ICIC Express Letters Volume 4, Number 5, October 2010 ICIC International ©2010 ISSN 1881-803X

pp. 1-6

A SECURITY LATTICE MODEL AND SIMULATION OF DATA PRIVACY IN **INA-CBG APPLICATION IN BADAN PENYELENGGARA JAMINAN SOSIAL** (BPJS)

Irvan Santoso¹, Ayu Hidayah Aslamiah² ¹Cyber Security Program, Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480 isantoso@binus.edu

> ²Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480 aaslamia@binus.edu

Received November 2019; accepted April 2010

ABSTRACT.

Data privacy is something that must be stored securely especially in a government system. The Indonesian government has introduced a universal coverage system that adopts Diagnosis Related Group (DRG), namely Indonesia Case Base Group (INA-CBG). INA-CBG is used to improve the service of healthcare providers and integrate user data into one database. Hence, there is a possibility of security gaps in protecting data from unauthorized users. Therefore, an analysis was performed based on the regulation published in Indonesia and the current INA-CBG system. A security lattice model of access control was produced to set constraints to prevent data privacy from being retrieved by unauthorized users. There are five security level categorizations, as follows Public, Research, Clinical, Financial, and Provider. Each security level was assigned to 26 attributes which are components of the current INA-CBG system. Furthermore, a simulation has been performed based on the model developed to ensure that the assigned level is correct. The results of the simulation will be used as a reference in determining the security lattice of each attribute if the result shows there are inappropriate levels.

Keywords: Security Level, Security Lattice, INA-CBG, DRG, Universal Coverage System.

1. Introduction.

The internet revolution has changed the world communication infrastructure with a lot of data exchange including data privacy [1]. Data privacy is something that must be kept confidential [2]. To maintain the data, we need a system and the right knowledge to store and determine which data can be accessed by people who have the authority [3]. Inevitably, this becomes a difficulty for system developers and users who use the system. The more secure a system makes the more difficult a system is used by its users. Developers must consider the extent to which they must create a system that is secure in preserving their data and is easy to use by users [4].

One system that needs serious consideration in its development is a government system that stores public data [5]. The government system must be very secure from all existing attacks from parties who do not have the authority. The Indonesian government has a Badan Penyelenggara Jaminan Sosial (BPJS) which provides a universal coverage system under the name Indonesia Case Base Group (INA-CBG) [6] which is regulated in [7]. INA-CBG adopted the Diagnosis Related Group (DRG) which grouped patients based on diagnoses and similar procedures to improve health service providers by changing the payment system [8]. Patients who are members and want to use the service must input their data into the system. Even though sensitive data has been regulated in [9] to ensure that data cannot be retrieved directly without complying with certain policies, the INA-CBG system integrates patient data into one database which makes it hard to ensure that there is no violation for privacy. There is a possibility of fraud such as aggregated patient illnesses that can be leaked to third parties such as pharmacies. If they know which area with illnesses that often occur, they can increase the medicine price in a particular area. This will be detrimental to people with illnesses without insurance. Besides, BPJS also provides a hospital information system named Sistem Informasi Manajemen Rumah Sakit (SIMRS) that regulated in [10] which has a bridging component to INA-CBG system.

Based on the problem, this research proposes a model of privacy assurance based on access control and partial order set which regulates confidentiality in a universal coverage system. This model will be used as a reference to make constraints. There are 63 constraints have been used to prevent violation transaction consisting of basic constraints, inference constraints, association constraints, and upper bound constraints. The developed constraints are made based on attributes in the healthcare provider. Then, secure information flow is applied to ensure information only can be read by authorized parties. Moreover, to prove the correctness of each constraint, algorithms and simulation using programming language will be performed to check whether the constraints have the right security level or vice versa.

2. Partial Order Set.

A binary relation can be said as partial ordering relation if the characteristics of that relation are reflexive, antisymmetric, and transitive [11]. Furthermore, it can be said as partial ordering relation if there are at least two items and one is larger or smaller. A Partial Order Set (Poset) can be described in the diagram, namely Hasse Diagram. Hasse Diagram has a pattern of tiers to the top.

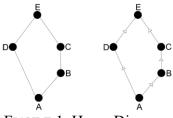


FIGURE 1. Hasse Diagram

Figure 1 shows an example of a Hasse Diagram. A is the lowest level and E is the highest level in the given diagram. A higher level may access the item below it, but the lower level cannot access the item above it. Furthermore, the Hasse Diagram has upper bounds and lower bounds for each node. A node is a Least Upper Bound (LUB) if it is upper bound for all nodes below it and there is no other upper bound. A node is a Greatest Lower Bound (GLB) if it is a lower bound of all nodes below it and there is no other lower bound. A lattice is a Poset (L, \geq) that has LUB and GLB.

3. Basic Definitions.

Access class AC is made to prevent violations in access control. AC is assumed as a pair of the form a security level L and a set of attributes A. The relation between that pair is limited by a Poset, namely dominance relation \geq . The \geq shows information which can be seen is information that has the same level as the subject or below. For instance, the expression $A \geq$ b means A dominates b. A may read the information on the level b or below. Besides, the partial order set (AC, \geq) is assumed as a classification lattice LT. Therefore, LT which will be used is partial order set (L, \geq) to specify A in LT may have dynamic security levels L. Furthermore, classification constraints C specify security levels on attributes based on confidentiality. A constraint c consists pair of *lhs* and *rhs*, where $c \in C$. In *lhs* and *rhs* can be a security level L or a mapping $\lambda: A \to L$, where (L, \geq) is a classification lattice. Constraints will be classified as basic constraints, inference constraints, association constraints, and upper bound constraints.

Basic constraints C_{basic} are constraints that specify the security level on each attribute. In C_{basic} , AC has a mapping of attribute $\lambda(A)$ and a security level L. Furthermore, C_{basic} map an attribute to a specific level, $\lambda(A) \ge L$, when $a \in A$ and $l \in L$. It can be assumed that attribute A has a level L. Inference constraints $C_{inference}$ are made to prevent bypassing of C_{basic} through data inference. Afterward, $C_{inference}$ maps attributes, $\lambda(a_x) \ge \lambda(a_y)$, when $a_x \in A$ and $a_y \in A$. It can be assumed that if the attribute in a_x is known, then the attribute in a_v can also be known. Association constraints $C_{association}$ are made to limit the information that may be known by a combination of attributes. It requires the least upper bound of the classifications to dominate the security level given by basic constraints. For instance, $lub\{\lambda(a_x), \lambda(a_y)\} \ge L$, where $a_x \in A$, $a_y \in A$, and $l \in L$. That constraint requires *lub* of the classifications between $\lambda(a_x)$ and $\lambda(a_y)$ to dominate a security level. Upper bound constraints C_{upper} are made to guarantee an attribute is always accessible to a specific level. These constraints prevent the classification of attributes being raised above certain levels. The mapping of C_{upper} is performed by setting a specific level and a specific attribute, $L \ge \lambda(A)$, where $a \in A$ and $l \in L$. In addition, upper bound constraints can indirectly affect other attributes to have the same security level. Base case constraints C_{base} will be created for induction method evaluation. In this constraint, all attributes will be assigned as the highest security level.

4. Constructed Model.

A new model is shown in Figure 2 by looking at the information flow process. The constructed model has a security lattice LT which will describe the relation of each attribute A from the lowest level Public to the highest level, Provider, where $LT = (L, \geq)$. Due to this

lattice, each attribute *a* will not have a static security level but may have a dynamic security level *l*.



FIGURE 2. The Lattice of Constructed Model

In the lowest level Public, some variables can be known by all of the people. Then, in the Research, some variables cannot be known by all of the people and should be given a higher level than Public. In addition, in Clinical, some variables can be revealed by combining the variables between the variable in Public and variable in Research. Then, at the highest level, Provider can read all of the variables. This level is usually owned by the government so that they can assess each patient.

5. Evaluation Method.

Attributes that will be submitted to the system consist of 26 attributes. The model uses these attributes which must be assigned to a certain security level. This assignment must comply with certain classification constraints. Therefore, this work solves the problem of lattice assignment which formulated as follows:

Problem 1 (Lattice-Assignment) given A, $LT = (L, \geq)$, C_{basic} , C_{base} , $C_{inference}$, $C_{association}$, and C_{upper} that satisfy λ : $A \geq L$.

To proof the correctness of our model, this work will use the induction method as an evaluation method.

Theorem 1 (Correctness) the algorithms in [12] solve **Problem 1 (Lattice-Assignment)**. Given A, $LT = (L, \geq)$, C_{basic} , C_{base} , $C_{inference}$, $C_{association}$, and C_{upper} that satisfy C or terminates with Failure if C are inconsistent.

Lemma 1 (Correctness) shows the correctness of algorithms by proofing the consistent of the constructed model and the termination of algorithms.

Proof: Let *c* be a set of classification constraints over a set of attributes and a classification lattice.

- A. If the result of **computeUpperBound** is *Failure*, then the sequence of constraints is inconsistent.
- B. If the result of **computeUpperBound** is *Success*, then the checked classification of mapping $\lambda: A \to L$ satisfies $c(\lambda \models c)$, then the sequence of constraints is consistent.
- C. computeUpperBound process always terminates.

To prove (A): for all mappings λ not dominate λ' , there exists a constraint $c \in c$ such that λ'

not equal with *c*. In addition, in the process of **computeUpperBound**, it is not possible to change the levels of the variable to higher.

To prove (B): mapping λ satisfies a sequence of constraints c by induction method. If mapping λ satisfies all immediate constraints on attributes in all constraints c, then $\lambda \ge \lambda'$, there exists a constraint $c \ge c$ such that $\lambda' \models c$.

To prove (C): **computeUpperBound** is arranged in a finite set based on the constraints. Each constraint will be called and checked by several functions. Each function will return a value in the process of checking, namely Success or Failure. If the result is Failure, then the process will be terminated. If the result is Success, then the process will be continued until the end of the program which will produce upper bounds.

6. Classified Constraints.

6.1 Basic Constraints.

Basic constraints assign security levels for each attribute. The purpose of basic constraints is to give basic security level and ascertain which attribute has a higher level in confidentiality. In our case, there are 26 basic constraints according to Table 1.

TABLE 1. Dasic Constraints	
C ₁ : λ (Healthcare provider code number) \geq	C ₁₄ : λ (Procedure) \geq Public
Public	
C ₂ : λ (Healthcare provider name) \geq Public	C ₁₅ : λ (Length of stay) \geq Public
C ₃ : λ (Healthcare provider class) \geq Public	C_{16} : λ (Main diagnosis) \geq Research
C ₄ : λ (Dates of discharge) \geq Public	C_{17} : λ (Secondary diagnosis) \geq Research
C ₅ : λ (Age (year)) \geq Public	C ₁₈ : λ (Activity daily living) \geq Research
C ₆ : λ (Age (day)) \geq Public	C ₁₉ : λ (INA-CBG code) \geq Research
C ₇ : λ (Dates of birth) \geq Public	C_{20} : λ (Special CMG) \geq Research
$C_8: \lambda(Gender) \ge Public$	C ₂₁ : λ (INA-CBG description) \geq Public
C ₉ : λ (Treatment class) \geq Public	C ₂₂ : λ (SEP number) \geq Clinical
C_{10} : λ (Dates of admission) \geq Public	C ₂₃ : λ (Medical record number) \geq Clinical
C_{11} : λ (Treatment type) \geq Public	C ₂₄ : λ (Top-up rates) \geq Financial
C_{12} : λ (Patient status) \geq Public	C ₂₅ : λ (Healthcare provider rates) \geq
	Financial
C_{13} : λ (Birth weight) \geq Public	C ₂₆ : λ (Total fare) \geq Financial

TABLE 1. Basic Constraints

Main diagnosis and Secondary diagnosis are assigned higher than Public. This is because the diagnosis of a patient may not be known by anyone other than the patient and healthcare provider. Moreover, all attributes which are related to patient and healthcare provider are assigned as Research. This is because those attributes may not be known generally. Furthermore, SEP number and Medical record number are assigned higher than Research, namely Clinical. This is because if someone knows SEP number or Medical record number, they may know about patient information, such as who and which disease.

6.2 Inference Constraints.

Inference constraints are made to prevent bypassing basic constraints through data inference. In our case, there are 22 inference constraints according to Table 2.

TABLE 2. Inference Constraints

C ₂₇ : λ (Healthcare provider code number) \geq	C ₃₈ : λ (Main diagnosis) $\geq \lambda$ (Procedure)
λ (Healthcare provider name)	
C_{28} : λ (Healthcare provider code number) \geq	C ₃₉ : λ (Secondary diagnosis) \geq
λ (Healthcare provider class)	λ (Procedure)
C ₂₉ : λ (Healthcare provider name) \geq	C_{40} : λ (Special CMG) $\geq \lambda$ (Procedure)
λ (Healthcare provider code number)	
C_{30} : λ (Healthcare provider name) \geq	C_{41} : λ (INA-CBG code) $\geq \lambda$ (Procedure)
λ (Healthcare provider class)	
C_{31} : λ (Length of stay) $\geq \lambda$ (Treatment type)	C ₄₂ : λ (INA-CBG description) \geq
	λ (Procedure)
C ₃₂ : λ (SEP number) $\geq \lambda$ (Gender)	C ₄₃ : λ (INA-CBG code) $\geq \lambda$ (Main
	diagnosis)
C ₃₃ : λ (SEP number) $\geq \lambda$ (Dates of birth)	C ₄₄ : λ (INA-CBG code) $\geq \lambda$ (Secondary
	diagnosis)
C ₃₄ : λ (SEP number) $\geq \lambda$ (Treatment class)	C ₄₅ : λ (INA-CBG description) $\geq \lambda$ (Main
	diagnosis)
C ₃₅ : λ (SEP number) $\geq \lambda$ (Treatment type)	C ₄₆ : λ (INA-CBG description) \geq
	λ (Secondary diagnosis)
C_{36} : λ (SEP number) $\geq \lambda$ (Main diagnosis)	C ₄₇ : λ (INA-CBG description) \geq
	λ (INA-CBG code)
C_{37} : λ (SEP number) $\geq \lambda$ (Secondary	C ₄₈ : λ (INA-CBG code) $\geq \lambda$ (INA-CBG
diagnosis)	description)

6.3 Association Constraints.

Association constraints are made to limit the information which may be exposed by a combination of attributes. It requires the least upper bound of the classifications to dominate the security level given by basic constraints. There are 10 association constraints according to Table 3.

TABLE 3. Association Constraints

C ₄₉ : lub{ λ (Healthcare provider code	C ₅₄ : lub{ λ (Healthcare provider name),
number), λ (Main diagnosis)} \geq Clinical	λ (Main diagnosis) $\} \ge$ Clinical
C ₅₀ : $lub \{\lambda(Healthcare provider code $	C ₅₅ : lub{ λ (Healthcare provider name),
number), λ (Secondary diagnosis)} \geq	λ (Secondary diagnosis)} \geq Clinical
Clinical	
C ₅₁ : lub{ λ (Healthcare provider code	C ₅₆ : lub{ λ (Healthcare provider name),
number), λ (INA-CBG code)} \geq Clinical	λ (INA-CBG code)} \geq Clinical
C ₅₂ : lub{ λ (Healthcare provider code	C ₅₇ : lub{ λ (Healthcare provider name),
number), λ (INA-CBG description)} \geq	λ (INA-CBG description) $\} \ge$ Clinical
Clinical	//
C ₅₃ : lub{ λ (Healthcare provider code	C ₅₈ : lub{ λ (Healthcare provider name),
number), λ (Special CMG)} \geq Clinical	λ (Special CMG)} \geq Clinical

6.4 Upper Bound Constraints.

Upper bound constraints are made to ensure the attribute is always accessible to a specific level. These constraints prevent the classification of attributes from being increased. In our case, there are five upper bound constraints according to Table 4.

C59: Clinical $\geq \lambda$ (Main diagnosis)C62: Clinical $\geq \lambda$ (INA-CBG description)C60: Clinical $\geq \lambda$ (Secondary diagnosis)C63: Clinical $\geq \lambda$ (Special CMG)C61: Clinical $\geq \lambda$ (INA-CBG code)C63: Clinical $\geq \lambda$ (Special CMG)

TABLE 4. Upper Bound Constraints

7. Result and Discussion.

A simulation will be performed for the proposed model by running the program which implemented the model. The results obtained are λ (Main diagnosis) = Clinical; λ (Secondary diagnosis) = Clinical; λ (INA-CBG code) = Clinical; λ (INA-CBG description) = Clinical; λ (Special CMG) = Clinical; λ (SEP number) = Clinical; λ (Healthcare provider name) = Provider; λ (Healthcare provider code number) = Provider. Furthermore, those results will be used to evaluate the consistency of results. The objective is to ensure the constructed model already provides consistent constraints.

Moreover, the induction method is applied to prove the correctness of the proposed model. As a base case, each attribute will be assigned as the highest level. Then it should be checked using the induction step. When **computeUpperBound** computes all attributes in *A* have already complied, then the constructed model is correct. If nothing changes, then the model is consistent. Because Provider dominates all attributes, constraints *C* are consistent.

In the base case, Main diagnosis, Secondary diagnosis, INA-CBG code, INA-CBG description, Special CMG are assigned as Provider. However, the level must comply with the C_{upper} . Hence, Main diagnosis, Secondary diagnosis, INA-CBG code, INA-CBG description, Special CMG will be assigned as Clinical. In addition, all related variables will be evaluated.

From the result of **computeUpperBound**, several analyses will be performed to check whether constraints are consistent. First, because these attributes dominate Procedure, then Procedure should be lowered to Clinical. Since Clinical \geq Public, the statement is still true. Second, the analysis will be performed for λ (Main diagnosis) \geq Research and the others. Since in C_{upper} , Main Diagnosis and the others are assigned as Clinical, then the result is Clinical \geq Research. It is still true because of Clinical dominates Research. Third, the analysis will be performed for λ (SEP number) $\geq \lambda$ (Main diagnosis) and λ (SEP number) $\geq \lambda$ (Secondary diagnosis). Since SEP Number is assigned as Clinical still dominates Main diagnosis and Secondary diagnosis that assigned as Clinical means it is still true. Fourth, the association of the following attributes will be checked such as $lub \{\lambda(Healthcare$ Provider code number), λ (Main diagnosis)} \geq Clinical; lub{ λ (Healthcare provider name), λ (INA-CBG code)} \geq Clinical. Since Healthcare provider name and Healthcare provider code number are assigned as Provider, then Main diagnosis and INA-CBG code are assigned as Clinical, the statement is true. This is because of $lub{Provider, Clinical} \ge$ Clinical. Therefore, the constraints C of the proposed model are consistent because no modification is needed.

9. Conclusion.

In the information system, the obligation to maintain users' privacy sometimes is not taken care of properly. This is because developers are only concerned about functionality. The Indonesian government has introduced BPJS system for universal coverage. BPJS can improve healthcare provider service by changing the payment system that adopts INA-CBG. In this work, a model had been proposed to ensure data privacy. The constructed model is created as a security lattice. Simulation has also been implemented using a programming language to check the consistency of classification constraints and to produce complaint upper bounds. The results of this simulation are correct and consistent because there are no changes in determining the security level.

For future development, a dynamic security lattice model will be created by using dynamic taint analysis to adapt to the condition that can be changed by government regulations, every healthcare provider's regulations, or the customer needs. Therefore, it will affect each healthcare provider to have their customized access control in competing and improving their services.

REFERENCES

- [1] Teik-Toe Teoh, Yok-Yen Nguwi, Yuval Elovici, Wai-Loong Ng and Soon-Yao Thiang, Analyst Intuition Inspired Neural Network Based Cyber Security Anomaly Detection, International Journal of Innovative Computing, Information and Control, vol.14, no.1, pp. 379-386, 2018.
- [2] D. Zhang, "Big Data Security and Privacy Protection," Proceedings of the 8th International Conference on Management and Computer Science (ICMCS 2018), 2018.
- [3] M. Jayabalan and T. O'daniel, "Access Control and Privilege Management In Electronic Health Record: A Systematic Literature Review," Journal Of Medical Systems, vol. 40, no. 12, Aug. 2016.
- [4] M. Green and M. Smith, "Developers Are Not The Enemy!: The Need For Usable Security Apis," IEEE security & privacy, vol. 14, no. 5, pp. 40–46, 2016.
- [5] J. Archenaa and E. M. Anita, "A Survey of Big Data Analytics in Healthcare and Government," Procedia Computer Science, vol. 50, pp. 408–413, 2015.
- [6] D. G. Tamtomo and B. Murti, "An Analysis of the Difference between INA-CBG Reimbursementand Medical Cost for Patients with Chronic Renal Disease: An Evidence from Kasih Ibu Hospital, Surakarta," Revitalizing Family Planning Program and Women's Empowerment for the Improvement of Population Well-being and Economic Development, 2018.
- [7] Undang-Undang No 24 Tahun 2011 tentang Badan Penyelenggara Jaminan Sosial. [Online]. Available: https://bpjs-kesehatan.go.id/bpjs/arsip/categories/Ng/undang-undang. [Accessed: 01-Dec-2019].
- [8] Peraturan Menteri Kesehatan Republik Indonesia Nomor 27 Tahun 2014. [Online]. Available: http://www.jkn.kemkes.go.id/attachment/unduhan/PMK No. 27 ttg Juknis Sistem INA CBGs.pdf. [Accessed: 01-Dec-2019].
- [9] Undang-undang Republik Indonesia Nomor 11 Tahun 2008. [Online]. Available: https://www.kpk.go.id/images/pdf/uu pip/UU ITE no 11 Th 2008.pdf. [Accessed: 01-Dec-2019].
- [10] Peraturan Menteri Kesehatan Republik Indonesia Nomor 82 Tahun 2013. [Online]. Available: https://www.kemhan.go.id/itjen/wp-content/uploads/2017/03/bn87-2014.pdf. [Accessed: 01-Dec-2019].
- [11] S. Bistarelli and F. Santini, "A Hasse Diagram for Weighted Sceptical Semantics With a Unique-Status Grounded Semantics," Logic Programming and Nonmonotonic Reasoning Lecture Notes in Computer Science, pp. 49–56, 2017.
- [12] ConstraintAlgorithms. [Online]. Available: https://github.com/raphael4/contraintAlgorithms [Accessed: 01-Dec-2019].

ICIC Express Letters (An International Journal of Research and Surveys)

Author's Warranty and Assignment of Copyright

Dear Contributor(s),

We are pleased that you are contributing a paper to the *ICIC Express Letters -- An International Journal of Research and Surveys* (ICIC-EL). State the complete title and all authors of the contribution, read the agreement, sign and return it to the ICIC-EL Office on the online submission system.

Paper ID: ICICEL-1912-003 Full title of paper: A SECURITY LATTICE MODEL AND SIMULATION OF DATA PRIVACY IN INA-CBG APPLICATION IN BADAN PENYELENGGRA JAMINAN SOSIAL (BPJS)

Author(s) (Full names): Irvan Santoso and Ayu Hidayah Aslamiah

Your submission of this work and our acceptance for publication are contingent upon your warranty of the following conditions which constitute the Agreement between the author(s) and the ICIC International.

You warrant that the above titled contribution is original, has not been published elsewhere or has not been submitted for publication elsewhere. You warrant that it does not contain libelous material. You agree to transfer copyright to the ICIC International upon acceptance for publication. If the article contains material copyrighted by others, the author warrants that written permission has been obtained from copyright holders who agree to transfer copyright to the extent allowable.

Please indicate your agreement with these terms by signing and dating below.

Author(s) Signature:

Date: January 12, 2020

Irvan Santoso

Please sign and return this form promptly

It is necessary to sign and return this Author's Warranty and Assignment of Copyright in order to publish your contribution

Return Copyright form on the submisson system



ICIC Express Letters

An International Journal of Research and Surveys

January 16, 2020

OFFICIAL RECEIPT (No. Receipt-ICICEL-1912-003)

Name: Irvan Santoso

Organization: BINUS University, Indonesia

Page Charges Paid: JPY48,000-

This is to acknowledge the receipt of the Publication Paper Charges for the paper "A Security Lattice Model and Simulation of Data Privacy in INA-CBG Application in Badan Penyelenggra Jaminan Sosial (BPJS)" in ICIC Express Letters – An International Journal of Research and Surveys (ISSN 1881-803X), Volume 14, Number 9 (tentative), September 2020.



ICIC-EL Editorial Office Tokai University, Kumamoto Campus 9-1-1, Toroku, Kumamoto 862-8652, Japan Tel: +81-96-386-2666 Fax: +81-96-381-7956 E-mail: office@icicel.org URL: http://www.icicel.org