həcmi nəzərə alınmalıdır. Belə ki, informasiya qarşılıqlaşdırmasına səbəb olan məqalələr aşağıda göstərilən əlamətlərə görə normal məqalələrdən fərqlənirlər :

1. Ensiklopedik məqalələrin müzakirə səhifələrinin həcmi məqalənin öz həcmindən daha böyük olur;
2. Məqalənin redaktəsi zamanı "silib-bərpa etmək" əməliyyatı həyata keçirilir. Vikipediya qaydalarına əsasən məqalədə üç və daha artıq "silib-bərpa etmək" əməliyyatı aparılıbsa anomaliya hesab edilir [13, 19].

Yuxarıda göstərilən şərtlərdən heç olmasa biri baş verərsə məqalə informasiya qarşılıqlaşdırmasına səbəb olan məqalələr siyahısına daxil edilməlidir.

Klasterləşmə alqoritmlərindən, o cümlədən qraflar nəzəriyyəsi ilə istifadə etməklə viki-səhifələri müxtəlif şərtlər daxilində qruplara ayırmaq mümkündür [20, 21]. İnformasiya qarşılıqlaşdırmasına səbəb olan məqalələr təyin edildikdən və qruplara ayrıldıqdan sonra bu məqalələrin yaradılmasında və redaktəsində iştirak edən sosial şəbəkələrin aşkarlanması məsələsi həll edilməlidir. Bunun üçün klasterlərə ayrılmış konfliktli məqalələrin loq-fayllarından istifadə etmək nəzərdə tutulmuşdur. Loq-faylları analiz etməklə məqalənin nə zaman, kim tərəfindən yaradılması və hansı redaktələrin edilməsi haqqında informasiya əldə etmək mümkündür [22].

Loq-faylın analizi zamanı məlum olur ki, istifadəçilərin bir hissəsi məqalədə yalnız bir və ya iki dəfə redaktə etdikləri halda, müəyyən sayda istifadəçilərin redaktələrinin sayı onlardır. İstifadəçi məqaləni redaktə etdikcə və hər dəfə “Save page” düyməsini vurduqca onun istifadəçi adı (*user name*) Vikipediyanın bazasında qeydiyyatda düşür.

Məqalədə etdikləri redaktələrin sayına görə viki-istifadəçiləri qruplara ayırmaq mümkündür [28]. İstifadəçilərin qruplaşdırılması nəticəsində informasiya qarşılıqlılaşmasına səbəb olan ensiklopedik məqalələrdə dəyişikliklər edən istifadəçiləri aşağıdakı siniflər üzrə təsnifatlandırmaq mümkündür:

- Təsədüfi istifadəçilər (redaktələrin sayı 2-dən çox deyil).
- Viki-səhifənin redaktəsini həyata keçirənlər (redaktələrin sayı 5-dən çox deyil).
- Viki-səhifələri nəzarətdə saxlayan gizli qruplar (redaktələrin sayı 5-dən artıq olanlar).

Gizli sosial şəbəkələri müəyyən etmək üçün viki-məqalədə yalnız maksimum redaktələri olan istifadəçilər nəzərə alınır.

V. NƏTİCƏ

Təklif edilən metod konfliktə səbəb olan ensiklopedik məqalələri nəzarətdə saxlayan gizli sosial şəbəkələri aşkar etməklə Vikipediya virtual ensiklopediyasında reallaşdırılan informasiya qarşılıqlılaşmalarının qarşısının alınmasına və viki-mühitdə informasiya təhlükəsizliyi probleminin həllinə yardım edə bilər. İnformasiya qarşılıqlılaşmasında istifadə olunan məqalələri, həmçinin, hər hansı ideologiyanın təbliği ilə məşğul olan qrupları aşkar etməklə ensiklopedik məqalələrin keyfiyyətini artırmaq, daha dolğun və düzgün informasiya ilə təmin olunmasına şərait yaratmaq mümkündür.

Viki-mühitdə gizli sosial şəbəkələrin aşkarlanması üçün təklif olunan metod yalnız viki-texnologiyası ilə yaradılan və idarə olunan layihələr üçün deyil, həmçinin web2.0 texnologiyası əsasında fəaliyyət göstərən bir çox sosial şəbəkələrin analizində də istifadə oluna bilər.

ƏDƏBİYYAT

- [1] Shachaf P., Hara N. Beyond vandalism: Wikipedia trolls // *Journal of Information Science*, 2010, vol. 36, no. 3, pp. 357–370.
- [2] Yasseri T., Sumi R., Rung A., Kornai A., Kertész J. Dynamics of Conflicts in Wikipedia // *PLoS ONE*, 2012, vol 7, no. 6, 2012, <http://arxiv.org/pdf/1202.3643v2.pdf>

- [3] West A.G., Kannan S., Lee I. STiki: an anti-vandalism tool for Wikipedia using spatio-temporal analysis of revision metadata / *Proceedings of the 6th International Symposium on Wikis and Open Collaboration (WikiSym '10)*, 2010, pp. 47–48.
- [4] Leuf B., Cunningham W. *The Wiki Way: Quick Collaboration on the Web*, Laflin, PA: Addison-Wesley, 2001, 200 p.
- [5] Luyt B., Tan D. Improving Wikipedia's Credibility: References and Citations in a Sample of History Articles // *Journal of the American Society for Information Science and Technology*, 2010, vol. 61, no. 4, pp. 715–722.
- [6] http://en.wikipedia.org/wiki/Wikipedia:What_Wikipedia_is_not
- [7] Holloway T., Božicevic M., Börner K. Analyzing and visualizing the semantic coverage of wikipedia and its authors // *Journal of Complexity*, 2007, vol. 12, no. 3, pp. 30-40.
- [8] Hara N., Shachaf P., Hew K.F., Cross-cultural analysis of the Wikipedia community // *Journal of the American Society for Information Science and Technology*, 2010, vol. 61, no. 10, pp. 2097–2108.
- [9] Moskaliuk J., Kimmerle J., Cress U. Collaborative knowledge building with wikis: the impact of redundancy and polarity // *Journal of Computers & Education*, 2012, vol. 58, no. 4, pp. 1049–1057.
- [10] Javanmardi S., Lopes C., Baldi P. Modeling user reputation in wikis // *Journal of Statistical Analysis and Data Mining*, 2010, vol. 3, no 2, pp. 126–139.
- [11] Javanmardi S., David W. McDonald, Lopes V. Vandalism detection in Wikipedia / *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, ACM Press., 2011. pp. 82–91.
- [12] Meishar-Tal H., Tal-Elhasid E. Measuring Collaboration in educational wikis – a methodological // *International Journal of Emerging Technologies in Learning*, 2008, vol. 3, pp. 46–49.
- [13] http://en.wikipedia.org/wiki/Wikipedia:Edit_warring
- [14] Adler B. T., Alfaro L. A content-driven reputation system for the Wikipedia / *Proceedings of the 16th International Conference on World Wide Web*, 2007, pp. 261–270.
- [15] Adler B. T., Alfaro L., Pye I. Detecting Wikipedia vandalism using WikiTrust / *Lab Report for PAN at CLEF 2010*, http://www.uni-weimar.de/medien/webi_s/research/events/pan-10/pan10-papers-final/pan10-vandalism-detection/adler10-detecting-wikipedia-vandalism-using-wikitrust.pdf
- [16] Adler B. T., Alfaro L., Mola-Velasco S.M., Rosso P., West A.G. Wikipedia Vandalism Detection: combining natural language, metadata, and reputation features / *Proceedings of the 12th International Conference on Intelligent Text Processing and Computational Linguistics, LNCS 6609*, Tokyo, 2011, pp. 277–288.
- [17] Alfaro L., Kulshreshtha A., Pye I., Adler B.T. Reputation systems for open collaboration // *Communications of the ACM*, 2011, vol. 54, no 8, pp. 81–87.
- [18] Алекперова И.Я., Анализ факторов влияющих на рейтинг статей в вики-среде // *Информационные Проблемы*, 2012, №2(6), s. 27–32.
- [19] <http://en.wikipedia.org/wiki/Wikipedia:Reviewing>
- [20] Ye N., Emran S. M., Chen Q., Vilbert S. Multivariate statistical analysis of audit trails for host-based intrusion detection / *IEEE, Transactions on Computers*, 2002, vol. 51, no 7, pp. 810-820.
- [21] Алгулиев Р.М., Алыгулиев Р.М., Алекперова И.Я., Кластерный подход к эффективному применению мультимедийных ресурсов в информационном противостоянии в вики-среде // *Автоматика и вычислительная техника*, 2014, №2, стр.48-63.
- [22] Iba T., Nemoto K., Peters B., Gloor P. A. Analyzing the creative editing behavior of Wikipedia editors: through dynamic social network analysis // *Procedia - Social and Behavioral Sciences*, 2010, vol. 2, issue 4, pp. 6441–6456.

Bu iş Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondunun maliyyə yardımı ilə yerinə yetirilmişdir –
Orant № EIF-RİTN-MQM-2/İKT-2-2013-7(13)-29/22/1