AUTONOMY IN INFORMATION FLOW CONTROL

CHANDRIKA BHARDWAJ

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2020

AUTONOMY IN INFORMATION FLOW CONTROL

by

CHANDRIKA BHARDWAJ

Department of Computer Science and Engineering

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2020

Certificate

This is to certify that the dissertation titled **AUTONOMY IN INFORMATION FLOW CONTROL** being submitted by **Ms. CHANDRIKA BHARDWAJ** for the award of **Doctor of Philosophy** in **Computer Science and Engineering** is a record of *bona fide* work carried out by her under my guidance and supervision at the Department of Computer Science and Engineering, Indian Institute of Technology Delhi. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma.

Sanjiva Prasad
Professor
Department of Computer Science and Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016

Acknowledgements

I wish to thank many people who have helped in their own ways, for making this work, and my degree, a reality.

First, I must thank my advisor, Sanjiva Prasad, for nurturing my curiosity and research interests, and providing me with continuous guidance, support, and inspiration. His profound wisdom and brilliant suggestions have guided my research in the right direction many times. If I am a better speaker or computer scientist due to my time at IIT Delhi, it is largely due to him. I hope that someday I approach the clarity of his thoughts and expression. Apart from research, I have learned the importance of enjoying the process, perseverance, and compassion, from him. I owe him thanks for so many things, but most of all, I thank him for his belief in my abilities and having patience with me.

I thank my thesis committee members, Huzur Saran, Ragesh Jaiswal, and Ponnurangam Kumaraguru, for their insightful questions, excellent suggestions, and feedback on this work, which helped me in refining this work. I can't express enough gratitude to S. Arun Kumar and Kolin Paul for making me a better researcher with their words of encouragement, constructive criticism, and feedback. I very much appreciate Sorav Bansal and Subodh Vishnu Sharma for engaging in insightful discussions and creating a lively research environment.

I thank the staff of CSE dept and SIT at IIT Delhi for their help in several administrative matters and their care: Manju Chopra, Rekha Rathee, Hemant Prasad, Shreshth Bagga, Vandana Ahluwalia, Rajesh Kumar, Suresh Mourya, Arun Kumar and M. M. Rathinam.

It has been a pleasure to be a graduate student in the VerTeCS research group, and I would like to thank the group members for creating such a thought-provoking environment. Jatin Batra, Madhulika Mohanty, Prajna Upadhyaya, Divyanshu Bagga, Vijay Kumar, Aruna Bansal, Geeta Yadav, Himanshu Gandhi, Swati Bhamu, Shubhrima Ghosh, and many others enlivened my graduate life as they were willing to share meaningful experiences and dinners, bounce off ideas each other, suggest good books, learn dance and swimming and play ultimate Frisbee.

I am grateful to my grandfathers, Late Pt. Nathan Singh Sharma and Pt. Ram Swaroop Sharma, and the rest of the family including Rajendra Kr Sharma, Pawan Sharma, and Smt. Saroj Sharma for being incredibly supportive of my aspirations for higher education.

My dearest friend and husband, Eshan Nanda, has patiently listened to my ideas, reviewed my drafts, gave some really good suggestions and feedback, and helped in practicing my talks. His optimism, patience, and unconditional love have been crucial in striving excellence in my work.

My adorable siblings, Prerna Bhardwaj and Avneesh Sharma, have been a constant source of blissfulness and solace in my life. Even though they are younger, both of them have inspired, motivated, and helped me their own ways.

I am most thankful to my loving parents, Kaushal and Bankey Lal Sharma, for being determined to raise excellent individuals out of us and nurturing our dreams and aspirations by providing every possible opportunity and continued support. They have taught us, led by example, made innumerable sacrifices, and pushed us to achieve the limits of our abilities. They have played a pivotal role in making this degree a reality, right from the pre-PhD era until its end. I am finding it difficult to list all the things I want to thank my parents for, therefore, I convey my deepest gratitude for the most important thing, that is, everything they have taught us.

Chandrika Bhardwaj

Abstract

Securing the data manipulated by computing systems is a challenge. Traditional mechanisms such as encryption and access control do not address all user requirements, which include *controlling information propagation*, instead of only prevention of unauthorised information access and release, and secure transmission. To ensure the end-to-end security and integrity of data in any system, information flow control (IFC) is required. IFC extends access control by not only regulating who is allowed to access what data, but also the subsequent use and *propagation* of the data accessed. In this thesis, we identify reasons for the lack of widespread adoption of existing IFC techniques, including decentralized ones (DIFC). We find that the existing IFC mechanisms are not intuitive or easy to understand and use. In addition, they do not allow autonomous organisations and individual users to formulate their security policies *independently* and collaborate later by sharing data (bilaterally) with other organisations while respecting one another's policies. We explore different dimensions of *autonomy* in information flow control and propose extensions to IFC/DIFC models in order to provide finer control and autonomy to individuals and organisations regarding the flow of information. The objective is to

develop models which are intuitive and easy to use, and provide individuals in an organisation finer control over the flow of information within their organisation while respecting the broad organisational flow policies, and also to provide autonomy to organisations when *post facto* collaborating with other organisations in a secure manner.

The first dimension that we explore is *autonomy across different organisations* when sharing information (bidirectionally) in a secure manner. Though the *lattice model* proposed by Denning in her seminal work provides secure information flow analyses with an intuitive and uniform mathematical foundation, it requires a principled extension to support the modular composition of secure flow analysis between organisations. This is because different organisations may employ quite different secure information flow lattices. In this thesis, we propose a *Lagois connection* framework that permits different organisations to exchange information while maintaining both security of information flow as well as their autonomy in formulating and maintaining security policies. The merit of this formulation is that it is simple, minimal, adaptable and intuitive, and provides a formal mathematical framework for establishing secure information flow across autonomous interacting organisations. We show that our framework is semantically sound, by proving that the connections proposed preserve standard correctness notions such as *non-interference* for a simple operational model with a "security type system". We then show how Lagois theory also provides a robust framework and methodology for negotiating and maintaining secure agreements on information flow between autonomous organisations, even when either or both organisations change their security lattices.

We additionally show that our secure connections framework extends naturally and con-

servatively to another dimension, namely *autonomy among individuals* by applying it to the Decentralised Labels Model of Myers *et al.*

The third dimension that we explore is *autonomy across different relationships*. We observed that data (and metadata) are generated by *relationships* between principals, or a principal and an entity, rather than solely by the principals. We propose a refinement of security classes in Denning's lattice-based framework to reflect *relationships* such as that between a patient and her doctor, and where finer-grained control over information flow based on a relationship hierarchy can be better expressed. The usefulness of the idea is illustrated with a substantial example from the healthcare domain.

We propose an Attribute-based model (ABCIF) where security classes are parameterised by values. The security lattice is kept height-bounded (irrespective of the number of individuals/entities in the organisation) by employing notions of over- and under-approximation. Flows *between* relationship classes are not *ad hoc*, but are based on *uniform policies* applicable within an organisation, which are easy for users to understand. The framework finds a formal basis in the idea of having sets of *credentials* for accessing information in a given parameterised security class. The use of credentials also ties the formulation of secure information flow back to its origins in *access control*. The main results include construction of a correct security lattice models with the desired properties, and showing that the flows permitted in it are sound with respect to the credentials-based semantics.

The ABCIF model does not, however, support features of delegation of authority and declassification. Thus, we propose a Decentralized Attribute Based Control of Information Flow (DABCIF) model, in which the authorisation is not unconditional but is contextual

and delimited by relationships within an organisation. The model employs the parameterised model of ABCIF in place of the principals hierarchy of the DLM model of Myers [53], and thus provides users within an organisation autonomy in formulating policies regarding data which they own but within the broad (uniform) organisational policies. By projecting out specific fine-grained authorities based on attributes possessed by a security principal, the DABCIF model supports computation in an environment with more precise (purpose-based and situation-based) modelling of mutual distrust.

Finally, the thesis discusses a framework, SIFT, for systematically and automatically annotating data with security classes so that users are relieved of having to create tags for each class of data and metadata that is collected in the system, thus making it user-friendly and scalable.

In summary, this thesis proposes elegant and rigorous frameworks for providing autonomy in formulating and enforcing decentralized fine-grained information flow security policies in a collaborative environment involving more than one administrative domain (organisation), several individuals and multiple relationships.

सारांश

संगणिक प्रणाली द्वारा हेरफेर किए गए डेटा को सुरक्षित रखना एक चुनौती है। एन्क्रिप्शन और अभिगम नियंत्रण जैसे पारंपरिक तंत्र सभी उपयोगकर्ताओं की आवश्यकताओं को संबो–धित नहीं करते हैं, क्योंकि इनमें सूचना प्रसार नियंत्रण की जगह केवल अनधिकृत सूचना की पहुंच और रिहाई को रोकना और सुरक्षित हस्तांतरण शामिल हैं। किसी भी संगणिक प्रणाली में डेटा की शुरू से अंत तक सुरक्षा और अखंडता सुनिश्चित करने के लिए, सूचना प्रवाह नियंत्रण (आई.एफ.सी.) की आवश्यकता है। आई.एफ.सी. अभिगम नियंत्रण को विस्तृत करता है, क्योंकि यह न केवल हस्तांतरण नियंत्रित करता है – कि किसे कौन से डेटा को उपयोग करने की अनुमति है, बल्कि डेटा को प्राप्त करने के बाद उसके उपयोग और प्रचार–प्रसार का विनि–यमन भी करता है। इस शोध प्रबंध में, हम विकेंद्रीकृत सूचना प्रवाह नियंत्रण (डी.आई.एफ.सी.) सहित मौजूदा आई.एफ.सी. तकनीकों को व्यापक रूप से न अपनाये जाने के कारणों की पहचान करते हैं। हम पाते हैं कि मौजूदा आई.एफ.सी. तंत्र न तो सहजज्ञ हैं और न ही समझने और उपयोग करने में आसान हैं। इसके अलावा, वे ऐसी व्यवस्था या तंत्र नहीं उपलब्ध कराते कि एक स्वायत्त संगठन या व्यक्तिगत उपयोगकर्ता स्वतंत्र रूप से अपनी सुरक्षा नीतियाँ बनाने के बाद, दूसरे संगठनों की नीतियों का सम्मान करते हुए उन संगठनों के साथ डेटा (द्विपक्षीय) साझा करते हुए मिलकर काम कर सकें। हम सूचना प्रवाह नियंत्रण में स्वायत्तता के विभिन्न आयामों का पता लगाते हैं और सूचना के प्रवाह के संबंध में व्यक्तियों और संगठनों को महीन नियंत्रण और स्वायत्तता प्रदान करने के लिए आई.एफ.सी. और डी.आई.एफ.सी. मॉडलों का विस्तार प्रस्तावित करते हैं। हमारा उद्देश्य ऐसे मॉडलों को विकसित करना है जो सहज और प्रयोग करने में आसान हों, और एक संगठन में व्यक्तियों को व्यापक संगठनात्मक प्रवाह नीतियों का सम्मान करते हुए उनके संगठन के भीतर सूचना के प्रवाह पर नियंत्रण प्रदान करें, और संगठनों को स्वायत्तता प्रदान करते हुए, अन्य संगठनों के साथ सुरक्षित तरीके से सहयोग करने में सहायता करें।

पहला आयाम जो हम खोजते हैं – सुरक्षित तरीके से जानकारी (द्विपक्षीय) साझा करते समय, सुरक्षा नीतियाँ बनाने और लागू करने में विभिन्न संगठनों की स्वायत्तता। यद्यपि डेनिंग के प्राथमिक कार्य में प्रस्तावित लैटिस मॉडल एक सहज और समान गणितीय आधार के साथ सुरक्षित सूचना प्रवाह विश्लेषण प्रदान करता है, लेकिन संगठनों के बीच सुरक्षित प्रवाह विश्लेषण के मॉड्यूलर संरचना का समर्थन करने के लिए एक सैद्धांतिक विस्तार की आवश्यकता है। ऐसा इसलिए है क्योंकि विभिन्न संगठन काफी अलग–अलग सुरक्षित सूचना प्रवाह जालकों को उपयोग कर सकते हैं। इस शोध–निबंध में, हम एक 'लैग्वा संयोजन तंत्र' का प्रस्ताव करते हैं जो विभि–न्न संगठनों को सूचना का आदान–प्रदान करते समय सूचना प्रवाह की सुरक्षा के साथ–साथ सुरक्षा नीतियों को बनाने और बनाए रखने में उन्हें स्वायत्तता प्रदान करता है। इस सूत्रीकरण की विशेषता यह है कि यह सरल, न्यूनतम, अनुकूलनीय और सहजज्ञ है, और साथ मिलकर काम करने वाले स्वायत्त संगठनों में सुरक्षित सूचना प्रवाह की स्थापना के लिए एक औपचारिक गणितीय तंत्र प्रदान करता है। यह साबित करते हुए कि प्रस्तावित तंत्र एक 'सुरक्षा प्रकार प्रणाली' के साथ एक सरल परिचालन प्रतिरूप के लिए, गैर–हस्तक्षेप जैसे मानक शुद्धता धारणाओं को संरक्षित करता है, हम दिखाते हैं कि हमारा तंत्र अर्थ विज्ञान के हिसाब से सही है। फिर हम

दिखाते हैं कि कैसे स्वायत्त संगठनों के बीच सूचना प्रवाह पर सुरक्षित समझौते निश्चय करने और बनाए रखने के लिए लैग्वा सिद्धांत एक मजबूत ढांचा और कार्यप्रणाली प्रदान करता है, तब भी जब या तो एक या दोनों संगठन अपनी सुरक्षा जाली बदलते हैं।

अतिरिक्त रूप से, प्रस्तावित सुरक्षित संयोजन तंत्र को डीआईएफसी पर लागू करके हम यह दिखाते हैं कि हमारा सुरक्षित संयोजन तंत्र स्वाभाविक रूप से और परंपरागत ढंग से दूसरे आयाम तक लागू होता है, जिसका नाम है – "व्यक्तियों के बीच स्वायत्तता"।

तीसरा आयाम जिस पर हमने शोध किया है – "विभिन्न रिश्तों में स्वायत्तता"। हमने देखा कि डेटा (और मेटाडेटा), केवल व्यक्तियों के बजाय, व्यक्तियों या व्यक्ति और एक इकाई के बीच के रिश्तों द्वारा उत्पन्न होते हैं। हम एक मरीज और उसके डॉक्टर के बीच के जैसे रिश्तों को प्रतिबिंबित करने के लिए डेनिंग के जाली–आधारित तंत्र में सुरक्षा वर्गों के शोधन का प्रस्ताव करते हैं – जहां संबंध पदानुक्रम के आधार पर, सूचना प्रवाह पर महीन व बारीकी से नियंत्रण और बेहतर ढंग से व्यक्त किया जा सके। इस सुझाव की उपयोगिता को दर्शाने के लिए स्वास्थ्य सेवा क्षेत्र का एक उपयुक्त एवं महत्वपूर्ण उदाहरण, इस शोध–निबंध में, चित्रित किया गया है।

हम एक विशेषता–आधारित मॉडल (ए.बी.सी.आई.एफ.) प्रस्तावित करते हैं, जहाँ सुरक्षा–कक्षाएं मानों द्वारा परिचालित की जाती हैं। संगठन में व्यक्तियों और इकाईयों की बहु–संख्या के बावजूद, सुरक्षा जालक को अधिक–अनुमान और कम–अनुमान के नियोजन के द्वारा ऊँचाई–सीमाबद्ध रखा गया है। रिश्तों द्वारा निर्धारित कक्षाओं के बीच सूचना प्रवाह निराधार या बेतरतीब नहीं है, किन्तु एक संगठन के भीतर लागू एकरूपक नीतियों पर आधारित है, जो कि उपयोग–कर्ताओं के लिए समझना आसान है। किसी दिए गए पैरामिट्रीकृत सुरक्षा वर्ग से जानकारी प्राप्त करने के लिए प्रत्यायकों के समूह को मुहैया कराने की आवश्यकता के प्रस्ताव के द्वारा प्रस्तावित तंत्र औपचारिक आधार पाता है। प्रत्यायक या क्रेडेंशियल का उपयोग, सुरक्षित जानकारी प्रवाह के निर्माण को वापस अपने मूल अभिगम नियंत्रण से संबंधित कराता है। मुख्य परिणामों में वांछित गुणों के साथ एक सही सुरक्षा जालक मॉडल का निर्माण शामिल है, और यह दिखाया गया है कि इसमें अनुमतिप्राप्त प्रवाह सुरक्षा संबंधी कक्षाओं के प्रत्यायक–आधारित अर्थों के संबंध में सही हैं।

हालांकि, ए.बी.सी.आई.एफ. मॉडल प्रतिनिधि की नियुक्ति और गैर–गोपनीयता (डीक्लासि–फिकेशन) की सुविधाओं को मुहैया नहीं कराता है। इसलिए, हम एक विशेषता आधारित वि–केंद्रीकृत सूचना प्रवाह नियंत्रण (डी.ए.बी.सी.आई.एफ.) मॉडल का प्रस्ताव करते हैं, जिसमें प्राधिकरण या प्राधिकृति बिना शर्त नहीं है, बल्कि एक संगठन के भीतर संबंधों द्वारा प्रासंगिक और सीमांकित है। यह मॉडल एबीसीआईएफ के पैरामिट्रीकृत मॉडल को मायर्स के डीएलएम मॉडल के सदस्य पदानुक्रम के स्थान पर नियोजित करता है और इस प्रकार उपयोगकर्ताओं को डेटा के संबंध में नीतियों के निर्माण में, एक संगठन की व्यापक (समान) संगठनात्मक नीतियों के अंतर्गत, व्यक्तिगत स्वायत्तता प्रदान करता है। किसी व्यक्ति के पास मौजूद विशेषताओं के आधार पर, विशिष्ट अधिकारों को दर्शा करके, डी.ए.बी.सी.आई.एफ. मॉडल आपसी अविश्वास के समक्ष बहुत ही सटीक (उद्देश्य–आधारित और स्थिति–आधारित) प्रतिरूपण करके सुरक्षित गणना संभव और सक्षम कराता है।

अंत में, एक तंत्र – 'सिफ्ट' विस्तृत किया गया है जो व्यवस्थित रूप से और स्वचालित रूप से

सुरक्षा कक्षाओं के साथ डेटा को जोड़ता है ताकि उपयोगकर्ताओं को सिस्टम में एकत्रित किये गए डेटा और मेटाडेटा के प्रत्येक वर्ग के लिए चिप्पी (टैग) बनाने से राहत मिले, और इस प्रकार यह आईएफसी तंत्रों को उपयोगकर्ताओं के अनुकूल और विस्तार योग्य बनाता है।

सारांश में, यह शोध–निबंध एक से अधिक प्रशासनिक कार्य क्षेत्रों (संगठनों), कई व्यक्तियों और कई रिश्तों को शामिल करते हुए सहयोगात्मक वातावरण में विकेंद्रीकृत, महीन व बारीकी से नियंत्रण और बेहतर ढंग से व्यक्त की जा सकने वाली सूचना–प्रवाह सुरक्षा नीतियों को तैयार करने और लागू करने में स्वायत्तता प्रदान करने के लिए, सुरुचिपूर्ण और गणितीय रूपरेखाओं का प्रस्ताव करता है।

Contents

List of Figures

List of Tables

Glossary

Non-Interference an attacker cannot distinguish between two executions that differ only in their private (high) inputs by observing only the public (low) outputs.

Medical Encounter generates medical data and metadata and is considered complete only when these get transformed into various EHR, EMR, etc.

Information Flows from object x to another object y when an operation, or series of operations use the value of x as input to derive a value for y.

Information Flow Control refers to the mechanisms of regulating the propagation of information among the objects in the system.

Acronyms

SIF Secure Information Flow.

IFM Information Flow Model.

IFC Information Flow Control.

DIFC Decentralised Information Flow Control.

DLM Decentralised Label Model.

LC Lagois Connection.

GC Galois Connection.

GI Galois Insertion.

DIFT Dependent Information Flow Types.

MoU Memorandum of Understanding.

SIFT Systematic Information Flow Tagging Framework.

ABCIF Attribute-based Control of Information Flow Model.

DABCIF Decentralized Attribute-based Control of Information Flow Model.

PII Personally Identifiable Information.

HIS Hospital Information System.

Symbols

SC Set of Security Classes.

$\sqsubseteq$ LHS argument is *permitted* to flow into RHS argument.

$\sqsupseteq$ RHS argument is *permitted* to flow into LHS argument.

$\sqcup$ Join operator for security classes.

$\succeq$ acts-for relation on principals/personae.

$\preceq$ symmetric inverse of acts-for relation on principals/personae.

$\curlyvee$ disjunction of principals/personae.

$\curlywedge$ conjunction of principals/personae.

$\cup$ set union.

$\cap$ set intersection.

$\subseteq$ subset.

α function mapping security classes from domain C to those in domain U, i.e., $\alpha : C \rightarrow U$.

γ function mapping security classes from domain U to those in domain C, i.e., $\gamma : U \rightarrow C$.

α^{-1} inverse of function α.

$P, Q, U, A, PC, D, N, R, PI, S, TA, \ldots$ Predicates for parameterised security classes or personae.

$\top$ Most restrictive security class in a security lattice.

$\bot$ Least restrictive security class in a security lattice.

$U(\top)$ Upper approximation of unary predicate security class, $U(x)$.

$U(\bot)$ Upper approximation of unary predicate security class, $U(x)$.

$\sum$ Dependent sum type.

$\prod$ Dependent product type.

$\mathfrak{Cr}$ The set of actual credentials declared in an administrative domain.

$\mathfrak{s}, \mathfrak{s}_i, \mathfrak{ind}, \mathfrak{subm}, \ldots$ Sorts of the arguments to Predicates.

$\mathsf{a}, \mathsf{b}, \mathsf{a1}, \mathsf{a2}, \mathsf{pc5}, \mathsf{pc6}, \mathsf{s91}, \mathsf{s92}, \ldots$ Concrete values from different sorts parameterizing the predicates.

D_{SC} Dataset of facts declaring ground parameterised security classes in an administrative domain.

$Q(a_1, \ldots, a_k)$ Ground parameterised security class with predicate Q and arguments $a_1, \ldots, a_k$.

$\mathbf{Cr}(Q, a_1, \ldots, a_k)$ Ground credential associated to ground parameterised security class $Q(a_1, \ldots, a_k)$.

$2^{\mathfrak{Cr}}$ Power set of actual credentials declared in an administrative domain.

$\downarrow s$ Downward Closure of a set of credentials, $\downarrow s = \{c \in \mathfrak{Cr} \mid (\exists c' \in s)\ H_C \vdash c \preccurlyeq c'\ \}$.

$\uparrow S$ Upward Closure of a set of sets of credentials, $\uparrow S = \{s' \subseteq \mathfrak{Cr} \mid (\exists s \in S)\ H_C \vdash s\ \hat{\sqsubseteq}\ s'\ \}$.

$\Uparrow S$ Superset Closure of a set of sets of credentials, $\Uparrow S = \{S' \mid S' \supseteq S\}$.

$\succcurlyeq$ dominates relation on credentials.

$\preccurlyeq$ symmetric inverse of dominates relation on credentials.

$\hat{\sqsubseteq}$ "is less able than" ordering on sets of credentials.

$\hat{\sqcup}$ Semantic join operator for sets of credentials.

$\hat{\sqcap}$ Semantic meet operator for for sets of credentials.

$[\![\ ]\!]$ Meaning.

$\breve{\sqsubseteq}$ "yields-to" ordering on parameterised security classes.

$\breve{\sqcup}$ Semantic join operator for parameterised security classes.

$\breve{\sqcap}$ Semantic meet operator for parameterised security classes.