

# Vikimetrik tədqiqatların informasiya təhlükəsizliyində rolu

İradə Ələkbərova

*AMEA İnformasiya Texnologiyaları İnstitutu*

depart17@iit.ab.az

**Xülasə**— məqalədə informasiya müharibəsi texnologiyalarının viki-mühitə təsiri problemləri analiz edilmiş, vikimetrik tədqiqatlar ilə əlaqədar yanaşmalar araşdırılmışdır. Həmçinin, Vikipediya virtual ensiklopediyasında həlli vacib olan məsələlər və problemlər göstərilmişdir. Viki-mühitədə reallaşdırılan informasiya müharibəsi texnologiyalarının təsnifatı aparılmış, tədqiqat obyektləri göstərilmişdir.

**Açar sözlər**— *viki-texnologiya; Vikipediya; viki-mühit; vikimetrik tədqiqatlar; Wikimedia Fondu; informasiya müharibəsi*

## I. GİRİŞ

İnternet texnologiyaları sırasında viki-texnologiyanın verdiyi imkanlar kollektiv şəkildə bilik yaratmaqla İnternet şəbəkəsinə yeni bir mühitin – viki-mühitin gəlməsinə və sürətli inkişafına şərait yaratdı. Viki-texnologiya hər bir İnternet istifadəçisinə geniş imkanlar verdi. Belə ki, istifadəçi onlayn şəraitdə veb-səhifə yarada, fayl yükləyə, artıq mövcud səhifələrdəki mətni dəyişdirə və adi sosial şəbəkələrdə olduğu kimi öz şəxsi səhifəsini yaratmaqla müzakirələrdə iştirak edə bilər.

XXI əsrin texnologiyası sayılan viki-texnologiyanın imkanları eyni zamanda istifadəçilərin virtual məkandan müxtəlif məqsədlər üçün istifadə etməsinə şərait yaradır. Bu imkanlara ilk növbədə reklam xarakterli məlumatlar, spamlar, dezinformasiyanın daxil etməsi, informasiya müharibələri aiddir.

İstənilən viki-səhifənin istifadəçilər tərəfindən dəyişdirilə bilməsi, səhifələrə müxtəlif multimedia resurslarının, hiperkeçidlərin və şərhlərin daxil edilməsi imkanının verilməsi viki-mühitdə informasiya təhlükəsizliyi sahəsində problemlər yaratmaqdadır [1].

Viki-texnologiya əsasında yaradılan ən nəhəng və populyar layihə 2001-ci ilin yanvar ayından fəaliyyət göstərən və bu gün İnternet mühitinin ən tanınan 10 saytı sırasında yer tutan Vikipediya ensiklopediyasıdır. 2013-cü ilin aprel ayına olan məlumata görə Vikipediya virtual ensiklopediyası 285 dildə fəaliyyət göstərir və ensiklopedik məqalələrin sayı 25 milyondan artıqdır [2].

## II. VİKİ-MÜHİTDƏ TƏDQIQAT OBYEKTƏLƏRİ

Vikipediya dünyada baş verən hadisələri, cəmiyyətdəki münasibətləri özündə əks etdirən nəhəng layihədir. Belə ki, virtual ensiklopediyaya hər gün müxtəlif mövzularda: sosial-iqtisadi, mədəni, elmi və başqa sahələrlə bağlı məlumatlar yerləşdirilməkdədir. Viki-mühitdə toplanan məlumatlardan istifadə sahələri də çoxalmışdır: təhsildə, mediada,

iqtisadiyyatda, elmi araşdırmalarda və s. sahələrlə bağlı Vikipediya serverlərinə hər ay yüz milyonlarla sorğu göndərilir.

Bu gün viki-mühitdə toplanan informasiyanın ictimai fikrə təsirində rolu artmışdır. Eyni zamanda cəmiyyətdə baş verən prosesləri öyrənmək üçün Vikipediya çox geniş imkanlara malikdir. Vikipediya səhifələrinin loq-fayllarını analiz etməklə onların nə zaman və kimlər tərəfindən yaradılması, dəyişdirilməsi və hətta hansı dəyişikliklərin həyata keçirilməsini öyrənmək mümkündür.

Vikipediya istifadəçiləri həm qeydiyyatdan keçməklə, həm də qeydiyyatdan keçmədən viki-mühitdə ensiklopedik məqalələrin (viki-səhifələrin) yaradılmasında və redaktə olunmasında iştirak edə bilərlər. Lakin nəzərə almaq lazımdır ki, qeydiyyatdan keçmədən viki-mühitdə fəaliyyət göstərən istifadəçilərin IP ünvanları da loq-fayllarda saxlanılır və viki-səhifələrin, istifadəçi səhifələrinin analizi ilə birlikdə viki-istifadəçilərin sosial-demografik portretinin öyrənilməsində önəmli rol oynayır [3].

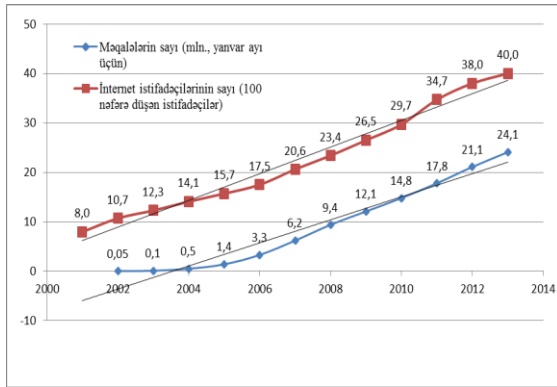
Vikipediyada yaradılan ensiklopedik məqalələr yarandığı şəraitə görə iki tipə bölünür: bir-neçə ekspert tərəfindən yaradılan və dar mövzunu əhatə edən və yüzlərlə istifadəçi tərəfindən yaradılan və geniş mövzunu əhatə edən məqalələr.

Vikipediya və onun layihələri Wikimedia Fondu (*Wikimedia Foundation, WMF*) tərəfindən idarə olunurlar. Fondun verilənlər bazasında (VB) saxlanılan istənilən informasiyanı bərpa etmək, digər versiyalarla müqayisə etmək və dəyişdirmək mümkündür. WMF-nin VB-da saxlanılan verilənlər hər bir İnternet istifadəçisi üçün açıqdır, yəni hər kəs bu verilənləri əldə etmək imkanına malikdir. Verilənlərin əlyətərliyi viki-mühitin müxtəlif istiqamətlərdə analizini aparmağa imkan yaradır:

- *Kontentin və istifadəçilərin artım dinamikası.*
- *Şəbəkədəki hiperkeçidlərin artım dinamikası və rolu.*
- *Linqvistik tədqiqatlar.*
- *Məqalələrin keyfiyyətinin analizi.*
- *Sosial şəbəkədə münasibətlərin və istifadəçilərin nüfuzunun (user reputation) analizi.*
- *Wiki-mühitdə həyata keçirilən informasiya müharibəsi texnologiyalarının analizi.*

Viki-mühitdə toplanan məlumatları analiz etməklə bütün sahələrdə olduğu kimi İKT sahəsində də müəyyən

araşdırmalar aparmaq mümkündür. Məsələn, İKT-nin inkişafının Vikipediya təsirini təyin etmək üçün ümumi İnternet istifadəçilərinin artım dinamikasını Vikipediya istifadəçilərinin artım dinamikası ilə müqayisə etmək. Şəkil 1-dən görüldüyü kimi 2001 və 2013-cü illər arasında İnternet istifadəçilərinin artım dinamikası Vikipediyanın bütün dil bölmələri üzrə ensiklopedik məqalələrin sayının artım dinamikası ilə təqribən eynidir. Şəkildən görünür ki, 2001 və 2004-cü illər arasında məqalələrin sayındakı artım tempi sonrakı illərlə müqayisədə xeyli az olmuşdur. Buna səbəb Vikipediyanın yarandığı ilk illər İnternet mühitində tanınmaması və 2004-cü ilə kimi cəmi 15 dildə fəaliyyət göstərməsi idi.



Şəkil 1. Vikipediya istifadəçilərinin ümumi İnternet istifadəçilərləndən asılılıq qrafı

Viki-mühitdə küllü miqdarda müxtəlif tipli informasiyanın toplanması, bu informasiyanın daima yenilənməsi və ya dəyişdirilməsi bir tədqiqat obyektini kimi analitikləri cəlb etməkdədir. İlk növbədə viki-mühitin cəmiyyətin bilik daşıyıcılarını özünə cəlb etməsi və insanların öz əmək və biliklərini təmənnaşız olaraq bu layihə vasitəsi ilə dünya ictimaiyyətinə bağışlaması indiyə kimi alimləri düşündürməkdədir [4].

Vikipediyanın əsas problemləri olan vandalizm, redaktələr müharibəsi və viki-istifadəçilər arasında baş verən konfliktləri araşdırarkən məlum olur ki, viki-texnologiyanın verdiyi imkanlar viki-mühitdə informasiya təhlükəsizliyi ilə əlaqədar çox ciddi problemlər yaradır. Problemlərin həlli üçün tədqiqatlar əsasən üç istiqamətdə aparılır:

- *Vikipediya və onun digər layihələrinin həcmnin və artım dinamikasının analizi.*
- *Viki-kontentin keyfiyyətinin analizi.*
- *Viki-mühitdə fəaliyyət göstərən sosial şəbəkənin analizi.*

Nəzərə almaq lazımdır ki, viki-mühit yalnız bilik mənbəyi deyil, həm də müxtəlif ideologiyalara və baxışlara malik qrupları birləşdirən çox nəhəng sosial şəbəkədir. Bu baxımdan sosial şəbəkələrin analizində istifadə edilən metod və nəzəriyyələr viki-mühitin analizində də çox geniş tətbiq olunur [5, 6].

Viki-mühitdə daima yenilənən kollektiv biliyin yaranma prosesini öyrənmək, Vikipediya ensiklopediyasında toplanan məqalələrin bütün sahələri nə dərəcədə əhatə etməsini təyin

etmək üçün, müqayisəli analiz metodları, təsadüfi altıoxluqlar nəzəriyyəsi üstünlük təşkil edir. Məqalələrin müəyyən əlamətlərə görə təsnifatlandırılmasında klasterləşmə, viki-mühitin analizində isə semantik analiz metodlarından istifadə edilir. Semantik analiz metodlarında HITS (*Hyperlink-Induced Topic Search*), PageRank, Random Forest alqoritmlərindən istifadəyə geniş yer verilir [7, 8].

Vikimetrik tədqiqatlarda istifadə olunan göstəricilər:

- a) *WMF VB-nin həcmi (bütün viki-səhifələrin həcmi, müzakirə, kateqoriya və istiqamətləndirmə səhifələri də daxil olmaqla).*
- b) *Viki-layihələrdə ensiklopedik məqalələrin sayı.*
- c) *Hər bir ensiklopedik məqalədə sözlərin sayı.*
- d) *Viki-səhifələrdə daxili keçidlərin sayı.*
- e) *Viki-layihələrdə qeydiyyatdan keçmiş istifadəçilərin sayı.*
- f) *Aktiv istifadəçilərin sayı (bir ay ərzində 5 redaktədən çox redaktəsi olan istifadəçilər).*
- g) *Çox aktiv istifadəçilərin sayı (bir ay ərzində 100 redaktədən çox redaktəsi olan istifadəçilər).*

Vikipediya beynəlxalq layihə olduğundan, hər hansı məqalənin yaradılmasında tam fərqli mövqelərdən çıxış edən qüvvələrin iştirakı təbii. Hər bir viki-səhifənin məzmununa və xarakterinə görə Vikipediya reallaşdırılan informasiya müharibəsi müstəvisində xalqlar və ölkələr arasındakı münasibətləri aydın görmək mümkündür [9]. Viki-mühitdə baş verən mübahisələr, informasiya qarşıdurmaları, həmçinin, istifadəçilər arasında əməkdaşlıq və bu münasibətlərin kontentə təsiri müxtəlif metod və alqoritmlərlə tədqiq edilməkdədir [10, 11]. Məsələn, vandalizmə məruz qalan məqalələrin aşkarlanması üçün metaverilənlərin məkan-zaman analiz (spatio-temporal analysis of metadata, STiki) metodundan, Lüman (Luhmann) modelindən və Piaget (Piaget) nəzəriyyəsi istifadə edilmişdir [12, 13].

Kaliforniya Universitetinin əməkdaşları S. Javanmardi və həmmüəllifləri viki-istifadəçilər arasında münasibətləri və hansının daha çox nüfuzə malik olmasını təyin etmək üçün təklif etdikləri model təsnifatlandırma metodları əsasında viki-mühitdə vandalları və təcrübəsiz istifadəçiləri təyin etməyə imkan verir. Təklif olunan model həmçinin, Data Mining metodlarına əsaslanaraq Vikipediya ensiklopediyasında istifadəçilərin davranışları haqqında da müəyyən proqnozlar vermək iqtidarındadır [14].

Viki-mühitdə münasibətləri, əməkdaşlığı ölçmək üçün digər yanaşma İsrail Universitetinin tədqiqatçıları tərəfindən təklif edilmişdir [15]. Bu metodda 3 ölçüdən istifadə edilir:

- a) *Viki-istifadəçilərin sayı.*
- b) *İnteraktivlik, yəni müəyyən vaxt intervalında istifadəçilərin viki-səhifələrdə etdikləri redaktələrin sayı.*
- c) *İntensivlik, yəni redaktələr nəticəsində viki-səhifədə olan dəyişikliklərin sayı.*

Təklif olunan metod əsasında redaktələr müharibəsi, viki-istifadəçilərin aktivliyi, etdikləri redaktələrin keyfiyyəti təyin edilir. Belə ki, bəzən viki-məqalələrdə edilən redaktələr digər istifadəçilər tərəfindən rədd edilir. Bu isə belə bir vəziyyət yaradır: məqalə üzərində işləyən istifadəçilərin sayı çoxdur və çoxlu sayda redaktələr olunmuşdur, lakin son nəticədə ensiklopedik məqalənin həcmi az dəyişmişdir.

Bəzi tədqiqatlarda viki-istifadəçilərin nüfuzu yalnız istifadəçilərin davranışları ilə təyin edilir. Məsələn, Adlerin təklif etdiyi alqoritmə əgər istifadəçilər tərəfindən edilən dəyişikliklər saxlanılırsa onların nüfuzları artır, etdikləri dəyişikliklər silinirsə nüfuzları azalır. Ensiklopedik məqalələrin keyfiyyətinin və müəlliflərin nüfuzunun analizində nüfuz əsaslanan sistem (a reputation-based system, Wiki-Trust) təklif edilmişdir [16].

Vikimetrik tədqiqatlar zamanı nüfuzun qiymətləndirilməsində “silib-bərpa etmək” (*delete – restore*) və saxta məlumatların yerləşdirilməsi kimi informasiya hücumlarını aşkarlayan modellərdən də istifadə edilir [17]. Modeldə təklif edilən “kontent-yonəldici nüfuz alqoritmi” (*algorithm for content-driven reputation*) yalnız yüksək nüfuzlu malik istifadəçilər tərəfindən razılaşdırılan kontenti əhəmiyyətli sayır. Yüksək nüfuzlu malik istifadəçilərin qeydiyyat adı səhifənin loq-faylında yoxdursa, kontent keyfiyyətsiz sayılır.

Viki-mühitin intellektual analizi tədqiqatların daha dərinə aparılmasına və gizli məlumatların əldə olunmasına şərait yaradır. İntellektual analiz üçün viki-mühit geniş imkanlara malikdir:

a) *Viki-mühit sənədləri idarə edir:*

- Milyonlarla viki-səhifələrin mövcudluğu.
- Səhifələrin müxtəlif tipli yüzlərlə hiperkeçidlərə malik olması (Dense link structure).

b) *Viki-mühit bilikləri idarə edir:*

- Fasiləsiz olaraq yeni kontentin yaradılması.
- Kontentin dinamik dəyişdirilməsi.
- Kontentin sistemləşdirilməsi.

c) *Viki-mühit nəhəng sosial şəbəkədir.*

Vikipediyanın yuxarıda göstərilən imkanları ilə əlaqədar Wiki Mining, Text Mining və Link Mining metodları əsasında viki-mühitdə səhifələr arasında semantik əlaqələrin analizini aparmaq, ensiklopedik məqalələrin əhatəliliyi və keyfiyyətini təyin etmək, gizli sosial şəbəkələri aşkar etmək mümkündür [18, 19].

### III. VIKI-MÜHİTDƏ İNFORMASIYA TƏHLÜKƏSİZLİYİ TƏDBİRLƏRİ

Yuxarıda deyilənlərdən göründüyü kimi viki-texnologiyanın imkanları viki-istifadəçilər arasında əməkdaşlığa şərait yaratdığı kimi informasiya müharibələrinə də şərait yaradır. Vikipediya İnternetin bir seqmenti olduğundan İnternetdə baş verən informasiya müharibələri və

informasiya təhlükəsizliyi ilə əlaqədar problemlər viki-mühitdə də mövcuddur. Viki-mühitdə reallaşdırılan informasiya müharibəsi texnologiyalarına aiddir:

- *Vandalizm (mətnin silinməsi (Blanking); kontentin başqası ilə əvəz edilməsi (Large-scale Editing); dezinformasiya).*
- *Psixoloji təsir (təhqir, yersiz şərhələr, təbliğat).*
- *Spam xarakterli məlumatların daxil edilməsi.*
- *“Kukla”larla (sockpuppets) təsir.*
- *Ziyanverici botların tətbiqi.*
- *Şəkillərlə hücum (Image Attack).*

Vikipediyanın dinamik inkişafına və problemlərin həllində çoxlu sayda təkliflərin verilməsinə baxmayaraq demək olar ki, ensiklopediyanın bütün layihələrində həlli vacib olan problemlər qalmaqdadır. İnformasiya təhlükəsizliyi ilə əlaqədar problemlərin həllində görülməsi vacib olan işlər bunlardır:

- *Vikipediyanın keyfiyyət göstəricisinə xələl gətirmədən viki-məqalələrin sayının artırılması.*
- *Vikipediya ensiklopedik məqalələr arasında hiperkeçidlərin dəqiqləşdirilməsi.*
- *Kateqoriyaların sistemləşdirilməsi.*
- *İnformasiya müharibələrində və redaktələr müharibəsində istifadə olunan məqalələrin aşkarlanması və yenidən işlənməsi.*
- *Qrafik kontentin keyfiyyətinin yaxşılaşdırılması və mövzuya uyğun yerləşdirilməsi.*
- *İstifadəçilər arasında əməkdaşlığın təşkili.*
- *İnformasiya qarşılıqlı əlaqələrində iştirak edən gizli sosial qrupların aşkarlanması.*

Vikipediya ensiklopedik biliyin təqdim olunmasında yeni bir modelin təzahürüdür. Vikipediyanın bu günkü müvəffəqiyyəti sübut etdi ki, müasir dövrdə İnternet istifadəçiləri virtual mühitin passiv istifadəçiləri olmaq istəmir, onlar veb-səhifələrdəki informasiyanı dinamik olaraq dəyişmək, yeni məlumatları daxil etmək, bilik və bacarıqlarını dünya ictimaiyyətinə göstərmək arzusundadırlar. Bu baxımdan Vikipediya bu gün çox ciddi informasiya mənbəyidir və onu digər ənənəvi informasiya mənbələrindən fərqləndirən ən böyük üstünlüyü operativliyidir.

### NƏTİCƏ

Bu gün dünyanın milyonlarla kompüterini birləşdirən global İnternet şəbəkəsinin ən nəhəng layihəsi olan Vikipediya özünün forma və tərkibi baxımından müxtəlif xüsusiyyətli, sərhəd tanımayan şəbəkə resursları ilə cəmiyyətin sosial təşkilinin ayrılmaz hissəsinə çevrilmişdir.

Ümumilikdə Vikipediya müxtəlif tipli tədqiqatlar üçün çox əlverişli mühitdir və bu layihədə sənədlərin semantik yaxınlığının analizindən başlayaraq İnternet istifadəçilərinin

aktivliyinə kimi müxtəlif sahələrdə tədqiqatlar aparmaq mümkündür.

#### ƏDƏBİYYAT

1. R. Sumi, T. Yasser, A. Rung, A. Kornai, J. Kertesz, “Edit Wars in Wikipedia”, Proceedings of the International Conference on Social Computing (SocialCom), IEEE, Boston, MA, USA, Oct. 2011, pp. 724-727.
2. [http://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](http://meta.wikimedia.org/wiki/List_of_Wikipedias)
3. İ.Y. Ələkbərova, “Viki-mühitin sosial-demografik portretinin analizi”, AMEA İTİ-nin “İnformasiya cəmiyyəti problemləri” jurnalı, 2012, s. 21-28.
4. J. Voss, “Measuring Wikipedia”, Proceedings of the International Conference of the International Society for Scientometrics and Informetrics: 10th, Stockholm (Sweden), 2005, pp. 24–28.
5. S. Blaschke, K. Stein, “Methods and Measures for the Analysis of Corporate Wikis: A Case Study”, Proceedings of the 58th Annual Conference of the International Communication Association ICA, Montréal, Canada, May, 2008, pp 1-24.
6. L. S. Buriol, C. Castillo, D. Donato, S. Leonardi, S. Millozzi, “Temporal Analysis of the Wikigraph”, Web Intelligence, Publisher: IEEE CS Press., 2006, pp. 45-51.
7. A. Halavais, D. Lackaff, “An Analysis of Topical Coverage of Wikipedia”, Journal of Computer-Mediated Communication, Vol. 13, Iss. 2, 2008, pp. 429–440.
8. U. Cress, J. Kimmerle, “A systemic and cognitive view on collaborative knowledge building with wikis”, International Journal of Computer-Supported Collaborative Learning, Vol. 3, Iss. 2, 2008, pp. 105-122.
9. İ.Y. Ələkbərova, “Viki mühitdə reallaşdırılan bəzi informasiya müharibəsi texnologiyalarının analizi”, “İnformasiya cəmiyyəti problemləri” jurnalı, №2(4), 2011, s. 18-28.
10. J. Moskaliuk, J. Kimmerle, U. Cress, “Collaborative knowledge building with wikis: The impact of redundancy and polarity”, Journal of Computers & Education, Vol.58, Iss. 4, 2012, pp. 1049-1057.
11. N. Hara, P. Shachaf, K.F. Hew, “Cross-cultural analysis of the Wikipedia community”, Journal of the American Society for Information Science and Technology, Vol. 61, Iss. 10, 2010, pp. 2097–2108.
12. B.T. Adler, L. de Alfaro, S. Mola-Velasco, P. Rosso, A. West, “Wikipedia vandalism detection: combining natural language, metadata, and reputation features”, Proceedings of the Computational Linguistics and Intelligent Text Processing. Lecture Notes in Computer Science, Vol. 6609, Springer, Berlin, 2011, pp. 277–288.
13. A.G. West, S. Kannan, I. Lee, “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). Article No. 32, July, 2010, pp. 47-48.
14. S. Javanmardi, C. Lopes, P. Baldi, “Modeling user reputation in wikis” Journal of Statistical Analysis and Data Mining, Vol. 3 Iss. 2, 2010, pp. 126-139.
15. H. Meishar-Tal, E. Tal-Elhasid, “Measuring Collaboration in Educational Wikis – A Methodological” International Journal of Emerging Technologies in Learning (IJET), Vol. 3, 2008, pp.46-49.
16. B. T. Adler, L. de Alfaro, “A content-driven reputation system for the Wikipedia”, Proceedings of the 16th international conference on World Wide Web, 2007, pp. 261-270.
17. K. Chatterjee, L. de Alfaro, I. Pye, “Robust content-driven reputation,” School of Engineering, University of California, Santa Cruz, USA, 2008. [Online]. Available: <http://www.soe.ucsc.edu/~luca/papers/08/ucsc-soe-08-09.pdf/>
18. K. Nakayama, T. Hara, S. Nishio, “Wikipedia Mining for An Association Web Thesaurus Construction”, Proceedings of the International Conference on Web Information Systems Engineering (WISE), 2007, pp, 322-334.
19. M. Ito, K. Nakayama, T. Hara, S. Nishio, “Semantic relatedness measurement based on Wikipedia link co-occurrence analysis”, International Journal of Web Information Systems, Vol. 7, Iss. 1, 2011, pp. 44 – 61.