

0 - 23 - Social network extraction based on Web - 3. the integrated superficial method.pdf



Date: 2018-07-04 13:01 UTC

* All sources 100 | Internet sources 42 | Own documents 36 | Organization archive 16 | Plagiarism Prevention Pool 6

- ✓ [0] https://www.researchgate.net/publication...d_superficial_method
90.7% 175 matches

- ✓ [1] "3028-3639-1-RV.pdf" dated 2017-10-30
77.7% 149 matches

- ✓ [2] https://www.researchgate.net/publication...d_Superficial_Method
15.8% 37 matches

- ✓ [3] https://www.researchgate.net/publication..._superficial_methods
14.5% 38 matches

- ✓ [4] "0 - 23 - Social network extraction...ot; dated 2018-07-04
13.9% 38 matches

- ✓ [5] "0 - 25 - Social-Network-Extraction...ot; dated 2018-07-04
13.4% 34 matches
⊕ 1 documents with identical matches

- ✓ [7] "CR-INT136-Social network extractio...ot; dated 2017-10-09
11.5% 31 matches

- ✓ [8] "0 - 22 - No research without publi...ot; dated 2018-07-04
9.3% 25 matches

- ✓ [9] "5295-7943-1-RV.pdf" dated 2018-04-06
9.1% 28 matches

- ✓ [10] "0 - 27 - Semantic interpretation o...ot; dated 2018-07-04
8.0% 23 matches

- ✓ [11] "CR-INT137-Enhancing to method for ...ot; dated 2017-10-09
7.4% 23 matches

- ✓ [12] https://www.researchgate.net/publication...from_the_Web_Snippet
7.0% 17 matches

- ✓ [13] "5296-7945-1-RV.pdf" dated 2018-04-06
7.0% 22 matches

- ✓ [14] "0 - 33 - SumutSiana.pdf" dated 2018-07-04
5.3% 18 matches

- ✓ [15] "CR-INT135-Information Retrieval on...ot; dated 2017-10-09
6.3% 16 matches

- ✓ [16] "5298-7949-1-RV.pdf" dated 2018-04-06
5.9% 17 matches

- ✓ [17] "CR-INT110-Semantic interpretation ...ot; dated 2017-10-09
5.9% 16 matches

- ✓ [18] https://www.researchgate.net/publication...s_Using_Web_Snippets
5.2% 10 matches

- ✓ [19] https://www.researchgate.net/publication...nformation_Retrieval
5.6% 15 matches

- ✓ [20] <https://ukm.pure.elsevier.com/en/publica...integrated-superfici>
5.1% 11 matches


- ✓ [21] "3472-4529-1-SM.pdf" dated 2017-10-30
4.9% 13 matches


- ✓ [22] "[doi 10.1109/stair.2011.5995766] N...ot; dated 2017-12-27
5.0% 14 matches


- ✓ [23] https://www.science.gov/topicpages/i/integrate_social_networking.html
4.5% 10 matches


- ✓ [24] "0 - 21 - Indonesia knowledge disse...ot; dated 2018-07-04
3.7% 13 matches


- ✓ [25] "0 - 32 - Research mapping in North...ot; dated 2018-07-04


- [25]  4.5% 13 matches


- [26]  "MahyuddinKMNASution_2017_IOP_Conf....ot; dated 2017-12-27" 4.3% 13 matches


- [27]  "Information Retrieval Based on the Extracted.pdf" dated 2017-12-27 4.0% 12 matches


- [28]  "[doi 10.1109/SAI.2016.7556125] Nas...ot; dated 2017-12-27" 4.3% 11 matches


- [29]  https://link.springer.com/chapter/10.1007/978-3-319-57261-1_31 4.2% 12 matches


- [30]  "0 - 31 - Redefining the magic squa...ot; dated 2018-07-04" 3.3% 11 matches


- [31]  "[doi 10.1109/CITSM.2016.7577513] N...ot; dated 2017-12-27" 3.9% 11 matches


- [32]  <https://www.springerprofessional.de/en/e...relation-be/12204946> 3.9% 12 matches


- [33]  "0 - 26 - Triangle number of Lauren...ot; dated 2018-07-04" 3.1% 11 matches


- [34]  "Social Network Mining (SNM) A Defi...ot; dated 2017-12-27" 3.5% 8 matches


- [35]  "MahyuddinKMNASution_2017_J._Phys....ot; dated 2017-12-27" 3.1% 9 matches


- [36]  "MahyuddinKMLubis_2017_J._Phys._Co...ot; dated 2017-12-27" 2.9% 7 matches


- [37]  https://www.researchgate.net/profile/Opim_Sitompul 2.5% 4 matches



- [38]  "0 - 20 - Information Retrieval Bas...ot; dated 2018-07-04" 3.2% 9 matches


- [39]  https://www.researchgate.net/scientific-contributions/9839813_Ron_Bekkerman 2.6% 7 matches


- [40]  https://www.researchgate.net/publication...traction_perspective 2.7% 6 matches



- [41]  "[doi 10.1109/iceei.2015.7352551] M...ot; dated 2017-12-27" 3.1% 6 matches


- [42]  www.science.gov/topicpages/i/integration_method_based.html 2.6% 4 matches


- [43]  "MahyuddinKMNASution_2017_J._Phys....ot; dated 2017-12-27" 2.3% 7 matches
 1 documents with identical matches



- [45]  "[doi 10.1109/infrkm.2012.6204999] ...ot; dated 2017-12-27" 2.9% 7 matches



- [46]  iopscience.iop.org/article/10.1088/1757-899X/220/1/012028/pdf 2.2% 7 matches


- [47]  "Lubis_2017_J._Phys.:_Conf._Ser._801_012014.pdf" dated 2018-04-16 2.2% 6 matches
 3 documents with identical matches


- [51]  https://www.researchgate.net/publication...demic_social_network 2.2% 6 matches


- [52]  "Ginting_2018_IOP_Conf._Ser.:_Mater...ot; dated 2018-04-03" 2.0% 6 matches


- [53]  "Maulida_2018_IOP_Conf._Ser.:_Mater...ot; dated 2018-04-16" 2.0% 6 matches
 4 documents with identical matches


- [58]  "16. IOP.pdf" dated 2017-12-07 2.0% 6 matches
 1 documents with identical matches


- [60]  "Ginting_2017_IOP_Conf._Ser.:_Mater...ot; dated 2018-04-03
1.9% 6 matches


- [61]  https://www.researchgate.net/publication...s_of_social_networks
1.9% 8 matches


- [62]  "Setyowidodo_2017_J._Phys.:_Conf._Ser._895_012020.pdf" dated 2018-03-27
1.8% 5 matches
⊕ 4 documents with identical matches


- [67]  "17. IOP.pdf" dated 2017-12-07
1.8% 6 matches
⊕ 1 documents with identical matches


- [69]  <iopscience.iop.org/article/10.1088/1755-1315/105/1/012041/pdf>
1.8% 5 matches


- [70]  "Maulida_2016_J._Phys.:_Conf._Ser._710_012012.pdf" dated 2018-04-16
1.8% 5 matches
⊕ 1 documents with identical matches


- [72]  "32. IOP.pdf" dated 2017-12-07
1.6% 5 matches
⊕ 1 documents with identical matches


- [74]  "AHMADPERWIRA_TEKNIK_USU_.pdf" dated 2018-02-11
1.5% 5 matches


- [75]  <iopscience.iop.org/article/10.1088/1742-6596/978/1/011003>
1.6% 4 matches


- [76]  "31. IOP.pdf" dated 2017-12-07
1.5% 4 matches
⊕ 2 documents with identical matches


- [79]  "The_Influence_of_Chicken_Egg_Shell_as_Fillers_on_B.pdf" dated 2018-04-03
1.5% 4 matches
⊕ 2 documents with identical matches


- [82]  https://www.researchgate.net/publication...om_email_and_the_Web
1.5% 3 matches


- [83]  www.science.gov/topicpages/s/social_networks_methods.html
1.1% 1 matches


- [84]  <https://www.sciencedirect.com/science/article/pii/S1570826805000089>
0.8% 3 matches


- [85]  "CR-INT009-Triangle Number of Laure...ot; dated 2017-10-08
0.8% 4 matches


- [86]  www.academia.edu/1004621/Extraction_of_academic_social_network_from_online_database
0.8% 2 matches


- [87]  "2475-2791-1-RV.pdf" dated 2017-09-26
0.7% 3 matches


- [88]  <https://www.researchgate.net/profile/Hon...llic-sulfide-ore.pdf>
0.7% 3 matches


- [89]  <https://ukm.pure.elsevier.com/en/organis...y-cait/publications/>
0.6% 3 matches


- [90]  toc.proceedings.com/34010webtoc.pdf
0.6% 2 matches



- [91]  https://en.wikipedia.org/wiki/Marketing_and_artificial_intelligence
0.6% 1 matches



- [92]  <https://www.sciencedirect.com/science/article/pii/S1877050917329010>
0.4% 2 matches



- [93]  https://link.springer.com/chapter/10.1007/978-3-319-04244-2_34
0.4% 2 matches



- [94]  "2668-3031-1-RV.pdf" dated 2017-09-26
0.3% 2 matches



- [95]  toc.proceedings.com/12500webtoc.pdf
0.4% 1 matches



- ✓ [96]  <https://www.semanticscholar.org/paper/Ne...8f7de3a2585486a216f8>
 1 matches



- ✓ [97]  iopscience.iop.org/year/1757-899X/Y2018
 2 matches



- ✓ [98]  "90-99Z_Article Text-204-1-4-20171227.pdf" dated 2018-01-23
 1 matches



- ✓ [99]  from a PlagScan document dated 2017-04-05 12:26
 1 matches



- ✓ [100]  https://www.researchgate.net/profile/Shahrul_Azman_Mohd_Noah
 1 matches



- ✓ [101]  <https://ocs.usu.ac.id/ICCAI/ICCAI2017>
 1 matches



- ✓ [102]  <https://www.facebook.com/usuofficial>
 1 matches



- ✓ [103]  from a PlagScan document dated 2017-04-06 09:41
 1 matches



- ✓ [104]  from a PlagScan document dated 2017-04-06 07:52
 1 matches



- ✓ [105]  from a PlagScan document dated 2017-04-05 07:51
 1 matches



- ✓ [106]  from a PlagScan document dated 2017-04-05 07:45
 1 matches



- ✓ [107]  from a PlagScan document dated 2017-04-05 06:55
 1 matches



- ✓ [108]  iopscience.iop.org/article/10.1088/0305-4624/12/5/302/pdf
 1 matches



- ✓ [109]  europepmc.org/articles/PMC4474359
 1 matches



- ✓ [110]  iopscience.iop.org/article/10.1088/2053-1583/2/4/044015/pdf
 1 matches



- ✓ [111]  www.vepub.com/conferences-view/Internati...1EyWEFWR2p3WEJtdz09/
 1 matches



- ✓ [112]  https://www.researchgate.net/publication...ion-related_webpages
 1 matches



- ✓ [113]  <https://twitter.com/usuofficial>
 1 matches



- ✓ [114]  <https://link.springer.com/article/10.1186/s12889-016-2722-9>
 1 matches



- ✓ [115]  "2018PTerapan(WPT)Proposal.pdf" dated 2018-02-11
 1 matches

- ✓ [116]  "Proposal Talenta An. Halim 2018.pdf" dated 2018-02-11
 1 matches


- ✓ [117]  "UlfiAndayani_FASILKOMTI_PDM.pdf" dated 2018-02-11
 1 matches

- ✓ [118]  "Ivan Indrawan_Fakultas Teknik_PDM.docx.pdf" dated 2018-02-11
 1 matches

- ✓ [119]  "3310-4262-1-RV.pdf" dated 2017-10-30
 1 matches

- ✓ [120]  "2660-3022-1-RV.pdf" dated 2017-09-26
 1 matches

8 pages, 3363 words

 A very light text-color was detected that might conceal letters used to merge words.

PlagLevel: selected / overall

175 matches from 121 sources, of which 43 are online sources.

Settings

Data policy: *Compare with web sources, Check against my documents, Check against my documents in the organization repository, Check against organization repository, Check against the Plagiarism Prevention Pool*

Sensitivity: *Medium*

Bibliography: *Consider text*

Citation detection: *Reduce PlagLevel*

Whitelist: --

PAPER • OPEN ACCESS

Social network extraction based on Web: ^[0]3. the integrated superficial method

To cite this article: ^[0]M K M Nasution et al 2018 J. Phys.: Conf. Ser. ^[0]978 012033

View the [article online](#) for updates and enhancements.

Social network extraction based on Web: 3. the integrated superficial method

M K M Nasution^{1*}, O S Sitompul¹ and S A Noah²

¹Technical Information, Fasilkom -TI, Universitas Sumatera Utara, Padang Bulan 20155
USU, Medan, Indonesia

²Knowledge Technology Research Group, Faculty of Information Science &
Technology, Universiti Kebangsaan Malaysia, Bangi 43600 UKM Selangor, Malaysia

E-mail: mahyuddin@usu.ac.id

Abstract. The Web as a source of information has become part of the social behavior information. Although, by involving only the limitation of information disclosed by search engines in the form of hit counts, snippets, and URL addresses of web pages, the integrated extraction method produces a social network not only trusted but enriched. Unintegrated extraction methods may produce social networks without explanation, resulting in poor supplemental information, or resulting in a social network of dumise laden, consequently unrepresentative social structures. The integrated superficial method in addition to generating the core social network, also generates an expanded network so as to reach the scope of relation clues, or number of edges computationally almost similar to $n(n-1)/2$ for n social actors.

1. Introduction

The Web as a source of information has a lot of potential usage in everyday life [1]. Information extraction relates to the structure of information to be generated [2, 3]. As with social networks, the extraction of social structures from the Web has a variety of different sources [4]. The sources can generally support each other, but there are fundamental differences in approach and outcomes. In general, supervised approaches always rely on more accurate explanations of the results of social network extractions, but are unable to grasp the meaning of changes that occur in information sources [5, 6, 7]. On the other hand, the unsupervised approach generally relies on the ability to capture change superficially, but can actually be enriched through the integration of different approaches to different sources of information from the Web.

So far, there are some superficial extraction methods. All of these methods have been developed differently by different researchers [8, 9, 10, 11]. One with other methods have been partially integrated, but the emphasis is to disclose the trusted information [6]. This paper intend to describe the integration of approaches into an enriched superficial method through different sources of information.

2. Related Works and Motivation



Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Although by involving the same queries, two steps for extracting social network from the Web, the queries produce different sources of information as an indicator of the relationship between social actors [12, 13]. For a pair of social actors a_i and a_j , through query $q \leftarrow t_{a_i}$ and $q \leftarrow t_{a_j}$

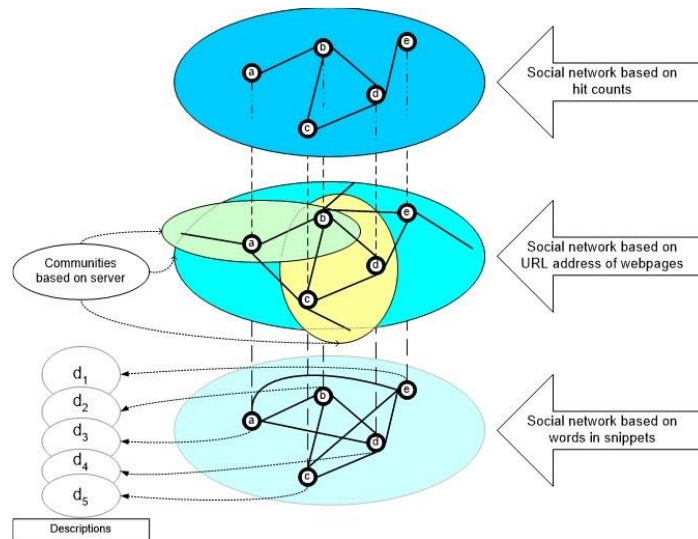


Figure 1. Levels of relationship in social networks

two singletons are generated: a hit count Ω_{a_i} and a hit count Ω_{a_j} . Through the search engine, with which t_{a_i} and t_{a_j} are two different social actors names as query content. By involving the same search engine, the relation clue in semantic is generated by the query $q \leftarrow t_{a_i}, t_{a_j}$ is a doubleton or hit count $\Omega_{a_i}, \Omega_{a_j}$. However, each query not only generates a hit count, but produces a snippet [5].

Each search engine produces a snippet as companion information from a hit count. Snippet serves as an explanation regarding query content. In each snippet there are a number of words and one identity uniquely for each source of information that is the URL address of the webpage [14]. Thus, for each query will result a set of s_i snippets and a set of URL addresses u_i if hit count $\Omega_{a_i} > 0$. For a collection of words from a non-empty snippet, it will contain the name of the social actor and also the words expressing an explanation of the activity of the social actor. For example, the affiliation of social actors, ideas or concepts contained in the paper title of social actor, performance targets to be planned by social actors, and so on [15, 16].

In general, the extraction of social networks from the Web uses a similarity measurement for a pair of social actors by involving three different hit counts [17]. This produces the strength relation. However, there is a very strong relationship up to the weakest relationship that can be generated through the similarity between the URL addresses [15]. This strongest relationship is based on the concept that one URL address as a representation of a webpage represents one event, and this results in social actors on the same webpage having a close relationship, which is generally described semantically as co-occurrence. However, the URL address of form is stratified as from directory to sub-directory, the lower part being the domain located at the base of the URL address. Furthermore, the similarity between URL addresses with each other is determined by the similarity between the parts separated by the slash from base to tip. While a collection of URL addresses are present via doubleton, it can be used as validation of the strength relation that arises from two sets of singleton-based URL addresses. In different cases, the number of URL addresses of doubleton becomes pure representation by co-occurrence [18].

[10]

In a different study, which also involves URL addresses, it has been revealed that a collection of words can express the relation clues of two social actors in addition to the description of the

relationship. On different occasions, this is reinforced by involving a collection of words from singleton or doubleton-based snippets [19].

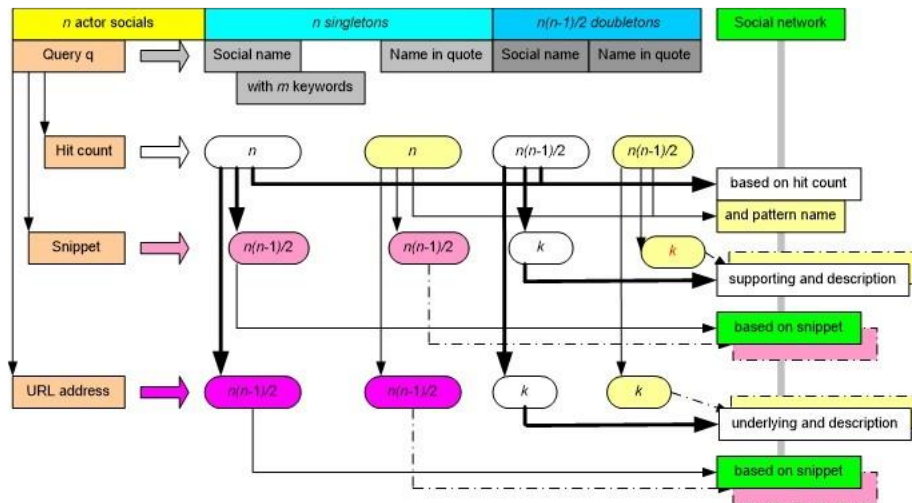


Figure 2. The cornerstone for integrating superficial methods

In general, the relationship between social actors derived from the same type of query will form layers of mutually supportive and explicit relationships, such as Fig. 1 for example.

3. The Proposed Approach

Based on the reviews have been made and the comparisons between different approaches in the extraction of social networks [20, 21], we construct an approach to the involvement of some superficial methods integratedly such as Fig. 2. Accordingly, different approaches are adopted to adjust to optimally implementation. That is adjustment to the limitations of search engine services for the number of queries and length of the query content [5].

When a query contains a social name in either singleton or doubleton, then the doubleton engagement in similarity is to show the function of reinforcing the existence of the relationship between two actors, since the doubleton arrangement of the two names for the two social actors causes a reduction of ambiguity, even though in singleton query processes by the search engine lifts all information based on similarity. For example, a social name in singleton such as $q = \text{Mahyuddin K. M. Nasution}$, while two social names in a doubleton like $q = \text{Mahyuddin K. M. Nasution, Opim Salim Sitompul}$. So if there is n social actors, then this process requires the involvement of $n+n(n+1)/2$ queries and $n(n+1)/2$ computations. Normally, each search engine returns a maximum of 10 snippets in each page for the associated query. Please note most of search engines will feature snippets of webpages containing the most information about query content first in the top of the snippet list [22]. Therefore, if all snippets inside the first page of doubleton are used to correctly describe the strength relation (as supporting and description), it no longer involve a query, but it only spend k computations for processing words in snippets. The next integration, involves to process a collection of URL addresses. For the same reason, the collection of URL addresses involved is a maximum of 10 first URL addresses based on doubleton (as underlying and description) [23]. Although the URL address uniquely represents the webpage, but it involves the URL address of the highest ranking webpage, so that it reveals a fundamental foundation for the establishment of a trusted social network in addition to being enriched with additional descriptions based on the webpage community as the origin of the server. Thus, this process also involve any additional queries, see Fig. 2. Significantly, this integrated approach involves $kn(n-1)/2$ computations and follows $n+n(n+1)/2$ queries [24].

By taking the name of social actors in the form of a pattern or exist in quotation marks, such as $q = \text{Mahyuddin K. M. Nasution}$ for singleton or $q = \text{Mahyuddin K. M. Nasution, Opim Salim Sitompul}$ for doubleton, the search engine produces the accurate information about hit count, snippet, and URL address. This is set on equality between the query content and the content of related webpages based on the pattern. However, the similar names of social actors can not be detected by search engines [25]. Thus, the second approach will automatically involve $kn(n-1)/2$ computations and follows $n + n(n-1)/2$ queries. If it involves also the keywords in query based on either singleton or doubleton, then we use the same process like above, i.e. $n(n-1)/2$ queries will be used and the number of computations is $kmn(n-1)/2$ where m for processing keywords [23, 26], see Fig. 2.

In addition, social network extraction may also involve singleton from two different social actor descriptions. The descriptions derived from the snippet will form the relationship between two social actors through the similarity between the description details [27]. It only requires n queries and $n(n-1)/2$ computations. Nevertheless, in general involvement of this description always further expanded the social network with the possibility that a non-existent relationship between two social actors will be formed from the description. Likewise, if it involves a collection of singleton-based URL addresses, it involves n queries and $n(n-1)/2$ computations, and will expand social network differently between other social actors [19]. Therefore, the use of singleton both snippets and URL addresses are integrated to form social networks opinions that may be wider than actual social networks [28], see Fig. 2.

By expressing some opinions of integration approaches, generally superficial methods can be integrated as follows:

- (i) Define a number of names of social actors.
- (ii) Submit a query from each actors name to the search engine.
 - (a) Record hit count based on singleton.
 - (b) Record snippets of singleton in the first page.
- (iii) Submit a query of a pair of actor names to a search engine.
 - (a) Record hit count based on doubleton.
 - (b) Record snippets of doubleton in the first page.
- (iv) Calculate the strength relation for each pair of social actors.
 - (a) Describe the description of the doubleton snippets to strengthen the evidence of the strength relation and calculate the comparison between the description weights and the hit count of doubleton.
 - (b) Describe the URL address of the doubleton snippets to strengthen the proof of strength relation and calculate the number of URL addresses composition against the hit count of doubleton.
- (v) Describe the resulting social network.

As for expanding social networks, an integrated approach may involve additional steps as follows:

- (i) Describe the description of URL address of the singleton snippets of any social actor.
 - (a) Calculate the clarity of description as additional strength relation between two social actors to expand the relations.
 - (b) Calculate the URL addresses similarity between two social actors as additional strength relation to expand the relations.
- (ii) Extend social networks through two additional strength relation

	1	2	3	4	5
1. DBLP					
2. Microsoft	0.8480				
3. FTSM	0.6937	0.6876			
4. Scribd	0.6460	0.6802	0.6667		
5. Myais	0.2381	0.2613	0.2754	0.3074	
6. Research	0.0814	0.0889	0.1956	0.1121	0.1863

Figure 3. Similarity of clusters based on URL addresses

Table 1. URL addresses as label of strength relation

No.	URL address	id	Vertices	Edges
1.	www.informatik.uni-trier.de	DBLP	347	10,825 (85.79%)
2.	adamic.research.microsoft.com	Microsoft	329	10,237 (81.13%)
3.	www.ftsm.ukm.iny	FTSM	291	9,876 (78.27%)
4.	www.scribd.com	Scribd	279	8,623 (68.34%)
5.	mjs.fsktm.um.edu.iny	Myais	124	3,000 (23.78%)
6.	research.ukm.iny	Research	63	1,043 (8.27%)

4. Experiment and Discussion

We conduct an experiment by involving 462 social actors. The social network extraction by using a method that revealed hit count singleton and hit count doubleton yielded 31, 623 strength relations between social actors, or 29.70% of 462 461/2 potential relations. Through the same experiment, the doubleton snippet and the URL addresses doubleton can be processed the alternative relationship between two actors. Through the collection of snippets, words and URL addresses in doubleton we generate the descriptions of each strength relation.

Thus, in this experiment the usability of the URL addresses for each social actor can be obtained, then we use for grouping strength relations based on the URL addresses as the social network clusters, see Table 1. The similarity between clusters shows that, although number of vertices and edges between clusters and other clusters indicate the interdependence between one social network with another, but no social network sequence based on that cluster becomes purely a part of another social sequence, or we cannot found that a community becomes part of another community as a whole based on the servers addresses, see Fig. 3.

Generating the alternative relationships integrately involves a singleton snippet containing the words and URL addresses. Although for 462 social actors, 12, 158 strength relations or 11, 41% of computation totality was done and the results of expansion was 11, 015 relations, while through the expanded URL addresses were 70, 604 relations. In an integrated manner, the social network expanded to 105, 168 edges or 98.76% of the potential relations.

5. Conclusion

The integrated superficial method is as an approach involving hit count, snippets or set of words and a collection of URL addresses either singleton or doubleton. This integration of information sources be an indicator of the relationships between social actors, not only provides the descriptions of strength relations but also results in the expansion of social networks in other forms. It is possible that the existence of built communities based on the underlying descriptions or the descriptions as explanation. With the information of clusters it is possible to see social behavior based on social structure, this be the target of subsequent research, as a follow-up in future.

References

- [1]^[0] Chakraborty T and Kearns M 2008 Bargaining solutions in a social network *WINE 2008 LNCS* **5385**.
- [2]^[0] Boyd D M and Roychowdhury V P 2008 Social network sites: Definition, History, and Scholarship *Journal of Computer-Mediated Communication* **13**.
- [3]^[0] Johnsen J W and Franke K 2017 Feasibility study of social network analysis on loosely structured communication networks *Procedia Computer Science* **108**.
- [4]^[0] Tang J, Zhang D, Yao L and Li J 2008 Extraction and mining of an academic social network *Proceedings of WWW 2008*.
- [5]^[0] Nasution M K M 2016 Social network mining (SNM): A definition of relation between the resources and SNA *International Journal on Advanced Science, Engineering and Information Technology* **6(6)**.
- [6]^[0] Matsuo Y, Mori J, Hamasaki M, Nishimura T, Takeda H, Hasida K and Ishizuka M 2006 POLYPHONET: An advanced social network extraction system *Proceedings of the 15th International Conference on World Wide Web (WWW 2006)*.
- [7]^[0] Nasution M K M, Sitompul O S, Sinulingga E P and Noah S A 2016 An extracted social network mining *Proceedings of 2016 SAI Computing Conference (SAI 2016)*.
- [8]^[0] Kautz H, Selman B and Shah M 1997 Referral Web: Combining social networks and collaborative filtering *Communication of the ACM* **40(3)**.
- [9]^[0] Mika P 2005 Flink: Semantic Web technology for the extraction and analysis of social networks *Web Semantics: Science, Services and Agent on the World Wide Web* **3(2-3)**.
- [10]^[0] Jin Y, Matsuo Y and Ishizuka M 2006 Extracting a social network among entities by web mining *Workshop on Web Content Mining with Human Language (ISWC 2006)*.
- [11]^[0] Jin Y, Matsuo Y, and Ishizuka M 2007 Extracting social networks among various entities on the Web *ESWC 2007 LNCS* **4519**.
- [12]^[0] Matsuo Y, Mori J, Hamasaki M, Nishimura T, Takeda T, Hasida K and Ishizuka M 2007 POLYPHONET: An advanced social networks extraction system from the Web *Journal of Web Semantics: Science, Services and Agents on the World Wide Web* **5**: 262-278.
- [13]^[0] Nasution M K M 2017 Social network extraction based on Web: 1. Related superficial methods *International Conference on Operational Research (InterIOR)*.
- [14]^[0] Nasution M K M and Noah S A 2017 Social network extraction based on Web: 2. Comparison of superficial methods *Procedia Computer Science*.
- [15]^[0] Nasution M K M and Noah S A 2010 Superficial method for extracting social network for academics using web snippets *Lecture Notes in Computer Science LNAI* **6401**.
- [16]^[0] Nasution M K M and Noah S A 2011 Extraction of academic social network from online database *2011 International Conference on Semantic Technology and Information Retrieval (STAIR 2011)*.
- [17]^[0] Nasution M K M, Sitompul O S, Nasution S, and Ambarita H 2017 New similarity *IOP Conference Series: Materials Science and Engineering* **180(1)**.
- [18]^[0] Nasution M K M 2017 Semantic interpretation of search engine resultant *International Conference on Operational Research (InterIOR)*.
- [19]^[0] Nasution M K M and Sitompul O S 2017 Enhancing Extraction Method for Aggregating Strength Relation Between Social Actors *Artificial Intelligence Trends in Intelligent Systems (AISC)* **573**.
- [20]^[0] Culotta A, Bekkerman R and McCallum A 2004 Extracting social networks and contact information from email and the Web *Computer Science Department Faculty Publication Series* **33**.
- [21]^[0] Mitra A, Paul S, Panda S and Padhi P 2016 A study on the representation of the various models for dynamic social networks *Procedia Computer Science* **79**.
- [22]^[0] Nasution M K M and Noah S A 2012 Information retrieval model: A social network extraction perspective *Proceedings - 2012 International Conference on Information Retrieval and Knowledge Management*.
- [23]^[0] Nasution M K M 2014 New method for extracting keyword for the social actor. *Lecture Notes in Computer Science LNAI* **8397 (PART 1)**.
- [24]^[0] Nasution M K M 2017 Modelling and Simulation of Search Engine *Journal of Physics: Conference*

Series 801(1).

- [25]^[0] Nasution M K M, Noah S A and Saad S 2011 Social network extraction: Superficial method and information retrieval *Proceedings of International Conference on Informatics for Development (ICID11)*.
- [26]^[0] Nasution M K M and Noah S A 2012 A methodology to extract social network form the Web Snippet *Cornell University Library arXiv:1211.5877*.
- [27]^[0] Nasution M K M, Hardi M and Syah R 2017 Mining of the social network extraction *Journal of Physics: Conference Series 801(1)*.
- [28]^[0] Nasution M K M, Syah R and Elfida M 2018 Information retrieval based on the extracted social network *Applied Computational Intelligence and Mathematical Method, Advances in Intelligent Systems and Computing 662*.