Significance Level of a Big Data Query by Exploiting Business Processes and Strategies

Loan Thi Ngoc Dinh, Gour Karmakar, Joarder Kamruzzaman, and Andrew Stranieri
Federation University Australia
{l.dinh; gour.karmakar; joarder.kamruzzaman;
a.stranieri}@federation.edu.au

Abstract. Querying data is one of the most frequent activities in business organisations. The tasks involving queries for big data collection, extraction and analysis have never been easy, because to obtain the high quality responses, the expected outcome from these tasks need to be more accurate and highly relevant to a business organisation. The emergence of big data era has further complicated the task. The enormous volume of data from diverse sources and the variety of queries impose a big challenge on business organisations on how to extract deep insight from big data within acceptable time. Determining significance levels of queries based on their relevance to business organisations is able to deal with such challenge. To address this issue, up to our knowledge, there exists only one approach in the literature to calculate the significance level of a query. However, in this approach, only business processes are considered by manually selecting weights for core and non-core business processes. As the significance level of a query must express the importance of that query to a business organisation, it has to be calculated based on the consideration of business strategic direction, which requires the consideration of both business processes and strategies. This paper proposes an approach for the first time where the significance level of a query is determined by exploiting process contributions and strategy priorities. The results produced by our proposed approach using a business case study show the queries that are associated with more important business processes and higher priority strategies have higher significance levels. This vindicates the application of the significance level in a query to dynamically scale the semantic information use in capturing the appropriate level of deep insight and relevant information required for a business organisation.

Keywords: Semantic similarity, Significance level, Business process, Business strategy, Query processing

1 Introduction

Under the influence of big data, query analysis has shown a huge difficulty for business organisations to handle their data processing systems [1-3]. To deal with such a challenge, a number of approaches and applications have been introduced. Some of

Lupeikiene A., Matulevičius R., Vasilecas O. (eds.):
Baltic DB&IS 2018 Joint Proceedings of the Conference Forum and Doctoral Consortium.

them have focused on how to increase the speed of data collection and analytics. The others are concentrating on the techniques to improve the accuracy of data processing and analysis.

To improve the speed of data analytics, several big data processing platforms and tools have already been introduced such as Hadoop, Spark, Kafka and Tableau (tableau.com). There also exist a number of studies to reduce redundant, irrelevant data during data collection and analysis for queries.

Nanchani et al. [4] introduced a solution for avoidance of duplication in storing customer records. Whereas, [5] and [6] presented the solutions in order to deal with redundant data. Besides that, since big data has led to the outcome delays and information overload, the techniques for big data collection and analytics have also been improved such as real-time data collection in [7, 8] and speed optimization during data processing in [4, 9]. Additionally, search engines (e.g., Google Trends and Baidu) have proven advantages in extracting useful business information for organisations [10, 11].

In addition to analytics, retrieving the deep insight of data also requires exhaustive data collection from all possible sources. Even with exhaustive data collection, a big data query cannot obtain deep insight from all relevant unstructured big data without semantic interpretation in both data collection and processing. This is because different big data sources can represent the same semantic information in many different ways. A huge number of data sources continuously generating an enormous amount of data and the requirement of semantic interpretation have also made big data collection and processing computationally expensive. To address these issues associated with obtaining deep insight and computational complexity, we need to dynamically scale the use of semantic interpretation in both data collection and processing and data sources based on the importance of a query.

The approach introduced in [1] can determine the significance level of a query. This approach claims that by considering the significance level of a query, the amount of time spent on less important queries can be reduced. For capturing deep insight, the data processing systems can focus more on the queries that have higher significance level. Nevertheless, in this approach, only business processes are considered by manually selecting weights for core and non-core business processes. Besides this, since the business strategies are not utilised at all, the significance level of a query calculated by this approach does not fully cover the objectives of a business organisation and hence cannot capture the deep insight of data relevant to fulfil the goals of an organisation.

As none of studies has worked on calculating significance level of a query based on the aspect of business strategy, thereby for the first time, we aim to introduce a method that determines the significance level of a query reflecting the business objectives. The contributions in this paper are highlighted as below:

- We determine the significance level of a query considering both business processes and organization strategies.
- The business process contributions are calculated based on the extent of the contribution of a process to business strategies, the number of strategies in which a process contributes to and the priority score of that strategy toward business organisation.

• The results produced by the proposed approach from our business case study have shown that the queries that are related to more important business processes and higher priority strategies, have higher significance levels.

The structure of this paper is as follows: Our proposed methodology is presented in Section 2. The results of the case study in a scenario of retail enterprise are described in Section 3. Section 4 concludes the paper.

2 Proposed Approach to Determine the Significance Level of a Query

As alluded before, a significance level of a query needs to reflect the importance of that query for a business organisation. Queries are normally very short and a business organisation is a combination of many activities and transactions. These raise an important question 'How to find a link between a query and a business organisation'. To answer this question, we have decided to choose business process and strategy as the main representation for a business organisation. This is because business processes and strategies are the two key entities that always play the main role in a business organisation [14]. Business processes represent all the activities and transactions in business organisations. Business strategies play vital roles in business development including earning more revenue, achieving strategic advantage and its expansion. A significance level of a query is thereby determined based on the relationship between query, process and strategy as seen in Fig. 1.

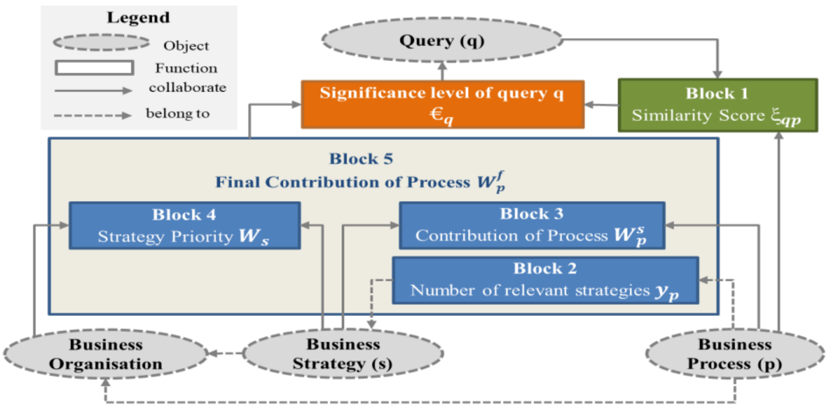

Fig. 1. The overall process of proposed approach to determine the significance level of a query

To our knowledge, the only approach available in the literature was introduced in [1]. As mentioned before, this approach determines the significance level of a query by considering only business processes. Moreover, the contribution of a process is intuitively assigned as two values, i.e., 0.5 or 1. The value 1 is for core business processes, while 0.5 for non-core business processes. Dividing processes into two groups (i) core and (ii) non-core does not fully reflect the strategic direction of a business organisation. Because some processes may have either very high or very low effect on a business organisation, while the others may have an effect that are fluctuating between the two extremes. This indicates that the contribution of a process should be a continuous value in [0, 1]. On the other hand, the business strategies that describe business tactics to achieve the business objectives and goals have not been taken into account in determining the significance level of query. Therefore, this demands the introduction of a new approach that is able to calculate the significance level of a query reflecting a process contribution more accurately and integrating the business strategies.

To embed the more accurate reflection of a business objective in a big data query, we aim to introduce an approach to calculate the significance level of a query based on the contribution weight of a process and the priority level of a strategy. Note, a contribution weight refers to how much impact a process p has on the strategy satisfaction [15]. For example, Table 1 shows a strategy S3 has three contributing processes P2, P4, P5, P6 and P12, whose contributions are 90%, 70%, 70%, 50% and 80%, respectively.

Table 1. Contribution weight of processes

	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**
S1	90%	90%	20%	90%	90%	50%	0%	0%	50%	65%	10%	65%	55%	0%
S2	50%	90%	45%	60%	90%	90%	70%	55%	50%	70%	20%	70%	70%	55%
S3	0%	90%	0%	70%	70%	50%	0%	0%	0%	0%	0%	80%	0%	0%
S4	0%	90%	0%	60%	70%	50%	0%	0%	0%	0%	0%	80%	0%	0%
S5	0%	90%	80%	75%	75%	80%	0%	0%	0%	60%	0%	80%	70%	0%
S6	0%	90%	90%	85%	85%	90%	0%	0%	0%	0%	80%	80%	70%	0%
S7	0%	90%	0%	70%	70%	50%	0%	0%	0%	0%	0%	80%	0%	0%
S8	0%	90%	0%	85%	90%	75%	0%	0%	0%	75%	0%	85%	90%	0%
S9	0%	90%	35%	10%	10%	10%	70%	50%	50%	65%	65%	0%	85%	40%
S10	0%	90%	70%	65%	65%	20%	85%	75%	80%	75%	75%	0%	75%	0%
S11	0%	90%	10%	50%	50%	75%	70%	50%	60%	70%	75%	0%	85%	0%
S12	0%	90%	10%	65%	65%	45%	0%	0%	0%	75%	65%	0%	70%	0%
S13	0%	90%	40%	85%	85%	65%	70%	40%	70%	80%	70%	0%	75%	0%
S14	0%	90%	70%	90%	90%	35%	70%	50%	0%	80%	75%	0%	80%	50%
S15	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90%
S16	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90%
S17	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90%
S18	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90%
S19	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	80%
S20	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	80%
S21	0%	90%	90%	50%	50%	55%	0%	0%	0%	0%	55%	0%	0%	0%
S22	70%	90%	75%	70%	90%	70%	0%	0%	0%	80%	55%	80%	85%	0%
S23	0%	90%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	90%

The term "strategy priority" refers to the task of ranking strategic objectives based on their importance in business organisation. By clarifying their roles and urgency, a business organisation has a better view of what to do first, what need to be focused more than the others [16, 17].

To consider process contributions to strategies and a strategy priority level in determining the significance level of a query, firstly, we need to find the process-strategy relationship. This can be done using a model introduced in [18] where the link between a strategy and a process is built by adopting a rule-based inference model.

Besides the importance of realising which strategies a process contributes to, secondly, we then need to indicate how successful a strategy will have when its related process is completely executed. This is called a contribution weight of a process for a specific strategy [15]. These contribution weights can be determined by the chief information officer or business analyst or the manager of a business organisation.

Thirdly, we need to consider how to determine the strategy priorities. A strategy priority represents the extent of importance and urgency of a strategy. This means the more importance and urgency, the higher priority the strategy has. By clarifying the strategy priority, a business organisation has a better view of what to do first and what needs to be focused on [16, 17].

In this approach, to determine the significance level of a query, five main values depicted in Fig. 1 need to be calculated: (i) the similarity scores between process p and query q, ξ_{qp} (Block 1 of Figure 1), (ii) the number of strategies y_p in which process p contributes to (Block 2), (iii) the contribution of process p to strategy s, W_p^s (Block 3), (iv) the strategy priority of strategy s, W_s (Block 4), and (v) the final contribution of process p, W_p^f using y_p, W_p^s and W_s (Block 5). Note, as described above, Block 5 represents the function for the calculation of final contribution process W_p^f. In Fig. 1, this block also shows that the calculation of W_p^f requires the execution of the functions represented by Blocks 2, 3 and 4.

The concept of semantic information is also valuable to model the business processes. The business processes that are semantically annotated have shown the enrichment in their process descriptions. A set of carefully selected semantic annotations over a business process can not only reduce the ambiguity to represent the logically connected activities of that business process for fulfilling a certain business goal but also provide business analysts with a better understanding about the business processes. Additionally, semantic annotations can also make a relationship between the business processes and the other characteristics of an organisation. Thus, the semantic annotation scheme allows us to link a query, business processes and business strategies semantically by avoiding mismatched and unstructured knowledge representation involved in a business process model [12, 13]. This has motivated us to calculate a similarity score between a query and a business process by exploiting the semantic similarity scores between the keywords of a query and the annotations of a business process.

Suppose query q consists of n keywords $\{k_1, k_2, \ldots, k_n\}$ and process p contains m annotations $\{a_1, a_2, \ldots, a_m\}$. The semantic similarity score ξ_{qp} between a process and a query is calculated as shown below:

$$\xi_{qp} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\xi(k_i, a_j)}{n\times m} \qquad (1)$$

where,

$$\xi(k_i, a_j) = e^{-\alpha l}\times\frac{(e^{\beta h} - e^{-\beta h})}{(e^{\beta h} + e^{-\beta h})}$$

Here, h is the depth of the sub-sumer of k_i and a_j in the hierarchical semantic tree of WordNet (https://wordnet.princeton.edu/). The value l is the shortest path between k_i and a_j in the hierarchical semantic tree of WordNet and this is calculated as mentioned in [19].

As mentioned before, the contribution of process p to strategy s, W_p^s can be assigned by chief information officer or business analyst. The strategy priority W_s of strategy s can be determined based on a business strategy prioritization tool [20].

We can then calculate the final contribution of a process is as follows:

(i) The contribution of a process should be determined by using a utility function that follows the diminishing property of microeconomic model [21], i.e., the contribution of a process should follow a logarithmic function. This means that similar to logarithmic curve, it should increase with increasing number of associated strategies up to a certain limit and after that the increment in contribution should slow down. Therefore, we can define the contribution of process p, W_p' as,

$$W_p' = \frac{\log(y_p^c + 1)}{\log(z^c + 1)} \qquad (2)$$

where, z is the total number of strategies considered, $0 \le W_p' \le 1$ for $y_p \le z$ and c represents the sensitivity of the process contribution with respect to the number of associated strategies. The process contribution (W_p') derived in (2) was plotted over the number of associated strategies (y_p) for different values of c=0.2, 0.5, 1.0, 2.0 and z=23, and is shown in Figure 2. Figure 2 shows for a constant value of z, the higher the value of c, the higher sensitivity of the process contribution with the number of business strategies in which a process contributes is. Therefore, for a particular business process, if it requires to quickly vary the process contribution over the number of associated strategies, the higher value of c needs to be used.

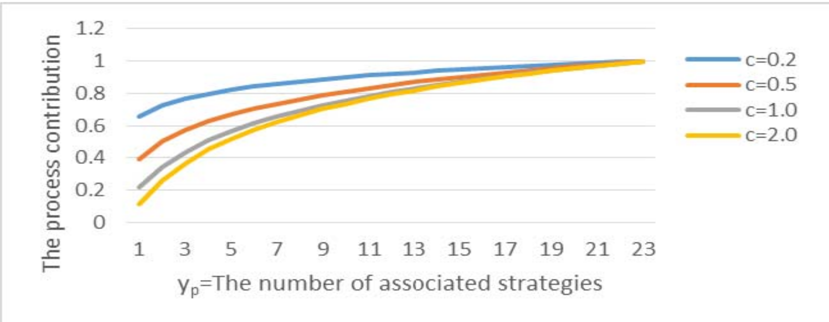

Fig. 2. The process contribution vs the number of associated strategies for different values of c=0.2, 0.5, 1.0, 2.0 and z=23

(ii) Considering the contribution of a process to all strategies, and the strategies' priority, we can define the following aggregation function to calculate the contribution of a process p, W_p^c using W_p^s and W_s as,

$$W_p^c = \frac{\sum_{s=1}^{z}(W_p^s \times W_s)}{z} \qquad (3)$$

Eq. (3) exhibits a consensus view that the higher the values of W_p^s and W_s, the higher value of W_p^c is.

Assuming the contribution of a process to the extent of covering the number of strategies, and the amount of contribution to all strategies and their priority levels as equally important, we can formulate an equation to determine the final contribution of process p in terms of a combination of W_p' and W_p^c defined in (2) and (3), respectively in the following way,

$$W_p^f = \frac{W_p' \times W_p^c}{\max(W_p^f)} \quad (4)$$

where, $\max(W_p^f)$ is the maximum value of all W_p^f and used to normalise the value in [0, 1].

Finally, similar to (2), assuming the significance level of a query varies non-linearly following a logarithmic curve with its similarity score with a process and the process contributions, the significance level of a query q can be determined using the similarity scored ξ_{qp} defined in (1) and the final contribution W_p^f derived in (4) as,

$$€_q = \frac{\log(\sum_{p=1}^{t}[(W_p^f \times \xi_{qp}) + 1]}{\log(t+1)} \quad (5)$$

with t is the total number of business processes.

3 Case Study

The proposed approach was implemented using Pyke package in Python programming language for establishing the relationship between a process and a strategy. The contribution weights of the processes were intuitively assigned and the strategy priorities shown in Table 2 were determined by using the strategy prioritization tool [20]. This case study consists of 18 queries - Q1 to Q18 as shown in [1], 14 business processes - P1 to P14 and 23 business strategies - S1 to S23 as described in [18].

Table 2. Priority score of strategies

Strategy	Priority	Strategy	Priority	Strategy	Priority
S1	7.1	**S9**	4.5	**S17**	3.5
S2	7.7	**S10**	7.8	**S18**	2.8
S3	4.8	**S11**	4.9	**S19**	5.3
S4	5.4	**S12**	4.8	**S20**	3.1
S5	7.1	**S13**	6.9	**S21**	5.7
S6	7.4	**S14**	5.4	**S22**	6.1
S7	5.0	**S15**	3.9	**S23**	4.8
S8	5.4	**S16**	3.2		

The strategy priorities were calculated using a software tool [20] and the priority value for each strategy are presented in Table 2. According to this tool, the strategy priority is determined by following three main criteria - (i) strategic fit, (ii) economic impact and (iii) feasibility. The elements of each criteria were weighted depending on their importance. The elements of each criteria for a specific strategy were then ranked between 1 to 10. The more important an element is, the higher rank it has. As a representative example, the calculated rank of each element of strategy S1 including S1's priority value (7.1) is presented in Table 3. Finally, the calculated priority value of each strategy is listed in Table 2. The results show S2 having priority value 7.7 is the highest priority, while S18 with priority value 2.8 is the lowest.

Table 3. An example for strategy priority of S1

Criteria	Element	Weight	Rank for S1
Strategic fit	Alignment with company goals	15%	9
	Market positioning	15%	9
	Core capabilities	10%	7
Economic impact	Revenue potential	10%	7
	Profitability & margin	15%	6
	Growth potential	15%	8
Feasibility	Technical risk	10%	1
	Resources - Financial	5%	8
	Resources - People	5%	7
Strategy priority for S1			**7.1**

The similarity score ξ_{qp} between each query and each process was calculated using (1). These scores are as shown in Table 4 that describes the highest (0.9994) similarity scores were obtained between Q1 and four processes {P3, P6, P7, P8} and the lowest (0.1505) similarity score was between Q8 and P9.

According to a rule-based inference model introduced in [18], we have applied this model to our scenario and found that the process P1 is associated with the strategies S1, S2 and S22. Similarly, the relationship of other processes with their relevant strategies was discovered using the same knowledge base. Then the contribution weights of all processes for each strategy were intuitively assigned and are given as described in Table 1. For example, the process P1 has the highest contribution weight on the strategy S1. This means after P1 is completely executed, the strategy S1 is expected 90% chance of success. For those places where the contribution weights are 0% such as the contribution weights of P1 on the strategies S3 to S21 and S23, this means there is no rule in knowledge base that reflects the link between P1 and these strategies or there is no effect of P1 on these strategies. Another example, strategy S7 has five contributing processes P2, P4, P5, P6 and P12, whose contributions are 90%, 70%, 70%, 50% and 80%, respectively. This means if the process P2 is completely executed, there is a 90% chance that strategy S7 is achieved. Similarly, if P4, P5, P6 or P12 is completely executed, there is a 70%, 70%, 50% or 80% of chance for the strategy S7 to be achieved.

Table 4. Similarity scores between queries and processes

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
Q1	0.86	0.26	1.00	0.20	0.74	1.00	1.00	1.00	0.19	0.38	0.32	0.30	0.17	0.45
Q2	0.34	0.69	0.46	0.57	0.57	0.48	0.55	0.55	0.29	0.56	0.53	0.61	0.56	0.43
Q3	0.29	0.42	0.28	0.73	0.50	0.33	0.46	0.46	0.31	0.46	0.42	0.56	0.40	0.60
Q4	0.25	0.41	0.34	0.49	0.45	0.29	0.44	0.44	0.17	0.41	0.30	0.45	0.37	0.51
Q5	0.42	0.56	0.39	0.56	0.56	0.43	0.47	0.45	0.18	0.56	0.39	0.56	0.47	0.51
Q6	0.60	0.39	0.46	0.40	0.57	0.56	0.43	0.43	0.25	0.37	0.41	0.40	0.29	0.57
Q7	0.43	0.41	0.43	0.47	0.62	0.39	0.56	0.56	0.22	0.42	0.28	0.49	0.40	0.38
Q8	0.36	0.49	0.41	0.43	0.54	0.35	0.52	0.52	0.15	0.67	0.21	0.42	0.40	0.38
Q9	0.36	0.44	0.44	0.44	0.52	0.37	0.71	0.62	0.26	0.44	0.32	0.40	0.38	0.36
Q10	0.52	0.45	0.47	0.43	0.64	0.49	0.59	0.59	0.28	0.44	0.28	0.45	0.41	0.34
Q11	0.45	0.42	0.41	0.46	0.62	0.46	0.54	0.54	0.31	0.41	0.30	0.48	0.37	0.42
Q12	0.40	0.42	0.46	0.40	0.63	0.45	0.57	0.57	0.22	0.45	0.29	0.43	0.42	0.28
Q13	0.47	0.45	0.44	0.50	0.53	0.37	0.39	0.39	0.23	0.53	0.41	0.38	0.41	0.70
Q14	0.25	0.34	0.28	0.37	0.31	0.21	0.34	0.28	0.23	0.38	0.36	0.42	0.28	0.40
Q15	0.39	0.49	0.26	0.42	0.46	0.35	0.53	0.50	0.21	0.40	0.35	0.45	0.30	0.46
Q16	0.21	0.29	0.24	0.34	0.45	0.15	0.29	0.29	0.21	0.32	0.28	0.29	0.29	0.29
Q17	0.37	0.27	0.37	0.36	0.42	0.28	0.33	0.33	0.23	0.37	0.26	0.38	0.27	0.45
Q18	0.32	0.50	0.25	0.43	0.39	0.38	0.30	0.28	0.24	0.33	0.36	0.48	0.27	0.49

Fig. 3 shows the significance level of all queries produced by our proposed approach using (5). To show the impact of c used in (2) on the significance level of all queries, we used three different values of c such as 0.5, 1 and 2 keeping the value of other parameters as the same. Figure 3 shows the lower the value of c, the higher the significance level of all queries is. This is because for a constant number of total strategies (e.g., 23 strategies have been used in this case study), the lower the value of c, the lower the effect of the number of strategies associated with a business process (y_p) is. Therefore, c provides a way to make a trade-off between W_p^r and W_p^c to a certain extent. For the sake of clarity, in the remainder of this paper, we will discuss the results only for c = 1. The significance levels with c = 1 are between 0.384 and 0.547. The results show that the highest significance level is 0.547. This is because, query Q2 has the highest similarity score 0.6864 with process P2. The contribution weights of process P2 with strategies S1 to S14 and S21 to S23 are 90%. The priority of strategies S1, S2, S5, S6, S10 and S13 are in the top of highest priority scores.

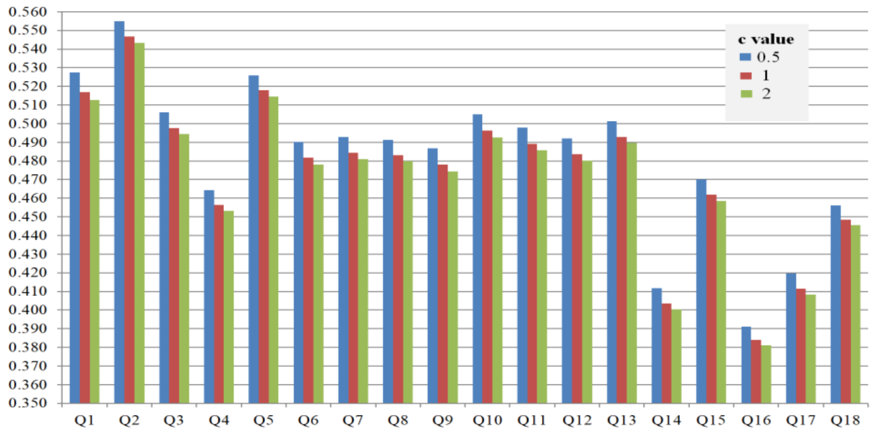

Fig. 3. Significance levels of queries

On the other hand, query Q16 has the lowest significance level as 0.384. This is mainly because this query has the highest similarity score 0.4519 with P5. Although, P5 also has contribution with similar group of strategies like P2 except S23 but the contribution weights of P5 are much lower than those of P2.

However, the significance level of the same queries computed by [1] without considering the contribution of a process to each strategy and the strategy priority is in [0.570, 0.806]. If we compare this range values with the values [0.384, 0.547] that are determined by our proposed approach for the same queries using the same processes, it shows that process contribution and strategy priority have a significant impact on accurately determining the significance level of queries.

In addition, in [1], the significance level of a query was calculated based on the contribution weight of a process which was only defined by two values, i.e., 0.5 (non-core) and 1 (core). Whereas, the contribution weight of the process in our proposed approach is not narrowed within those two values, it is a continuous value within [0, 1]. This captures the process contribution more accurately in our proposed approach than [1]. Furthermore, our proposed approach also exploits the contribution of a process to strategies and the strategy priority directly. Therefore, it is expected that the significance level of a query reflects the importance of a query for a business organisation more accurately.

4 Conclusions

In this paper, a method for calculating the significance level of a query has been introduced. The significance levels have reflected not only the importance of a query toward business organisation but also emphasised the relevance between queries and business strategies. The results from the implementation have demonstrated that, the more relevance a query has, the higher score it owns. Based on this, the enterprise information system in business organisation can understand which query is more important to focus and spend more time to extract deep insight of data. Therefore, this can improve the efficiency and effectiveness of the information system in order to manage big data better.

References

1. L. T. N. Dinh, G. Karmakar, J. Kamruzzaman, and A. Stranieri, “Significance level of a query for enterprise data,” in 30th International Conference of Business Information Management Association (IBIMA), 2017.
2. D. G. Andersen and N. Feamster, “Challenges and Opportunities in Internet Data Mining,” Parallel Data Laboratory, Carnegie Mellon University, Research Report CMU-PDL-06-102, 2006.
3. M. H. ur Rehman, C. S. Liew, A. Abbas, P. P. Jayaraman, T. Y. Wah, and S. U. Khan, “Big data reduction methods: a survey,” Data Science and Engineering, vol. 1, no. 4, pp. 265-284, 2016.
4. P. Nachnani, D. Milbrath, S. Levine, and E. Levine, “Method and system for cleaning data in a customer relationship management system,” Nov. 15 2016, US Patent 9,495,403.

5. T.-H. Chen, W. Shang, Z. M. Jiang, A. E. Hassan, M. Nasser, and P. Flora, “Finding and evaluating the performance impact of redundant data access for applications that are developed using object-relational mapping frameworks,” IEEE Transactions on Software Engineering, vol. 42, no. 12, pp. 1148-1161, 2016.
6. M. H. ur Rehman, V. Chang, A. Batool, and T. Y. Wah, “Big data reduction framework for value creation in sustainable enterprises,” International Journal of Information Management, vol. 36, no. 6, pp. 917-928, 2016.
7. M. R. Fellenz and M. Brady, “RFID and data capture technologies in global service supply chains: meeting the information management challenge,” Communications of the IBIMA, vol. 4, pp. 41-49, 2008.
8. P. Lubell-Doughtie, P. Pokharel, M. Johnston, and V. Modi, “Improving data collection and monitoring through real-time data analysis,” in 3rd ACM Symposium on Computing for Development, p. 28, 2013.
9. M. Jarke, M. Lenzerini, Y. Vassiliou, and P. Vassiliadis, “Query processing and optimization,” in Fundamentals of Data Warehouses, Springer, pp. 107-122, 2003
10. L. Vaughan, “Discovering business information from search engine query data,” Online Information Review, vol. 38, no. 4, pp. 562-574, 2014.
11. L. Vaughan and Y. Chen, “Data mining from web search queries: A comparison of Google trends and baidu index,” Journal of the Association for Information Science and Technology, vol. 66, no. 1, pp. 13-22, 2015.
12. D.M. Riehle, S. Jannaber, P. Delfmann, O. Thomas, and J. Becker, “Automatically Annotating Business Process Models with Ontology Concepts at Design-Time,” in International Conference on Conceptual Modeling, Springer, pp. 177-186, 2017.
13. C. Di Francescomarino, M. Rospocher, C. Ghidini, and A. Valerio, “The role of semantic annotations in business process modelling,” In IEEE International Conference on Enterprise Distributed Object Computing (EDOC), pp. 181-189, 2014.
14. E. D. Morrison, A. K. Ghose, H. K. Dam, K. G. Hinge, and K. Hoesch-Klohe, “Strategic alignment of business processes,” in International Conference on Service-Oriented Computing, Springer, pp. 9-21, 2011.
15. R. Guizzardi and A. N. Reis, “A method to align goals and business processes,” in International Conference on Conceptual Modeling, Springer, pp. 79-93, 2015.
16. A. Nieto-Rodriguez. (2016) How to prioritize your company’s projects. [Online]. Available: https://hbr.org/2016/12/how-to-prioritizeyour-companys-projects
17. D. Lidow. (2017) A better way to set strategic priorities. [Online]. Available: https://hbr.org/2017/02/a-better-way-to-set-strategic-priorities
18. L. T. N. Dinh, G. Karmakar, J. Kamruzzaman, and A. Stranieri, “A rule based inference model to establish strategy-process relationship,” in 30th International Conference of Business Information Management Association (IBIMA), 2017.
19. Y. Li, D. McLean, Z. A. Bandar, J. D. O’shea, and K. Crockett, “Sentence similarity based on semantic nets and corpus statistics,” IEEE transactions on knowledge and data engineering, vol. 18, no. 8, pp. 1138-1150, 2006.
20. High performance marketing – demand metric. Demand Metric. [Online]. Available: https://www.demandmetric.com/content/businessstrategy-prioritization-tool
21. D. Niyato and E. Hossain, “A microeconomic model for hierarchical bandwidth sharing in dynamic spectrum access networks,” IEEE Transactions on Computers, vol. 59, no. 7, pp. 865-877, 2010.