Exploring association among quality indicators of the In-service Education Information Service

LUNG-HSING KUO, HUNG-JEN YANG, TSUNG-JUNG TSAI, FONG-CHING SU
National Kaohsiung Normal University
No.116, Heping 1st Rd., Lingya District, Kaohsiung City 80201, Taiwan (R.O.C.)
admi@nknucc.nknu.edu.tw, hjyang@nknucc.nknu.edu.tw , davidttz@gmail.com
&
WEN-CHEN HU
Department of Computer Science
University of North Dakota
3950 Campus Road, University of North Dakota, Grand Forks, ND 58202, USA
wenchen@cs.und.edu

Abstract: -The purpose of this study was to evaluate association level among system quality indicators of the In-service Education Information Service in Taiwan. The system was first established by National Kaohsiung Normal University based the trust of Ministry of Education on the year of 2003. There are 192035 teachers using this service. An investigation research method was applied to examine the data of quality indicators. The research structure in this study included dependent variables of connection quality indicators, and independent variables of system hardware quality indicators. A system monitoring tool called PRTG was used for collecting data. The sample collected period was from 2016/9/30 to 2017/9/30. There were overall 6735 records of each indicators. The sample size was 629 and confidence interval was 4.9 at confidence level of 99%. A canonical correlation analysis procedure was applied to reveal the relationship between system hardware and connection quality indicators. Based upon verified statistical analysis results, three major conclusions were presented. The are four significant canonical correlation pairs between connection set and hardware set.

Key-Words: - System Quality Relationship, In-Service Education Information Service, Canonical Correlation Analysis

1 Introduction

The purpose of this study was to evaluate association level among system quality indicators of the In-service Education Information Service in Taiwan. The system was first established by National Kaohsiung Normal University based the trust of Ministry of Education on the year of 2003. Since then, the system has been maintained for near 200 thousands users to access in-service education information service.

1.1 In-service Education Information Service

This service provides end-users to access information about in-service education courses. Teachers could to register course through the system and check with their personal in-service education records. Teachers also could search courses offered by nationally authorized institutions.

In-service course providing institutions create course record on this system. After proved by higher rank administrator, the course information could be circulated nationally on the system. Courses provided could be searched via course search interface by any user.

Fig. 1 National In-service Education Information Service Webside

Course offered institutions would get registry information and know who would be in the class beforehand. The course providing institution would upload learning record of each course members

1.2 Users' Characteristics of Service

Users of our service are teachers around the whole nation. They may access the service from their institutions and their home also. There are 192035 teachers as mentioned in the 2015 yearbook of teacher Education Statistics Education [1]. The service users are not only teachers, but also supervisors and administrators of institutions which offer in-service education courses.

General users might request information about what courses they could take, when the course would be conducted, where the course would be taught, and even register a course.

For the course providers or institutions, they require the service of creating course, editing course, announce a course, and recording attendance of a course.

2 Methodology

The purpose of this study was to evaluate system quality of the In-service Education Information Service in Taiwan. An investigation research method was applied to examine the data of quality indicators.

In this section, research structure, research objects, research steps, research tools, data analysis, and statistical hypothesis would be reported.

2.1 Research Structure

The research structure in this study included dependent variables of quality indicators such as service uptime, and independent variables of locations, month, and weekday.

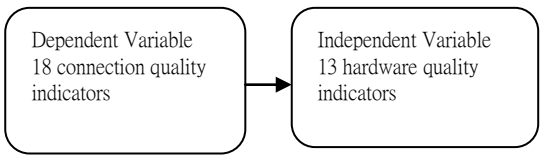

Fig. 2 Research Structure

There are 31 quality indicators which could be grouped into two sub-sets. The first set is connection quality indicators and the second set is hardware quality indicators. There are eighteen indicators in the first set. There are thirteen indicators in the second set.

Table 1Connection quality indication at NKNU

Indicator Name	
C2WEB1	C01
C2WEB2	C02
C2WEB4	C03
C2WEB3	C04
C2NCHC	C05
C2SciTechVista	C06
C2KNOWLEDGE	C07
C2KUAS	C08
C2NSYSU	C09
C2LightProfDeveWeb	C10
C2MOEelearn	C11
C2MOEProfDevIntegrationWeb	C12
C2MOE	C13
C2Hinet	C14
C2Yahoo	C15
C2Google	C16
C2CNN	C17
C2Fb	C18

In Table 1, eighteen connection quality indicators located at NKNU were listed. Those are indicators used to measure point to point connecting status including uptime/downtime, response time, and time stamp.

Table 2 Connection quality indicators at NCHC

Indicator Name	
NKNU System Health	H01
NKNU Probe Health	H02
NKNU MSSQLserverStatistics	H03
NKNU % Disk Write Time	H04
NKNU % Disk Read Time	H05
NKNU % Disk Time	H06
NKNU Disk Bytes/Sec Total	H07
NKNU Disk Read Bytes/sec Total	H08
NKNU Disk Reads/sec Total	H09
NKNU Disk Write Bytes/sec Total	H10
NKNU Disk Writes/sec Total	H11
NKNU Disk Transfers/sec Total	H12
NKNU Intra Net Total	H13

In Table 2, thirteen hardware quality indicators located at NKNU were listed. Those are indicators used to measure database/SQL related status including system health, probe health, disk speed, disk read speed, disk transfer speed, disk write speed, batch speed, and intra-net volume.

2.2 Research Objects

The purpose of this study was to evaluate system quality of the In-service Education Information Service in Taiwan. In this study, the research objects

are quality indicators of the system. The research data had collected since 2016.

The data collected period was from 2016/9/30 to 2017/9/30. The population of monitored data was 6735. The relationship between two sets of quality indicators would be based upon a whole year random sampled 629 records.

According to the population and sample size, the confidence interval is 4.9 at 50 percentage and confidence level of 99%.

2.3 Data analysis

In this section, CCA would be review in focus of general form, limitation, and fundamental equations for explaining data analysis of this study.

2.3.1 Canonical Correlation in General

The goal of canonical correlation is to analyze the relationships between two sets of variables. Canonical correlation provides a statistical analysis for research in which each subject is measured on two sets of variables and the researcher wants to know if and how the two sets relate to each other[2, 3].

Sets of variables on each side are combined to produce, for each side, a predicted value that has the highest correlation with the predicted value on the other side. The combination of variables on each side can be thought of as a dimension that relates the variables on one side to the variables on the other[2].

2.3.2 Limitations of Canonical Correlation

In theory, the most critical limitation is interpretability. Canonical solutions are often mathematically elegant but uninterpretable.

The algorithm used for canonical correlation maximizes the linear relationship between two sets of variables. If the relationship is nonlinear, the analysis misses some or most of it.

2.3.3 Fundamental Equations for Canonical Correlation

There are several ways to write the fundamental equation for canonical correlation—some more intuitively appealing than others[2]. The equations are all variants on the following equation:

$$R = R_{yy}^{-1} R_{yx} R_{xx}^{-1} R_{xy} \quad (1)$$

The canonical correlation matrix is a product of four correlation matrices, between DVs (inverted), between IVs (inverted), and between DVs and IVs.

Although computing eigenvalues and eigenvectors is best left to the computer, the relationship between a canonical correlation and an eigenvalue is listed as following.

$$\lambda_i = r_{ci}^2 \quad (2)$$

Each eigenvalue, λ_i, is equal to the squared canonical correlation, r_{ci}^2, for the pair of canonical variates.

Two sets of canonical coefficients (analogous to regression coefficients) are required for each canonical correlation, one set to combine the DVs and the other to combine the IVs. The canonical coefficients for the DVs are found as follows:

$$B_y = (R_{yy}^{-1/2})' \overline{B}_y \quad (3)$$

3 Findings

In this section, research findings would be reported according to investigation results. First, descriptive results of investigation would be presented. Second, verified statistical analysis results would be reported.

Those thirty-one quality indicators were investigated mainly focused on the performance.

3.1 Descriptive Analysis

In the following, quality indicators would be reported based upon descriptive statistics.

3.1.1 Quality Indicators of Connection

There were eighteen connection quality indicators for monitoring system quality in different ways.

In Table 3, their N, Minimum, Maximum, Mean, and Std. Deviation were listed under ID.

Table 3 Connection Quality Indicators' N, Min, Max, Mean, Std. Deviation, & Unit

ID	N	Minimum	Maximum	Mean	Std. Deviation
1.	619	8.2500	13929.9833	171.461273	1263.989613
2.	619	6.8667	114.5500	15.743763	8.9810229
3.	619	8.6333	17283.2833	181.823502	1384.498811
4.	619	68.2167	16446.3667	283.483286	1546.355359
5.	619	79.3000	3247.9333	144.915545	248.8617985
6.	619	68.7667	10811.0769	995.490218	819.831977
7.	619	2923.8167	12923.5946	4705.658904	996.2406879
8.	619	148.8000	2096.9667	308.315330	198.3648539
9.	619	17.8167	2944.0345	29.890810	123.745990
10	619	173.6500	1420.2292	230.821350	133.3181393
11	619	45.5833	30815.1290	760.590021	1446.033558
12	619	137.8667	1181.6333	202.114589	93.9847714
13	619	120.6316	2423.5500	743.115688	578.4719182
14	619	44.1000	1555.5208	62.659402	71.8122368
15	619	451.9000	3851.3833	701.841477	363.9108182
16	619	86.1000	6847.7833	162.753757	383.0950572
17	619	195.5667	2608.4500	568.636883	345.7645276
18	619	280.9500	2185.3500	423.918629	153.2278803

3.2.2 Quality Indicators of Hardware

There were thirteen connection quality indicators for monitoring system quality in different ways. In Table 4, their N, Minimum, Maximum, Mean, and Std. Deviation were listed under ID.

Table 4 Hardware Quality Indicators' N, Min, Max, Mean, Std. Deviation, & Unit

ID	N	Minimum	Maximum	Mean	Std. Deviation
19	619	53.6667	100.0000	69.676400	5.8066243
20	619	0	100.0000	99.451799	6.6454506
21	619	0.3619	646.8956	41.460682	88.0867864
22	619	0.0346	13172.8274	148.281855	1116.848908
23	619	0.0302	17625.05554	161.508197	1310.167416
24	619	0.0802	26528.2554	224.610040	2073.766567
25	619	150266.0300	502662830.8	14360936.75	34924133.09
26	619	28499.9540	21924317.95	8830505.036	4.807E+14
27	619	1.0491	4536.3640	119.855906	273.0865887
28	619	102302.8062	252910557.5	5543645.130	14409062.36
29	619	5.4635	4012.7895	112.740351	262.2988147
30	619	6.6631	8545.1563	232.500412	502.3049698
31	619	3610868.677	1.05471E+11	4076079157	1.31210E+10

3.3 Verified Analysis

There are three verified analysis reported in this section. Those are

- Significance of canonical correlations
- Correlations of variables and variants
- Variance accounted for
  - By canonical correlations
  - By same-set canonical variates
  - By other-set canonical variates (redundancy).

3.3.1 Evaluation of Assumptions

To improve linearity of relationship between variables and normality of their distributions, transformation techniques were applied to variables.

No outliers were identified. Assumptions regarding within-set multicollinearity were met.

3.3.2 Canonical Correlation

In Table 5, the first canonical correlation was 0.755; the second was 0.641; the third was 0.455; the fourth was 0.386; the fifth was 0.294.

In Table 12, all five canonical correlations included, $X^2(234, N = 619) = 1236.71$, $p < 0.001$, with the first canonical correlation removed, $X^2(204, N = 619) = 727.66$, $p < 0.001$, with the second canonical removed, $X^2(176, N = 619) = 408.63$, $p < 0.001$, with third canonical correlation removed, $X^2(150, N = 619) = 269.02$, $p < 0.001$, and with the fourth canonical correlation removed, $X^2(150, N = 619) = 171.71$, $p < 0.05$. Subsequent X^2 tests were not statistically significant. The first five pairs of canonical variates, therefore, accounted for the significant relationships between the two sets of variables.

Table 5 Canonical correlations

	Correlation
1	0.755
2	0.641
3	0.455
4	0.386
5	0.294
6	0.256
7	0.209
8	0.164
9	0.156
10	0.131
11	0.077
12	0.069
13	0.041

The fifth pair would not be dropped because of the correlations less than 0.3. Data on the first four pairs of canonical variates appear in Table 13. Shown in the table are correlations between the variables and the canonical variates, standardized canonical variate coefficients, within-set variance accounted for by the canonical variates (proportion of variance, redundancies, and canonical correlations.

Table 6 Test that remaining correlations are zero

	Wilk's	Chi-SQ	DF	Sig.
1	0.128	1236.713	234	0
2	0.299	727.657	204	0
3	0.507	408.631	176	0
4	0.64	269.017	150	0
5	0.752	171.713	126	0.004
6	0.823	117.099	104	0.179
7	0.881	76.205	84	0.715
8	0.921	49.246	66	0.939
9	0.947	32.746	50	0.972
10	0.971	17.856	36	0.995
11	0.988	7.487	24	0.999
12	0.993	3.932	14	0.996
13	0.998	1.019	6	0.985

4 Conclusion

The purpose of this study was to evaluate relationship between quality indicators sets, hardware set and connection set, of the In-service Education Information Service in Taiwan. Based upon an investigation method, quality indicators were identified and established probe to collect long term data for evaluation.

According to the research findings, there are three major conclusions.

Table 7 Correlations, Standardized Canonical Coefficients, Canonical Correlations, Proportions of Variance, and Redundancies Between Attitudinal and Health Variables and Their Corresponding Canonical Variates

Quality Indicator	First canonical variate	coefficient	2 correlation	coefficient	3 correlation	coefficient	4 correlation	coefficient
Connection Set								
C2WEB1	-0.53	-0.05	-0.46	-0.34		-0.37		0.00
C2WEB2		0.07	0.64	0.38		0.05		-0.21
C2WEB4	0.89	0.77		-0.61		-0.30		-0.10
C2WEB3	-0.66	-0.21		0.02		-0.16	-0.39	-0.50
C2NCHC		-0.19	0.40	-0.06		0.16	0.31	0.14
C2SciTechVista		0.01	0.35	0.08		-0.01	0.42	0.26
C2KNOWLEDGE		0.16	0.61	0.28		0.15	0.33	-0.02
C2KUAS		0.04		0.25	0.65	0.57		0.18
C2NSYSU		0.01		-0.09		-0.05		-0.09
C2LightProfDeveWeb	-0.37	-0.25	-0.48	-0.37	0.33	0.33	0.48	0.48
C2MOEelearn		0.28	-0.61	-0.17		0.30		-0.23
C2MOEProfDevIntegrationWeb		0.08	0.39	0.10		0.09		-0.25
C2MOE		-0.12		-0.10		0.03		-0.33
C2Hinet		-0.05		-0.04		-0.05		0.09
C2Yahoo		0.12		-0.02	0.61	0.52		-0.09
C2Google		0.04		0.01		0.12		-0.20
C2CNN		0.05		0.09		-0.17	0.34	0.21
C2Fb		-0.06		0.14		0.09	0.37	0.18
Proportion of variance	0.11		0.12		0.07		0.09	Total =.39
Redundancy	0.06		0.05		0.01		0.01	Total =.14
Hardware Set								
NKNU System Health		-0.18	-0.60	-0.43	-0.47	-0.60	-0.52	-0.57
NKNU Probe Health	0.79	0.45		-0.44		-0.32		0.72
NKNU MSSQL serverStatistics	0.77	0.40		0.16	0.32	0.37	-0.32	-0.59
NKNU % Disk Write Time		0.22	0.64	0.48	-0.53	-0.41		-0.49
NKNU % Disk Read Time		0.40	-0.37	0.58	0.32	-0.08		-0.34
NKNU % Disk Time		-0.22	-0.56	-0.64	0.45	0.68		0.26
NKNU Disk Bytes/Sec Total		-0.28	-0.54	0.14	0.33	1.09		-1.41
NKNU Disk Read Bytes/sec Total	0.31	0.02	-0.44	-0.60		-1.47		1.48
NKNU Disk Reads/sec Total	0.44	-0.32	-0.40	0.46		0.91		-0.50
NKNU Disk Write Bytes/sec Total		-0.20	-0.51	-0.04	0.47	0.01		0.39
NKNU Disk Writes/sec Total		0.00	-0.53	0.04	0.43	0.33		-0.78
NKNU Disk Transfers/sec Total		0.55	-0.48	-0.14	0.38	-1.41		0.86
NKNU Intra Net Total	0.78	0.27	-0.39	-0.08		0.12		-0.36
Proportion of variance	0.19		0.23		0.13		0.06	Total =.61
Redundancy	0.11		0.10		0.03		0.01	Total =.24
Canonical correlation etc.	0.76		0.64		0.46		0.39	

4.1 Relationship between Sets

The are four significant pairs between connection set and hardware set.

In Fig. 7, each pair is illustrated quality indicators of both connection set and hardware set.

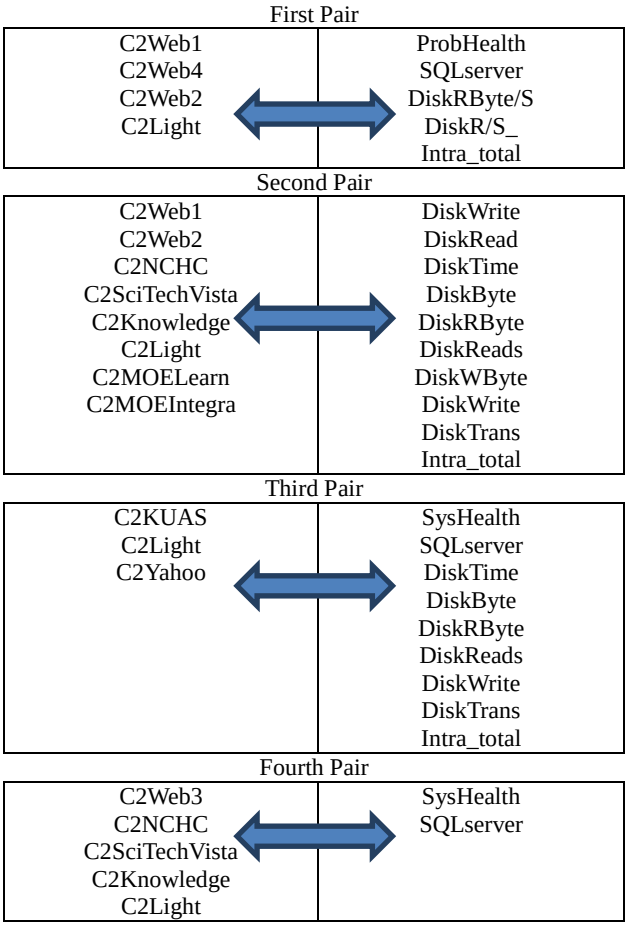

Fig. 3 Canonical correlation pairs between connection set and hardwar set

4.2 Canonical Variates

Total proportion of variance and total redundancy indicate that the first two pair of canonical variates was highly related, the next two pairs were moderately related.

With a cutoff correlation of .3, the variables in the connection set that were correlated with the first canonical variate were (invert of) C2Web1, (log of) C2Web4, (invert of) C2Web3, and (invert of) Light Prof. Dev. Integration Web. Among the hardware variables, (reflect & Log of) Probe health, (invert of) MS Sql Server, (invert of) Disk Read Bytes, (invert of) Disk Read, (invert of) Intra Net correlated with the first canonical variate. The first pair of canonical variates indicates that those with C2Web1(-0.53), C2Web4(0.89), C2Web3(-0.66), and Light Prof. Dev. Integration Web.(-0.37) are associated with

Probe health (0.79), MS Sql Server (0.77), Disk Read Bytes (0.31), Disk Read (0.44), intra Net (0.78).

In Table 13, the variables in the connection set that were correlated with the second canonical variate were (invert of) C2Web1, (square root of) C2Web2, (log of) C2NCHC, (log of) C2SciTechVista, (square root) of C2Knowledge, (invert of) Light Prof. Dev. Integration Web, (invert of) MOE elearn, and (square root of) MOEProDevIntegrationWeb. Among the hardware variables, (square root of) system health, (log of) %disk write time, (invert of) %disk read time, (invert of) %disk time, (invert of) disk bytes, (invert of) disk read bytes, (invert of) disk read, (invert of) disk write bytes, (invert of) disk writes, (invert of) disk transfer, and (invert of) intra net correlated with the second canonical variate. The second pair of canonical variates indicates that those with C2Web1 (-0.46), C2Web2 (0.64), C2NCHC (0.40), C2SciTechVista (0.35), C2Knowledge (0.61), Light Prof. Dev. Integration Web (-0.48), MOE elearn (-0.61), and MOEProDevIntegrationWeb (0.39) are associated with system health (-0.60), %disk write time (0.64), %disk read time (-0.37), %disk time (-0.56), disk bytes (-0.54), disk read bytes (-0.44), disk reads (-0.40), disk write bytes (-0.51), disk writes(-0.52), disk transfer (-0.48), and intra net (-0.39).

In Table 13, the variables in the connection set that were correlated with the third canonical variate were (invert of) C2KUAS, (invert of) C2LightProfDeveWeb, and (invert of) C2Yahoo. Among the hardware variables, (square root of) system health, (reflect & log of) MS Sql server, (log of) %disk write time, (invert of) %disk read time, (invert of) %disk time, (invert of) disk bytes, (invert of) disk write bytes, (invert of) disk writes, and (invert of) disk transfer correlated with the third canonical variate. The third pair of canonical variates indicates that those with C2KUAS (0.65), C2LightProfDeveWeb (0.33), and C2Yahoo (0.61) are associated with system health (-0.47), MS Sql server (0.32), %disk write time (-0.53), %disk read time (0.32), %disk time (0.45), disk bytes (0.33), disk write bytes (0.47), disk writes (0.43), and disk transfer (0.38).

In Table 13, the variables in the connection set that were correlated with the third canonical variate were (invert of) C2Web3, (log of) C2NCHC, (log of) C2SciTechVista, (square root of) C2Knowledge, (invert of) C2LightProfDeveWeb, (square root of) C2CNN, and (log of) C2Fb. Among the hardware variables, (square root of) system health and, (reflect & log of) MS Sql server correlated with the fourth canonical variate. The fourth pair of canonical variates indicates that those with C2Web3 (-0.39), C2NCHC (0.31), C2SciTechVista (0.42), C2Knowledge (0.33), C2LightProfDeveWeb (0.48), C2CNN (0.34), and C2Fb (0.37) are associated with system health (-0.52) and, MS Sql server (-0.32).

4.3 Implications

Based upon the first pair of canonical variates, connection C2web1, C2web3 and Light prof. dev. Integration web with less traffic but higher traffic on C2Web4 are likely to have more quality of probe health, MS SQL service, Disk Read Butes, Disk Read , and intra net flow

Based upon the second pair of canonical variates, that is, connection C2web2, and Light prof. dev. Integration web with less traffic but higher traffic on C2Web2, C2NCHC, C2SciTechVista and C2Knowledge are likely to have more quality of disk write time, but lower quality of system health, disk read time, % disk read time, % disk time, disk bytes, disk read bytes, disk read, disk write bytes, disk writes, disk transfer and intra net.

Based upon the third pair of canonical variates, that is, connection C2KUAS, Light prof. dev. Integration web and, C2Yahoo with higher traffic are likely to have more quality of MS Sql, % disk read time, % disk time, disk bytes, disk write bytes, disk writes and, % disk transfer but lower quality of system health, and % disk write time.

Based upon the fourth pair of canonical variates, that is, connection C2NCHC, C2SciTechVista, C2Knowledge, Light prof. dev. Integration web, C2CNN and, C2Fb with higher traffic but lower traffic on C2Web3 are likely to have less quality of system health and MS Sql.

References:

[1] M. o. Education, "In-service Teachers," *Yearbook of Teacher Education Statistics, The Republic of China,* 2015.

[2] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 6 ed. Boston: Pearson, 2013.

[3] E. VAICIUKYNAS *et al.*, "Towards Voice and Query Data-based Non-invasive Screening for Laryngeal Disorders," in *The 17th International Conference on Automatic Control, Modelling & Simulation (ACMOS '15)*, Tenerife, Canary Islands, Spain, 2015, pp. 32-39: WSEAS Press.