خوارزمية اكتشاف دودة الانترنت معتمدة على سلوكها

محمد ماهر رشيد
وزارة العلوم والتكنولوجيا / دائرة تكنولوجيا المعلومات
بغداد –العراق

الخلاصة

تنتشر دودة الانترنت بشكل آلي عبر الإنترنتَ في مدة قصيرة جداً. تُلحقُ الديدانُ أضراراً في الشبكةِ باستهلاك مصادرِها مثل سرعة نطاق الانترنت على حساب كفاءته. كما هو معروف ان الإنترنتَ له وظيفةُ مؤثرةُ في الاقتصاد ويعد ركنا أساسا في الحياةِ. وبتوقفه ولو لساعاتُ فسوف يسبب خسائر اقتصادية كبيرة. اقترح في هذا البحث خوارزمية تقوم على اكتشاف ديدان الانترنت من خلال سلوكها. ان البحث يقوم بالكشف عن الديدان السريعة والبطيئة الانتشار. اذ ترسل الدودة نفسها الى الحاسبات الأخرى بواسطة الانترنت، ويترتب على ذلك الارسال، فشل كثير من الاتصالات. إن معدلَ الاتصالات الفاشلة هو العاملُ الرئيس الذي يعتمدُ عليه في هذا البحث لاكتشاف الدودةِ. أظهرت النتائج أن الخوارزمية المقترحة تستطيعُ أَن تَكتشفَ أنواعَ جديدةَ مِن الديدانِ. فضلا عن كونها أسرعُ مِن الخوارزميةِ التقليديةِ في اكتشاف ديدان الانترنت.

الكلمات المفتاحية: اكتشاف الديدان، فشل الاتصال وجدار ناري.

Internet Worm Detection Algorithm Based on its Behavior

Mohammad Maher Rasheed
Ministry of Science and Technology/ Information Technology Directorate
Baghdad- Iraq
E-mail: mohmadmhr@yahoo.com

Abstract

Internet worm are spread automatically and can spread on its the hosts in a very short time. Worms may cause congestion in the network, which leads to large queuing delays, and high packet loss. The Internet is an influential function in the economy and reckon mainstay to life. Once the internet is broken down, it will cause a huge economic loss. Internet worm itself to another computer, this way generates a lot of failure connections. The average of failure connections is the main factor that the research technique depends on in detecting Internet worm. In this research the reswlts showed that proposed algorithm can detect a new types of worm that were fast or slow spread. Moreover, this research showed that proposed algorithm detection result was faster than traditional algorithm for detecting the worms.

Keywords: Worms Detection, Failure Connection and Firewall.

المقدمه

ان دودة الانترنت هي عبارة عن برنامج حاسوبِ ذاتي الاستنساخِ حيث تستعملُ الشبكة لإرسال نسخِ من نفسها إلى محطات حاسوبِ طرفية عن طريق الشبكةِ ويَعملُ ذلك بدون أي تدخل للمستخدم. حالياً، ديدان الانترنت تشكل تهديدًا على أمنِ المعلومات الذي قَد تُتسببُ بتباطؤ الشبكةِ والذي يؤدي إلى الانتظار والتأخير في استجابة عمل الانترنت، ويؤدي إلى خسارة حزمةِ عاليةِ من الانترنت. تعد دودتي Codered و Nimda التي نُشِرتا في 2001، من الديدان التي تسبب أضرارَ مادية كبيرة. أن اكتشاف الدودةِ واحتوائها يَجِبُ أَن يُكونَ سريع ولا يجعل لها أي فرصة للنجاحِ لأن ديدان الانترنت انتشارها سريع جداً (Costa *et al*., 2005) و يؤثر الإنترنتَ كوظيفةُ في الاقتصاد كونه ركنا أساسيا في الحياةِ. إن أي ضرر في خدمة الإنترنت ولو لساعاتُ، سوف يتسبب بخسارة اقتصادية كبيرة Schechter *et al*., (2014).

على خلاف الفيروس، الديدان لَيست بِحاجةٍ إلى أَن تَربطَ نفسها ببرنامجِ حالي. الديدان يُمكِنُ أَن تعمل بالكامل بشكل مستقل وخلال الشبكة، بينما الفيروس يَحتاجُ ملف مضيفِ حيث يستعمله للمُكَاثَرة (Computer Worms Information, 2008).

تعتبر العتبة العامل الرئيسي الذي يعتمد علية في اكتشاف دودة الانترنت ومعناه هو اكتشاف عدد محدد من فشل الاتصال في وقت محدد. حيث ان زيادة العتبة يعطي بطا في اكتشاف دودة الانترنت وتقليله يعطي إنذارات كاذبة.

تَكتشفُ الطرقُ التقليديةُ دودةُ الإنترنتِ المجهولةِ بسلوكيتها ولكنها ما زالَت تَتباطأُ في اكتشاف دودةِ الإنترنتِ ولا تَستطيعُ اكتشاف كُل أنواع الديدان. Chen and Tang, (2004) بَنى الخوارزميةَ لاكتشاف الدودةِ المجهولة، لكن نسبةَ العتبة المحسوبة تَأخذ وقت طويلَ لاكتشافها.

Schechter et al., (2004) قَدم نظرة هجينة لاكتشاف ديدان الانترنت ولكن الخوارزمية لا تَعملُ بشكل صحيح لاكتشاف الدودةِ بطيئة الانتشار.

Yang *et al*., (2006) بَنى خوارزمية لاكتشاف الدودةِ والمتكونة من خوارزميتينِ، الخوارزمية الأولى " خوارزمية المدى القريب " لاكتشاف الدودة التي هي سريعة الانتشار نوعا ما، لكن الخوارزميةَ الثانيةَ " خوارزمية المدى الأطولِ " لاكتشاف لا تَستطيعُ اكتشاف بَعض أنواعِ الديدان "الخلسة". الخوارزمية أيضاً لا تَستطيعُ حساب عتبة الوصول وهي ليست مرنة مع الوقت اعتمادا على عدد فشل الاتصال في كل دقيقة لذلك هناك عتبة واحدة ويجب الوصول لها لاكتشاف دودة الانترنت والتي يكون اكتشافها أكثر من دقيقة.

Zou *et al*., (2005) قدم " كشف اتجاه " للاكتشاف دودة في مرحلةِ توليدِها المبكرةِ باستعمال " مرشِح Kalman " ويعتمد هذا المرشح على الانظمة الديناميكية الخطية.

Chen and Tang, (2004) اقترحا نظاما معتمدا على فشلِ طلبِ الاتصال الذي يتم استلامه من خلال الموجهة (Router) في الشبكة. وكانا قد اقتراحا نظام يسمى (DAW) والذي هو نظام مضاد لديدان الانترنت والذي يعمل على تباطأُ الدودة آلياً أَو تُوقفُ انتشارها. ولكن فترة اكتشافها طويلة قد تصل الى أكثر من شهر وأخيرا Rasheed *et al*. (2012) صمم خوارزمية تعمل على اكتشاف دودة الانترنت التي تستخدم في المسح بروتوكول TCP ولكنها لا تستطيع ان تكتشف الدودة البطيئة الانتشار.

يهدف البحث الحالي على اقتراح معادلة جديدِة لحساب العتبة أحد عناصرها معدل الفشل فضلا عن اعتماد تغير مدى العتبةَ على اختلاف الزمن لاكتشاف الأنواع الجديدةِ مِن الديدان. وصولا الى ان الخوارزمية المقترحة يُمكِنُ أَن تَكتشفَ دودة الانترنت بطيئة الانتشار.

الية عمل الخوارزمية المقترحة

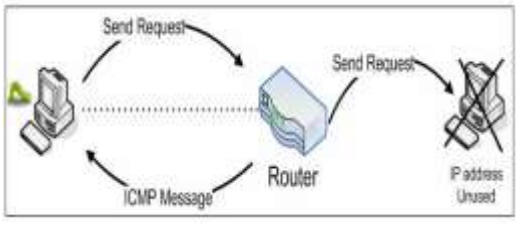
تعتمد خوارزميتنا على نوعيين من البروتوكولات لاستلام رسائل الفشل الأول هو Transmission Control Protocol (TCP) والثاني هو Internet Control Message Protocol (ICMP).

التقنية المتبعة في هذا البحث هو جمع رسالة الخطأ التي يتم استلامها من الموجهة مثل رسالة الفشل في بروتوكول TCP حيث يستلم Reset /Acknowledgment (RST /ACK) وفي رسالة فشل اخرى يتم استقبالها من بروتوكول اخر هو ICMP يستلم رسالة الفشل "Destination Unreachable" ورسالة أخرى هي "Time Exceeded" والتي تستقبل هذه الرسائل من الموجه (Router).

ICMP "Destination Unreachable"

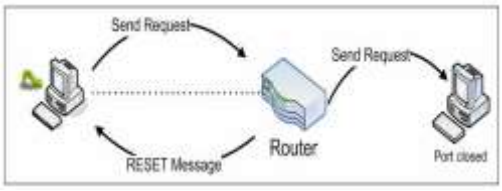
تستلم هذه الرسالة عندما ترسل الدودة الطلب من الحاسبة الضحية الى الحاسبة المقصودة ويكون عنوان الحاسوب المقصود غير موجود او غير مستعمل وكما ذكر (Ellis, 2004).

و"ICMP "Time Exceeded هي رسالة يتم استلامها عندما لا يستجيب الحاسوب المقصود الى الدودة فيقوم الموجهة بأرسال هذه الرسالة الى الحاسبة التي تحتوي على دودة الانترنت، انظر إلى الشكل (1).



شكل (1) فشل الاتصال ICMP

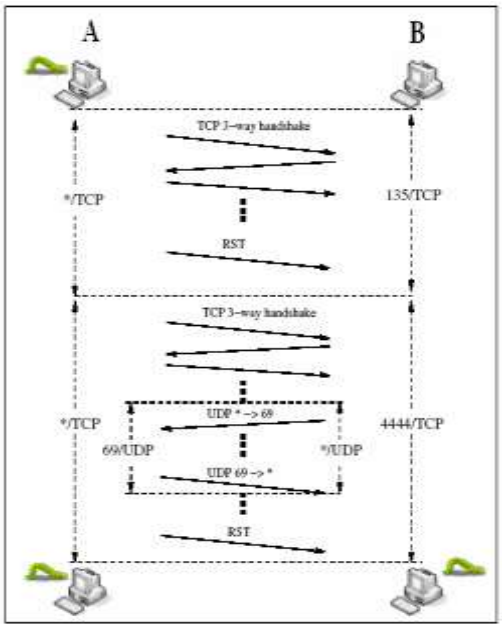
عندما تُرسَلُ حزمة إلى عناوين IP وهذه العناوين مستعملة ولكن المنفذ Port مغلق للحاسبة الضحية، ان الحاسبة المصابة تستلم رسالة Reset (Ellis, 2004)، (انظر إلى الشكل 2).



شكل (2) فشل الاتصال Reset

تَجعلُ هذه الخوارزميةِ اختلافا بين الاتصال المنتظمِ واتصال الدودةِ. تَمسحُ الدودةُ عناوين IP المختلفةَ في كل ثانية. تَعتمدُ الخوارزميةُ على رسائل فشل الاتصال على العناوينِ العشوائيةِ المختلفةِ التي يتم توليدها من الدودة وبذلك يكون هناك عدد كبير مِن فشل الاتصال إذا كان الحاسوبَ يحتوي على دودة الإنترنت.

الاتصال الناجح لدودةِ الانترنت كما في الشكل (3) قد تَقبلُ من الحاسوبِ الآخرِ لأن الدودة تُولدُ عنوان IP عشوائيَ وهذا العنوان هو في الحاسوب B. ولكن هناك شرطَ أخر مهمُ جداً لنسخ الدودةِ مِن الحاسوبِ A إلى الحاسوبِ B عندما يكون المنفذ (Port) للحاسوب B مفتوحُ.



شكل (3) سلسلة عدوى دودة الانترنت Jiang and Xu, (2006

تُسجل الخوارزمية المقترحة في هذا البحث عددُ الاتصالات الفاشلِة في العداد مثل رسائلICMP فشل الاتصال أو Reset وهذه الرسائل يستقبلها الموجه (Router) عندما لا يستطيع الوصول إلى العناوين أي إن العنوان غير مستخدم أو المنفذ مغلق. الخوارزمية يَجِبُ أَن تَكُونَ مكانها في الموجهة Router (انظر إلى الشكل 4). لكشف فشل الاتصال، الخوارزمية تعتمد على (العنوان المصدري، المنفذ المصدري، العنوان المقصود، المنفذ المقصود) من الحزمة.

العداد في الخوارزمية يسجلُ رُزَمَ الاتصالات الفاشلةِ التي عادة من الاتجاه الخارجي للعناوين IPإلى العناوين IP المصدري الداخلي وجمع الاتصالات خلال دقيقة واحدة. لكي يكون الاكتشاف أسرع.

حيث تقوم الخوارزمية بحساب عدد الفشل في الاتصال وبالتالي تقوم الخوارزمية بزيادة "العداد" ويأخذ بنظر الاعتبار إذا كان هنالك نجاح في الاتصال Synchronize/Acknowledgment (SYN/ACK) فتقوم الخوارزمية بنقصان ال عداد"CFC". ويسجل العداد عنوان المقصود Destination IP الذي تم الفشل فيه ما عدا ذلك، تقوم الخوارزمية بإهمال تسجيل الفشل أذا كان عنوان المقصود مسجلُ فيه أكثر من فشل لان دودة الانترنت تبحث دائما عن عنوان أخر مختلف. فالفشل الذي يكون بنفس العنوان يكون فشل طبيعي وبعيد عن تهديد دودة الانترنت.

الخوارزمية المقترحة تفترض: –

CFC = عداد فشل الاتصال.

HC= العنوان المؤرخ للاتصال الناجح والفاشل.

AFC= عدد فشل الاتصال مقسوم على الوقت من بدا اكتشاف فشل اتصال الى الوقت الحالي (وبالدقائق).

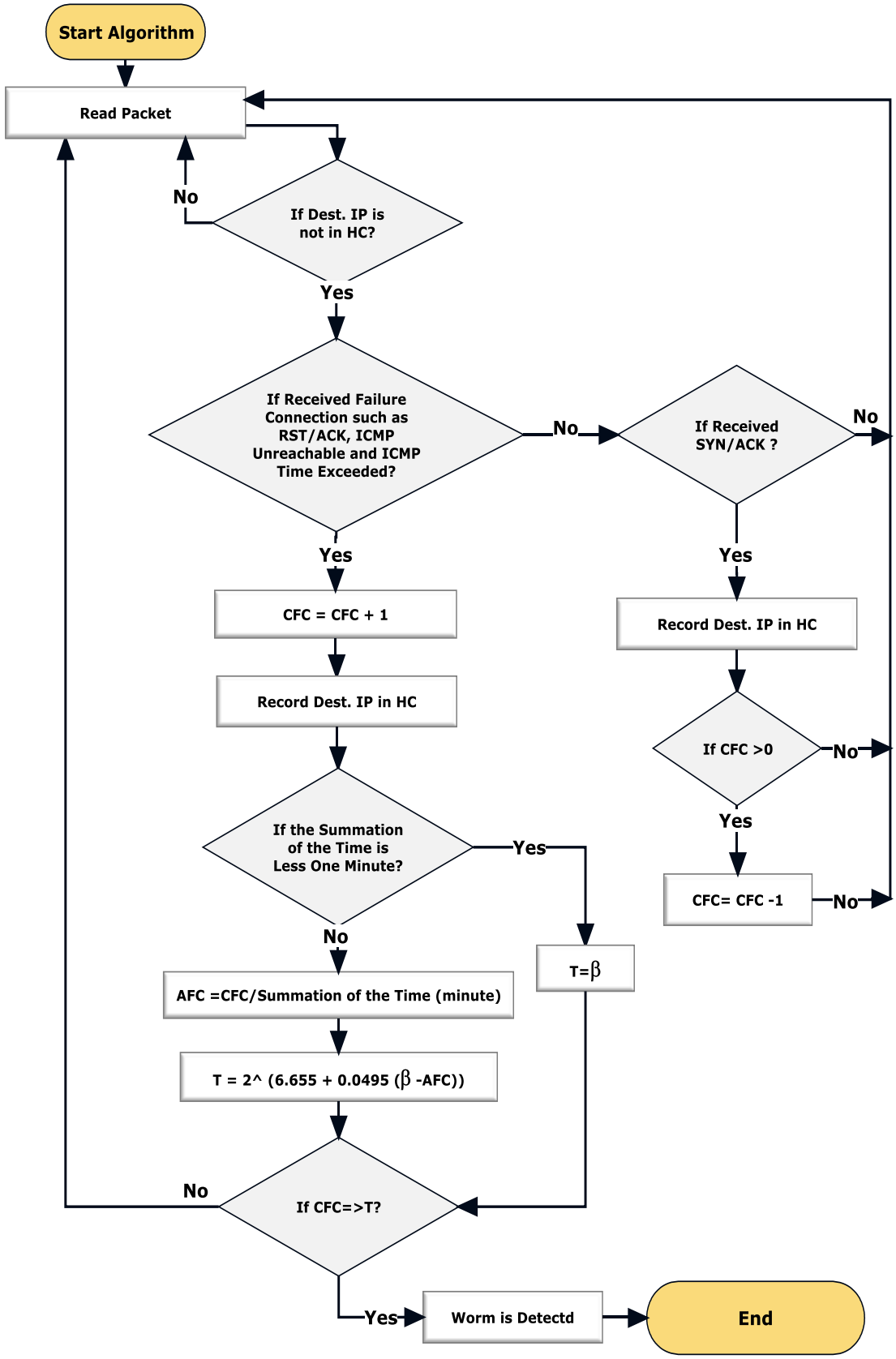
β= 101 فشل اتصال.

مجموع العتبة T في "خوارزمية المدى البعيدِ كما في المعادلةِ التالية: –

$$T=2^{(6.65+0.0495(\beta+AVC))}$$

أن الاعداد الثابتة في الخوارزمية تعتمد على عتبة الخوارزمية للبحث (Yang *et al.*, 2006) ولكن تم انشائها بطريقة الاعتماد على عتبات مختلفة فتزداد سرعة الخوارزمية في كشف الديدان.

إن المعادلةَ تعتمدُ على معدل الاتصال الفاشلِ لحِساب العتبةِ. الخوارزمية المقترحة يمكن أَن تَكتشفَ الدودةَ مبكراً في الوقتِ العاديِ. لكن إذ لم تستطيعُ الخوارزمية الاكتشاف في المرحلةِ المبكرةِ، تُزودُ الخوارزميةَ وقتَ أكثرَ لاكتشاف الدودة كما في الشكل(4).



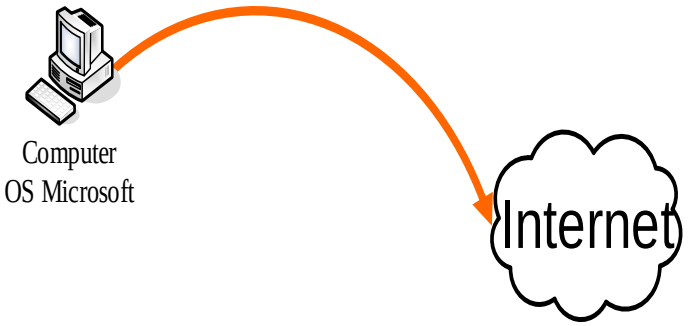
شكل(4) المخطط الانسيابي للخوارزمية المقترحة

منهجية البحث

في المواد وطرائق العمل سوف يتم اعداد البيئة للتقييم والتحقق من صحة الخوارزمية وبعد ذلك سوف يستعرض البحث نتيجة الخوارزمية في الإنذارات الكاذبة ومن ثم مقارنة النتائج في اكتشاف ديدان الانترنت. وقد اعتمد الاختبار على دودةَ واحدة فقط واختبرت على سرعات مختلفة للانتشار لأن كُل ديدان الإنترنتِ لَها نفس خصائص مِن اتصال الفشلِ.(Ellis, 2004)

المواد وطرائق العمل

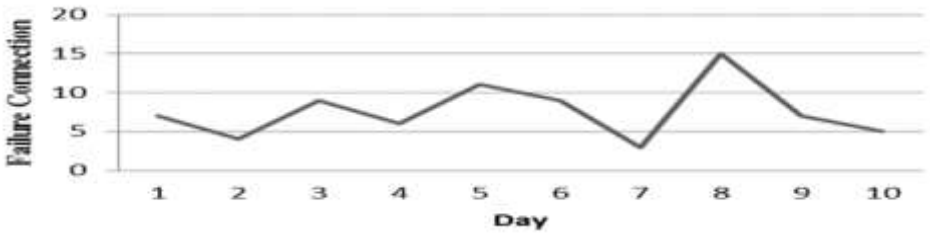
لأعداد البيئة لتقييم والتحقق من صحة الخوارزمية فقد تم تنصيب نظام التشغيل مايكروسوفت 2000 نسخة 4 Service Pack. وكانت الحاسبة متصلة بشبكة الإنترنت بسرعة نطاق 3.6 ميجابايت في الثانية، وكما هو مبين في الشكل(5).



شكل(5) أعداد بيئة الخوارزمية المقترحة

نتائج الخوارزمية المقترحة في الإنذارات الكاذبة

جربت الخوارزمية المقترحة في بيئة غير مصابة بأي فايروس وكان مستخدم الحاسبة يمارس التصفح في المواقع كالفيسبوك واليوتيوب والياهو ماسنجر وكانت نتيجة الخوارزمية المقترحة 15 فشل لليوم الواحد وكان هذا الرقم هو الحد الأقصى، وبلغ المجموع لمدة عشرة أيام 76 اتصال فاشل. علاوة على ذلك، كانت عتبة الخوارزمية المقترحة للكشف عن الفشل في شبكة الإنترنت تعادل 101 . اتصال فاشل بالدقيقة وهنالك عتبات اخرى. تم فحص الخوارزمية المقترحة لمدة عشرة أيام، ولم يتم إطلاق أي تحذير من قبل الخوارزمية المقترحة. وكان متوسط الاتصال الفاشل في اليوم الواحد 7.6 فشل يوميا، وكان الفشل منخفض مقارنة مع الخوارزميات الأخرى. والشكل (6) يوضح عدد الفشل اليومي.



شكل(5) فشل الاتصال اليومي في الخوارزمية المقترحة وعدم وجود أي انذار كاذب

مقارنة النتائج في اكتشاف ديدان الانترنت

قورنت الخوارزمية المقترحة مع الدراسة Yang *et al.*, (2006) وأظهرت المقارنة أن الخوارزمية المقترحة كانت أسرع من خوارزمية Yang *et al.*, (2006) حيث كانت الخوارزمية المقترحة تحتوي على عتبات مختلفة كما في الجدول(1). علاوة على ذلك، تعتمد الخوارزمية المقترحة على ثلاث أنواع من رسائل الفشل وهي Reset وICMP Destination Unreachable وICMP Time Exceeded، ولكن خوارزمية Yang et al., (2006) تعتمد على نوعيين اثنين من رسائل الفشل هما Destination Unreachable ICMP و / RST ACK .

النتيجة أن الخوارزمية المقترحة أسرع من خوارزمية Yang et al. (2006)، في جانبين. الأول، الخوارزمية المقترحة لديها عتب متعددة، لذلك السبب يعتبر الوصول أسرع، وكذلك ان الخوارزمية المقترحة تعتمد على ثلاثة أنواع من رسائل الفشل فهي الأسرع للوصول الى العتبة هذه الميزتان جعلت الخوارمية اسرع مقارنة مع الخوارزميات الاخرى.

جدول (1) الفرق في سرعة اكتشاف دودة الانترنت بين Yang et al., (2006) والخوارزمية المقترحة

الوقت اللازم لاكتشاف دودة الانترنت في خوارزمية Yang *et al. (2006)*	العتبة في خوارزمية Yang *et al.* (2006)	الوقت اللازم لاكتشاف دودة الانترنت من قبل الخوارزمية المقترحة	عتبة الخوارزمية المقترحة	معدل الفشل التي تولدها الدودة وبالدقائق
1801 Second	3001	62 Second	104	100
1819 Second	3001	65 Second	108	99
1837 Second	3001	69 Second	112	98

References

Alagna, T. and Schmidt, H., (2005) The Black Book on Corporate Security, Chapter 7, Larstan Publishing.

Chen, S. and Tang, Y., (2004) Slowing Down Internet Worms. In Proceedings of 24th International Conference on Distributed Computing Systems (ICDCS'04). Tokyo, Japan.

Costa, M.; Crowcroft, J.; Castro, M.; Rowstron, A.; Zhou, L.; Zhang, L. and Barham, P., (2005) Vigilante: End-to-end Containment of Internet worms. In Proc. of the 20th ACM Symp. on Operating Systems Principles (SOSP). Brighton, UK.

Computer Worms Information, (2008) http://virusall.com/worms.shtml Accessed Jan. 2nd.

Ellis, D. R.; Aiken, J. G.; Attwood, K. S. and Tenaglia, S. D. A, (2004) Behavioral Approach to Worm Detection. In Proceedings of the Second ACM Workshop on Rapid Malcode (WORM).

Jiang, X. and Xu, D. (2006) Profiling Self-propagating Worms via Behavioral Footprinting. In Proceedings of ACM Workshop on Recurring Malcode.

Rasheed, M.; Ghazali, O. and Budiarto, R., (2012) SYN Scanning Worm Detection. Trends in Applied Sciences Research, (7), 859-871.

Schechter, S.; Jung, J.; and Berger, A. W., (2004) Fast Detection of Scanning Worm Infections. In 7th International Symposium on Recent Advances in Intrusion Detection (RAID). France.

Selvaraj, D.; and Ganapathi, P., (2014) Packet Payload Monitoring for Internet Worm Content Detection Using Deterministic Finite Automaton with Delayed Dictionary Compression. Journal of Computer Networks and Communications, (9).

Yang, X.; Lu, J. ; Zhu, Y. and Wang, P., (2006) Simulation and Evaluation of A New Algorithm of Worm Detection and Containment. In Proceedings of the Seventh International Conference on Parallel and Distributed Computing. Taiwan, 448-453.

Zou, C.; Gong, W. and Towsley, D., (2005) The Monitoring and Early Detection of Internet Worms. ACM Trans. on Networking.