Generating Luck from Weak Ties in Social Networks

Iaakov Exman, Omer Ganon and Asaf Yosef
Software Engineering Department
The Jerusalem College of Engineering – JCE - Azrieli
Jerusalem, Israel
iaakov@jce.ac.il, ganon136@gmail.com, asafjo23@gmail.com

***Abstract*— One often assumes that for online Social Networks of related people, relations with strong ties better characterize the person one is looking for. However, a paradox already stated by Granovetter is the opposite assumption that weak ties to other people may be the more significant in certain contexts. This paper investigates this latter contrarian hypothesis as a novel tool to extract knowledge and systematically generate luck in the given contexts. Similarly to interestingness, luck is modeled relative to the context, by combining two functions – Relevance and Surprise. The Surprise expresses the importance of weak ties. A Luck-Generator software tool has been developed as an experimental testbed to interact with any social network. Its inputs, chosen by the Luck-Generator customer, are a context, a social network, and the customer's network page. The hypothesis is validated by results showing that relevance alone is not enough to actually generate all the potential luck: the weak ties' surprise contribute essentially to optimize success in the context task. Preliminary results are illustrated by 'Getting a Job' case study.**

Keywords: Luck calculation; Luck-Generator; Weak ties; Context; Social Network; Interestingness; Relevance; Surprise; Software Architecture; Knowledge Discovery; Knowledge Extraction.

I. Introduction

Social Networks, besides being a huge source of searchable information, have the potential to significantly enhance performance of a variety of tasks, not necessarily related to the explicitly declared purpose of any particular network.

Concerning information search, we have previously defined and demonstrated the usefulness of an Interestingness [10] measure – composed of *Relevance* and *Surprise* functions – focusing search outcomes, beyond the capabilities of neutral search engines provided by social networks.

Regarding enhancing task performance, this work proposes a new kind of ***knowledge extraction***, by means of ***Luck measurement***, where by luck we mean systematically reaching goals with low apparent probability. Similarly to Interestingness, Luck is also obtained by a couple of functions, but now calculated upon different input types, with the special *Surprise* role, to overcome the low apparent probability.

A. Systematic Generation of Luck

Our working hypothesis is the assumption stated long ago by Granovetter [16] that weak ties to other members of a – real or virtual – social network may be surprisingly more significant than strong ties in certain circumstances. Given a certain context, defining a task to be performed, one computes a Luck measure for *relevant* social network members, with a *Surprise* function quantitatively expressing the weak ties of network members. These were inspired by the generic definition of interestingness:

$$Interestingness = Relevance * Surprise \tag{1}$$

The operator ***** in this equation in its most generality, when not commutative, is a kind of composition. In the simpler commutative cases it is just plain multiplication (see [10]).

The rationale, actual functionalities in the analogous equation for calculating Luck, the input variables and additional motivation are formulated in the more theoretical section III of this paper.

B. Weak Ties in Social Networks

A natural representation of a social network is a graph in which vertices stand for network members and edges represent their ties to other network members. The tie strength – or rather tie weakness – can be a function of a few different variables, e.g. distance in terms of counting weighted graph edges, content similarity and communication frequency.

The goal of this paper is to validate the working hypothesis by evaluating the calculated Luck with respect to the contribution of surprisingly weak ties and its effective results for the context task.

C. Paper Organization

The remaining of the paper is organized as follows. Section II concisely reviews Related Work. Section III formulates the Luck generation underlying theory. Section IV describes the Luck-Generator software tool architecture and implementation. Section V illustrates the Luck generation task by means of a case study. Section VI concludes the paper with a discussion.

DOI: 10.18293/SEKE2020-156

II. RELATED WORK

We concisely review the literature related to Luck characterization, Interestingness concepts, and practical applications of weak ties within social networks.

A. Luck Characterization

We refer to *Luck* in a positive context of systematic generation, in order to succeed in concrete tasks performance, in contrast to random uncontrollable situations, in which sometimes one achieves "by chance" desirable outcomes. An interesting extended example of the latter negative meaning is the book by Clayton Christensen and co-authors [6] entitled "Competing against Luck". It advocates causality as opposed to the frustration of hit-and-miss innovation, viz. leaving your fate to luck.

Dowding [9] deals mostly with moral aspects of luck; he also suggests a simplistic measure of luck as a relationship between expected value of outcome (EV) and the actual outcome (AV), thus $Luck = AV - EV$, where in a serial of trials one would expect that AV approaches EV.

Liechti et al. [23] use a more sophisticated definition of luck as the unexpected component of performance. It is a sum of three terms: a- the actual deviation from expected performance; b- an overconfidence bias; c- a look back bias (a difference of subjective expectation at a certain time ***t*** and at a previous time). This definition is closer to our own definition, which involves a surprise (or unexpectedness) factor.

B. Interestingness Concepts and Applications

It is worthwhile to be acquainted with the literature on Interestingness, as the calculation of this quantity shown in equation (1), inspired the proposed calculation of *Luck*, in particular the *Surprise* factor, as explained in section III.

Overviews of Interestingness measures for typical purposes, such as Data Mining and Knowledge Discovery are found in Geng et al. [14] and McGarry [25]. For instance, criteria on how to determine interesting rules/patterns generated in data mining are described by Lenca et al. [22].

There are several differing approaches to interestingness as described e.g. in the Klosgen and Zytkow Handbook [20], especially by Tuzhilin [28]. Exman, defined Interestingness as a product of relevance and surprise in 2009 [10]. This definition has been implemented with successful Web search results, in software tools such as the one described in [11].

C. Social Networks Weak Ties and Applications

Granovetter [16], [17] is the pioneer of asserting significance to Weak Ties in social networks. He also was one of the first researchers that actually made concrete application of the theory in his book [18] originally published in 1974, in the context of "Getting a Job". A generic analysis of networks from an historical viewpoint is the book by Ferguson [12], which includes chapter 6, explicitly dealing with weak ties.

The importance of weak ties in social networks triggered a variety of studies. Many of them supported the theory – such as Brown and Konrad [4], DeMeo et al. [7]. In contrast, some of them rather emphasized the importance of strong ties – such as Gee et al. [10], Krackhardt and co-authors [17]. Others, extended the theory to different applications, – such as Baer [2], Centola [5] – or provided general appraisals e.g. Sinan [25].

Specifically concerning the "Getting a Job" context, besides Granovetter, one finds Gee et al. [13] and the paper by Tassier on "Labor Market implications of Weak Ties" [27]. Of significance for this work is the statement by Tassier that weak ties in a person's social network grows with network distance exponentially faster than strong ties, which is reasonable.

Finally, the technical issue of measuring the strength of a tie is dealt with e.g. in the paper by Marsden and Campbell [24].

III. LUCK IN CONTEXT

This section's goal is to formulate the theoretical basis for Luck calculation for any given context data set. It is the result of Luck mathematical modeling, based upon assumptions following experimental results, ours and in the literature on social network ties' strength. It starts from an abstract scheme reflecting actual experiments with (non-virtual) networks.

A. The Abstract Scheme

Our idea, on how to generate Luck, avoids the controversy on the relative importance of strong ties vs. weak ties, in a straightforward way by involving both strong and weak ties.

Our abstract scheme, in the next text-box has two actions, not necessarily in a fixed order, which may occur concurrently.

Abstract Scheme: Luck Generation

1st action: a relevant ***strong tie*** – determines the task to be performed, within the chosen context;
2nd action: a surprising ***weak tie*** – obtains a pointer to the desired outcome.

This abstract scheme is illustrated by 3 stories that actually occurred in human (not virtual) networks.

The first story task was to "*find a scientific collaborator*". The relevant ***strong tie*** was to participate in a conference whose main topic fits the researcher's scientific interests. The Conference was held in China. The ***weak tie*** was to find among the many conference participants a Spanish researcher with whom a vivid conversation of mutual interest developed. The surprising aspect was to travel a long distance to China to find a Spanish researcher.

The second story task was to "*find a job in the profession*". The ***strong tie*** was to be an active member in relevant professional interest groups in the internet. This story referred to a Java programming language interest group. The ***weak tie*** was, in response to an inquiry, to get an answer from an old acquaintance in the past, but disconnected for several years. The acquaintance enabled a successful information exchange, leading to a concrete job, which was actually taken.

The third story task was to "*find a candidate for a vacant position*" in our institution. In this story, the ***weak tie*** occurred first. A certain candidate presented himself to the candidates' recruiter, to show his credentials, and by the way mentioned members of his family. The ***strong tie*** was that the candidate's brother learnt years ago in the same class and was well-known to the recruiter, being a strong implicit recommendation.

B. Modeling Luck Calculation

Given the literature on social network ties' strength and the previous abstract scheme, we make the following assumptions:

1. ***Complementary Exponential decay of ties*** – strong ties decay exponentially with the network distance, while weak ties increase exponentially and vice-versa (see e.g. Tassier [27]);
2. ***People Matching with strong ties*** – strong ties bond similar people to each other (see e.g. Granovetter [16], and Krackhardt [21]) and vice-versa mismatching for weak ties;
3. ***Time Commutativity of strong/weak ties*** – sometimes the strong tie action precedes the weak tie action, other times the order is reversed (as illustrated by the above stories of the abstract scheme).

We now formulate the necessary equations to model Luck calculation, based upon the above assumptions. In terms of notation we define two functions that calculate the contribution of strong and weak ties as follows:

- ***Relevance*** – calculates the strong ties' contribution;
- ***Surprise*** – calculates the weak ties' contribution.

By the 1st assumption on "Complementary Exponential decay" each of these functions is an exponential, with complementary signs. By the 2nd assumption on "People Matching" with strong ties and "People Mismatching" with weak ties one has:

$$\mathrm{Re}\,levance = \exp(Match) \qquad (2)$$

$$Surprise = \exp(Mismatch - Match) \qquad (3)$$

By the 3rd assumption on "Time Commutativity" one has:

$$Luck = \mathrm{Re}\,levance + Surprise \qquad (4)$$

The "plus" operator is certainly commutative. A "multiplier" operator in this equation is obviously unsuitable, as the exponential nature of these terms would cause undesirable exponents addition.

Finally, substituting equations (2) and (3) into equation (4) one obtains:

$$Luck = \exp(Match) + \exp(Mismatch - Match) \qquad (5)$$

In practice, to use this equation in calculations, one still needs to make adjustments to normalize the expressions of Match and Mismatch, in order to eliminate dependence on set sizes.

C. Luck Calculation with Keyword Sets

In this paper we restrict Luck calculation due to the representation of social network members by their respective Keyword Sets.

First, additional notations are introduced:

- ***Context*** – is the keyword set defining a task, e.g. "*find a job in a specific profession*", such as software engineering;
- ***Customer*** = ***C*** – is the person, member of a social network, who demands the performance of the Context task; it also designates the keyword set of this person;
- ***Follower = F –*** is another member of the same Customer's social network, which is a follower (in the social network sense) of the Customer; it also designates the keyword set of the Follower; ***F*** is generalizable to a Follower of a Follower of the Customer, or to any distance from the Customer.

The keyword set of the Context is determined before any computation starts. The keyword set of the Customer and of each Follower are sub-sets of the Context keyword set. These are determined by extracting sets from the person pages in the Social Network, and finding the intersection of the extracted sets with the Context keyword set.

In this work ***Match*** and ***Mismatch***, are keyword set operations necessary to obtain respectively the Relevance and Surprise functions, by comparing keyword sets for each Customer ***C*** with the keyword set for a Follower ***F***. *Match* calculates a similarity measure of the input sets, i.e. keywords appearing in the intersection $\bigcap$ of these sets:

$$Match = C \bigcap F \qquad (6)$$

The output is the number of intersection elements of ***C*** and ***F***.

Mismatch calculates the sets' dissimilarity, viz. a symmetric difference Δ between ***C*** and ***F***. It is the union $\bigcup$ of the relative complements of these sets:

$$Mismatch = C \Delta E = (C - F) \bigcup (F - C) \qquad (7)$$

Match and Mismatch diagrams are seen in Fig. 1.

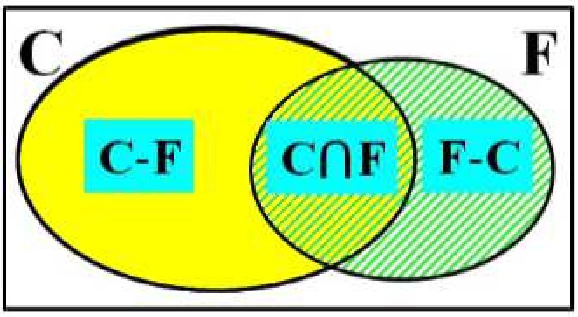

Figure 1. Schematic Match and Mismatch diagram – ***C*** represents the Customer keyword set (yellow). ***F*** represents the Follower keyword set (hatched green). Match is the intersection ***C***$\bigcap$***F***. Mismatch is the union between the relative complements ***C-F*** and ***F-C***.

IV. LUCK-GENERATOR: SOFTWARE ARCHITECTURE AND IMPLEMENTATION

In this section we describe the Luck-Generator software tool software architecture and implementation. The tool enabled testing of the Luck calculations and the Case study in section V. Its output is a list of candidates: a number of Customer followers with the highest calculated Luck values.

A. *Luck-Generator Architecture*

The Luck-Generator software architecture has the following roles, as shown in Fig. 2:

- *Front-End* – for input and output;
- *APIs* – for interaction with any chosen social networks;
- *Keyword handler* – to extract and collect keyword sets;
- *Local Storage* – to avoid repeated networks access;
- *Inquirer* – to retrieve necessary data from storage;
- *Calculators* – of Tie Strength and Luck;
- *Analysis tool* – for system maintenance.

The Luck-Generator architecture was designed to be generic, and not fitting any particular Social Network. One only needs to insert the needed specific API.

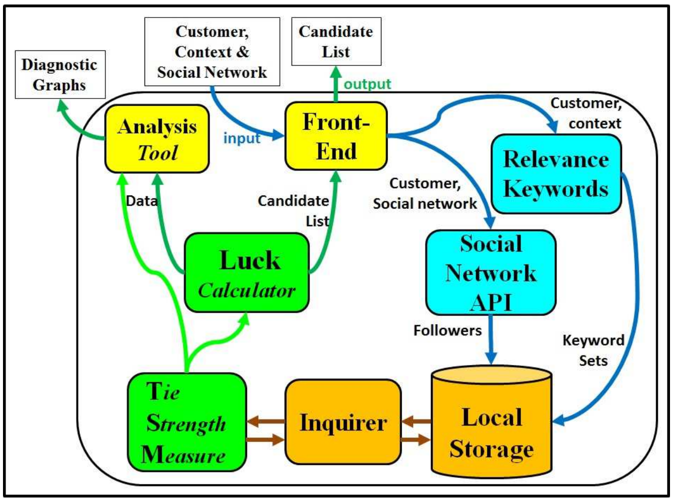

Figure 2. Luck-Generator Software Architecture – The Front-End (yellow) in the figure upper part inputs data and outputs the resulting candidate list. The upper-right modules (in blue) get the inputs (customer, social network, context) and obtain keyword sets and followers to be stored locally. The Inquirer (orange) retrieves data to calculators (green) for Tie Strength and Luck. The Analysis tool obtains diagnostic graphs for system maintenance.

B. *Luck-Generator: Implementation*

As far as possible the system is composed of existing software modules. For instance, extraction of keywords to characterize the context, the Customer and followers is done with the help of Datamuse – a semantic network with a word-finding query engine for system developers – through its API.

Similarly, access to Social Networks is done by specific available APIs.

V. CASE STUDY: GETTING A JOB

The chosen Context task for our case study is "*find a job in the profession*". The context was defined, data from a Social Network extracted and calculations performed, as reported here.

A. *Context Definition: Getting a Job*

The chosen profession was "Software Engineering". The Context diverse keyword set is in the next text-box: it has 'word pairs' and even keywords not exactly belonging to software.

Context Keyword Set

Software, engineering, developer, DevOps, computers, algorithm, TechOps, python, programmer, java, 'computer science', 'data science', 'data analyze', C++, web, framework, embedded 'alpha version', API, api, app, application, beta, version, bios, QA, automation, agile, scrum, demo, development, device, emulator, freeware, 'open source', interface, 'operating systems', workflow 'machine learning', 'deep learning', startup, innovation, internet, IoT, VR, code, coding.

The social network was dictated by an available API. We started testing with a couple of initial "customers" searching for the job. According to their extracted keywords characterization these have been involved previously with software, to assure that testing is realistic.

Normalization of both the Match and Mismatch functions in equation (5) was done by a sum of the intersection of the Context and Costumer keyword sets with the intersection of the Context and each Follower keyword sets.

B. *Calculation Results: Relevance vs. Surprise*

Calculation results were obtained with input data extracted from the social network, for each Customer, and a small number of followers and all the available followers of followers.

The next fig. 3 shows an inverse exponential relation between Relevance and Surprise for the data-set of a certain Customer.

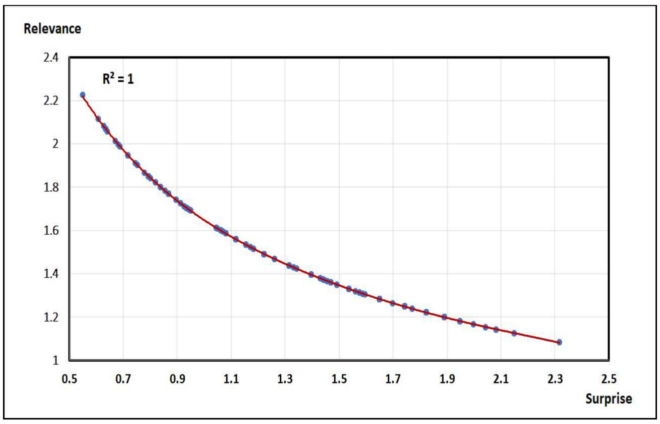

Figure 3. Graph of Calculated Relevance against Surprise for Customer C.D. – There is an inverse relation between these two quantities: when Relevance – expressing Strong Ties – decreases exponentially, Surprise – expressing Weak ties – increases and vice-versa, as predicted by our Model.

C. *Empirical Validation*

We have used several validation techniques to increase confidence on the obtained results. These included:

- Results Consistency – Variation of customers;
- Robustness – Variation of Context keyword sets;
- Comparison with results shown in the literature.

As an example of Results Consistency, Figure 4 shows the same calculation of Luck vs. Surprise for four Customers (L.M., C.D., M.M. and S.C.) Although the Customers and their followers' data sets are totally independent, the functional behavior is very similar.

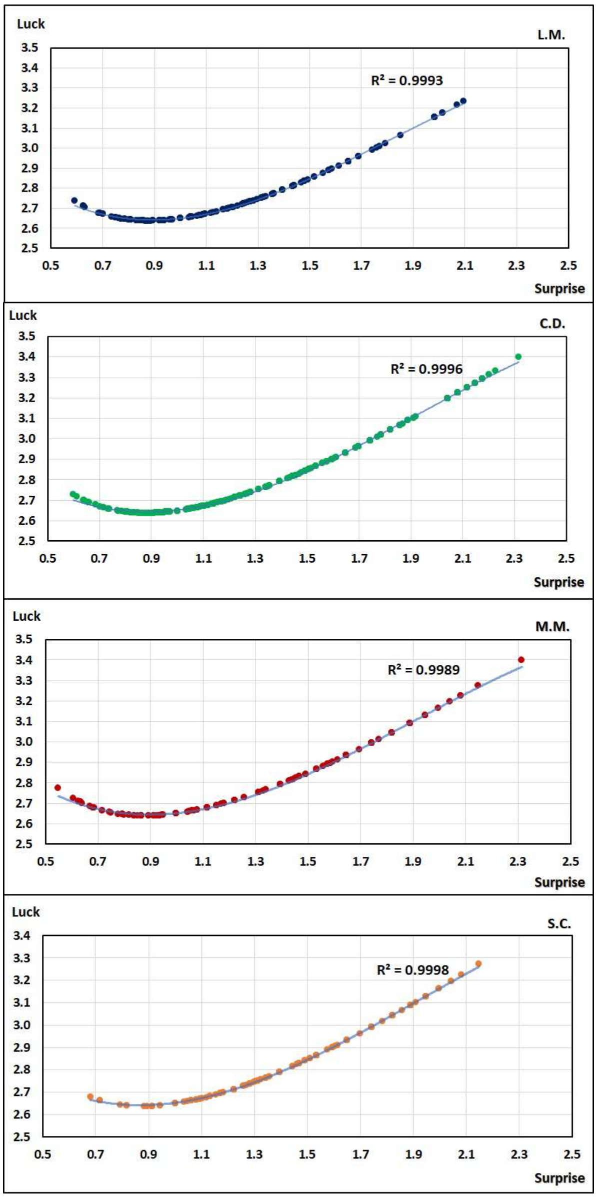

Figure 4. Graphs of Calculated Luck against Surprise for diverse Customers – As Surprise – expressing Weak Ties – increases, our Model predicts that calculated Luck also increases. A smaller increase of Luck at the left-hand-side of the graph, corresponds to a Relevance increase – expressing Strong Ties. Dots show results calculated for actual data from the Social Network. Trend-lines are very good polynomial fittings. All four graphs have the same scales.

VI. DISCUSSION

We deal here with open issues triggered by the preliminary results of this work.

A. *The Functionality of Luck Calculation*

The functionality of Luck calculation in this paper is based upon empirical assumptions. These have been validated to be reasonable and self-consistent.

Nonetheless, it would be desirable to formulate a more rigorous derivation of the equations we have used.

A few possible starting points are as follows:

- *Maximum entropy approach* – it is well-known that such an approach, i.e. maximum entropy under constraints, obtains probability distributions with exponential functionality, where the exponent is a negative quantity. This would be suitable to explain the exponential expressions in equation (5) of this paper;
- *Hyperbolic Modeling of Probability Distributions*- for example, one may perceive that the functionality of Luck calculation in the same equation (5) has an obvious similarity to a Hyperbolic Cosine. An example of probability modeling involving hyperbolic functions is found in the work of Hanaki et al. from Tsukuba University [19].

B. *Systematic Luck vs. Irrationality*

From the very beginning of this work we adopted a positive constructive view of Luck, in other words "Systematic Look". This paper is a contribution in this direction. This is not an esoteric point of view. There is a non-negligible number of works with this approach.

We mention here Dowding [8] which argues for the utility of ideas of luck and "systematic luck". Hanaki et al. [19] suggest that people learning from experience leads them to make choices with "luckier" outcomes than others. Contrast these with Adaval [1].

C. *Other Variables for Tie Strength Measure*

Besides keyword sets, we are aware of other important variables to characterize Tie Strength, which were not considered in this work. These include among others, topology measures such as relationships among edges and vertices in the social network and communication intensity between members of the social network, such as frequency and the nature of the communication, either generic such as 'like' or more personal contents.

We are currently working to integrate these other variables in the same generic equations of our model – described in section III *B*.

D. *Future Work*

In addition to the interpretation issues and the number of variables to characterize Tie Strength, important directions for further investigation are:

- Extensive application to a variety of Customers and their followers;
- Application to different contexts, besides "*finding a job*" that has already been intensively researched in the literature;
- Usage of different functions to calculate Relevance and Surprise, such as Tf*Idf, and compare their results with those of match and mismatch;

E. Main Contributions of this Paper

The main contributions of this paper are: 1- the idea of systematic generation of Luck in a constructive sense, within contexts of practical tasks, exploring social networks; 2- to model the significant contribution of Weak Ties for Luck generation in terms of a mathematical expression of Surprise.

REFERENCES

[1] R. Adaval, "Culture and Cognition: the Case of Irrational Beliefs about Luck", *Advances in Consumer Research*, Vol. 33, pp. 623-628 (2006).

[2] M. Baer, "The Strength-of-Weak-Ties Perspective on Creativity: A Comprehensive Examination and Extension", J.App.Psychology, 95, pp-592-601, (2010).

[3] M. Bertrand and S. Mullainathan, "Are CEOs Rewarded for Luck? The Ones with Principals are", The Quarterly Journal of Economics, Vol. 116, pp. 901-932 (Aug. 2001). http://www.jstor.org/stable/2696421

[4] D.W. Brown and A.M. Konrad, "Granovetter Was Right: The Importance of Weak Ties to a Contemporary Job Search", Group & Organization Management, 26, 434-462, (2001).

[5] D. Centola, "Complex Contagions and the Weakness of Long Ties", Am.J.Sociology, 113, pp. 702-734, (2007). DOI: **https://doi.org/10.1086/521848**

[6] C.M. Christensen, T. Hall, K. Dillon and D.S. Duncan, Competing Against Luck – The Story of Innovation and Customer Choice, Harper Business, New York, NY, USA, October 2016.

[7] P. DeMeo, E. Ferrara, G. Giumara and A. Provetti, "On Facebook, Most Ties are Weak", Comm. ACM, 157, pp. 78-84 (October 2014) https://doi.org/10.1145/2629438; See also: arXiv:1203.0535, 2012.

[8] K. Dowding, "Resources, Power and Systematic Luck: A Response to Barry", *Politics, Philosophy & Economics*, pp. 305-322 (October 2003) DOI: https://doi.org/10.1177/1470594X030023002

[9] K. Dowding, "Luck and Responsibility", in M. Matravers and L. Meyer, (eds.) Democracy, Equality and Justice, Routledge, London, UK, 2008.

[10] I. Exman, "Interestingness – A Unifying Paradigm – Bipolar Function Composition", in Proc. KDIR Int. Conf. on Knowledge Discovery and Information Retrieval, pp. 196-201, 2009.

[11] I. Exman, G. Amar and R. Shaltiel, R., "The Interestingness Tool for Search in the Web", in Proc. SKY'2012 Int. Workshop on Software Knowledge, pp. 54-63, 2012.

[12] N. Ferguson, *The Square and the Tower* – Networks, Hierarchies and the Struggle for Global Power", Penguin Books, UK, (2018).

[13] L.K. Gee, J. Jones and M. Burke, "Social Networks and Labor Markets: How Strong Ties Relate to Job Finding On Facebook's Social Network", J.Labor Economics, 35, pp. 485-518 (April 2017). DOI: https://doi.org/10.1086/686225

[14] L. Geng and H.J. Hamilton, "Interestingness Measures for Data Mining: A Survey", ACM Computing Surveys, Vol. 38, (3), Article 9, 2006.

[15] Z. Gilani, R. Farahbakhsh, G. Tyson, L. Wang and J. Crowcroft, "An in-depth characterization of Bots and Humans on Twitter", 2017.

[16] M.S. Granovetter, "The Strength of Weak Ties", Am.J.Sociology, 78, pp. 1360-1380, (May 1973).

[17] M.S. Granovetter, "The Strength of Weak Ties: A Network Theory Revisited", Sociological Theory, 1, pp. 201-233, (1983).

[18] M.S. Granovetter, *Getting a Job: A Study of Contacts and Careers*, 2nd ed. University of Chicago Press, Chicago , IL, USA, (1995).

[19] N. Hanaki, A. Kirman and M. Marsili, "Born Under a Lucky Star?", Tsukuba Economics Working Papers, No. 2009-003, (March 2009), Tsukuba, Japan.

[20] W. Klosgen and J.M. Zytkow, (eds.), *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press, Oxford, UK, 2002.

[21] D. Krackhardt, N. Nohria and B. Eccles, "The Strength of Strong Ties: The Importance of Philos in Organizations", in Networks in the Knowledge Economy , (2003).

[22] P. Lenca, P. Meyer, B. Vaillant and S. Lallich, "On selecting interestingness measures for association rules: user oriented description and multiple criteria decision aid", European J. Operational Res., Vol. 183, pp. 610-626, 2008.

[23] D. Liechti, C. Loderer and U. Peyer, "Luck and Entrepreneurial Success", Political Science, Psychology, Economics, 2012. DOI:10.2139/ssrn.1954560

[24] P.V. Marsden and K.E. Campbell, "Measuring Tie Strength", Ann. Meeting Am. Sociological Assoc., pp. 482-499, (1984).

[25] K. McGarry, "A survey of interestingness measures for knowledge discovery", Knowledge Engineering Review J., 20 (1), 39-61, 2005.

[26] A. Sinan, "The Future of Weak Ties", Am.J.Sociology, 121. pp. 1931-1939, (2016). DOI: http://dx.doi.org/10.1086/686293

[27] T. Tassier, "Labor Market Implications of Weak Ties", Southern Econ.J., 72, pp. 704-719, (2006). DOI: 10.2307/20111842.

[28] A. Tuzhilin, "Usefulness, Novelty, and Integration of Interestingness Measures", chapter 19.2.2 in ref. [16], pp. 496-508, 2002.