Cent. Eur. J. Comp. Sci. • 2(3) • 2012 • 214-221
DOI: 10.2478/s13537-012-0014-8

VERSITA

Central European Journal of **Computer Science**

Intrusion detection system episteme

Research Article

Daniel Mihályi*, Valerie Novitzká†, Martina Ľaľová‡

Department of Computers and Informatics,
Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Letná 9, 04200 Košice, Slovakia

Received 01 February 2012; accepted 17 August 2012

Abstract: In our paper we investigate the possibilities of modal logics in coalgebras of program systems. We deal with a simplified model of an intrusion detection system. We model an intrusion detection system as a coalgebra and construct its Kripke model of coalgebraic modal linear logic using powerset endofunctor. In this model we present our idea how a fragment of epistemic linear logic can provide knowledge and belief of intrusion attempt.

Keywords: coalgebra • modal logic • epistemic logic • linear logic • Kripke model

1. Introduction

In programming systems we are interested not only in their construction, but also in their observable behavior. Observable behavior can be modeled by coalgebras [1, 9, 10] using modal logic [3]. Coalgebras can model various types of transition systems. Within behavioral observations some events can repeat and they can provide us with some interesting knowledge about program systems. Following the results in [11] we can be sure that objective knowledge implies rational belief. Knowledge and belief are fundamental notions of epistemic logic.

In our approach we investigate possibilities of obtaining objective knowledge and rational belief for a simplified model of an intrusion detection system (IDS). Incoming packets form infinite streams and some of them can contain some intrusion attempts. These attemps can be recognized through characteristic symptoms. Specific combinations of these symptoms give us knowledge about some kind of incoming intrusion. Moreover, if it comes from the same IP address and repeatedly, then we obtain a certitude that it is a real intrusion attempt and we can make our decision about competent reactions. We use a fragment of epistemic linear logic with explicit logical operators for objective knowledge and rational belief as a suitable logical system for reasoning about intrusion attempts.

In this paper we extend our results published in [4] where IDS is modeled as a coalgebra over appropriate polynomial endofunctor. In our previous work a coalgebra was used as a model investigating observable behavior of IDS. In this paper we extend our approach and showing how knowledge and belief can be formulated in Kripke model of possible worlds over this coalgebra.

* *E-mail: daniel.mihalyi@tuke.sk (Corresponding author)*
† *E-mail: valerie.novitzka@tuke.sk*
‡ *E-mail: martina.lalova@tuke.sk*

2. Coalgebraic modal linear logic for IDS

Typically, coalgebraic approaches use a modal logic called coalgebraic modal logic [3] with two modal operators ($\Box$ for necessity and $\Diamond$ for possibility). In our approach we choose as a basic logic linear logic because of the causality of its linear implication. We extend it by modal operators. The syntax of our modal linear logic fragment is as follows:

$$\varphi ::= \ a_i \mid \varphi_1 \multimap \varphi_2 \mid \varphi_1 \otimes \varphi_2 \mid \Box\varphi \mid \Diamond\varphi \mid \mathbf{1}, \tag{1}$$

where

- a_i are atomic propositions,
- $\varphi_1 \multimap \varphi_2$ means linear implication. This implication ensures that the action φ_2 follows after the action φ_1,
- $\varphi_1 \otimes \varphi_2$ is multiplicative conjuction expressing that φ_1 and φ_2 are both executed,
- $\Box\varphi$ means application of the necessity operator to formula φ,
- $\Diamond\varphi$ means application of the possibility operator to formula φ,
- $\mathbf{1}$ is a neutral element of the multiplicative conjuction.

We illustrate coalgebraic modal linear logic with an example of IDS. We consider only two types of possible intrusions, A or B. Let O be a sender indentification (e.g. the IP address). Then we construct the category $\mathcal{P}acket$ of incoming packets as follows:

- objects are significant packet fragments for identification of intrusion attempts,

$$p = (A + B) \times O \tag{2}$$

- morphisms are mappings $next$ between objects

$$next : p_i \to p_{i+1}, \tag{3}$$

where $i \in \mathbb{N}$. It is clear that the category $\mathcal{P}acket$ has special sets as objects. Now we define polynomial endofunctor $T : \mathcal{P}acket \to \mathcal{P}acket$ on this category as follows:

$$T(p) = X \times p \quad and \quad T\big(next(p)\big) = X \times next(p), \tag{4}$$

where X denotes state space, in this case a stream of incoming packets. Then the coalgebraic specification for polynomial endofunctor T is

$$\langle hd, tl \rangle : \rho_p \to T\rho_p. \tag{5}$$

In [4] we modeled the IDS system as a coalgebra

$$\Big(\rho_p, \langle hd, tl \rangle\Big) \tag{6}$$

for infinite packet stream ρ_p. The operations hd resp. tl are obvious operations returning head resp. tail of a given stream.

Contemporary experiences in the area of system behavior have shown the importance of selecting an appropriate modal logical language as a specification language for various transition systems. Formulae of this language are used for logical reasoning over states of dynamic system that are captured by the coalgebra of corresponding polynomial (powerset) endofunctor. We formulated coalgebraic linear logic based on a multimodal language that is suitable for behavioral

description of infinite, non trivial heterogenous data structures, i.e. packets at the coalgebra as intrusion detection system in [5].
In the following text let $Prop$ be the set of propositions.
As a model of our logic we use the Kripke model of possible worlds [12] that is characterized by Kripke frame

$$(W, \leq, w_0), \tag{7}$$

with satisfaction relation $\models$ as a tuple

$$(W, \leq, \models, w_0), \tag{8}$$

where

- W is a set of possible worlds,
- $\leq$ is accessibility relation $\leq \subseteq W \times W$,
- $\models$ is satisfaction relation

$$\models: W \times Prop \to \{0, 1\}, \tag{9}$$

  where 1 means satisfaction and 0 means non satisfaction,

- w_0 is a designated world.

Notation $w_1 \leq w_2$ is as follows: a possible world w_2 is reachable (accessible) from w_1. According to the philosophy of possible world semantics: "it is possible what is reachable together" [12].
A coalgebra can be seen as a general form of Kripke frame for modal logics. An interpretation of a formula in coalgebra is given by predicate lifting. Predicate lifting is a natural transformation i.e. morphism (λ) between functors ($\mathcal{P}^-$,T) as follows

$$\lambda : \mathcal{P}^- \Rightarrow \mathcal{P}^- \circ T, \tag{10}$$

where $\mathcal{P}^-$ is a contravariant powerset functor $\mathcal{P}^- : Set \to Set$ between sets (Fig. 1).

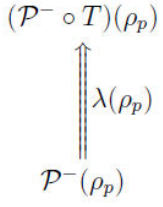

Figure 1. Predicate lifting.

$\lambda(\rho_p)$ is a class of morphisms defined by

$$\lambda(\rho_p) : \mathcal{P}^-(\rho_p) \to (\mathcal{P}^- \circ T)(\rho_p). \tag{11}$$

Now we can interpret the formulae in any T-model as follows:

$$(\lambda(\rho_p), \langle hd, tl \rangle : \rho_p \to T\rho_p), \tag{12}$$

where

- for every formula φ we define validity set $[\![\varphi]\!] \subset \rho_p$ by induction on the structure of φ,
- for modal operator $\Box$ we define the validity as

$$[\![\Box\varphi]\!] = \mathcal{P}^-\big(\langle hd, tl\rangle : \rho_p \to T\rho_p\big) \circ \lambda\big([\![\varphi]\!]\big). \tag{13}$$

The operator of possibility $\Diamond$ is dual to the operator of necessity $\Box$. In the following we use the traditional notation for Kripke models. It is clear that the set W of possible worlds corresponds with the stream of packets ρ_p.

- Every world $w \in W$ corresponds with a packet $p \in \rho_p$,
- the reachability relation

$$\leq \ \subseteq \ W \times W \tag{14}$$

gives a $\mathcal{P}$-coalgebra

$$\big(W, \langle hd, tl\rangle : \rho_p \to T\rho_p\rangle\big), \tag{15}$$

where

$$\big(\langle hd, tl\rangle : \rho_p \to T\rho_p\big)_{\leq}(w) = \big\{w' \in W \mid (w, w') \in \leq\big\}. \tag{16}$$

Then the formulae of our language are expressed as the infinite sequence

$$\begin{aligned} &(\mathbf{1})\\ &(p_1, \mathbf{1})\\ &\big(p_1, (p_2, \mathbf{1})\big)\\ &\Big(p_1, \big(p_2, (p_3, \mathbf{1})\big)\Big)\\ &\dots\\ &\bigotimes \Big\{\big(p_0, p_1, p_2, p_3, \dots, (true)\big)\Big\}, \end{aligned} \tag{17}$$

where the raw (**1**) denotes the empty formula and $\bigotimes$ denotes infinite linear multiplicative conjuction.

3. Epistemic linear logic for IDS

Traditionally, epistemic logic is characterised as intensional logic with modalities refering to objective knowledge and rational belief [11]. Again, we choose linear logic as the basic logic and we extend it with epistemic operators. The syntax of our epistemic linear logic has the following form [2]:

$$\varphi ::= \ a_i \mid K_c\varphi \mid B_c\varphi \mid \varphi_1 \otimes \varphi_2 \mid \varphi_1 \multimap \varphi_2 \mid !\varphi, \tag{18}$$

where

- a_i are atomic propositions (i.e. pieces of knowledge),
- $K_c\varphi$ denotes that a rational agent c knows that φ,
- $B_c\varphi$ denotes that a rational agent c believes that φ,
- $\varphi_1 \otimes \varphi_2$ realizes the linear conjunction of two formulas φ_1,φ_2,
- $\varphi_1 \multimap \varphi_2$ realizes the linear implication of two formulas φ_1,φ_2,
- $!\varphi$ is empiric modal operator expressing repeated objective knowledge.

Table 1. Particular types of network intrusions.

Type A	Type B
(ICMP Ping NMAP)	*(TCP Portscan)*
IP Protocol == icmp	MAC Addr == MACDAD
dsize == 0	IP Protocol == 255
itype == 8	IP TTL == 0

Table 2. Intrusion type *A - ICMP Ping NMAP*.

Type A	a_1	a_2	a_3
w_8	0	0	0
w_7	0	0	1
w_6	0	1	0
w_5	0	1	1
w_4	1	0	0
w_3	1	0	1
w_2	1	1	0
w_1	1	1	1

4. Motivation example

In accordance with the example mentioned in [4], an infinite stream of packets ρ_p can be realised by the following infinite sequence

$$\begin{aligned}
&(p_1, p_2, p_3, p_4, p_5 \dots) \mapsto \\
&\mapsto \big(A \times O, (B \times O, \mathbf{1} \times O, A \times O, B \times O, \dots)\big) \ \mapsto \\
&\mapsto \big(A \times O, B \times (O, \mathbf{1} \times O, A \times O, B \times O, \dots)\big) \ \mapsto \\
&\mapsto \big(A \times O, B \times O, \mathbf{1} \times O, (A \times O, B \times O, \dots)\big) \ \mapsto \\
&\mapsto \big(A \times O, B \times O, \mathbf{1} \times O, A \times O, (B \times O, \dots)\big) \ \mapsto \\
&\mapsto \dots,
\end{aligned} \tag{19}$$

where p_1, p_2, p_3, p_4, p_5 are treated packet fragments from any pattern of network traffic and A resp. B are specifications of particular intrusion attempt ICMP Ping NMAP resp. TCP Portscan mentioned in Table 1.
For our fragment of epistemic linear logic we define its model as Kripke model of possible worlds. We can use the same Kripke frame

$$(W, \leq, w_0) \tag{20}$$

constructed for modal linear logic in the previous section.
Let $AP = \{a_1, a_2, a_3, b_1, b_2, b_3, \dots\}$ be the set of atomic propositions. Every atomic proposition denotes one symptom of an appropriate intrusion attempt. According to Table 1 we denote the knowledge about an intrusion attempt of "Type A" i.e. *ICMP Ping NMAP* by the tuple (a_1, a_2, a_3) where

- a_1 : $IPProtocol$ is equal to $icmp$,
- a_2 : $dsize$ is equal to 0,
- a_3 : $itype$ is equal to 8.

If a symptom a_i is present then we assign the value 1 to it. Otherwise we assign to a_i the value 0. According to the Table 2 we will work with eight possible worlds.

Table 3. Intrusion type B - $TCP\ Portscan$.

Type B	b_1	b_2	b_3
w_{16}	0	0	0
w_{15}	0	0	1
w_{14}	0	1	0
w_{13}	0	1	1
w_{12}	1	0	0
w_{11}	1	0	1
w_{10}	1	1	0
w_9	1	1	1

The intrusion attempt of "Type A" occurs only if all a_i have the value 1. Therefore we consider w_1 as designated world ($w_1 \equiv w_0$).
We do the same for the next intrusion attempt of "Type B", i.e. $TCPPortscan$ we consider the following pieces of knowledge (b_1, b_2, b_3)

- b_1 : $MACAddr$ is equal to $MACDAD$,
- b_2 : $IPProtocol$ is equal to 255,
- b_3 : $IPTTL$ is equal to 0.

If a symptom b_i is present then we assign the value 1 to it. Otherwise we assign to b_i the value 0. According to the Table 3 we will work also with eight possible worlds.
The intrusion attempt of "Type B" occurs only if all b_i have the value 1. Therefore we consider w_9 as designated world ($w_9 \equiv w_0$).
In our simple example we consider only one (rational) agent 007. This agent can be a part of communication interface between the human and computer system.
From the behavioral sequence of packets (19) our agent can achieve particular knowledge (e.g. based on visible alerts on monitor). We denote by

- $K_{007}a_i$ that our agent 007 has the particular piece of knowledge about a_i,
- $K_{007}b_i$ that our agent 007 has the particular piece of knowledge about b_i.

By induction

- $K_{007}\varphi$ is an epistemic linear formula that denotes

$$K_{007}a_1 \otimes K_{007}a_2 \otimes K_{007}a_3 \tag{21}$$

- $K_{007}\psi$ is an epistemic linear formula that denotes

$$K_{007}b_1 \otimes K_{007}b_2 \otimes K_{007}b_3 \tag{22}$$

In Fig. 2 we illustrate the process of achieving knowledge about intrusion attempts of "Type A" and "Type B". Then

- $K_{007}\varphi$ denotes objective knowledge that on the packet p_1 was captured intrusion attempt of "Type A" i.e. ICMP Ping NMAP,
- $K_{007}\psi$ denotes objective knowledge that on the packet p_3 was captured intrusion attempt of "Type B" i.e. TCP Portscan.

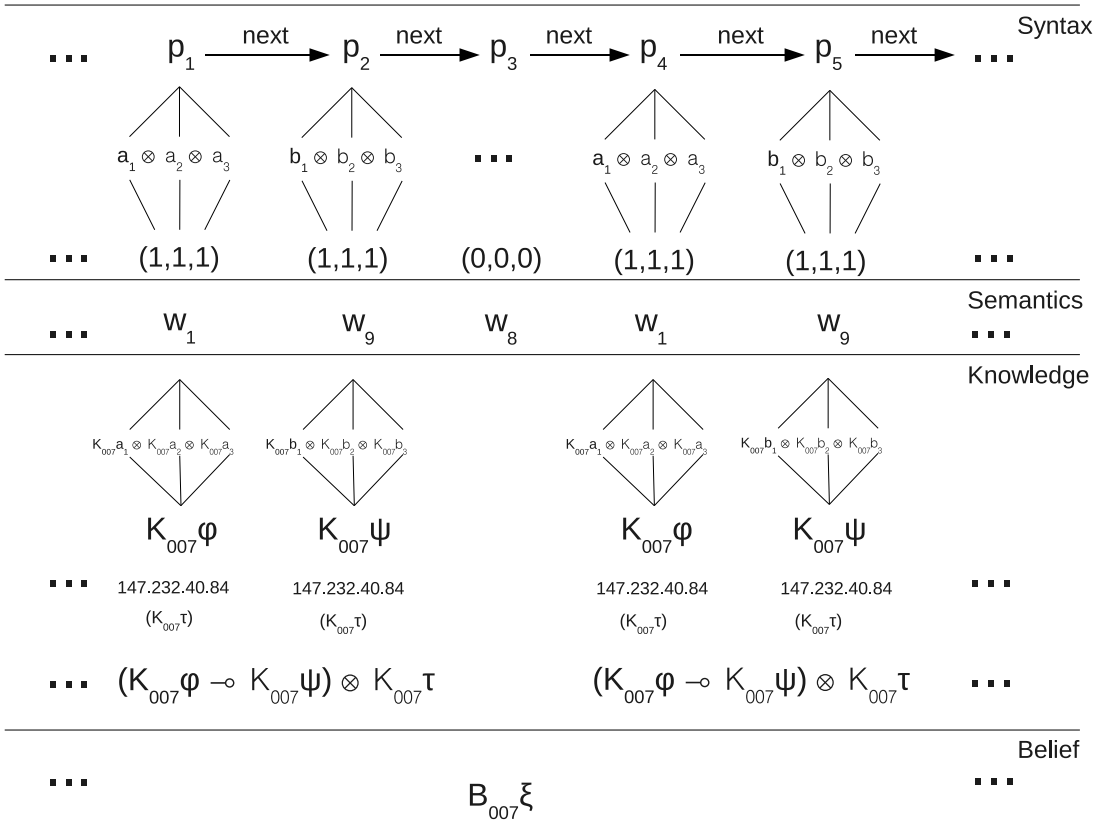

Figure 2. Episteme.

Here we need to acquire another piece of knowledge about the sender of a given packet e.g. its IP Address. This knowledge can be achieve from the attribute *srcIP* in packet header. We denote by $K_{007}\tau$ the objective knowledge about sender identification of the "catched packet".
The formula

$$(K_{007}\varphi \multimap K_{007}\psi) \otimes K_{007}\tau \tag{23}$$

describes the situation that an intrusion attempt of "Type A" followed by an intrusion attempt of "Type B" have occured from an intruder with the same identification. This kind of attempt is known as vertical portscan.
We achieve empirical rational belief denoted $B_{007}\xi$ from repeatedly occuring knowledge about intrusion attempts of "Type A" and "Type B" incomming from the same sender that is given by its identification O.

$$\begin{aligned} \text{if} \quad & K_{007}\varphi \otimes K_{007}\psi \otimes K_{007}\tau \text{ and} \\ & K_{007}\varphi \otimes K_{007}\psi \otimes K_{007}\tau \text{ and} \\ & \ldots \\ \text{then} \quad & B_{007}\xi \end{aligned} \tag{24}$$

Using the empiric modality operator we can denote the process from objective knowledge to rational belief as the following linear implication

$$!(K_{007}\varphi \otimes K_{007}\psi \otimes K_{007}\tau) \multimap B_{007}\xi \tag{25}$$

5. Conclusion

In this paper we presented our ideas about achieving knowledge and belief from the observable behavior of program systems. We illustrated our approach on the simplified intrusion detection system and we showed how the pieces of knowledge can be achieved from some symptoms, how its combination gives us the knowledge about some intrusion detection and how the repeating of some knowledge leads to belief about some concrete intrusion attempt. Our approach is based on coalgebraic modeling of system behavior. Instead of obvious correspondence with modal logic we constructed a model of a simple fragment of epistemic linear logic suitable for our purposes and showed the process from pieces of knowledge to knowledge and belief on the model.

Our approach used only IP protocol version IPv4 and only two possible intrusion attempts. In this paper we did not investigate the expressive power of our model, which is the aim of our further research. Our idea could be generalized for other types of intrusion attempt (viruses, hardware and software overloads, etc.) and we would like to investigate achieving knowledge and belief for IP protocol version ipv6, too. In further research we would like to investigate distributed intrusion attempts using groups of agents.

Acknowledgements

This work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120030) supported by the Research & Development Operational Program funded by the ERDF.

References

[1] Gumm H.P., Functors for Coalgebras. Algebr. Univ., 45, 135-147, 2001
[2] Kamide N., Linear and affine logics with temporal, spatial and epistemic operators. Theor. Comput. Sci., 353, 165207, 2006
[3] Kurz A., Coalgebras and Modal Logic. CWI, Amsterdam, Netherlands, 2001
[4] Mihályi D., Novitzká V., a Coalgebra as an Intrusion Detection System, Acta Polytech. Hung., 7, 71-79, 2010
[5] Mihályi D., Novitzká V., Princípy duality medzi konštruovaním a správaním programov, Equilibria, 2010
[6] Moss L.S., Coalgebraic Logic. Ann. Pure Appl. Logic, 96, 1999
[7] Schröder L., Pattinson D., Coalgebraic Modal Logic: Forays Beyond Rank 1. IFIP WG 1.3 meeting, Sierra Nevada, 2008
[8] Schubert Ch., Topo-Bisimulations are Coalgebraic. Rendiconti dell' Istituto di Matemàtica dell'Universitá di Trieste, 42, 257-270, 2010
[9] Slodičák V., Macko, P., New approaches in functional programming using algebras and coalgebras, In: European Joint Conferrences on Theory and Practice of Software - ETAPS 2011 (March 2011), Workshop on Generative Technologies, Universität des Saarlandes, Saarbrücken, Germany, 1323
[10] Slodičák V., Some useful structures for categorical approach for program behavior, J. Inform. Organ. Sci., 35, 99-109, 2011
[11] Voorbraak F., Generalized Kripke Models For Epistemic Logic. Proceedings of Fourth Conference on Theoretical aspects of reasoning about knowledge, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1992, 214-228
[12] Zouhar M., Základy logiky. Proceedings of Fourth Conference on Theoretical aspects of reasoning about knowledge, Veda SAV, Bratislava, 2008