



PREFACE

ISICO 2017 is the fourth edition of the Information Systems International Conference (ISICO). ISICO 2017 has taken place at Sanur Paradise Plaza Hotel, Bali Indonesia, between 6th and 8th November, 2017. The theme of the conference is "Innovation of Information Systems – visions, opportunities and challenges". ISICO 2017 has been hosted by Department of Information Systems, Institut Teknologi Sepuluh Nopember (ITS). In 2013, ISICO becomes the official Association for Information Systems (AIS) Indonesia Chapter (named AISINDO) affiliated conference.

The main philosophy of ISICO's presence is that to widen the research collaboration among IS researchers worldwide. Especially (but not limited) those that of developing countries. To this point, ISICO does not aim to replace existing IS Conferences such as PACIS, AMCIS, ICIS, or ECIS but rather it complements their presences.

The salient features of ISICO are the keynotes and Scopus-index Elsevier publication support. In terms of Keynotes, we regularly invites Global AIS President such as Prof. Douglas Vogel, Hong Kong (ISICO 2013), Prof Jaekyu Lee, Korea (ISICO 2015), and Prof Matti Rossi, Finland (ISICO 2017).

This year ISICO was able to attract more than 153 submissions from 16 different countries. From those submissions only 93 were selected for publication, so the acceptance rate this year was 60.7%. This Conference Proceedings volume contains all papers accepted for publication in ISICO 2017. ISICO 2017 comprises a series of independent tracks that are relevant to Information Systems discipline. The conference cover enterprise systems track, information systems management track, data acquisition and information dissemination track, data engineering and business intelligence track, and IT infrastructure and security track.

We would like to thank Department of Information Systems, Institut Teknologi Sepuluh Nopember for hosting the conference, as well as all participants for their contributions. We would also like to thank our distinguished program committee members for the efforts they have put in reviewing the papers. Special thanks to Prof. Matti Rossi, Dr. Ahmed Imran and Prof. Caroline Chan for the keynote speech.

We are looking forward to the Fifth Information Systems International Conference (ISICO 2019).

Conference Chair:

Faizal Mahananto, PhD (Institut Teknologi Sepuluh Nopember, Indonesia)

GUEST EDITOR

Khin Lwin, PhD

LIST OF REVIEWER

Ahmad Muklason	Amalia Utamima
Amna Shifia Nisafani	Amy Connolly
Angelia Melani Adrian	Anisah Herdiyanti
Anushia Inthiran	Apol Pribadi
Aravind Sesagiri Raamkumar	Arif Wibisono
Aris Tjahyanto	Bekti Cahyo Hidayanto
Cecil Donald	Dedi Iskandar Inan
Dr. Mohamad Taha Ijab	Dr. Ruchi Nanda
Dwi Yuli Rakhmawati	Eko Wahyu Tyas Darmaningrat
Erma Suryani	Faizal Johan Atletiko
Faizal Mahananto	Febriyian Samopa
Feby Artwodini	Flavio Horita
Gali Naveh	Hatma Suryotrisongko
Henning Titi Ciptaningtyas	Hudan Studiawan
Irmasari Hafidz	Jamal Elden
Kauser Ahmed	Keng Hoon Gan
Khakim Ghozali	Komarudin
Mahendrawathi Er	Muhammad Hafidz Fazli Bin Md Fauadi
Muhammad Nazrul Islam	Nisfu Asrul Sani
Nuno Laranjeiro	Nur Aini Rakhmawati
Nurlida Basir	R.S. Ajin
Rahmat Trialih	Rajamohana SP
Ratna Sari Dewi	Renny Pradina Kusumawardani
Retno Aulia Vinarti	Roslina Ibrahim
Rully Agus Hendrawan	Samiaji Sarosa
Satria Fadil Perdana	Sholiq
Sudipta Roy	Syed Nasirin
Tony Dwi Susanto	Tse Guan Tan
Utku Kose	Wan Mohd Nazmee Wan Zainon
Wira Redi	Wiwik Anggraeni

Yanti Andriyani
Yuliani Dwi Lestari

Yong Liu
Yusraini Muharni

Table of Contents

Preface	
Faizal Mahananto.....	1
Engaging with Customer Using Social Media Platform: A Case Study of Malaysia Hotels	
Kamarul Faizal Hashim, and Nawar Abbood Fadhil.....	4
Self-Branding on Social Media: An Analysis of Style Bloggers on Instagram	
Rendan Liu, and Ayoung Suh.....	12
Indonesia local government information completeness on the web	
Fajara Kurniawan, Nur Aini Rakhmawati, Abi Nubli Abadi, Muhammad Zuhri, and Wisnu Tri Sugiyanto.....	21
Community Detection On Citation Network Of DBLP Data Sample Set Using LinkRank Algorithm	
Satrio Baskoro Yudhoatmojo, and Muhammad Arvin Samuar.....	29
The Utilization of Filter on Object-based Opinion Mining in Tourism Product Reviews	
Aris Tjahyanto, and Bonda Sisephaputra.....	38
The Performance of Ant System in Solving Multi Traveling Salesmen Problem	
Eka N. Kencana, IDa Harini, and K. Mayuliana.....	46
Surveying LinkedIn Profiles of Data Scientists: The Case of the Philippines	
Jerina Jean Ecleo, and Adrian Galido.....	53
Nusantara: A New Model of Knowledge Management in Government Human Capital Management	
Elin Cahyaningsih, Dana Indra Sensuse, Aniati Murni Arymurthy, and Wahyu Catur Wibowo.....	61
Spatial data utilization for location pattern analysis	
Dyah Lestari WIDaningrum, Isti Surjandari, and Aniati Murni Arymurthy.....	69
Effects of Word Class and Text Position in Sentiment-based News Classification	
June Ling Ong Hui, Gan Keng Hoon, and Wan Mohd Nazmee Wan Zainon.....	77
Social Network Extraction Based on Web. A Comparison of Superficial Methods	
Mahyuddin K.M. Nasution, and Shahrul Azman Noah.....	86
Data Warehouse with Big Data Technology for Higher Education	
Leo Willyanto Santoso, and Yulia.....	93
Teenstagram TimeFrame: A Visualization for Instagram Time Dataset from Teen Users (Case Study in Surabaya, Indonesia)	
Irmasari Hafidz, Alvin Rahman Kautsar, Tetha Valianta, and Nur Aini Rakhmawati.....	100
A Study on the Effectiveness of Tree-Maps as Tree Visualization Techniques	
Lim Kian Long, Lim Chien Hui, Gim Yeong Fook, and Wan Mohd Nazmee Wan Zainon.....	108
Disaster Knowledge Management Analysis Framework Utilizing Agent-Based Models: Design Science Research Approach	
Dedi Iskandar Inan, and Ghassan Beydoun.....	116

An Adjustable Autonomy Management Module for Multi-agent Systems Salama A. Mostafa, AIDa Mustapha, Mohd Shariffuddin Ahmad, and Moamin A Mahmoud.....	125
Coupled HIDDEN Markov Model for Process Discovery of Non-Free Choice and Invisible Prime Tasks Riyanarto Sarno, and Kelly R. Sungkono.....	134
Modified Regression Approach for Predicting Number of Dengue Fever Incidents in Malang Indonesia Wiwik Anggraeni, Rafika Nurmasari, Edwin Riksakomara, Febriliyan Samopa, Radityo Prasetyanto Wibowo, Lulus Condro T., and Pujiadi.....	142
Killer Whale Algorithm: An Algorithm Inspired by the Life of Killer Whale Totok R. Biyanto, Matradji, Sonny Irawan, Henokh Y. Febrianto, Naindar Afdanny, Ahmad H. Rahman, Kevin S. Gunawan, Januar A.D. Pratama, and Titania N. Bethiana.....	151
Application of Killer Whale Algorithm in ASP EOR Optimization Totok R. Biyanto, Matradji, Sawal, Ahmad H. Rahman, Arfiq I. Abdillah, Titania N. Bethiana, and Sonny Irawan.....	158
Tracking People by Detection Using CNN Features Dina Chahyati, Mohamad Ivan Fanany, and Aniati Murni Arymurthy.....	167
Advanced Traveler Information System: Itinerary Optimization as an Orienteering Problem Using Iterative Local Search-Hill Climbing Algorithm Jockey Satria Wijaya, Wiwik Anggraeni, Ahmad Muklason, Faizal Mahananto, Edwin Riksakomara, and Arif Djunaidy.....	173
Estimating Gas Concentration using Artificial Neural Network for Electronic Nose Shoffi Izza Sabilla, Riyanarto Sarno, and Joko Siswantoro.....	181
The Performance of ARIMAX Model and Vector Autoregressive (VAR) Model in Forecasting Strategic Commodity Price in Indonesia Wiwik Anggraeni, Kuntoro Boga Andri, Sumaryanto, and Faizal Mahananto.....	189
Simple Symbolic Dynamic of Heart Rate Variability Identify Patient with Congestive Heart Failure Faizal Mahananto, and Arif Djunaidy.....	197
A Framework for Knowledge Based Software Service Supply Chain (SSSC): A Comparative Analysis with Existing Frameworks Ali Baydoun, and Jamal El-Den.....	205
ERP Post Implementation Review with Process Mining: A Case of Procurement Process Mahendrawathi ER, Shania Olivia Zayin, and Firman Jati Pamungkas.....	216
Scalable indexing algorithm for multi-dimensional time-gap analysis with distributed computing Riska Asriana Sutrisnowati, Bernardo Nugroho Yahya, Hyerim Bae, Iq Reviessay Pulshashi, and Taufik Nur Adi.....	224
The Development of Photovoltaic Power Plant for Electricity Demand Fulfillment in Remote Regional of Madura Island using System Dynamics Model Lilia Trisyathia Quentara, and Erma Suryani.....	232
Developing Salesperson Performance Indicators on Instant Messaging Platform Amna Shifia Nisafani, Arif Wibisono, Safrina Kharisma Imandani, and Radityo Prasetianto Wibowo.....	239
Public Sector Accountants' Opinion on Impact of a New Enterprise System Zaini Zainol, Dahlia Fernandez, and Hawa Ahmad.....	247
Structural Similarity Measurement of Business Process Model to Compare Heuristic and Inductive Miner Algorithms Performance in Dealing with Noise Ifrina Nuritha, and Mahendrawathi ER.....	255
Curriculum Assessment of Higher Educational Institution Using Aggregate Profile Clustering Satrio Adi Priyambada, ER Mahendrawathi, and Bernardo Nugroho Yahya.....	264

Analyzing the Effectiveness of Public e-Marketplaces for Selling Apparel Products in Indonesia Amna Shifia Nisafani, Arif Wibisono, and Muchammad HaIDar Tegar Revaldo.....	274
Evaluation of E-Commerce Product Reviews Based on Structural, Metadata, and Readability Characteristics Rully Agus Hendrawan, Erma Suryani, and Rani Oktavia.....	280
Motivational Factors for Knowledge Sharing using Pedagogical Discussion Cases: Students, Educators, and Environmental Factors Narumon Sriratanaviriyakul, and Jamal El-Den.....	287
The impact of Knowledge Management on Organizational Productivity: A Case Study on Koosar Bank of Iran Fatemeh Torabi, and Jamal El-Den.....	300
Multiview Similarity Assessment Technique of UML Diagrams Alhassan Adamu, and Wan Mohd Nazmee Wan Zainon.....	311
Effects of Technology Readiness Towards Acceptance of Mandatory Web-Based Attendance System Mahendra Adhi Nugroho, and M. Andryzal Fajar.....	319
Exploratory Study of SMEs Technology Adoption Readiness Factors Mahendra Adhi Nugroho, Arief Zuliyanto Susilo, M. Andryzal Fajar, and Diana Rahmawati.....	329
An Overview of Software Functionality Service: A Systematic Literature Review Masrina A. Salleh, Mahadi Bahari, and Nor Hidayati Zakaria.....	337
Analysis on Factors Influencing Textile Cyberpreneur’s Intention to Adopt Cloud-Based m-Retail Application Wan Safra Diyana Wan Abdul Ghani, Nik Zulkarnaen Khidzir, Tan Tse Guan, and Mohammad Ismail.....	345
The Organization Factors as Barrier for Sustainable Health Information Systems (HIS) – A Review Noor Azizah Mohamadali, and Nurul Aqilah Zahari.....	354
Beyond Organizational Motives of e-Government Adoption: The Case of e-Voting Initiative in Indonesian Villages Manik Hapsara, Ahmed Imran, and Timothy Turner.....	362
The Technology Factors as Barriers for Sustainable Health Information Systems (HIS) – A Review Noor Azizah Mohamadali, and Nur Faizah Ab Aziz.....	370
User Satisfaction and Intention to Use Peer-to-Peer Online Transportation: A Replication Study Adhi Setyo Santoso, and Liza Agustina Maureen Nelloh.....	379
The study on negative eWOM and its relationship to consumer’s intention to switch Mobile Service Provider Geetha Nadarajan, Jamil Bojei, and Haliyana KhalID.....	388
Millennials’ Perception on Mobile Payment Services in Malaysia Yeow Pooi Mun, Haliyana KhalID, and Devika Nadarajah.....	397
The Use of Analytic Hierarchy Process for Software Development Method Selection: A Perspective of e-Government in Indonesia Melisa Helingo, Betty Purwandari, Riri Satria, and Iis Solichah.....	405
A Modification Complexity Factor in Function Points Method for Software Cost Estimation Towards Public Service Application Renny Sari Dewi, Apol Pribadi Subriadi, and Sholiq.....	415
The Effect of Social Media to Cultural Homecoming Tradition of Computer Students in Medan Arif RIDho Lubis, Ferry Fachrizal, and Muharman Lubis.....	423
Understanding the Total Value of Information Technology Services from the Perspective of Students and Academic Staffs Anisah Herdiyanti, Nanda Restanena Listyawati, and Hanim Maria Astuti.....	429

WebGIS for Asset Management of Land and Building of Madiun City Government R.V. Hari Ginardi, Wawan Gunawan, and Septiawan Rosetya Wardana.....	437
A New Approach of Indonesian University Webometrics Ranking Using Entropy and PROMETHEE II Handaru Jati, and Dhanapal Durai Dominic.....	444
Understanding the Implementation of Telerehabilitation at Pre-Implementation Stage: A Systematic Literature Review Tiara Izrinda Jafni, Mahadi Bahari, WaIDah Ismail, and Abduljalil Radman.....	452
A HybrID Cuckoo Optimization and Harmony Search Algorithm for Software Cost Estimation Alifia Puspangrum, and Riyanarto Sarno.....	461
A Comparative Study of Software Development Size Estimation Method: UCPabc vs Function Points Sholih, Renny Sari Dewi, and Apol Pribadi Subriadi.....	470
Cyclomatic Complexity for Determining Product Complexity Level in COCOMO II Muhammad Asep Subandri, and Riyanarto Sarno.....	478
Service Quality Analysis for Online Transportation Services: Case Study of GO-JEK Shilvia L. Br. Silalahi, Putu W. Handayani, and Qorib Munajat.....	487
Multi Methods for Knowledge Management Strategy Roadmap of Government Human Capital Management Elin Cahyaningsih, Dana Indra Sensuse, and Handrie Noprisson.....	496
Factors that Affecting Behavioral Intention in Online Transportation Service: Case study of GO-JEK Rizky Septiani, Putu Wuri Handayani, and Fatimah Azzahro.....	504
Usability Evaluation to Enhance Software Quality of Cultural Conservation System Based on Nielsen Model (WikiBudaya) Feby Artwodini Muqtadiroh, Hanim Maria Astuti, Eko Wahyu Tyas Darmaningrat, and Fenty Rizky Aprilian.....	513
Determinants of CAATT acceptance: Insights from public accounting firms in Indonesia Muhammad Rifki Shihab, Nina Meilatinova, Achmad Nizar HIDayanto, and Herkules.....	522
Users' Motivation in Sharing Information on Social Media Afira Putri Ghaisani, Putu Wuri Handayani, and Qorib Munajat.....	530
The Moderation Effect of Age on Adopting E-Payment Technology Anggar Riskinanto, Bayu Kelana, and Deliar Rifda Hilmawan.....	536
Barriers to Electronic Health Record System Implementation and Information Systems Resources: A Structured Review Jaillah Mae Gesulga, Almarie Berjame, Kristelle Sheen Moquiala, and Adrian GalIDo.....	544
E-Government Integration through Implementation of web-based GIS on Community Health monitoring in Jembrana Regency, Bali Jatmiko Wahyu Nugroho Joshua, I Putu Agus Swastika, and Tri Oktin Windha Daniaty.....	552
User Acceptance of e-Government Citizen Report System (a Case Study of City113 App) Tony Dwi Susanto, Made Mira Diani, and Irmasari HafIDz.....	560
Risks Assessment of Information Technology Processes Based on COBIT 5 Framework: A Case Study of ITS Service Desk Hanim Maria Astuti, Feby Artwodini Muqtadiroh, Eko Wahyu Tyas Darmaningrat, and Chitra Utami Putri.....	569
Designing an Effective Collaboration using Information Technology Towards World Class University Linda Salma Angreani, and Annas Vijaya.....	577

Understanding of Public Behavioral Intent to Use e-Government Service: An Extended of Unified Theory of Acceptance Use of Technology and Information System Quality Berlilana, Taqwa Hariguna, and Nurfaizah.....	585
The Development of Work Instruction as a Solution to Handle IT Critical Incidents in Units within an Organization Febriliyan Samopa, Hanim Maria Astuti, and Mahesti Ayu Lestari.....	593
Understanding the Quality Gap of Information Technology Services from the Perspective of Service Provider and Consumer Anisah Herdiyanti, Alitya Novianda Adityaputri, and Hanim Maria Astuti.....	601
Challenges and Solutions for Applications and Technologies in the Internet of Things Saad Albishi, Ben Soh, Azmat Ullah, and Fahad Algarni.....	608
“Four-Hospitality: Friendly Smart City Design for Disability” Hatma Suryotrisongko, Reginia Cindy Kusuma, and RV Hari Ginardi.....	615
Usable Security: Revealing End-Users Comprehensions on Security Warnings Ammar Amran, Zarul Fitri Zaaba, Manmeet Mahinderjit Singh, and Abdalla Wasef Marashdih.....	624
Information Privacy Concerns on Teens as Facebook Users in Indonesia Ari Kusyanti, Dita Rahma Puspitasari, Harin Puspa Ayu Catherina, and Yustiyana April Lia Sari.....	632
A Review on Cloud Computing Acceptance Factors Mohd Talmizie Amron, Roslina Ibrahim, and Suriayati Chuprat.....	639
Cross Site Scripting: Removing Approaches in Web Application Abdalla Wasef Marashdih, and Zarul Fitri Zaaba.....	647
Security Strategies for Hindering Watering Hole Cyber Crime Attack Khairun Ashikin Ismail, Manmeet Mahinderjit Singh, Norlia Mustaffa, Pantea Keikhosrokiani, and Zakiah Zulkeffi.....	656
Typosquat Cyber Crime Attack Detection via Smartphone Zakiah Zulkeffi, Manmeet Mahinderjit Singh, Azizul Rahman Mohd Shariff, and Azman Samsudin.....	664
A Study on Intrusion Detection Using CentroID-Based Classification Bambang Setiawan, Supeno Djanali, and Tohari Ahmad.....	672
Analysis the Performance of Vehicles Ad Hoc Network Saed Tarapiah, Kahtan Aziz, and Shadi Atalla.....	682
Developing an Information Security Policy: A Case Study Approach Fayez Hussain Alqahtani.....	691
Design and Implementation of Real-Time Mobile-based Water Temperature Monitoring System Paul B. Bokingkito, and Orven E. Llantos.....	698
Mobile Web Energy Monitoring System Using DFRduino Uno Kristine Mae E. Galera, and Orven E. Llantos.....	706
A performance evaluation for assessing registered websites Nur Aini Rakhmawati, Valliant Ferlyando, Febriliyan Samopa, and Hanim Maria Astuti.....	714
The Existence Of Cryptography: A Study On Instant Messaging Vania Beatrice Liwandouw, and Alz Danny Wowor.....	721
Development of mobile electronic nose for beef quality monitoring Dedy Rahman Wijaya, Riyanarto Sarno, Enny Zulaika, and Shoffi Izza Sabila.....	728
Design and Development of Backend Application for Public Complaint Systems Using Microservice Spring Boot Hatma Suryotrisongko, Dedy Puji Jayanto, and Aris Tjahyanto.....	736

APEX System: An Integration of Management Information Concept Aries Muftie, Djoko Budhi Setyawan, Supardi, Iwan Fuad, and Andre Parvian Aristio.....	744
Network Intrusion Detection Systems Analysis using Frequent Item Set Mining Algorithm FP-Max and Apriori Bekti Cahyo HIDayanto, Rowi Fajar Muhammad, Renny P Kusumawardani, and Achmad Syafaat.....	751
Development of AndroID Application for Courier Monitoring System Faizal Johan Atletiko.....	759



4th Information Systems International Conference 2017, ISICO 2017, 6-8 November 2017,
Bali, Indonesia

Social Network Extraction Based on Web. A Comparison of Superficial Methods

Mahyuddin K. M. Nasution^{a,*}, Shahrul Azman Noah^b

^aInformation Technology Department, Universitas Sumatera Utara, Medan 20155, Sumatera Utara, Indonesia

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

Abstract

The Web as a source of information has many potentials which allow to use the different treatments on social network extraction methods. The approaches generally we identified as superficial methods in unsupervised stream. However, the same resources of social networks, i.e. based on a community of social actors, reveal many of different approaches to produce social networks. Therefore, based on a treatment to another treatments, from the given treatments until the different social networks generated and it has been declared different methods. It requires comparison to reveal the properties of social networks and their methods in this paper. It is revealed that there is a core social network has similarity with other social networks is more than 1% as general property of the extracted social networks, whereby there is a social network for different methods has the common edges in graph.

© 2018 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 4th Information Systems International Conference 2017.

Keywords: Superficial method, occurrence, co-occurrence, hit count, snippet, URL address, Web, vertices, social actors, edges, strength relations

1. Introduction

The Web as an information source has a lot of potential to be extracted into consideration for decision making [1]. Extracting social network from Web not only considers the available potential only but systematically is also the way to gain structural social behavior [2]. Therefore, the methods applied for extraction also vary, and each involves the different potentialities of information source, although it involves the same social actors. However, since the method externally is outside the search engine system [3], although the methods in unsupervised stream heavily depends on the search engine [4], consequently the method can't fully utilize the information resources optimally [5], moreover the methods in supervised stream [6] such as the use of: Latent Dirichlet Allocation (LDA) [7] or Hidden Markov Model (HMM) [8], for example.

* Corresponding author. Tel.: +62-061-821-1633 ; fax: +62-061-821-9411.

E-mail address: mahyuddin@usu.ac.id

The different information involvement of the Web in disclosing social networks has resulted in the different approaches for extracting social network from Web [9, 10]. However, in the same potential, it has not been revealed that there is a difference between these approaches. Therefore, this paper will reveal the performance of each approach in the superficial methods involving the experiment of social network extraction.

2. A Review

In the formal definition of social networks, expressed in graph theory $G(E, V)$, it has been disclosed that there is a set of vertices V and a set of edges E , with which $v_i \in V \ i = 1, \dots, n$ denotes entities in social networks and $v_i v_j \in E$ denotes relationship between entities $v_i, v_j \in V$ in a social network [11, 12]. This definition reveals that in the extraction of social networks using the basic superficial method (BSM) occurs the process (a) determines the social actors and (b) builds relationships between them [13, 14]. Next we consider some basic characters related to the superficial methods.

In the first process, the use of the social name a_i (without quotes) in the query q is to represent a social actor or

$$\mathbf{a}_i = q \leftarrow a_i, \quad (1)$$

then generally search engines generate ambiguous information about the actor social [15]. However, with the addition of the keyword kw , in general it can reduce naturally the default property of used social name [13, 16], i.e. consequent of

$$\mathbf{aw}_i = q \leftarrow a_i, kw \quad (2)$$

happen reduction ambiguity, with which $|aw_i| \leq |a_i|$, $|a_i| \in \mathbf{a}_i$ is a cardinality of a_i and $|aw_i| \in \mathbf{aw}_i$ is a cardinality of a_i, kw [17]. While using the well-defined name of social actor (in quotes) in the query q will raises the entire social actors related information or

$$\mathbf{a}^{i'} = q \leftarrow "a_i". \quad (3)$$

In last case, $|a^{i'}| \leq |a_i|$, and $|a^{i'}| \in \mathbf{a}^{i'}$ is a cardinality of " a_i " [18]. As well as with

$$\mathbf{aw}^{i'} = q \leftarrow "a_i", "kw". \quad (4)$$

is about one of information concentrations of a social actor, $|aw^{i'}| \leq |a^{i'}|$, and $|aw^{i'}| \in \mathbf{aw}^{i'}$ is a cardinality of " a_i ", " kw " [19, 20, 17].

In the second process, the relationship between two social actors is based on the concept of co-occurrence [21, 22, 23]. Thus,

$$\mathbf{a}_i \mathbf{a}_j = q \leftarrow a_i, a_j, \quad (5)$$

is a process to elevate the clue of relation be relationship between two actors, with which $|a_i \cap a_j| \leq |a_i|$ and $|a_i \cap a_j| \leq |a_j|$, and $|a_i \cap a_j| \in \mathbf{a}_i \mathbf{a}_j$ is a cardinality of a_i, a_j . The addition of a keyword towards the co-occurrence will usually reduce the number of information presented, that is

$$\mathbf{aw}_i \mathbf{aw}_j = q \leftarrow a_i, a_j, kw, \quad (6)$$

but it should meet that $|aw_i \cap aw_j| \leq |a_i \cap a_j|$, $|aw_i \cap aw_j| \in \mathbf{aw}_i \mathbf{aw}_j$ is a cardinality of a_i, a_j, kw [10, 23]. Likewise, the use of the well-defined name of social actor in the query will reveal the relationship between two social actors appropriately, that is

$$\mathbf{a}^{i'} \mathbf{a}^{j'} = q \leftarrow "a_i", "a_j". \quad (7)$$

whereby $|a^{i'} \cap a^{j'}| \leq |a^{i'}|$ and $|a^{i'} \cap a^{j'}| \leq |a^{j'}|$, $|a^{i'} \cap a^{j'}| \in \mathbf{a}^{i'} \mathbf{a}^{j'}$ is a cardinality of " a_i ", " a_j ". Thus, if $|a^{i'} \cap a^{j'}| = 0$, then it means exactly there is no relationship between two social actors [14]. Whereas the addition of a keyword towards the co-occurrence based on pattern as follows

$$\mathbf{wa}^{i'} \mathbf{wa}^{j'} = q \leftarrow "a_i", "a_j", "kw", \quad (8)$$

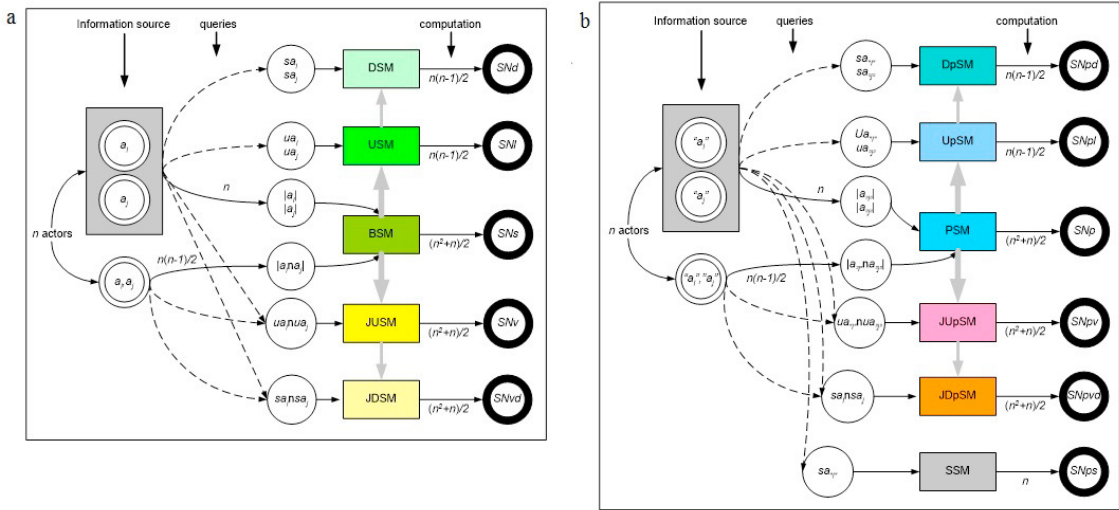


Fig. 1. A variety of social network extraction methods

produces the relationship of a pair of social actors, whereby $|aw^{a_i} \cap aw^{a_j}| \leq |a^{a_i} \cap a^{a_j}|$, and $|aw^{a_i} \cap aw^{a_j}| \in wa^{a_i} wa^{a_j}$ is a cardinality of "a_i", "a_j", "kw" [24].

In this case of occurrence, if each of a_i, "a_i", a_i, kw, and "a_i", "kw" has a cardinality greater than 0, then each of them contains one or more snippets and URL addresses [17], i.e. s_i, sa^{a_i}, saw^{a_i}, and saw^{a_i} respectively is a snippet of them, so ua_i, ua^{a_i}, uaw_i, and uaw^{a_i} respectively is a collection of URL addresses of them [4, 25]. Similar to co-occurrence, each of a_i, a_j, "a_i", "a_j", a_i, a_j, kw, and "a_i", "a_j", "kw" has a list of snippets or a list of URL addresses, if they have a cardinality greater than 0. By involving Eq. (1) and $ua_i \in \mathbf{a}_i$, for a pair of social actors a social network SNl can be generated. We call this approach the underlying superficial method (USM) whereby the established relationship can be validated by involving $ua_i, ua_j \in \mathbf{a}_i, \mathbf{a}_j$ or using Eq. (5) in the joint underlying superficial method (JUSM) for generating a social network SNv as validation of SNl. Similar to this concept, because $sa_i \in \mathbf{a}_i$ used in the descriptive superficial method (DSM) and generates descriptive social network SNd, whereby $sa_i, sa_j \in \mathbf{a}_i, \mathbf{a}_j$ that used in the joint descriptive superficial method (JDSM) generates the descriptive social network SNvd as validation of SNd, see Fig. 1a. Likewise, if the concept of pattern is used in a query, various social networks will be generated. But on the use of pattern, the possibility of returning information can be more appropriate so that it is used as an access point to more deep information, for example accessing the online database [26, 18]. We call the latter approach the seed based superficial method (SSM) that is sometimes it is used as the basis for development of social network in a different way than usual [27, 28], see Fig. 1b. In this case, it is a social network based on seed SNps.

3. The Proposed Approach

Through the above reviews have been generated some treatments that may be used in methods for extracting social network from Web. Each treatment comes from the same potential and with different approaches [29]. The potential used in the extraction of a social network is the number of social actors as vertices in social network. The number of social actors is equal to the number of the required occurrence, whereby occurrence as a treatment, but the realization in the number of hit counts may be different or all hit counts must be greater than 0 [30, 31]. Thus, hit counts equal to 0 doesn't present snippet and URL addresses so no treatment may be applied there.

Involvement of occurrence and/or co-occurrence in each method, results in the presence of different treatments as different approaches. This different treatments, not only changes the complexity but also the presence of pseudo similarity that helps in solving of resource constraints. The resource limitations associated with using search engines to respond to the queries submitted. Therefore, BSM [15, 32, 33], USM [4] and DSM [18, 34] differ in resource engagement iterations, but they still are the same computationally. In addition, in certain methods it can also be expanded

Table 1. List of methods and their complexity and realization

Method	Computation	occurrence	co-occurrence	vertices	realization	edges	realization
BSM	106,491	462	106,491	462	100.00%	31,623	29.70%
BSMv1	106,491	462	106,491	373	80.74%	16,388	15.39%
BSMv2	106,491	462	106,491	389	84.20%	17,132	16.09%
PSM	106,491	462	106,491	462	100.00%	22,836	21.44%
PSMv1	106,491	462	106,491	369	84.20%	16,371	15.38%
PSMv2	106,491	462	106,491	382	82.68%	17,120	16.07%
USM	106,491	462	0	462	100.00%	99,150	93.11%
DSM	106,491	462	0	434	93.94%	12,158	11.41%
SSM	-	76	0	462	607.89%	1,297	1.22%

with different approaches, such as BSM converted into BSMv1 and / or BSMv2 by adding different or equal keywords, so the possibility of extraction results will be different [19, 10]. Therefore, the comparison of the complexity of performance on each method based on the reviews will show the feasibility and reliability of methods for the same purpose in the extraction of social networks from the web [35].

The outcome of the method also shows the feasibility and reliability of the method. Output is about how much relations in a social network, or number of edges as realization. The density of the social network as resultant of method is determined by the many relationships of two social actors as outcomes. Therefore, if the relation between two social actors is not limited to emotional closeness but also ideas and concepts or other relationships, then the outcome of the used method requires another consideration either from the point of density or description [12].

So, the way to see the performance of social network extraction methods is to compare some of the following,

1. Measurement of complexity: Computation (based on iteration), number of occurrences, and number of co-occurrences.
2. Percentage of output to number of relations potentially: realization of occurrences be vertices, and realization of co-occurrences be edges.
3. Similarities between the social networks, i.e. by using Jaccard coefficient based on edges in social networks [36, 11]:

$$sim_e = \frac{|E_1 \cap E_2|}{|E_1| + |E_2| - |E_1 \cap E_2|} \quad (9)$$

$|E_1 \cap E_2|$, $|E_1|$, and $|E_2|$ are number of edges in social networks: $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$.

4. Experiment and Discussion

By using 76 social actors (professor) as seeds, as much as 385 other social actors is generated as potential for building social network. The first approach as an initiation we use the SSM method by involving the source of information from the online database (such as DBLP) directly through a URL address based on each seed and extracting the names of other social actors in co-occurrence formation. It is to get current information about relations between social actors [18], and as contra against manual effort and not up-to-date information of DBpedia [38]. So, in this experiment there are 462 social actors. We use this collection of social actors to form social networks both with other superficial methods and different sources of information from the Web.

Each method produces different number of edges, although by using same number of vertices. In addition, methods produce the different number of actors who have no relationship with others [37]. In order to produce the edges in the network, each method produces unequal strength relations between the pairs of social actors, there are several strength relations between social actors can be recognized by one method but cannot be disclosed by other methods, see Table 1. Therefore, each method uses different factors and concepts that serve as a basis for extracting relationships between social actors from information sources [39].

PSM	0.2252					
BSMv1	0.1793	0.7189				
BSMv2	0.1868	0.7574	0.7637			
DSM	0.0940	0.1389	0.1315	0.1367		
USM	0.2879	0.2216	0.1627	0.1703	0.1192	
SSM	0.0137	0.0401	0.0360	0.0339	0.0187	0.0417
	BSM	PSM	BSMv1	BSMv2	DSM	USM

Fig. 2. Similarity of edges in the social networks

Based on online database, such as DBLP, SSM generates new actors and their relations based on concept of co-author relationship [18]. 76 professor (as academic actors) be seeds to get 385 other academic actors, but the social network density based on relations only 1.22% of 106,491 possible relations potentially [40].

For 462 social actors there are 106,491 potential relationships in symmetry. Through BSM, realized 462 occurrences and 106,491 co-occurrences into 31,623 strength relations. For the same case, but involving keywords:

1. Related to affiliation (in basic superficial method with keyword (BSMv1)), the approach produces 373 vertices and 16,388 edges (as a realization).
2. While, by using keyword for developing a community (in basic superficial method with other keyword (BSMv2)), the approach produces 389 vertices and 17,132 edges.

By using occurrences only, USM produces 462 vertices and 99,150 edges, USM produces 462 vertices and 99,150 edges, and DSM produces 434 vertices and 12,158 edges. In general, another approach developed from its basic form in realization results in lower social network density than the original form. In this case, BSM, BSMv1, and BSMv2 sequentially produce 29.70%, 15.39% and 16.09% strength relations of 106,491 relations potentially. As well as with BSM, three methods: PSM, PSMv1, PSMv2 produce 21.44%, 15.38%, and 16.07% strength relations of 106,491 relations potentially, respectively. However, all approaches used as validation depend on the occurrences and co-occurrences of the related method. For example, JUSM and JDpSM depends on occurrences and co-occurrences of BSM, or JUpSM and JDpSM depends on occurrences and co-occurrences of PSM. Thus the USM on its own is different from the related approach, USM contains a pseudo-strength relations. It based on the domain similarity of the website (the domains of URL addresses).

The similarity between social networks on the edges shows that there are methods that imply so close performance, namely PSM, BSMv1 and BSMv2. While other methods show social networks as results that are different from the others. In particular, that the similarity of social networks from SSM to other method shows the difference due to the resulting actors having stronger relationships other than co-author relationship. Furthermore, the existence of relationships within the online database is not necessarily detectable directly through queries in other methods, except that it only reveals some of the content of the online database. So there is the core of all social networks of more than 1% as the same result of all methods.

5. Conclusion

Unequal treatments result in different approaches to information resources, although involving potential same social actors, and this also results in several methods of extraction of social networks in the unsupervised stream. The superficial methods that have similar treatment show the similarity of results. Different results show that the resulted social networks have different characteristics. Nevertheless, the similarity between the edges of social networks shows the existence of the core of social networks as the foundation for the development of a social network. Therefore, different results from different methods suggest the needing next study for getting an integrity approach, while for getting trusty information is by involving dataset as standard evaluation.

References

- [1] F. Fontaine. (2008) "Why are similar workers paid differently? The role of social networks." *Journal of Economic Dynamics & Control* **32**: 3960-3977.
- [2] Mahyuddin K. M. Nasution and S. A. Noah. (2012) "Information retrieval model: A social network extraction perspective." *Proceedings - 2012 International Conference on Information Retrieval and Knowledge Management (CAMP'12)*, 6204999: 322-326.
- [3] Mahyuddin K. M. Nasution. (2017) "Modelling and Simulation of Search Engine." *Journal of Physics: Conference Series* **801** (1), 012078.
- [4] Mahyuddin K. M. Nasution and S. A. Noah. (2010) "Superficial method for extracting social network for academics using web snippets." *Lecture Notes in Computer Science* **6401** LNAI: 483-490.
- [5] R. Bekkerman, and A. McCallum. (2005) "Disambiguating Web appearances of people in a social network." *Proc. WWW 2005*: 463-470.
- [6] A. McCallum, X. Wang, and A. Corrada-Emmanuel. (2007) "Topic and role discovery in social networks with experiments on Enron and Academic Email." *Journal of Artificial Intelligence Research* **30**: 249-272.
- [7] A. McCallum, A. Corrada-Emmanuel, and X. Wang. (2004) "The author-recipient-topic model for topic and role discovery in social networks: Experiments with Enron and Academic Email. *Technical Report*, University of Massachusetts, Amherst, UM-CS-2004-096,
- [8] Clifford Weinstein, William Campbell, Brian Delaney, and Gerald O'Leary. (2009) "Modeling and detection techniques for counter-terror social network analysis and intent recognition." *IEEE Aerospace conference*.
- [9] Y. Matsuo, J. Mori, M. Hamasaki, T. Nishimura, T. Takeda., K. Hasida, and M. Ishizuka. (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.
- [10] YingZi Jin, Y. Matsuo, and M. Ishizuka. (2007) "Extracting social networks among various entities on the Web." *ESWC 2007* **4519** LNCS: 251-266.
- [11] P. Mika. (2007). *Social Networks and the Semantic Web*. Springer-Verlag: Berlin.
- [12] Mahyuddin K. M. Nasution. (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): 975-981.
- [13] Yutaka Matsuo, Junichiro Mori, Masahiro Hamasaki, Takuichi Nishimura, Hideaki Takeda, Koiti Hasida, and Mitsuru Ishizuka. (2007) "POLYPHONET: An advanced social network extraction system from the Web." *Web Semantics: Science, Services and Agent on the World Wide Web* **5** (4): 262-278.
- [14] Mahyuddin K. M. Nasution, S. A. M. Noah, and S. Saad. (2011) "Social network extraction: Superficial method and information retrieval." *Proceedings of International Conference on Informatics for Development (ICID11)*: c2-110 c2-115.
- [15] Henry Kautz, Bart Selman, and Mehul Shah. (1997) "ReferralWeb: Combining social networks and collaborative filtering." *Communication of the ACM* **40** (3).
- [16] Mahyuddin K. M. Nasution. (2014) "New method for extracting keyword for the social actor." *Lecture Notes in Computer Science* **8397** LNAI (PART 1): 83-92.
- [17] Mahyuddin K. M. Nasution. (2012) "Simple search engine model: Adaptive properties." *Cornell University Library* arXiv:1212.3906v1.
- [18] Mahyuddin K. M. Nasution and S. A. Noah. (2011) "Extraction of academic social network from online database." *2011 International Conference on Semantic Technology and Information Retrieval(STAIR 2011)*, 5995766: 64-69.
- [19] YingZi Jin, Y. Matsuo, and M. Ishizuka. (2006) "Extracting a social network among entities by web mining." *Workshop on Web Content Mining with Human Language (ISWC 2006)*.
- [20] Mahyuddin K. M. Nasution. (2012) "Simple search engine model: Selective properties." *Cornell University Library* arXiv:1303.3964v1.
- [21] YingZi Jin, Y. Matsuo, and M. Ishizuka. (2007) "Extracting inter-Firm networks from World Wide Web." *The 9th IEEE International Conference on E-Commerce Technology and the 4th International Conference on Enterprise Computing E-Commerce and E-Services (CEC-EEE 2007)*: 635-642.
- [22] YingZi Jin, Y. Matsuo, and M. Ishizuka. (2008) "Extracting inter-Firm networks from the World Wide Web using a general-purpose search engine." *Online Information Review* **32**(2): 190-200.
- [23] Mahyuddin K. M. Nasution. (2012) "Simple search engine model: Adaptive properties for doubleton." *Cornell University Library* arXiv:1212.4702v1.
- [24] Mahyuddin K. M. Nasution. (2017) "Semantic interpretation of search engine resultant." *InteriOR*.
- [25] Mahyuddin K. M. Nasution, O. S. Sitompul, Sawaluddin Nasution, and H. Ambarita. (2017) "New Similarity." *IOP Conference Series: Materials Science and Engineering* **180** (1), 012297.
- [26] D. M. Boyd, and V. P. Roychowdhury. (2008) "Social network sites: Definition, History, and Scholarship." *Journal of Computer-Mediated Communication* **13**: 2010-230.
- [27] L. A. Adamic, and A. Adar. (2003) "Friends and neighbours on the Web. *Social Networks* **25**: 211-230.
- [28] Mahyuddin K. M. Nasution, M. Hardi, and R. Syah. (2017) "Mining of the social network extraction." *Journal of Physics: Conference Series* **801** (1), 012020.
- [29] Mahyuddin K. M. Nasution, R. Syah, and M. Elveny. (2017) "Studies on behaviour of information to extract the meaning behind the behaviour." *Journal of Physics: Conference Series* **801** (1), 012022.
- [30] Peter Mika. (2005) "Social network and the semantic Web the next challenge." *IEEE Intelligent Syst.* **20**(1) (2-3): 82-85.
- [31] Mahyuddin K. M. Nasution and O. S. Sitompul. (2017) "Enhancing Extraction Method for Aggregating Strength Relation Between Social Actors." *Artificial Intelligence Trends in Intelligent Systems (AISC)* **573**: 312-321.
- [32] P. Mika. (2005) "Flunk: Semantic Web technology for the extraction and analysis of social networks." *Journal of Web Semantics* **3**: 211-223.
- [33] Yutaka Matsuo, Junichiro Mori, Masahiro Hamasaki, Takuichi Nishimura, Hideaki Takeda, Koiti Hasida, and Mitsuru Ishizuka. (2006) "POLYPHONET: An advanced social network extraction system." *Proceedings of the 15th International Conference on World Wide Web 2006 (WWW*

- 2006): 397-406.
- [34] J. Tang, D. Zhang, L. Yao, and J. Li. (2008) "ArtnetMiner: Extraction and mining of an academic social networks." *Proceedings of KDD'08*: 990-998.
 - [35] Mahyuddin K. M. Nasution, R. Syah, and M. Elfida. (2018) "Information retrieval based on the extracted social network." *Applied Computational Intelligence and Mathematical Method, Advances in Intelligent Systems and Computing* **662**.
 - [36] M.-M. Deza and E. Deza. (2006) *Dictionary of distance*. Elsevier.
 - [37] Mahyuddin K. M. Nasution and S. A. Noah. (2012) "A methodology to extract social network form the Web Snippet." *Cornell University Library* arXiv:1211.5877.
 - [38] Sebastian Hellmann, Clause Stadler, Jens Lehmann, and Sören Auer. (2009) "DBpedia live extraction." *OTM 2009, Part III, LNCS* **6871**: 1209-1223.
 - [39] Mahyuddin K. M. Nasution, O. S. Sitompul, E. P. Sinulingga, and S. A. Noah. (2016) "An extracted social network mining." *Proceedings of 2016 SAI Computing Conference (SAI 2016)*, 7556125: 1168-1172.
 - [40] Mahyuddin K. M. Nasution, Marischa Elveny, Rahmad Syah, and Shahrul Azman Noah. (2015) "Behavior of the resources in the growth of social network." *5th International Conference on Electrical Engineering and Informatics: Bridging the Knowledge between Academic, Industry, and Community*, ICEEI 2015, 7352551: 496-499.