Sistemas Inmunes Artificiales:
Criterios de selección de tecnología para la seguridad de la información

Acevedo, Pablo Farías, Christian Taft, Yari Ivan Muscatiello, Juan Orlando, Leonardo

ABSTRACT

En vista de la masividad de uso de los sistemas y la criticidad de la información, surge la necesidad de estar protegidos ante las amenazas conocidas y desconocidas que comprometen la seguridad de los usuarios. Los sistemas inmunes biológicos presentan características, como la capacidad de aprendizaje, memoria y adaptación, que resultan efectivas en la protección de un organismo. El estudio de dichas características ha permitido la creación de Sistemas Inmunes Artificiales, cuya aplicación brinda nuevas soluciones a problemas relacionados con la integridad de la información.

En el presente trabajo, se realiza una introducción a estos sistemas y se comparan diferentes implementaciones existentes para que cada usuario pueda analizar qué producto se ajusta más a sus necesidades y también para intentar saber sobre qué tecnologías se debería seguir avanzando en trabajos futuros.

Keywords: AIS, invasive software, computer security, machine learning.

1. Introducción

Hoy en día, la protección de la información tiene una relevancia masiva. La creciente interconexión de sistemas de información, dada su extensa aceptación y adopción, han agilizado y facilitado el comercio y las telecomunicaciones. Incluso estrategias de negocios enteras dependen del funcionamiento correcto de estos sistemas que favorecen, en gran medida, el acceso a la información necesaria. Como contraparte, estos sistemas también tienen debilidades al exponer los datos y permitir su manipulación [1].

El acceso ubicuo a nuestra información, ha despertado el interés de aquellos que buscan obtener alguna ventaja de las debilidades de los sistemas involucrados. Dada la criticidad que significa que todas estas actividades comerciales y sociales sean soportadas por sistemas de información es que día a día aumenta la cantidad de amenazas y código malicioso, como los virus, troyanos, gusanos y ransomware, motivo por el cual es cada vez más importante contar con sistemas de seguridad para resguardar la integridad de los datos almacenados.

Los métodos tradicionales para proteger la información son reactivos y dependen de la intervención humana para ser efectivos

[1]. Los tiempos entre la aparición de una nueva amenaza y la aplicación de soluciones que protejan esos sistemas, en muchos casos, varía de días a semanas, ya que especialistas en seguridad informática deben cumplir la labor de analizar manualmente el código para detectar si se trata de malware[1] y en tal caso deben actualizar la base de firmas para que dicho programa en un próximo análisis sea detectado por el antivirus. Durante todo este tiempo, los usuarios se encuentran desprotegidos y a merced de estos ataques [2]. Por ello, es necesario buscar soluciones que brinden un mejor tiempo de respuesta ante los peligros actuales.

Sumado al alto tiempo de reacción mencionado, las aplicaciones antivirus tienen que actualizarse constantemente debido a la creciente aparición de nuevos virus informáticos y códigos maliciosos. Mientras que hace algunos años, las actualizaciones eran necesarias de forma mensual, hoy se realizan cada pocas horas [2].

Para hacer frente a estos problemas, surgen los Sistemas Inmunes Artificiales (AIS por sus siglas en inglés, Artificial Immune Systems) como una alternativa eficaz para proteger la información, estos sistemas prometen acelerar la reacción incluso posibilitando la detección temprana de amenazas nunca antes encontradas.

El presente estudio se enfoca en los sistemas de seguridad que protegen aquello que consideran propio, contra ataques externos o desórdenes internos, aplicando AIS. Se exponen distintas soluciones existentes y se comparan utilizando una serie de criterios propuestos.

En la Sección 2 se desarrolla una introducción a las características de la tecnología que da soporte a los sistemas inmunes artificiales y el estado actual de su investigación. En la Sección 3 se explica el funcionamiento de cuatro sistemas basados en AIS. En la Sección 4 se proponen criterios de selección aplicables a los sistemas presentados y se desarrolla una comparación entre estos sistemas. Finalmente, en la Sección 5 se presentan las conclusiones y en la Sección 6, un apartado de trabajos futuros.

2. Sistemas Inmunes Artificiales

Un sistema biológico inmune (BIS por sus siglas en inglés, Biological Immune Systems) es un sistema de autodefensa que involucra distintos procesos que protegen un organismo contra enfermedades y mantienen la homeostasis. Llamamos enfermedad a cualquier desorden o condición anormal en un ser vivo; ésta puede ser causada por disfunciones internas o agentes externos, como patógenos. Por otro lado, la homeostasis es la regulación interna de diferentes variables que deberían mantenerse en un estado constante o con la menor variación posible, como la temperatura interna del organismo [3].

El campo de estudio de los Sistemas Inmunes Artificiales tiene dos ramas principales que Cohen denomina *literal y metafórica* [4], la diferenciación se debe a cuan aproximada está la solución AIS a un BIS. El estudio literal de AIS apunta a la construcción de sistemas que simulan *in*

[1] Del inglés “malicious software”.

silico[2] un sistema inmune biológico, mientras que el estudio metafórico está inspirado en las características de un BIS y permite analizar y resolver diferentes problemas, imitando el comportamiento de los sistemas inmunes. Existe una tercera rama que busca entender mejor la biología inmunitaria mediante el uso de AIS como modelo artificial.

Este estudio se enfoca en la construcción de sistemas de seguridad que protegen aquello que el mismo sistema define como el ser (*self*) de los agentes externos (*non-self*), siendo Self lo que forma parte del Sistema de Información y Non-Self los agentes externos al mismo. El propósito de un sistema inmune es poder analizar el *non-self* y entender si es potencialmente dañino o malicioso para el *self* [5]. Esto resulta de aplicar una aproximación metafórica a los AIS.

Los sistemas inmunológicos artificiales incorporan mecanismos conocidos de los organismos vivos y usan metáforas para desarrollar una solución a un problema particular de seguridad.

Los mecanismos que se han analizado son distintas implementaciones del algoritmo AIS, donde cada uno utiliza una técnica diferente:

Solución por hardware: Un dispositivo físico analiza cada una de las instrucciones que el procesador recibirá con el fin de determinar qué instrucciones son seguras para ejecutar.

Solución basada en la nube: Un agente intermediario se comunica con un servidor en la nube para consultar la seguridad de cada archivo sospechoso.

Solución colaborativa: Un conjunto de computadoras conectadas en red comparten la información de las nuevas amenazas que han detectado con el propósito de aprender más rápido en conjunto.

Solución basada en BIS humano: Emulación del sistema inmunológico del ser humano para defenderse ante las posibles amenazas.

Todas estas soluciones tienen en común que usan un algoritmo de selección negativa, el cual consiste en discriminar lo propio del sistema de lo que es externo, y sobre lo que el sistema detecta como externo se hace una evaluación para determinar si es seguro o no ejecutarlo. Si se diagnostica que no es seguro, se genera un detector que servirá en un futuro para no tener que volver a analizar el archivo y poder definir automáticamente si es o no seguro. El beneficio de generar estos detectores es que pueden compartirse de modo que se acelerara el aprendizaje de otras terminales sin necesidad de que cada una de ellas deba analizar el archivo para determinar que no es seguro y generar el detector [6].

Como se ha mencionado, los BIS están compuestos por múltiples subsistemas y procesos que protegen al organismo. Se conocen como patógenos a los agentes que invaden el sistema y antígenos a las sustancias que, al ingresar al sistema, generan respuestas por parte del BIS. Todo BIS puede dividirse en dos ramas principales: el sistema inmunitario innato y el adaptativo.

El sistema innato no conoce invasores específicos sino que provee protección general al sistema contra patógenos que ingresan al cuerpo [3]. Si bien no es un

[2] Simulado por computadora.

sistema completo, tiene la capacidad de detectar amenazas y generar respuestas, siendo principalmente estático.

Por el contrario, el sistema adaptativo (también llamado adquirido o específico) permite al sistema inmunitario lanzar ataques contra agentes invasores que el sistema innato no puede combatir [7]. El sistema adquirido está dirigido a esos agentes, a los que aprende a reconocer bajo exposición gracias a los linfocitos (un tipo de glóbulos blancos). Se considera que las respuestas del sistema adaptativo son específicas de cada antígeno pues responden diferente a cada uno, una vez detectado.

En general, los estudios sobre AIS están enfocados en el sistema adaptativo, por su capacidad para aprender y memorizar respuestas a diferentes agentes peligrosos, además de tener la capacidad de reaccionar rápidamente a nuevos ataques a los que nunca había sido expuesto.

En la próxima sección haremos referencia al malware que representa lo que para un BIS llamamos patógeno: un tipo de software que tiene como objetivo infiltrarse en un sistema de información para extraer datos o para corromper a los mismos.

3. AIS Aplicado

A continuación se describe el funcionamiento de cuatro aplicaciones diferentes de AIS: Colaborativo, Basado en Seguridad en la Nube, Implementado por Hardware y Detección de Código Basada en el BIS Humano.

3.1. AIS Colaborativo

Esta solución se conoce como Cuerpo Inmunitario Colaborativo (ICB, del inglés Immune Collaborative Body) y cada computadora que forma parte del sistema se denomina Cuerpo Inmunitario (IB, del inglés Immune Body). La implementación consiste en un grupo de computadoras con ambientes similares comunicadas entre sí mediante un canal inmunitario, como se puede apreciar en la Figura 1.

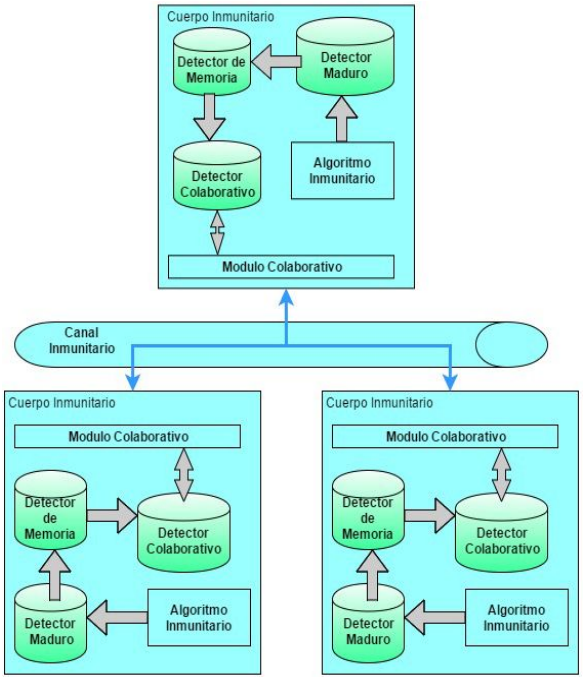

Fig. 1 - Cuerpo Inmune Colaborativo

El canal inmunitario permite a cada IB compartir sus linfocitos generados a lo largo de su ejecución para mejorar la eficiencia y la tasa de detección de todo el sistema, sin perjudicar la diversidad y auto adaptabilidad de cada sistema individual. Por lo tanto, si una computadora detecta, con el algoritmo de selección negativa, malware que no es parte de la base de firmas, el AIS generará detectores y los compartirá con el resto de las

computadoras del sistema para reducir el tiempo de entrenamiento de cada IB conectada a la red. Esto se traduce en una reacción más rápida ante el próximo intento de intrusión por parte del mismo malware, lo que permite una mejora en la cobertura ante las epidemias que se propagan rápido en los sistemas [8].

3.2. Seguridad en la Nube Basada en AIS

Este modelo utiliza un agente anfitrión en cada computadora, el cual no posee una base de firmas local como los antivirus tradicionales. El agente anfitrión ("host agent" en inglés), se encarga de categorizar el código ejecutable como sospechoso o no sospechoso, enviando únicamente los sospechosos a un servidor en la nube, que evalúa según distintas bases de firmas, si el ejecutable es malware. Si no es parte de ninguna base de firmas, se revisa la lista blanca (o "whitelist", donde se guardan las excepciones establecidas por el usuario) y si tampoco se encuentra allí, se ejecuta un algoritmo AIS de selección negativa con el objetivo de determinar si el ejecutable es potencialmente dañino, a partir de su comportamiento.

En caso de encontrar el código malicioso en las bases de firmas o por el algoritmo de AIS, se envía una notificación al anfitrión para que el mismo se encargue del tratamiento de dicho ejecutable, ya que el sistema presupone que es una potencial amenaza tal como se muestra debajo, en la Figura 2.

El anfitrión también puede disponer de una memoria caché para almacenar las últimas detecciones realizadas y evitar recurrir al servidor en la nube cada vez que el mismo malware intente ingresar al sistema.

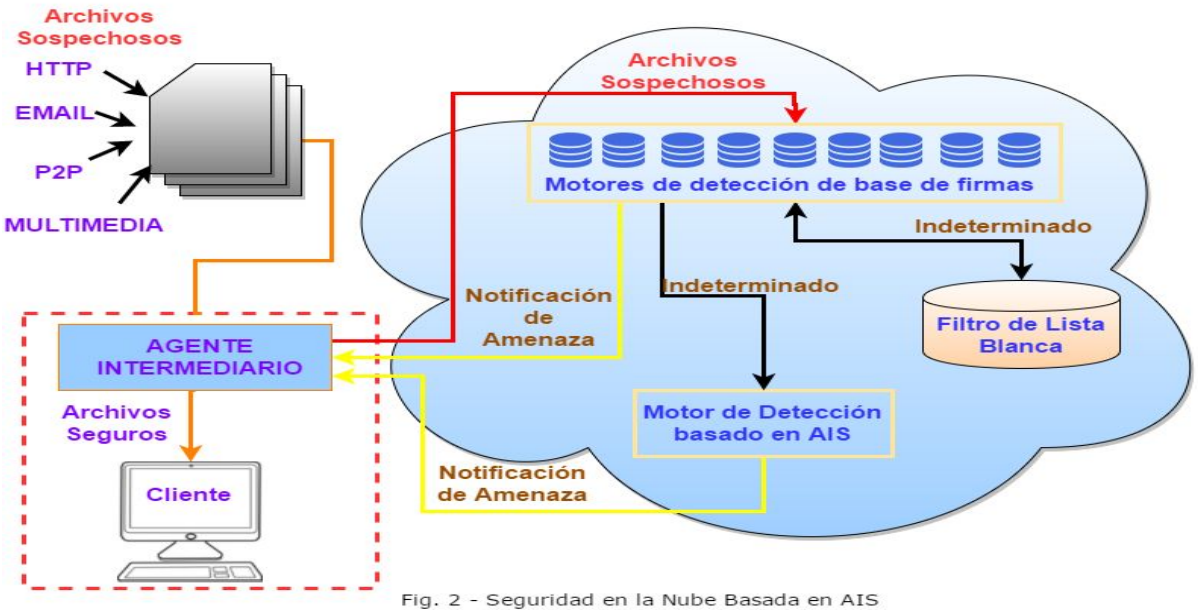

Fig. 2 - Seguridad en la Nube Basada en AIS

En este método, la ventaja principal es que el usuario no debe preocuparse por mantener la base de firmas actualizada, ya que la misma reside en un servidor. Además, se disminuyen los tiempos de procesamiento al delegar el análisis a la nube y no realizarlo de forma local. Como contraparte, si no se dispone de una conexión estable a una red, este tipo de solución no puede implementarse [9].

3.3. Implementación Hardware de AIS

Esta solución de AIS añade al procesador principal de una computadora, un dispositivo de hardware que actúa como mecanismo inmunitario para defender al sistema contra los ataques maliciosos. Esta tecnología permite delegar el procesamiento de AIS a otro dispositivo de hardware independiente para no disminuir el rendimiento general del sistema.

El dispositivo hardware dedicado intercepta las instrucciones que deberá ejecutar el procesador principal y obtiene la información necesaria para detectar comportamiento malicioso. La intercepción de instrucciones se realiza durante los tiempos de acceso a memoria, haciendo transparente para el procesador principal, la actividad del dispositivo AIS.

El código de operación de las instrucciones y la dirección objetivo transmitidas por el decodificador de direcciones del sistema son unidas en una única secuencia binaria. Las secuencias son enviadas al hardware AIS, que realiza una evaluación del comportamiento y, en el caso de detectar comportamiento de tipo malware, genera una advertencia al procesador de propósito general [10].

3.4. Detección de malware inspirada en el BIS humano

La detección de código malicioso inspirada en el BIS humano fue analizada en múltiples investigaciones. Esta sección se enfoca, en particular, en una implementación llamada MaCDI [11].

MaCDI propone dos fases claramente diferenciadas: La fase adolescente (identificada con el sistema inmunológico innato) y la fase adulta (relacionada con el sistema inmunológico adquirido).

Cuando un nuevo ejecutable ingresa al sistema, se lo considera en fase adolescente (ver Figura 3) y todas las llamadas al Sistema Operativo[3] que invoca son comparadas con un perfil de malware. Si alguna llamada o invocación coincide con alguno de los patrones presentes en dicho perfil, la ejecución finaliza, el ejecutable cambia a estado de cuarentena y se actualiza el perfil de malware.

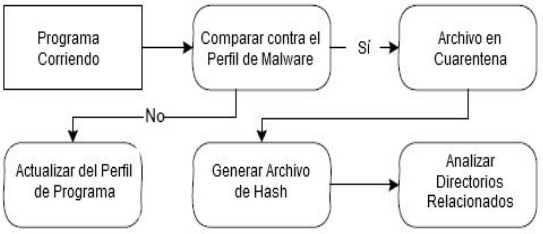

Figura 3. Fase Adolescente.

El programa ejecutable cambia a fase adulto dependiendo de las siguientes variables:

[3] En inglés System call o SysCall, es el mecanismo usado por una aplicación para solicitar un servicio al Sistema Operativo.

a. Cantidad de tiempo que el programa ha estado ejecutándose.
b. El porcentaje de código que ha sido ejecutado (naturalmente no todo el código de un programa se ejecuta siempre pero cuanto más haya sido analizado, mayor será el valor de esta variable).
c. El número de intentos fallidos para agregar un nuevo patrón de llamada a sistema excede un límite.

En la fase adulta (ver Figura 4) no se realiza la comparación sobre el perfil de malware, sino que se hace sobre un perfil de programa. Todas las llamadas al sistema que realice serán cotejadas con el perfil del programa y, si el programa comienza a comportarse de una forma que no es la esperada, vuelve a la fase adolescente.

Figura 4. Fase adulta.

Si un proceso se cambia a estado de cuarentena durante la fase adolescente, el escáner se activa durante intervalos para ejecutar una operación de limpieza. Además, se encarga de buscar los archivos relacionados con el programa y reubicarlos en un espacio de cuarentena. Por último, se realiza una copia de seguridad de los archivos infectados antes de eliminarlos. Si el proceso en cuarentena no cumple con los patrones de malware, se restauran los archivos y eliminan las copias de seguridad.

4. Comparación entre los sistemas relevados

En la sección anterior se han desarrollado las principales características de cuatro implementaciones distintas de sistemas que aplican AIS.
En la siguiente tabla se puede apreciar la comparación entre dichos sistemas, desde los siguientes enfoques:

- La conectividad necesaria para su funcionamiento.
- El costo de implementación.
- El uso de base de firmas para poder realizar una comparación con las amenazas existentes.
- El tipo de código que analiza (instrucciones individuales o ejecutables completos).
- La velocidad con la que se adaptan a nuevas amenazas.
- La capacidad de almacenamiento requerida de forma local.
- La posibilidad de procesar en paralelo, sin interrupción de la ejecución local de procesos.
- Impacto en el rendimiento general del sistema.

Tabla I: Comparativa entre Implementaciones AIS

	Colaborativo	Basado en la nube	Implementación por hardware	Basado en BIS humano
Red necesaria	Intranet	Internet	Ninguna	Ninguna
Costo	Bajo	Alto	Muy Alto	Bajo
Utiliza base de firmas	Sí	Sí	No	Sí
Analiza	Ejecutables	Ejecutables	Instrucciones	Instrucciones
Velocidad de aprendizaje	Rápida	Muy Rápida	Lenta	Lenta
Almacenamiento utilizado	Mucho	Poco	Poco	Mucho
Procesamiento paralelo	No	Sí	Sí	Sí
Impacto en el Rendimiento del Sistema	Medio	Bajo	Muy Bajo	Alto

La implementación colaborativa tiene como ventaja ser una solución que no requiere acceso a internet, sino que únicamente necesita una red de área local para funcionar, esto lo define como una solución económica. Al ser varias computadoras trabajando en conjunto, la velocidad de aprendizaje se vuelve rápida. Pero también conlleva la desventaja de tener que guardar la información generada a lo largo de la ejecución en unidades de almacenamiento internas por lo que requiere una importante cantidad de espacio lógico para funcionar.

La solución Basada en la Nube tiene como ventaja una alta velocidad de aprendizaje ya que el servidor recibe archivos sospechosos de muchos dispositivos (inclusive dispositivos correspondientes a otros clientes o empresas) y esa característica hace que la velocidad de aprendizaje aumente drásticamente. Como desventaja se puede mencionar que hace permanente uso de conexión de internet por lo que se verá una penalización en la velocidad de subida y de bajada así como también se sufrirá una imposibilidad de acceder al servicio en situación donde no se disponga de conexión a internet. La otra desventaja es que el servidor en la Nube tiene un costo asociado que puede resultar bastante elevado comparado con otras soluciones que únicamente constan de software. La tecnología AIS basada en hardware también tiene el costo del dispositivo físico. La implementación por hardware tiene como ventajas una alta velocidad de detección, ya que en tiempo real analiza instrucción por instrucción y que no necesita ningún tipo de conexión. Esta última característica es una desventaja también ya que al trabajar individualmente

el tiempo de aprendizaje es más lento que los dos esquemas anteriormente nombrados. Otro inconveniente es que habría que agregar a los procesadores un dispositivo hardware lo cual incrementa los costos de la solución en gran medida. Otra desventaja es que solo dispone de un análisis AIS mientras que el resto de las soluciones también se apoyan en un análisis de base de firmas. La implementación inspirada en el BIS humano tiene como ventaja no depender de internet para funcionar, pero esto genera la utilización de mucho espacio de almacenamiento. Otra desventaja que posee es que cuenta con una velocidad de aprendizaje lenta ya que si un ejecutable no se comporta de la manera esperada mientras está en la fase adulto, regresa a la fase adolescente, donde nuevamente se lo compara con un perfil de malware.

Entonces, se podría afirmar que no hay una solución que sea mejor que otra, sin embargo hay soluciones que son más adecuadas al problema que tiene un usuario en particular. Si nos enfocamos en seguridad para empresas, estimamos que las dos mejores alternativas sería implementar AIS Basada en la Nube o Colaborativa porque garantizan una alta velocidad de aprendizaje, que es fundamental para dicho segmento de mercado ya que cualquier pérdida o robo de información puede resultar en pérdidas millonarias. Evaluando el caso del usuario particular, creemos que la mejor alternativa es AIS inspirada en el BIS humano ya es una solución de bajo costo, que funcionan individualmente y que probablemente trabajará en simultáneo con el antivirus local de esa computadora.

5. Conclusión

Las amenazas a las que están expuestos los sistemas se incrementan constantemente a gran escala y es muy importante contar con tecnología que pueda dar respuesta. La sofisticación de los atacantes ha ido incrementando, incluso a una velocidad superior que la de creación de mecanismos de defensa tradicionales. Por eso, es de vital importancia desarrollar e implementar técnicas cuyo objetivo sea contrarrestar ataques y defenderse a sí mismos de nuevas amenazas, sobre todo de aquellas que no podrían ser previstas antes de su primera aparición.

Para lograr esos objetivos, una solución emergente son los Sistemas Inmunes Artificiales para permitir a los sistemas informáticos defenderse, no sólo de amenazas ya conocidas, sino de ataques nunca antes considerados.

En este trabajo se realizó un relevamiento de cuatro sistemas de seguridad que implementan AIS de formas distintas. Se elaboró un cuadro siguiendo una serie de criterios que permiten apreciar las principales características de las soluciones presentadas.

Por último, es crucial remarcar la necesidad de tomar una actitud proactiva frente a la situación, subrayando la importancia de trabajar para lograr un desarrollo eficiente y así reducir los riesgos inherentes a la interconectividad.

6. Trabajo Futuro

Consideramos necesario continuar con la búsqueda de tecnologías AIS, tanto de futuras como otras no relevadas en esta investigación, con el fin de ampliar la base de conocimiento.

Las implementaciones más prometedoras para las empresas son Colaborativa y Basada en la Nube debido a su rápida velocidad de aprendizaje. Es por ello que seguir avanzando sobre dichas alternativas parece ser la decisión más acertada debido a que las empresas son organizaciones que necesitan poner especial énfasis en la protección de sus datos.

Referencias

[1] Paul K. Harmer, Paul D. Williams, Gregg H. Gunsch, and Gary B. Lamont, "An Artificial Immune System Architecture for Computer Security Applications", IEEE transactions on evolutionary computation, vol. 6, no. 3, 2002.

[2] M. Leon, "Internet virus boom", Infoworld, vol. 22, no. 3, pp. 36–37, Jan. 2000.

[3] C.A. Janeway Jr., R. Medzhitov, "Innate immune recognition", Annual Review of Immunology 20 (2002) 197–216.

[4] Irun Cohen, "Real and artificial immune systems: computing the state of the body", 2007. Nature reviews immunology.

[5] Gilang Ramadhan, Yusuf Kurniawan, Chang-Soo Kim, "Design of TCP SYN Flood DDoS Attack Detection using Artificial Immune Systems", 2016. IEEE 6th ICSET.

[6] J. Timmis, A. Hone, T. Stibor, and E. Clark, "Theoretical advances in artificial immune systems", Theoretical Computer Science, vol. 403, no. 1, pp. 11 – 32, 2008.

[7] Charles A. Janeway Jr., Travers Paul, "Immunobiology: The Immune System in Health and Disease", third ed., Garland Publishing, New York, 1997.

[8] Yang He, Liang Yiwen, Li Tao, Zhou Ting, "A model of Collaborative Artificial Immune System", IEEE 2nd International Asia Conference on Informatics in Control, Automation and Robotics, 2010.

[9] Xufei Zheng, Yonghui Fang, "An AIS-based cloud security model", International Conference on Intelligent Control and Information Processing, August 13-2015.

[10] Zhaojun Lu, Gen Pei, Bojun Liu, Zhenglin Liu, "Hardware implementation of negative selection algorithm for malware detection", IEEE International Conference on Electron Devices and Solid-State Circuits, 2015.

[11] Mohd Fadzli Marhusin, David Cornforth, Henry Larkin, "Malicious Code Detection Architecture Inspired by Human Immune System", IEEE International Conference on Intelligent Control and Information Processing, 2008.