Research Article

A Study of K-ISMS Fault Analysis for Constructing Secure Internet of Things Service

Hwankuk Kim, 1 Jongin Lim, 2 and Kyungho Lee 2

1 Korea Internet & Security Agency (KISA), Seoul 138-803, Republic of Korea
2 Center for Information Security Technologies (CIST), Korea University, Seoul 136-701, Republic of Korea

Correspondence should be addressed to Hwankuk Kim; rinyfeel@kisa.or.kr

Received 18 November 2014; Accepted 16 December 2014

1. Introduction

According to a survey by Gartner, things connected to the Internet are expected to grow to 26 billion, and market size is forecast to reach USD 1 trillion by 2020 [1]. IoT is widely applied in various areas closely related to daily lives such as smart home appliance, smart car, and healthcare. Since multiple networks can be controlled even with a sensor, hacking of an area can be fatal since it can cause security threats in a chain reaction. With the IoT service broadly utilizing the sensor information to provide a wide range of information, the risk of information leak will increase. Malfunction or suspension of IoT devices will also pose very serious threat even to social infrastructure that the economic damage is predicted to reach KRW 17.7 trillion by 2020 [1]. Therefore, there is a need to consider the technical and administrative vulnerabilities of each element, such as sensor, wired and wireless network, and platform of IoT, from the design stage and to review and study the application of the existing standard as the key tool for continuously inspecting them even at the operating stage.

Korean government operates the information security management system “K-ISMS” certification system to assess if an organization has established and managed appropriate information security environment. Therefore, the K-ISMS, which is similar to “ISO27001” as the international standard for information security management systems, is designed to improve the information security management level of enterprises and protect them from various security threats [2]. K-ISMS evaluates whether an enterprise has set up a comprehensive information security management system including administrative, technical, and physical protective measures to protect the safety of their information assets, using the 104 certification criteria; it then issues the certification if the enterprise meets the requirements. K-ISMS was introduced in Korea in 2002, and 332 enterprises have acquired the certification. Since the acquisition of the certification by an
enterprise over a certain size became mandatory in 2013, demand for and interest in K-ISMS certification has been gradually increasing [3].

With Internet security incidents (e.g., hacking) occurring frequently in K-ISMS certified enterprises, however, there are growing concerns about the quality deterioration of K-ISMS certification assessment these days. Since the information security management systems of enterprises are different, and fault cases vary, specialization and extensive experiences in certification assessment are required. Moreover, there are limits in maintaining objective and accurate assessment quality, since the 104 criteria should be evaluated within a short period of time. To solve these problems, there were various studies including the case study of faults in K-ISMS [4], the economic effect analysis of K-ISMS certification [5], and the analysis about process of security management for various IT services [6, 7]. However, a study using data mining between faults was not performed yet.

Thus, we studied to provide a guide to extracting preferred assessment items during the limited assessment period by analyzing the fault pattern that occurs frequently and association through the K-ISMS fault data. For this purpose, we apply the data mining analysis technique in order to analyze the association relationship among the fault items of the K-ISMS certification criteria. The paper is organized as follows. Section 2 introduces association rule mining the most known and used unsupervised algorithms. In Sections 3 and 4, the experiments are performed on K-ISMS fault data. The conclusion is given in Section 5.

2. Theoretical Background

Data mining is a knowledge-finding process that extracts unknown useful information by analyzing a large quantity of accumulated data. Among the research studies identifying the hidden pattern in the data, the association rule finding area was studied most widely in many areas of market forecast, medical and IT engineering research [8, 9]. The association rule analysis refers to a technique that finds a useful pattern, which is expressed as a “condition-result” formula among data items. The list of association rules extractable from a given data set is compared in order to evaluate their importance level. The measures commonly used to assess the strength of an association rule are the indexes of support, confidence, and interest [10].

The problem of finding association rules $X \rightarrow Y$ was first introduced in 1993 by Agrawal et al. [11] as the data mining task of finding frequently cooccurring items in a large Boolean transactional database D [12]. Typical applications include retail market basket analysis, item recommendation systems, cross selling, and loss-leader analysis. In the classical framework, an association rule is considered to be interesting if its support (s) and confidence (c) exceed some user-defined minimum thresholds [13]. Support is defined as the percentage of transactions in the data that contain all items in both the antecedent and the consequent of the rule; that is, $P(X \cap Y) = |X \cap Y|/|D|$ [14]. Confidence on the other hand is an estimate of the conditional probability of $Y$ given $X$; that is, $P(X \cap Y)/P(X)$ [13].

Association rule finding consists of the process of identifying an association rule that has the threshold value of predefined support and confidence. This process broadly includes two processes [15, 16]. One is “frequent itemset finding,” which finds the itemset that satisfies the support threshold value “minimum support” only as a technique of finding the items that occur concurrently in the transaction. The other is “association rule generation,” which adopts the rule satisfying the confidence threshold value “minimum confidence” only among association rules created from the found frequent itemsets.

2.1. Frequent Itemset Finding. While finding the frequent itemset, a combination of the itemset that can be created from the given item is created, and the transaction data is searched for the individual itemset that has been created in this manner to check whether the minimum support can be satisfied.

When a set of frequent items in the transaction database is $I = \{i_1, i_2, \ldots, i_n\}$ and a transaction set composed of $m$ transactions $T = \{t_1, t_2, \ldots, t_m\}$ is given, transaction $t_i$ is defined as a subset of frequent set $I (t_i \subseteq I)$ [17]. If $I$ includes $k$ items, it is defined as $k$-itemset [17, 18]. For example, {Beer, Diapers, Milk} are 3 itemsets; the null item has no items at all.

If transaction $t_i$ includes all items in itemset $X$, $t_i$ is said to support $X$, expressed as Supp($X$). At this time, support count, $\sigma(X)$ can be regarded as a number of transactions including the itemset in question:

$$\sigma(X) = \left| \{t_i \mid X \subseteq t_i, t_i \in T\} \right|. \quad (1)$$

When the user defines minimum support (minsup) and $\sigma(X) \geq$ minsup is satisfied, itemset $X$ is called frequent itemset [19].

2.2. Association Rule Generation. Rule generation is a process of creating an association rule from the frequent itemset found during the “frequent itemset finding” process. Suppose $X$ and $Y$ are a set of items that do not contain the same element: $X \subset I$, $Y \subset I$ and $X \cap Y \neq \emptyset$ and $Y \neq \emptyset$ [16]. The association rule expresses an association among frequently occurring data in the form of “condition $\rightarrow$ result (Rule: $X \Rightarrow Y$)” rule. At this time, $X$ is called LHS (Left Hand Side), and $Y$ is called RHS (Right Hand Side) [19].

Support and confidence are used as a statistical criterion to verify the validity of the association rule. “Support” is a ratio of the transactions that satisfy items $X$ and $Y$ among all transactions and is expressed as Supp($X \Rightarrow Y$). At this time, since “support” implies the frequency of the frequently occurring pattern or rule, “support” should have a big value to increase the usefulness of the pattern or rule:

$$\text{Supp}(X \Rightarrow Y) = \frac{n(X \cap Y)}{n(\text{Total})}. \quad (2)$$

“Confidence” is a criterion for implying the strength of the rule. If the rule $X \Rightarrow Y$ exists, “confidence” refers to a ratio of the transaction that includes $Y$ at the same time, among the transactions that include $X$. “Confidence” is expressed as $\text{conf}(X \Rightarrow Y)$. This "confidence" becomes an
index that can measure the accuracy of the conclusion Y’s rule, if condition X is true. Therefore, high “support” enables accurate prediction:

\[
\text{Conf}(X \implies Y) = \frac{\Pr(X \land Y)}{\Pr(X)} = \frac{n(X \land Y)}{n(X)}.
\]

Finding an association rule in the data item involves finding an itemset that has higher support and confidence than the user-defined minimum support and minimum confidence. For example, let us assume a situation where in the following association rule candidates are identified from the \{bread, egg, milk\} itemset [18]. At this time, if the minimum confidence is 70%, the second and third association rules will be selected. In this way, possible combinations of all itemsets are created, and some of them are selected as association rule depending on whether the minimum confidence is satisfied or not.

- (Rule 1) bread ⇒ egg, milk confidence = 0.3/0.5 = 60%.
- (Rule 2) bread, egg ⇒ milk confidence = 0.3/0.3 = 100%.
- (Rule 3) bread, milk ⇒ egg confidence = 0.3/0.4 = 75%.

A strong association rule can be filtered out using the support and confidence criteria. However, there are weakness of support and confidence. Support suffers from the “rare item problem” [20]: infrequent items not meeting minimum support are ignored which is problematic if rare items are important [21]. On the other hand, if the minimum support is low, the find area becomes larger. In addition, high minimum confidence and minimum support do not necessarily mean strong association, and they can occur accidently. Therefore, there is a need for statistical correlation analysis, such as lift and conviction, to solve these problems and find a strong association rule [21]. Interest (or lift) is another statistic which attempts to correct this weakness [22].

Confidence tends to rate rules highly where the consequent Y is frequent [23]. For example, if 80% of transactions in a database contain Y, then the expected confidence of any rule X → Y is 80%, even before taking the influence of X on Y into account.

The interest(X → Y) is defined as the confidence(X → Y) divided by the proportion of all transactions that contain Y. This scales the confidence to account for the commonality (or rarity) of Y [22].

The interest measure [13] is defined over \([0, \infty]\) and its interpretation is as follows:

\[
\text{Interst}(X \implies Y) = \text{lift} = \frac{\Pr(X \land Y)}{\Pr(X) \Pr(Y)} = \frac{\text{Supp}(X \cup Y)}{\text{Supp}(X) \cdot \text{Supp}(Y)}.
\]

(i) If \text{interest}(\text{lift}) < 1, then X and Y appear less frequently together in the data than expected under the assumption of conditional independence. X and Y are said to be negatively interdependent.

(ii) If \text{interest}(\text{lift}) = 1, then X and Y appear as frequently together as expected under the assumption of conditional independence. X and Y are said to be independent of each other.

(iii) If \text{interest}(\text{lift}) > 1, then X and Y appear more frequently together in the data than expected under the assumption of conditional independence. X and Y are said to be positively interdependent.

3. Association Rule Analysis of the K-ISMS Fault

3.1. Analysis Data. In this paper, we analyzed the fault data of 76 enterprises that received certification assessment in 2013 (uses only the statistical results) and applied the representative “Apriori algorithm” [17–19] for association rule mining. The average fault rate of those enterprises was found to be 15%. (The explanation about terms of K-ISMS control items used in this paper is described in appendix.)

Among the 104 K-ISMS certification assessment items, the frequent itemset whose value was higher than the minimum support was created. A total of 825 rules were found using the brute-force method. When a number of fault items included in these rules were analyzed, 58 one-itemsets, 494 two-itemsets, 312 three-itemsets, and 20 four-itemsets were created. Figure 1 illustrates the K-ISMS fault items that occur frequently, listed in order of CL-1-1 (asset identification), AC-3-3 (access control), OS-2-2 (security system operation), AC-2-3 (access right review), and CC-1-1 (encryption policy establishment).

If a number of items increase when generating a frequently occurring itemset candidate, the computation workload increases exponentially. To solve this problem, the “pruning” method [24] is used to make the unnecessary part concise. “Pruning” involves getting rid of the combination that does not satisfy the threshold criterion in each phase. The most universal pruning method is MSP (minimum support pruning). In other words, if support of the itemset combination is smaller than the threshold value, the item is no longer added. To remove the duplicated association rule, 307 association rule candidate groups are created through support-based pruning.

3.2. Association Rule Analysis. In the stage of support analysis, a 10% support of a certain rule (X → Y) means that the ratio of the rule followed by the rule in question is 10% among all faults. Figure 2(a) shows a support distribution graph of all fault items. The average of support is 13.7%, and 2 rules have more than 30% support value. On the other hand, 19 rules have over 20% support value and 286 rules have over 10% support value. (We used arules package in R tools [19, 25].)

Support is designed to measure how frequently those two faults occur among all transactions, whereas confidence implies that the possibility of fault “Y” occurrence is 30% if “X” has occurred when the confidence of a certain rule (X → Y) is 30%. Figure 2(b) shows a confidence distribution graph of all fault items. The average of confidence is 51%.
11 rules have more than 80% confidence and 98 rules have over 60% confidence and 160 rules have over 50% confidence.

Figure 3 shows a visualization of the correlation among fault items that were found by the measure criterion (support, confidence, and lift) of association analysis.

Table 1 shows the top 5 association rules sorted by the measure of support value. Number 44 of rule can be analyzed as follows. The support level is 34.2%, which is the ratio of finding a fault in the CL-1-1 and OS-2-2 control items at the same time. There is 59% probability that a fault occurs in the OS-2-2 control item if a fault occurs in the CL-1-1 control item. In addition, since the lift is over 1, which shows a correlation between the two control items, the correlation of the association rule is strong. On the other hand, Number 443 of rule is the association rule with low correlation because the lift is under 1, even though the support and confidence of this rule are 30.2% and 52.2%, respectively.

Table 2 shows the top 5 association rules sorted by the measure of confidence.

Number 40 of rule can be analyzed as follows. The ratio that a fault occurs in control items CC-2-1, DR-2-1, and AC-2-3 at the same time “support” is low (10.5%). Note, however,
Figure 3: (a) This graph focuses on how the rules are composed of individual items and shows which rules share items. (b) Scatter plot with three measures (e.g., support and confidence and lift).

Table 1: Top 5 association rules sorted by the measure of support.

| Rules | Support | Confidence | Lift |
|-------|---------|------------|------|
| 442   | {{(C)CL-1-1} ⇒ {(C)OS-2-2}} | 34.2% | 59.0% | 1.247 |
| 443   | {{(C)AC-3-3} ⇒ {(C)CL-1-1}} | 30.2% | 52.2% | 0.902 |
| 369   | {{(C)PH-1-4} ⇒ {(C)AC-3-3}} | 26.3% | 74.0% | 1.279 |
| 423   | {{(C)CC-1-1} ⇒ {(C)CL-1-1}} | 26.3% | 60.6% | 1.046 |
| 429   | {{(C)AC-2-3} ⇒ {(C)AC-3-3}} | 26.3% | 60.6% | 1.046 |
that the ratio of a fault occurring in control item AC-2-3—when fault occurs in control items AC2-2 and OS-2-2 "confidence"—is 100%. In addition, since the lift of this rule is over 1, the correlation of the association rule seems strong rule.

4. Results and Discussion

We have performed the process of figuring out the association rule that has the predefined support and confidence threshold value, in order to carry out relation analysis among K-ISMS faults. The first process is “frequent itemset finding,” which finds the itemset that satisfies the support threshold value only. The other process is “association rule generation” that adopts the rule satisfying the confidence threshold value only among association rules, which were created from the found frequent itemsets. Table 3 shows the summary of strong association rules within the range of the minimum support and minimum confidence.

However, all strong association rules are not always useful. The support-confidence framework can induce a rule $X \rightarrow Y$ as an interesting rule, even though $X$ occurrence actually does not affect $Y$ occurrence. To solve this issue, interest analysis is required that indicates the level of rules’ correlation. This paper selects 43 association rules by applying the following three conditions (see Table 4):

1. minimum support $> 30\%$ and minimum confidence $> 50\%$ and lift $> 1$,
2. minimum support $> 20\%$ and minimum confidence $> 30\%$ and lift $> 1$,
3. minimum support $> 10\%$ and minimum confidence $> 80\%$ and lift $> 1$.

From the result of analysis, we can forecast that the association of fault occurrence among control items is high, as the rule \{CL-1-1 (Information asset identification) $\Rightarrow$ OS-2-2 (Security system operation)\} is the one that has over 30% support level and 50% confidence level. The “information asset identification” is a control item that should classify and identify all information assets of the organization. A fault occurs, if those information assets are not identified periodically, or some assets are skipped. In other words, if a fault is found in “information asset identification,” there is a high possibility of fault occurrence in "security system operation.” The "security system operation" fault refers to the case that the security system operating procedure is not complied, or the blocking rule management log of the security system (e.g., firewall) is not available or lost.

| Table 3: Summary of the association rules found by interesting measure (support, confidence, and lift). |
|---|
| Support | Confidence | Number of rules | Lift $>1$ |
|---|---|---|---|
| 10% | 302 | 258 |
| 20% | 300 | 252 |
| 30% | 287 | 246 |
| 40% | 233 | 212 |
| 50% | 160 | 152 |
| 60% | 98 | 98 |
| 70% | 30 | 30 |
| 80% | 11 | 11 |
| 90% | 4 | 4 |
| 100% | 2 | 2 |
| 10% | 21 | 15 |
| 20% | 21 | 15 |
| 30% | 21 | 15 |
| 40% | 18 | 13 |
| 50% | 13 | 11 |
| 60% | 7 | 7 |
| 70% | 2 | 2 |
| 80% | 0 | 0 |
| 90% | 0 | 0 |
| 100% | 0 | 0 |
| 10% | 2 | 1 |
| 20% | 2 | 1 |
| 30% | 2 | 1 |
| 40% | 2 | 1 |
| 50% | 2 | 1 |
| 60% | 0 | 0 |
| 70% | 0 | 0 |
| 80% | 0 | 0 |
| 90% | 0 | 0 |
| 100% | 0 | 0 |

There were two rules that have over 20% support level and 70% confidence level. The first rule was \{PH-1-4 (Access control of physical area) $\Rightarrow$ AC-3-3 (User password management)\} and the second rule was \{AC-4-6 (access control of internet connection) $\Rightarrow$ CL-1-1 (Information asset identification)\}. The first rule—"Access control of physical” control item, refers to the requirement that only the authorized person should be allowed to access the major systems inside the security area, and the access log should be reviewed periodically. A fault occurs, if access control of the outsiders is not sufficient, or
Table 4: The result of associated relationship among K-ISMS faults (interest \( \geq 1 \)).

| Probability by support, confidence | Association rules | Associated relationship among K-ISMS faults |
|-----------------------------------|-----------------|------------------------------------------|
| **Support > 30%**                 | \((C)CL-1-1 \Rightarrow (C)OS-2-2\) | If a fault occurs in “Information asset identification” ⇒ there is a probability of a fault concurrently in “System and service operation security” |
| **Confidence > 70%**              | \((C)PH-1-4 \Rightarrow (C)AC-3-3\) | If a fault occurs in “Access control of physical area” ⇒ there is a probability of a fault concurrently “User password management” |
|                                  | \((C)AC-4-6 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Access control of internet connection” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)CC-2-1 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Encryption key generation and use” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)CC-2-1 \Rightarrow (C)AC-3-3\) | If a fault occurs in “Encryption key generation and use” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | \((C)CC-1-1 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Encryption policy establishment” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)AC-2-3 \Rightarrow (C)AC-3-3\) | If a fault occurs in “Access permission review” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | \((C)OS-6-2 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Logging and archiving” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
| **Confidence > 60%**              | \((C)PH-1-4 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Access control of physical area” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)OS-6-2 \Rightarrow (C)OS-2-2\) | If a fault occurs in “Logging and archiving” ⇒ there is a probability of a fault concurrently in “Performance and capacity management” |
|                                  | \((C)AC-2-3 \Rightarrow (C)OS-2-2\) | If a fault occurs in “Access permission review” ⇒ there is a probability of a fault concurrently in “Performance and capacity management” |
|                                  | \((C)AC-3-3 \Rightarrow (C)AC-4-6\) | If a fault occurs in “User password management” ⇒ there is a probability of a fault concurrently in “Access control of internet connection” |
| **Confidence > 50%**              | \((C)AC-3-3 \Rightarrow (C)AC-4-6\) | If a fault occurs in “User password management” ⇒ there is a probability of a fault concurrently in “Network access control” |
|                                  | \((C)OS-4-6 \Rightarrow (C)CL-1-1\) | If a fault occurs in “User password management” ⇒ there is a probability of a fault concurrently in “Training and evaluation” |
| **Confidence > 40%**              | \((C)CC-2-1, (C)AC-2-3 \Rightarrow (C)AC-3-3\) | If two faults occur in “Encryption key generation and use” and “Recovery measure establishment based on impact analysis” ⇒ there is a probability of a fault concurrently in “Access permission review” |
|                                  | \((C)OS-2-7 \Rightarrow (C)CL-1-1\) | If two faults occur in “Administrator and special permission management” and “Performance and capacity management” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
| **Confidence > 30%**              | \((C)CC-2-1, (C)AC-2-3 \Rightarrow (C)AC-3-3\) | If two faults occur in “Access control of internet connection” and “Performance and capacity management” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)OS-4-1 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Information system storage media management” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)DR-2-1, (C)OS-2-2 \Rightarrow (C)CL-1-1\) | If two faults occur in “Recovery measure establishment based on impact analysis” and “Performance and capacity management” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | \((C)OS-2-7 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Open server security” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
| **Confidence > 20%**              | \((C)AC-3-3 \Rightarrow (C)ET-2-1\) | If a fault occurs in “Logging and archiving” ⇒ there is a probability of a fault concurrently in “Training and evaluation” |
|                                  | \((C)AC-3-3 \Rightarrow (C)ET-2-1\) | If a fault occurs in “User password management” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | \((C)AC-3-3 \Rightarrow (C)ET-2-1\) | If a fault occurs in “Server access control” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | \((C)OS-6-2 \Rightarrow (C)CL-1-1\) | If two faults occur in “Security grade and handling” and “requirement of security in contract with external parties” ⇒ there is a probability of a fault concurrently in “Logging and archiving” |
|                                  | \((C)OS-5-2 \Rightarrow (C)OS-2-2\) | If two faults occur in “Information asset identification” and “Patch management” ⇒ there is a probability of a fault concurrently in “Performance and capacity management” |
|                                  | \((C)IH-1-1 \Rightarrow (C)CL-1-1\) | If a fault occurs in “Intrusion incident response procedure establishment” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
### Table 4: Continued.

| Probability by support, confidence | Association rules | Associated relationship among K-ISMS faults |
|-----------------------------------|-------------------|-------------------------------------------|
| **Confidence > 70%**              |                   |                                           |
|                                   | (C)AC-2-3, (C)PH-1-4-1 ⇒ [(C)CC-2-1] | If two faults occur in “Access permission review” and “Access control of physical area” ⇒ there is a probability of a fault concurrently in “Encryption key generation and use” |
|                                   | (C)CL-1-1, (C)ET-2-1 ⇒ [(C)OS-2-2] | If two faults occur in “Information asset identification” and “Training and evaluation” ⇒ there is a probability of a fault concurrently in “Performance and capacity management” |
|                                   | (C)DS-1-1 ⇒ [(C)CL-1-1] | If a fault occurs in “Definition of security requirement during system development” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                   | [(C)AC-2-3 (M)RM-2] ⇒ [(C)CC-2-1] | If two faults occur in “Access permission review” and “(M) Risk identification and evaluation” ⇒ there is a probability of a fault concurrently in “Encryption key generation and use” |
|                                   | [(C)CC-1-1 (C)OS-6-2] ⇒ [(C)CL-1-1] | If two faults occur in “Encryption policy establishment” and “Logging and archiving” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                   | [(C)CC-1-1, (C)OS-2-2] ⇒ [(C)CL-1-1] | If two faults occur in “Encryption policy establishment” and “Performance and capacity management” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                   | {(C)PH-3-1} ⇒ [(C)CL-1-1] | If a fault occurs in “Personal business environment security” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                   | (C)AC-4-4 ⇒ [(C)AC-3-3] | If a fault occurs in “Database access control” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                   | (C)AC-2-3, (C)CC-1-1 ⇒ [(C)CC-2-1] | If two faults occur in “Access permission review” and “Encryption policy establishment” ⇒ there is a probability of a fault concurrently in “Encryption key generation and use” |
|                                   | (C)CL-1-1, (C)OS-4-2 ⇒ [(C)OS-2-2] | If two faults occur in “Information asset identification” and “Portable storage media management” ⇒ there is a probability of a fault concurrently in “Performance and capacity management” |
|                                   | (C)AC-2-3 ⇒ [(C)AC-3-3] | If two faults occur in “Access permission review” and “Encryption policy establishment” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                   | (C)CC-1-1 ⇒ [(C)AC-3-3] | If two faults occur in “Patch management” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                   | (M)RM-2 ⇒ [(C)OS-6-3] | If two faults occur in “User password management” and “(M) Risk identification and evaluation” ⇒ there is a probability of a fault concurrently in “Access and use monitoring” |
|                                   | (C)SO-1-3 ⇒ [(C)AC-3-3] | If a fault occurs in “Information security committee” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                   | (M)PM-3 ⇒ [(C)AC-3-3] | If a fault occurs in “Internal audit” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                   | (M)PM-3 ⇒ [(C)CL-1-1] | If a fault occurs in “Internal audit” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                   | (C)OS-6-3 ⇒ [(C)AC-3-3] | If a fault occurs in “Access and use monitoring” ⇒ there is a probability of a fault concurrently in “User password management” |

| Support > 10%                    |                   |                                           |
|----------------------------------|-------------------|-------------------------------------------|
|                                  | {(C)CL-1-1} ⇒ [(C)OS-2-2] | If two faults occur in “Access permission review” and “Access control of physical area” ⇒ there is a probability of a fault concurrently in “Encryption key generation and use” |
|                                  | {(C)DS-1-1} ⇒ [(C)CL-1-1] | If a fault occurs in “Definition of security requirement during system development” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | [(C)AC-2-3] ⇒ [(C)CC-2-1] | If two faults occur in “Access permission review” and “(M) Risk identification and evaluation” ⇒ there is a probability of a fault concurrently in “Encryption key generation and use” |
|                                  | [(C)CC-1-1] ⇒ [(C)CL-1-1] | If two faults occur in “Encryption policy establishment” and “Logging and archiving” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | [(C)CC-1-1, (C)OS-2-2] ⇒ [(C)CL-1-1] | If two faults occur in “Encryption policy establishment” and “Performance and capacity management” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | {(C)PH-3-1} ⇒ [(C)CL-1-1] | If a fault occurs in “Personal business environment security” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | (C)AC-4-4 ⇒ [(C)AC-3-3] | If a fault occurs in “Database access control” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | (C)AC-2-3, (C)CC-1-1 ⇒ [(C)CC-2-1] | If two faults occur in “Access permission review” and “Encryption policy establishment” ⇒ there is a probability of a fault concurrently in “Encryption key generation and use” |
|                                  | (C)CL-1-1, (C)OS-4-2 ⇒ [(C)OS-2-2] | If two faults occur in “Information asset identification” and “Portable storage media management” ⇒ there is a probability of a fault concurrently in “Performance and capacity management” |
|                                  | (C)AC-2-3 ⇒ [(C)AC-3-3] | If two faults occur in “Access permission review” and “Encryption policy establishment” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | (C)CC-1-1 ⇒ [(C)AC-3-3] | If two faults occur in “Patch management” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | (M)RM-2 ⇒ [(C)OS-6-3] | If two faults occur in “User password management” and “(M) Risk identification and evaluation” ⇒ there is a probability of a fault concurrently in “Access and use monitoring” |
|                                  | (C)SO-1-3 ⇒ [(C)AC-3-3] | If a fault occurs in “Information security committee” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | (M)PM-3 ⇒ [(C)AC-3-3] | If a fault occurs in “Internal audit” ⇒ there is a probability of a fault concurrently in “User password management” |
|                                  | (M)PM-3 ⇒ [(C)CL-1-1] | If a fault occurs in “Internal audit” ⇒ there is a probability of a fault concurrently in “Information asset identification” |
|                                  | (C)OS-6-3 ⇒ [(C)AC-3-3] | If a fault occurs in “Access and use monitoring” ⇒ there is a probability of a fault concurrently in “User password management” |
**Table 5:** The abbreviations of K-ISMS certification criteria (104 control activities).

| Management process | Information security countermeasures (92 control activities) | Abbreviations |
|--------------------|---------------------------------------------------------------|---------------|
| Establishment of IS polices and setting ISMS scope | Establishment of information security polices | (M)EP-1, (M)EP-2 |
| Management responsibility and information security organization | Participation of the management | (M)MR-1, (M)MR-2 |
| Risk management | Setting up the risk management method and plan | (M)RM-1, (M)RM-2, (M)RM-3 |
| Implementation of information security countermeasures | Selecting information security countermeasures and establishing an implementation plan | (M)IC-1, (M)IC-2 |
| Postmanagement process | Effective implementation of information security measures | (M)PM-1, (M)PM-2, (M)PM-3 |
| Information security policies | Policy approval and announcement | Policy approval, Policy announcement, Connectivity to the upper-level policy, Policy review, Policy document management |
| Policy system | Writing a policy document, Policy review, Policy document management |
| Policy maintenance | Policy review, Policy document management |
| Information security organization | Designating the CIO, Composition of the working organization, Information security committee, Roles and responsibilities | (C)SO-1-1, (C)SO-1-2, (C)SO-1-3, (C)SO-1-4 |
| Roles and responsibilities | Information security committee, Roles and responsibilities |
| Security of external parties | Requirement of security in contract with external parties, Security when the outsider’s contract is expired | (C)EP-1, (C)EP-2-1, (C)EP-2-2 |
| Information asset classification | Information asset identification and responsibilities, Information asset classification and handling | Information asset identification, Responsibility allocation by information asset, Security grade and handling |
| Education and training on information security | Setting up training programs, Training and evaluation | Training plan, Training target, Training details and methods, Training implementation and evaluation |
| Personal security | Information security responsibility | Designating and supervising main workers, Job separation, Nondisclosure agreement, Resignation and job change management, Reward and punishment regulations |
| Personnel policy | | |
| Physical security | Physical security area | Designating a security area, Protecting facility, Work within the security area, Access control, Mobile device taking in/out, Cable security |
| System protection | System arrangement and management |
| Office security | | Shared business environment security |
Table 5: Continued.

| Management process | Information security countermeasures (92 control activities) | Abbreviations |
|--------------------|-------------------------------------------------------------|---------------|
| Security management during analysis and design for system | Security requirement definition (C) DS-1-1 | |
| | Authentication and encryption function (C) DS-1-2 | |
| | Security log function (C) DS-1-3 | |
| | Access control function (C) DS-1-4 | |
| | Implementation and test (C) DS-2-1 | |
| System development security | Separation of the development and operating environment (C) DS-2-2 | |
| | Separation of the operating environment (C) DS-2-3 | |
| | Test data security (C) DS-2-4 | |
| | Source program security (C) DS-2-5 | |
| Implementation and transfer security | Outsourced development security (C) DS-3-1 | |
| Outsourced R&D security | | |
| Cryptography security | Encryption policy establishment (C) CC-1-1 | |
| | Encryption key management (C) CC-2-1 | |
| Access control policy | Access control policy establishment (C) AC-1-1 | |
| Access control management | User registration and authorization (C) AC-2-1 | |
| | Administrator permission management (C) AC-2-2 | |
| | Access permission review (C) AC-2-3 | |
| Access control and identification | User authentication (C) AC-3-1 | |
| | User identification (C) AC-3-2 | |
| | User password management (C) AC-3-3 | |
| Access control area | Network access (C) AC-4-1 | |
| | Server access (C) AC-4-2 | |
| | Application access (C) AC-4-3 | |
| | Database access (C) AC-4-4 | |
| | Mobile device access (C) AC-4-5 | |
| | Internet connection (C) AC-4-6 | |
| Operating procedure and change control | Operating procedure establishment (C) OS-1-1 | |
| System and service operation security | Change control (C) OS-1-2 | |
| | Information system acceptance (C) OS-1-3 | |
| | Security system operation (C) OS-2-1 | |
| | Performance and capacity management (C) OS-2-2 | |
| | Fault management (C) OS-2-3 | |
| Operations security | Remote operation management (C) OS-2-4 | |
| | Smart work security (C) OS-2-5 | |
| | Wireless network security (C) OS-2-6 | |
| | Open server security (C) OS-2-7 | |
| | Backup management (C) OS-2-8 | |
| | Vulnerability check (C) OS-2-9 | |
| | e-Commerce security (C) OS-3-1 | |
| | Information transmission policy establishment and contract conclusion (C) OS-3-2 | |
| Media security | Information system storage media management (C) OS-4-1 | |
| Malicious code management | Portable storage media management (C) OS-4-2 | |
| | Malicious code control (C) OS-5-1 | |
| | Patch management (C) OS-5-2 | |
| | Time synchronization (C) OS-6-1 | |
| Log management and monitoring | Logging and archiving (C) OS-6-2 | |
| | Access and use monitoring (C) OS-6-3 | |
| | Intrusion attempt monitoring (C) OS-6-4 | |
| Intrusion incident handing | Incident response procedure establishment (C) IH-1-1 | |
| | Incident response procedure establishment (C) IH-1-2 | |
| | Intrusion incident training (C) IH-2-1 | |
| | Intrusion incident reporting (C) IH-2-2 | |
| | Intrusion incident handling and recovery (C) IH-2-3 | |
| | Intrusion incident analysis and sharing (C) IH-2-4 | |
| Postmanagement process | Recurrence prevention (C) IH-3-1 | |
| | | (C) IH-3-2 | |
| IT disaster recovery planning | System development (C) DRR-1-1 | |
| Countermeasures implementation | Recovery measure establishment based on impact analysis (C) DRR-2-1 | |
| | Test and maintenance (C) DRR-2-2 | |
the mobile device (e.g., USB) can be brought to it. In addition, the fault of the “user password management” control item occurs when the password of major systems is not changed periodically, or the password use requirements are not met. Therefore, if the enterprise does not perform proper “access control” on major facilities and systems, there is a possibility that the “user password management” of major systems (e.g., server, network) can also be insufficient.

5. Conclusions

In this study, we used the association mining applied with the “Apriori” algorithm in order to analyze the correlation among K-ISMS faults and could find 43 association rules. The result of this study suggests having a high correlation among faults as if the organization identifies and manages their information asset carelessly, then it can also affect security system operation.

Therefore, the result of those association rule may be referred to as the useful information for decision-making of organization’s security activities and can be be utilized as a guide to the assessment method during K-ISMS certification assessment. If any fault occurs among K-ISMS certification criteria, those items related to the association rule can be checked intensively. Also, it can be a guidance of analyzing the level of the Plan-Do-Check-Act activity (organization’s security management phase) from the perspective of the correlation among faults.

However, finding a useful rule can be different according to the size of the data set because the adoption of the useful association rule depends on the occurrence frequency of the analysis data. Therefore, we need various studies of K-ISMS fault analysis such as association in accordance with the scope of organization’s certification (a number of employees and system). Based on the association rule results obtained in this study, decision-making tree analysis to forecast the fault status, and fault factor analysis using the structural equation model will be studied as a subsequent study.

Because the paradigm of IT environment is changing from conventional PC and mobile to IoT, new approach is needed in terms of range of protection targets, characteristics of protection targets, and protection subjects to strengthen the security level of IoT continuously. In other words, the protection target should be expanded from the existing PC and mobile devices to all objects such as home appliances, automobiles, and medical supplies. It is also necessary to break away from the conventional method of protection with separate security system and software implementation and interface to establish the security policy, procedure, and standard to control and manage efficiently the administrative security, physical security, and technical security.

Moreover, there is a need for the application of information security management system suitable for the IoT environment to maintain the continuous security level of IoT services and prevent the spread of risk of intrusion incidents, including the identification of key assets to be protected and threats as well as assessment of the current security level to establish policies for coping with threats.

Appendix

See Table 5.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1004) supervised by the NIPA (National IT Industry Promotion Agency).

References

[1] Ministry of Science; ICT and Future Planning (MSIP), “The security roadmap of internet of things,” Report of MSIP, 2014.
[2] S. Park and K. Lee, “Advanced approach to information security management system model for industrial control system,” The Scientific World Journal, vol. 2014, Article ID 348305, 13 pages, 2014.
[3] Korea Information Security Agency, “The guide of information security management system in Korea,” 2014, http://isms.kisa.or.kr.
[4] D. Lee, “Introduction of ISMS certification system,” a presentation material in KISA, 2013.
[5] C. Park, S. Jang, and Y. Park, “A study of effect of ISMS certification on organization performance,” International Journal of Computer Science and Network Security, vol. 10, no. 3, pp. 10–20, 2010.
[6] K. Kim and S. Kim, “Evaluation criteria for Korean smart grid based on K-ISMS,” Journal of Korea Institute of Information Security & Cryptology, vol. 22, no. 6, pp. 1375–1391, 2012.
[7] K. Haufe, S. Dzombieska, and K. Brandis, “Proposal for a security management in cloud computing for health care,” The Scientific World Journal, vol. 2014, Article ID 146970, 7 pages, 2014.
[8] Q. L. Hai, A.-I. Somjit, and A.-I. Ngamnij, “Association rule hiding based on intersection lattice,” Mathematical Problems in Engineering, vol. 2013, Article ID 210405, 11 pages, 2013.
[9] N. M. Khairudin, A. Mustapha, and M. H. Ahmad, “Effect of temporal relationships in associative rule mining for web log data,” The Scientific World Journal, vol. 2014, Article ID 813983, 10 pages, 2014.
[10] R. kenett and S. Salini, “Relative linkage disequilibrium: a new measure for association rules,” in Advances in Data Mining, Medical Applications, E-Commerce, Marketing, and Theoretical Aspects, vol. 5077 of Lecture Notes in Computer Science, pp. 189–199, Springer, Berlin, Germany, 2008.
[11] R. Agrawal, T. Imielinski, and A. Swami, “Mining association rules between sets of items in large databases,” in Proceedings of the ACM SIGMOD International Conference on Management of Data, pp. 207–216, May 1993.
[12] I. N. Shaharanee, T. S. Dillon, and F. Hadzic, “Ascertainment data mining rules using statistical approaches,” in Proceedings of International Symposium on Computing, Communication, and Control (ISCCC ’11), vol. 1, pp. 182–183, IACSIT Press, 2011.
[13] T. Brijs, K. Vanhoof, and G. Wets, "Defining interestingness for association rule," *International Journal Information Theories & Applications*, vol. 10, pp. 370–372, 2010.

[14] V. Jaiswal and J. Agarwal, "The evolution of the association rules," *International Journal of Modeling and Optimization*, vol. 2, no. 6, pp. 726–729, 2012.

[15] C. S. Kanimozhi Selvi and A. Tamilarasi, "An automated association rule mining technique with cumulative support thresholds," *International Journal of Open Problems in Computer Science and Mathematics*, vol. 2, no. 3, pp. 427–438, 2009.

[16] M. J. Zaki, "Scalable algorithms for association mining," *IEEE Transactions on Knowledge and Data Engineering*, vol. 12, no. 3, pp. 372–390, 2000.

[17] J. Hipp, U. Güntzer, and G. Nakhaeizadeh, "Algorithms for association rule mining—a general survey and comparison," in *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining*, vol. 2, pp. 58–59, SIGKDD Explorations, July 2000.

[18] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," in *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB ’94)*, pp. 487–499, September 1994.

[19] M. Hahsler, B. Grün, and K. Hornik, "Arules—a computational environment for mining association rules and frequent item sets," *Journal of Statistical Software*, vol. 14, no. 15, pp. 11–22, 2005.

[20] B. Liu, W. Hsu, and Y. Ma, "Mining association rules with multiple minimum supports," in *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD ’99)*, pp. 337–341, San Diego, Calif, USA, August 1999.

[21] M. Hahsler, *A Probabilistic Approach to Association Rule Mining*, Southern Methodist University, Dallas, Tex, USA, 2008.

[22] A. Wright, E. S. Chen, and F. L. Maloney, "An automated technique for identifying associations between medications, laboratory results and problems," *Journal of Biomedical Informatics*, vol. 43, no. 6, pp. 891–901, 2010.

[23] P.-N. Tan, V. Kumar, and J. Srivastava, "Selecting the right interestingness measure for association patterns," in *Proceedings of the 8th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD ’02)*, pp. 32–41, July 2002.

[24] T. Pang-Ning, M. Steinbach, and V. Kumar, "Association analysis, basic concepts and algorithms," in *Introduction to Data Mining*, chapter 6, pp. 327–414, Addison-Wesley, 2005.

[25] M. Hahsler and S. Chelluboina, "Visualizing Association Rules: Introduction to the R-extension Package arulesViz," The Comprehensive R Archive Network, http://cran.r-project.org/web/packages/arulesViz.
