Journal of KONBiN 44(2017)
ISSN 1895-8281
DOI 10.1515/jok-2017-0067 ESSN 2083-4608

IDENTIFICATION METHOD OF CONTENT CHANGES OF INFORMATION CONTAINED INSIDE FILES

METODA IDENTYFIKACJI ZMIAN TREŚCI INFORMACJI ZAWARTYCH W PLIKACH

Artur Szleszyński

Wyższa Szkoła Oficerska Wojsk Lądowych
The General Tadeusz Kościuszko Military Academy of Land Forces

Abstract: *The thesis shows the use of message hash function for integrity attribute change identification in sensitive information asset. Changes introduced to file content may suggest confidentiality attribute violation. It has been verified in what way random changes introduced into file content affect function. The minor number of bytes changed in comparison to file size that was assumed, derives from potential benefits for the attacker. Large number of changes in the file may suggest the situation of encrypting it. Such an action leads to accessibility attribute violation*

Keywords: *hash function, sensitive information asset*

Streszczenie: *Artykuł opisuje wykorzystanie funkcji haszującej do wykrywania zmian atrybutu integralności wrażliwych zasobów informacyjnych. Zmiany te mogą oznaczać naruszenie atrybutu poufności zasobu. W pracy sprawdzono, jak przypadkowe zmiany w treści plik lub plików wpływają na wartość funkcji haszującej. Niewielka liczba zmienionych bajtów, w porównaniu z rozmiarem pliku, sugeruje, że zmiany wprowadzone w jego treści są użyteczne dla atakującego. Duża liczba wykrytych zmian może sugerować zaszyfrowanie treści pliku w celu naruszenie atrybutu dostępności.*

Słowa kluczowe: *funkcja haszująca, wrażliwy zasób informacyjny*

IDENTIFICATION METHOD OF CONTENT CHANGES OF INFORMATION CONTAINED INSIDE FILES

1. Introduction

Information assets are one of the key elements of modern organizations. The information plays the leading role in information services. It is a key feature in decision making process. The proper information delivered in proper time enables undertaking proper decisions.
Apart from deliberations on the quality of information used in decision making process we shall focus on determining changes introduced to this information. Taking after ISO/IEC 27001, the information is secured whenever it consists of three unchanged attributes: confidentiality, integrity and accessibility [1]. An interesting matter is the use of evaluation methods of changes in integrity attribute of information assets that are parts of IT system for estimation of confidentiality attribute violation. Before any other consideration takes place it is essential to define the meaning of information assets. In case of IT system these assets are data gathered in files. These files may contain configuration data of computer operating system or network devices, business data or steering data for controllers such as SCADA.
According to Bell - La Padula safety model, files in IT system are the subjects accessible only by the authorized object [2,3]. Object may be the user or operating system process executed with object authorization properties [7,9]. If the object was authorized to access the subject, therefore, according to the rules of access to information assets, it may be assumed that changes in the file content were made by authorized object.
The problem occurs when a suspicion arises that changes in the examined file were introduced by unauthorized object. Such an unauthorized object may be the intruder or malicious software (malware). There is a possibility a source of the changes of file content may be the noise inside transmission channel. Sources of the noise can be a natural (i.e: lightnings) or artificial (other radio transmitters, electrical engines). Mentioned events are common for radio frequency channel. Radio frequency channels operate in very harmful environment so that users have to be sure received message is reliable. Described request is crucial for domains of the military and the crisis management actions. Several clues need to be examined to prove the information asset confidentiality attribute violation. One of them is asset integrity attribute violation. Another problem that may occur in this task is the effect of large number of files that need to be examined as well as the time necessary for this examination. The simplest way of determining the changes introduced by an object in the examined subject is storing the backup copy of a file that is the pattern to be compared with the changed version.

However, in this approach, it is necessary to keep the pattern file that is considered to be reliable. Another problem is the modification of a pattern made in the subject by authorized object. In case of modification in pattern content, the estimation of changes introduced in the file is corrupted. It is necessary to keep several patterns from different moments of time. Maintaining several patterns require making and storing backup copies. Storing several copies of pattern file requires maintaining the versions of it. This task complicates the detection process of integrity attribute violation and consequently unauthorized modifications in the file content. The number of tasks connected to estimation of possible violation of confidentiality attribute increases, which may be an example of scale effect related to the number of pattern file copies and the number of files being examined.
It is assumed that the safe way of verification of integrity attribute violation is making the hash of file. For this purpose such techniques as md-5, sha-224 or newer may be used. Methods described above enable storage of the hash files and not the original data. The hash file of examined object, according to the method, is as large as 128 bits (md-5) or 256 bits (sha-256) [4,5].

2. Problem definition

The suggested method of identification changes introduced to the examined file based on comparison of its content to pattern file has a major drawback. In case of large number of files that are of a large size (over 1MB) the time necessary for comparison of its content may be very long. The simplification of this comparison is a binary checkout of a file content. Exemplary time of binary comparison of two files depending on its size is shown in Table 1.

Tab. 1 Time of binary comparison content of two files (source: author's own work)

No	File size [kB]	Time of content comparison [s]
1	97	1.1
2	97 / 152	0.83
3	310	0.69
4	619	5.76
5	1273	1.1

Mean time necessary for comparison of two files equals 1,9s in the test range from 0,69s to 5,67s. Therefore, comparison of i.e. 100 files requires from 69 to 576 s according to expected mean value of 190 s. If the number of files is larger, time required for comparison of their content is much longer. It should be, therefore, considered another method of identification of file integrity attribute change. This method may be the file hash function such as md-5, sha-1 or sha-256.

The problem that needs to be solved is the answer to these questions: what kind of examination of integrity attribute in sensitive data files will ensure faster verification of modified files? How reliable is that solution?

3. Proposition of problem solution

Abandoning the comparison of file contents every time, there is an option to use the hash function made for the file of interest. Considering the comparison of both files, generating the hash function and comparing it to hash function of pattern file allows to shorten the time necessary to complete the task. It may be assumed, therefore, that the first criteria of estimation is the minimum time needed to discover the violation of integrity attribute in the subject file and finally possible violation of its confidentiality attribute. There is a lot of software that generate hash files. In Linux Operating System there are programs that generate hash files md-5, sha-1 to sha-256 that are maintained in bash terminal. Using bash terminal can be more convenient for information security officer. The tool allows storage outcomes of work inside file what gives a possibility to observe changes of tested files over the time. Windows Operating System uses md5checksum tool. This tool enables generating and comparing hash files of these types: md-5, sha-1 to sha-512.
For test procedure two files have been chosen: hash function for the pattern file and its modified version were calculated. Values of both hash functions are shown in table 2.If there are any differences between hash function values of two examined files there is a first clue that content one of them was changed. That means necessity careful use of the message included inside received file. The receiver of the message should to check if changes were made by authorized personnel.
Number of changed bytes in both files are shown in table 3. Number of changed bytes in comparison to file size is minor. It is in range from 0.0008% to 0.0101%. Changes in hash file values appear form the first byte, which illustrates Table 2.

Tab. 2 Files and their modification digest (source: author's own work)

File	Md-5 digest of pattern file	Modification file content	Md-5 digest of modified file
file1.doc	30E1DF1B05AC757537159C3CEFBA0094	file1m.doc	8BDCB6B2FD55E41D60B9D07C48C65FB6
file2.pdf	664AB1BCD90CA42975264 3DC3E0A6CB6	file2m.pdf	CC690EB99939022E6EA933AE8E510B3A
file3. mpp	65F5B187D8BDCD7B30A98391EFDD231A	file3m. mpp	2190246F6E44C800BF02E099C90CE4C1
file4. pcapng	594CDE5672FC3A49903F9BE346C2C79F	file4m. pcapng	CA9DBC8EC4F7857EC357A6AF5443E437

Tab. 3 Number of changed bytes between pattern and modified files (source: author's own work)

File	Number of changed bytes	Percentage value of change [%]
file1.doc / file1m.doc	10	0.0101
file2.pdf / file2m.doc	11	0.0034
file3.mpp / file3.mpp	5	0.00078
file4.pcapng / file4m.pcapng	11	0.00084

A minor change in file content caused the change in hash file value. It may be assumed that bigger change in file content will also be indicated in changed file value. Therefore, file integrity attribute violation was properly identified. The research made by author shows that if the change in file content is to gain profit for the intruder, the number of changed bytes will be minor. Large number of bytes changed in the file may suggest the encryption of its content. That means that main attribute of information asset security violated by taken actions is accessibility. It is less probable that intruder was interested in intentional content violation what means violation security attribute of information asset. However that event discredits organization's internal information security management system (ISMS). The value of this variable is shown in Table 4.

Tab. 4 Number of changed bytes inside the files as result of security incidents (source: author's own work)

Security incident	Size of pattern file [B]	Number of changed bytes	Percentage value of change [%]
Intentional change of invoice content (changed bank account number)	100	8	8
Encryption of file content	66	58	88

For what is above the checkout on reliability of changes identification was done on minor number of changed bytes. The next step is to check, which object (person or process) introduced changes that were discovered. One example of decision making system is described in A. Szleszyński thesis. The thesis presents a decision making mechanism that identifies confidentiality attribute violation of a file on the basis of integrity attribute change [8]. The fuzzy-logic decision making mechanism used to estimate possible confidentiality attribute change in a file that consists sensitive data is described [8]. The parameter that needs to be considered is the significance of introduced changes in the file content.

Large number of changes indicates violation of accessibility attribute of information assets. Minor number of changes suggests deliberate change in file content. Such a modification implies that the unauthorized object acquainted file content, which means the file lost confidentiality attribute. This change will be the basis for verification. A minor change in file content will be revealed in hash file value. Such kind of modification means that content of the file is known for unauthorized object. This means loss of confidentiality attribute by the file. In order to evaluate exact number of changed bytes inside test file and they significance there is necessary to have trustworthy version of file. The trustworthy copy of file is a pattern use to comparison the content of both files.
In case of a small number of changed bytes inside the tested file the importance of changes must be check. The importance means a modification influence on a message content included inside the file.

4. Conclusions

The use of hash function in order to identify modification of integrity attribute of file is justified. A method has not use of pattern file to preliminary classification of the files with changed content. This activity decreases a potential vulnerability related to unauthorized access to pattern file. An attacker cannot replace pattern file in order to hide their destructive activity. Described method is used by computer operating systems to store users' access passwords [7,9].
Depicted solution is reliable because modification a small number of bytes inside the file implies hash function value change. This selection saves the time required for classification of the significance of the changes. In the literature there is information that undermines the use of certain functions shortcuts because of the possibility of such conduct modification which is not shown by the abbreviation [6]. This applies to hash functions such as md-5 or sha-256. Used in the experiment, the program allows generating a hash function such as sha-256. A longer string of bytes requires longer time to make a comparison. A disadvantage of the described method comparison is the only binary statement that modify the contents of the file were occurred. The hash function does not contain information about the number of bytes changed in the test file. It is difficult to select these files in which potential changes have the highest significance and they have to be compared with pattern files. Disadvantage the method is lack of possibility to detect scale of changes, inside the file content, that were made. Furthermore, it is hard to state which one security attribute of file was violated. If the goal of analysis is determining which one security attribute was changed process of analysis has to be extended. One from possible technique that can be used is an expertise made by expert system. Expert system ought to diagnose which security attribute of file was changed and what is the importance of this incident. Then Information Security Management team should to prepare the plan of actions that drives on higher level or at least keep an acceptable level of information security inside the organization [1,3].

5. References:

[1] International Standard ISO/IEC 27001:2005 *Information technology – Security techniques – Information security management systems – Requirements*, Geneva 2005.

[2] Jóźwiak I.J., Szleszyński A., *Study of the security of process running in computer operating systems*, Safety and Reliability: Methodology and Application, pp. 651 – 654, ESREL 2014, CRC Press, 2014.

[3] Liderman K., *Information security*, PWN, Warszawa 2012 (in polish language)

[4] RFC 1321, *The MD-5 Message – Digest Algorithm*, Network Working Group, (access on-line) www.rfc-editor.org, 1992.

[5] RFC 3874, *A 224-bit One-way Hash Function: SHA-224*, Network Working Group, (access on-line) www.rfc-editor.org, 2004.

[6] RFC 4270, *Attack on Cryptographic Hashes in Internet Protocols*, Network Working Group, (access on-line) www.rfc-editor.org, 2005.

[7] Stallings W., *Operating System Internals and Design*, Prentice Hall, Upper Sadle River, 2012.

[8] Szleszyński A., *The Method of Evaluation Degree of Changes Confidentiality Attribute of Information Asset Inside ITC System*, Journal of KONBiN 33(1), pp. 159 – 168, Warszawa 2015.

[9] Tanenbaum A., Boss H., *Modern Operating Systems*, Pearson Education, 2015.

Artur Szleszyński Msc *Artur Szleszyński MsC, he was graduated at Faculty of Electronics Military University of Technology. He studied at Faculty Informatics and Management Wrocław University of Technology. He has got Microsoft Certified Professional certificate. His scientific research area is information and ICT systems security.*

METODA IDENTYFIKACJI ZMIAN TREŚCI INFORMACJI ZAWARTYCH W PLIKACH

1. Wprowadzenie

Informacje należą do kluczowych zasobów współczesnych organizacji. Informacje odgrywają wiodącą rolę w usługach informacyjnych. Są kluczowym elementem procesu decyzyjnego. Właściwe informacje dostarczane we właściwym czasie umożliwiają podejmowanie właściwych decyzji. Oprócz rozważań nad jakością informacji wykorzystywanych w procesie podejmowania decyzji należy się skupić na określeniu zmian wprowadzonych w informacji. Zgodnie z normą ISO/IEC 27001 informacja uważana jest za zabezpieczoną, wtedy gdy trzy atrybuty jej bezpieczeństwa są niezmienione. Do atrybutów należą: poufność, integralność i dostępność [1]. Interesującym jest wykorzystanie metod oceny zmian w atrybutu integralności zasobu informacyjnego do oszacowania naruszenia atrybutu poufności. Przed podjęciem jakichkolwiek rozważań konieczne jest określenie znaczenia aktywów informacyjnych dla systemu teleinformatycznego. W przypadku systemu teleinformatycznego te aktywa informacyjne są danymi zgromadzonymi w plikach. Pliki mogą zawierać dane dotyczące konfiguracji systemu operacyjnego komputerów lub urządzeń sieciowych, mogą one być danymi biznesowymi lub danymi wykorzystywanymi przez urządzania sterujące takie jak kontrolery typu SCADA.

Zgodnie z modelem bezpieczeństwa systemów teleinformatycznych Bell`a - La Padul`i pliki w systemie teleinformatycznym[1] dostępne są tylko dla uprawnionego obiektu [2,3]. Obiektem może być proces użytkownika lub systemu operacyjnego wykonywany z uprawnieniami typowi (przydzielonymi) dla niego [7,9]. Jeśli obiekt uzyskał dostęp do zasobu, to zgodnie z przyjętymi zasadami dostępu do zasobów informacyjnych można założyć, że zmiany w zawartości pliku zostały wykonane przez upoważniony obiekt.

Problemem jest wystąpienie podejrzenia, że zmiany w badanym pliku zostały wprowadzone przez nieupoważniony obiekt. Takim nieuprawnionym obiektem może być intruz lub złośliwe oprogramowanie (ang. malware). Należy zatem zbadać dostępne przesłanki, aby w móc udowodnić naruszenie atrybutu poufności aktywu informacyjnego. Jedną z przesłanek jest naruszenie atrybutu integralności aktywów. Problemem, jaki może wystąpić w realizacji tego zadania, jest efekt skali wynikający z dużej liczby plików, które należy zbadać. Dodatkowo czas, który będzie potrzebny do przeprowadzenia tego badania, gdyż zależy nam na jak najmniejszej inercji rozwiązania. Najprostszym sposobem określania zmian wprowadzonych przez obiekt w badanym obiekcie jest przechowywanie kopii zapasowej pliku, która będzie wzorcem wykorzystanym do porównania zawartości obu plików.

[1] Określane terminem podmiot

Jednakże w tym podejściu konieczne jest przechowywanie pliku wzorcowego co tworzy nową podatność w systemie. Wzorzec uważany jest za wiarygodny tak więc stanowi swoisty etalon pomiarowy. Innym problemem jest modyfikacja pliku wzorca dokonanego przez uprawniony obiekt. W przypadku modyfikacji zawartości wzorca, oszacowanie zmian wprowadzonych w pliku możliwe jest dopiero po odzyskaniu go z kopii bezpieczeństwa. Konieczne jest przechowywanie kilku wersji pliku wzorca z różnych momentów czasowych. Utrzymywanie kilku wersji pliku wzorcowego wymaga opracowania zasad zarządzania wersjami. Pojawienie się wersji plików komplikuje zadanie polegające na wykrywaniu naruszenia atrybutów integralności, a tym samym nieautoryzowane modyfikacje zawartości pliku. To z zwiększa liczbę zadań związanych z oszacowaniem ewentualnego naruszenia atrybutu poufności, a co związane jest z liczbą kopiowanych plików wzorcowych oraz liczbą badanych plików.
Można przyjąć, że bezpiecznym sposobem weryfikacji naruszenia atrybutu integralności jest wykorzystanie funkcji skrótu pliku (funkcja haszująca). W tym celu można użyć takich funkcji jak md-5, sha-224 lub nowsza. Wymienione techniki umożliwiają gromadzenie plików haszujący w miejsce oryginalnych dane. Funkcja mieszająca badanego obiektu, posiada rozmiar od 128 bitów (md-5) lub 256 bitów (sha-256) [4,5].

2. Określenie problemu

Opisaną metodą identyfikacji zmian wprowadzonych w badanym pliku jest porównanie jego zawartości z zawartością pliku wzorcowego. Jednakże metoda ta posiada poważną wadę. W przypadku dużej liczby plików, które są dużych rozmiarów (powyżej 1 MB), czas niezbędny do porównania ich treści może być bardzo długi. W uproszczeniu polega ono na binarnym porównaniu zawartości badanego pliku z jego wzorcem. Czas potrzebny do wykonania binarnego porównania zawartości dwóch plików, w zależności od jego wielkości pokazano w tabeli 1.

Tab. 1 Czas potrzebny do binarnego porównania zawartości dwóch plików (źródło: opracowanie własne)

Lp	**Rozmiar pliku [kB]**	**Czas potrzebny do porównania zawartości plików [s]**
1	97	1.1
2	97 / 152	0.83
3	310	0.69
4	619	5.76
5	1273	1.1

Średni czas potrzebny do porównania zwartości dwóch plików wynosi 1,9 s w przedziale od 0,69 do 5,67 s dla próby testowej. Zatem porównanie 100 plików będzie wymagać od 69 do 576 s, zaś średnia wartość czasu potrzebnego do wykonania tej czynności wyniesienie 190 s. Jeśli liczba plików będzie większa, czas potrzebny do porównania ich zawartości jest również się wydłuży. Należy zatem rozważyć inną metodę identyfikacji zmiany atrybutu integralności plików. Tą metodą może być użycie funkcji skrótu pliku, na przykład md-5, sha-1 lub sha-256. Problemem, który trzeba rozwiązać, jest znalezienie odpowiedzi na następujące pytania: jaki rodzaj badania zmian atrybutu integralności wrażliwych plików danych zapewni szybką identyfikację wprowadzonych zmian? Jak niezawodne jest przyjęte rozwiązanie?

3. Propozycja rozwiązania

Rezygnując z porównania zawartości dwóch plików, za każdym razem kiedy istnieje podejrzenie naruszenia integralności, można posłużyć się funkcją skrótów dla badanego pliku oraz pliku wzorcowego. Wygenerowanie obu funkcji a następnie ich porównywanie skróci czas potrzebny do wykonania tego zadania, gdyż porównuje się dwa ciągi danych o długości od 16 do 32 bajtów. W związku z tym można założyć, że pierwsze kryterium oszacowania -- minimalny czas potrzebny do wykrycia naruszenia atrybutu integralności w pliku zostanie spełniony. Jeżeli obie funkcje skrótów będą rożne jest to przesłanką do możliwego naruszenia atrybutu poufności. Weryfikacja przesłanki będzie wymagać dalszego zbadania. Zadanie to można wykonać korzystając z odpowiedniego oprogramowania. Istnieje wiele programów, które generują funkcje skrótów. W systemie operacyjnym Linux programy, generujące funkcje skrótów md-5, sha-1 do sha-256, są dostępne z poziomu terminala np. bash. System operacyjny Windows korzysta z narzędzia md5checksum. Jest to narzędzie umożliwiające generowanie funkcji skrótów oraz ich porównywanie. Generowanymi funkcjami skrótów mogą funkcje typu: md-5, sha-1 do sha-512.

Dla potrzeb procedury testowej przygotowano dwa rodzaje plików, dla których wygenerowano funkcje haszujące. Jeden z plików został celowo zmodyfikowany. Wartości funkcji haszujących dla obu plików użytych w badaniu przedstawiono w tabeli 2. Liczbę zmienionych bajtów w plikach – wykorzystanych do porównania – pokazano w tabeli 3. Liczba zmienionych bajtów w porównaniu do rozmiaru pliku jest niewielka. Zawiera się w przedziale od 0,0008% do 0,0101%. Zmiany w wartościach pliku mieszania pojawiają się już w pierwszym bajcie, co ilustruje tabela 2.

Tab. 2 Pliki wraz z ich zmodyfikowaną funkcją skrótu (źródło: opracowanie własne)

Plik	**Funkcja skrótu md-5**	**Zmodyfikowany plik**	**Zmodyfikowana funkcja skrótu md-5**
file1.doc	30E1DF1B05AC757537159C3CEFBA0094	file1m.doc	8BDCB6B2FD55E41D60B9D07C48C65FB6
file2.pdf	664AB1BCD90CA429752643DC3E0A6CB6	file2m.pdf	CC690EB99939022E6EA933AE8E510B3A
file3. mpp	65F5B187D8BDCD7B30A98391EFDD231A	file3m. mpp	2190246F6E44C800BF02E099C90CE4C1
file4. pcapng	594CDE5672FC3A49903F9BE346C2C79F	file4m. pcapng	CA9DBC8EC4F7857EC357A6AF5443E437

Tab. 3 Liczba zmienionych bajtów w plikach testowych w stosunku do wzorca (źródło: opracowanie własne)

Plik	**Liczba zmienionych bajtów**	**Procentowa wartość wprowadzonej zmiany [%]**
file1.doc / file1m.doc	10	0.0101
file2.pdf / file2m.doc	11	0.0034
file3.mpp / file3.mpp	5	0.00078
file4.pcapng / file4m.pcapng	11	0.00084

Niewielka zmiana zawartości pliku spowodowała zmianę wartości funkcji haszującej. Można przyjąć, że większa zmiana zawartości plików również będzie wykazywana w zmienionej wartości funkcji. Zatem naruszenie atrybutu integralności plików zostało poprawnie zidentyfikowane. Badanie przeprowadzone przez autora wykazało, że jeśli zmiana zawartości plików ma przynieść zyski intruzowi, to liczba zmienionych bajtów będzie niewielka. Duża liczba bajtów w pliku może sugerować szyfrowanie jej zawartości. Wartość tej zmiennej przedstawiono w tabeli 4.

Tab. 4 Liczba zmienionych bajtów wewnątrz pliku będąca wynikiem incydentów w bezpieczeństwie informacji (źródło: opracowanie własne)

Incydent w bezpieczeństwie	**Rozmiar pliku wzorcowego[B]**	**Liczba zmienionych bajtow**	**Procentowa wartość zmian [%]**
Celowa zmiana treści faktury (zmieniony numer konta bankowego)	100	8	8
Zaszyfrowanie treści pliku	66	58	88

Zatem sprawdzenie wiarygodności wykrywania wprowadzonych zmian, dokonano na małej liczbie zmienionych bajtów. Następną czynnością jest sprawdzenie, jaki obiekt (osoba lub proces) wprowadził wykryte zmiany. Przykład systemu diagnozującego opisano w pracy A. Szleszyńskiego. W artykule przedstawiono mechanizm decyzyjny, który identyfikuje naruszenie atrybutu poufności pliku na podstawie zmian atrybutu integralności [8]. W rozwiązaniu zastosowano mechanizm rozmyto - logicznego wnioskowania wykorzystywany do oszacowania możliwej zmiany atrybutu poufności w pliku [8]. Parametrem, który należy uwzględnić w diagnostyce, jest znaczenie wprowadzonych zmian w zawartości pliku. Duża liczba zmian wskazuje na naruszenie atrybutu dostępności zasobów informacyjnych. Mniejsza liczba zmian sugeruje przemyślaną zmianę zawartości plików. Taka modyfikacja oznacza, że nieautoryzowany obiekt zapoznał się z treścią pliku, co oznacza utratę atrybutu poufności. Zatem wielkość zmiany będzie przesłanką do dalszej weryfikacji. Na podstawie przeprowadzonych badań wykazano, że mała liczba zmienionych bajtów w pliku zostanie odzwierciedlona w wartości obliczonej funkcji skrótu pliku.
W celu zidentyfikowania zmienionych bajtów wewnątrz pliku oraz ich znaczenia niezbędne jest posiadanie wiarygodnej wersji wzorcowej pliku. Wiarygodna wersja wzorcowa pliku służy do porównania zawartości obu plików.
W przypadku niewielkiej liczby zmienionych bajtów w badanym pliku istotność zmian musi zostać sprawdzona. Istotność oznacza wpływ modyfikacji zawartości wiadomości przechowywanej w pliku.

4. Wnioski

Użycie funkcji skrótu w celu zidentyfikowania modyfikacji atrybutu integralności pliku jest uzasadnione. Metoda nie korzysta z pliku wzorcowego do wstępnej klasyfikacji plików o zmienionej zawartości. Czynność ta zmniejsza potencjalną lukę związaną z nieautoryzowanym dostępem do pliku wzorcowego.

Podmiot nieuprawniony nie może zastąpić pliku wzorcowego w celu ukrycia swojej destrukcyjnej działalności. Opisana metoda jest używana w systemach operacyjnych komputerów do przechowywania haseł użytkowników [7,9].
Przedstawione rozwiązanie jest niezawodne, ponieważ modyfikacja niewielkiej liczby bajtów wewnątrz pliku oznacza zmianę wartości funkcji haszowania. Ten wybór pozwala zaoszczędzić czas niezbędny do klasyfikacji znaczenia zmian. W literaturze znajdują się informacje, które naruszają użycie pewnych skrótów funkcji ze względu na możliwość takiej modyfikacji zachowania, której nie pokazano skrótem [6]. Dotyczy to hasz md-5 lub sha-1.
Używany w eksperymencie program umożliwia generowanie funkcji mieszania, takiej jak sha-256. Dłuższy ciąg bajtów wymaga dłuższego czasu na porównanie. Wadą opisanego porównania jest jedynie binarna informacja o modyfikacji zawartości pliku. Funkcja skrótu nie zawiera informacji o liczbie bajtów zmienionych w pliku testowym. Trudno jest wybrać te pliki, w których potencjalne zmiany mają największe znaczenie i muszą być porównane z plikami wzorcowymi. Wadą metody jest brak możliwości wykrycia skali zmian, wewnątrz zawartej w pliku zawartości. Ponadto trudne jest określenie, który z atrybutów bezpieczeństwa został naruszony. Jeśli celem analizy jest identyfikacja, który z atrybutów bezpieczeństwa został zmieniony, proces analizy musi zostać przedłużony. Jedną z możliwych technik, którą można użyć, jest ekspertyza wykonana przez system ekspercki. System ekspercki powinien zdiagnozować, który atrybut bezpieczeństwa został zmieniony i jakie jest znaczenie tego zdarzenia. Następnie zespół ds. Zarządzania bezpieczeństwem informacji wewnątrz organizacji powinien przygotować plan działań, które prowadzą na wyższym poziomie lub przynajmniej utrzymują akceptowalny poziom bezpieczeństwa informacji wewnątrz organizacji [1,3].

5. Literatura:

[1] International Standard ISO/IEC 27001:2005 *Information technology – Security techniques – Information security management systems – Requirements*, Geneva 2005.
[2] Jóźwiak I.J., Szleszyński A., *Study of the security of process running in computer operating systems*, Safety and Reliability: Methodology and Application, pp. 651 – 654, ESREL 2014, CRC Press, 2014.
[3] Liderman K., *Information security*, PWN, Warszawa 2012 (in polish language)
[4] RFC 1321, *The MD-5 Message – Digest Algorithm*, Network Working Group, (access on-line) www.rfc-editor.org, 1992.
[5] RFC 3874, *A 224-bit One-way Hash Function: SHA-224*, Network Working Group, (access on-line) www.rfc-editor.org, 2004.

[6] RFC 4270, *Attack on Cryptographic Hashes in Internet Protocols*, Network Working Group, (access on-line) www.rfc-editor.org, 2005.
[7] Stallings W., *Operating System Internals and Design*, Prentice Hall, Upper Sadle River, 2012.
[8] Szleszyński A., *The Method of Evaluation Degree of Changes Confidentiality Attribute of Information Asset Inside ITC System*, Journal of KONBiN 33(1), pp. 159 – 168, Warszawa 2015.
[9] Tanenbaum A., Boss H., *Modern Operating Systems*, Pearson Education, 2015.

Mgr inż. Artur Szleszyński, *absolwent Wydziału Elektroniki Wojskowej Akademii Technicznej. Ukończył studia podyplomowe na Wydziale Informatyki i Zarządzania Politechniki Wrocławskiej. Jest posiadaczem tytułu Microsoft Certified Professional. Obszar zainteresowań naukowych dotyczy bezpieczeństwa informacji i niezawodności systemów teleinformatycznych.*