Wiórki **prawdy**

BADANIA | Gdy korzystamy z niszczarki do papieru, najczęściej jesteśmy pewni, że zniszczonych za jej pomocą danych nie da się już odzyskać. Okazuje się jednak, że rzeczywistość jest zupełnie inna. Prezentujemy wyniki badań przeprowadzonych w PAN w zakresie odtwarzania pociętych dokumentów.

Szymon Piłat, Kamil Kulesza

Ważnym wydarzeniem w czasie rewolucji islamskiej w 1979 roku było zajęcie przez „studentów" ambasady USA w Teheranie. Niedługo przed tym wydarzeniem rząd Stanów Zjednoczonych podjął decyzję o zniszczeniu poufnych dokumentów (np. raportów CIA), znajdujących się w tej placówce. Ostatecznie zrobiono to, wykorzystując niszczarki do papieru. Niestety, Amerykanom nie wystarczyło już czasu na utylizację ścinków. Kiedy ambasada została zajęta, Irańczycy przechwycili zniszczone dokumenty. Ich rekonstrukcją zajęli się zręczni perscy tkacze dywanów. W ciągu kilkunastu lat udało się im odtworzyć około 60 tomów dokumentów, które następnie stopniowo upubliczniano.

Podobnie było w wypadku akt należących do STASI. W 1989 roku niemieckie tajne służby zniszczyły w pośpiechu około 45 milionów dokumentów. Po latach odnaleziono uszkodzone akta i postanowiono je odtworzyć. Ta skomplikowana, kosztująca 40 mln euro operacja stała się nawet inspiracją do nakręcenia filmu. Rekonstrukcji wymagały tysiące zniszczonych dokumentów. Początkowo odzyskiwano je ręcznie. Natomiast później wykorzystano zaawansowaną technikę (np. skanery Lufthansy, skanujące dwustronnie 10 000 stron na godzinę) oraz skomplikowane oprogramowanie stworzone przez naukowców z Fraunhofer Institut für Produktionsanlagen und Konstruktionstechnik w Berlinie.

W obu wypadkach, pomimo usilnych prób, nie udało się doszczętnie zniszczyć informacji zapisanych na papierze. Podobnych historii jest oczywiście więcej, dlatego przeprowadzone w PAN badania miały pomóc w znalezieniu odpowiedzi na pytanie, jak bardzo zniszczenie dokumentów jest skuteczne oraz jak bardzo skomplikowana jest ich rekonstrukcja. Analizę problemu podzielono na dwa wątki: teoretyczny oraz praktyczny. W pracach trwających trzy miesiące wzięło udział około 20 osób, z których zdecydowaną większość stanowili studenci rozmaitych kierunków technicznych.

Klasy tajności niszarek

Poziom tajności niszczarek do papieru najczęściej podaje się, stosując klasyfikację DIN. Rozróżnia ona 6 klas urządzeń w zależności od wielkości ścinków, powstałych po zniszczeniu kartki:

- klasa 1 – paski o szerokości 12 mm,
- klasa 2 – paski o szerokości 6 mm,
- klasa 3 – paski o szerokości 2 mm,
- klasa 4 – prostokąty o wym. 2 × 15 mm,
- klasa 5 – prostokąty o wym. 0,8 × 12 mm,
- klasa 6 – prostokąty o wym. 0,8 × 4 mm.

Istnieją również normy amerykańskie:

- US NSA/CSS 02-01 – prostokąty o wym. 1 × 4 mm
- US Department of Defense: Top Secret – prostokąty o wym. 0,8 × 11,1 mm

Niszczarki tnące papier na paski oferują niewielki poziom bezpieczeństwa. Przy rekonstrukcji dokumentu za pomocą komputera urządzenie o czwartej klasie tajności również może okazać się mało efektywne. Dopiero sprzęt spełniający normy piątej klasy DIN i normy amerykańskie daje nam 100% pewności skutecznego oraz bezpiecznego zniszczenia poufnych danych.

Matematyczne puzzle

Wątek teoretyczny rozwijany był przez matematyków. Potraktowali oni niszczarkę do papieru jako maszynę szyfrującą lub generator liczb losowych. Dzięki temu udało się matematycznie opisać procedurę niszczenia dokumentów. Zadaniem grupy teoretycznej było także opracowanie możliwych scenariuszy ataku na urządzenie z wykorzystaniem technik matematycznych. Prace tego zespołu pozwoliły stworzyć skuteczne i praktyczne metody rekonstruowania pism.

Podczas badań praktycznych przez większość czasu zajmowano się analizą skuteczności niszczarek o klasie tajności 1–3. Według klasyfikacji DIN jest to sprzęt, który tnie papier na paski (patrz: **ramka „Klasy tajności niszczarek"**). Kartka zniszczona w urządzeniu trzeciej klasy może zostać złożona na 50! (silnia) różnych sposobów. Tak więc liczba możliwych kombinacji jest tak duża, że ich zastosowanie w praktyce znacznie przekracza moc obliczeniową najpotężniejszych komputerów. Okazuje się jednak, że istnieją sposoby, które pozwalają liczbę tę znacząco zmniejszyć. Najprostsze metody zakładają:

- wykorzystanie różnego rodzaju symetrii występujących w dokumencie (np. okrągłych pieczęci),
- odnoszenie się do stałych punktów występujących w formularzu (np. do nagłówków z miejscowością i datą),
- eliminację marginesów czy też analizę długości paragrafów oraz akapitów.

Spostrzeżenia te pozwoliły na opracowanie zestawu metod i algorytmów, które są pomocne w odzyskiwaniu „pociętych" informacji. Dzięki nim kartkę papieru można było złożyć na nowo w kilka lub w kilkadziesiąt minut. Podczas doświadczeń okazało się, że najtrudniejsza jest rekonstrukcja prostych figur geometrycznych.

W kontekście bezpieczeństwa urzędu lub firmy warto zwrócić uwagę na wyniki badań dotyczących rekonstrukcji dokumentów zawierających tekst. Według naszych analiz odtworzenie kartek z bardzo

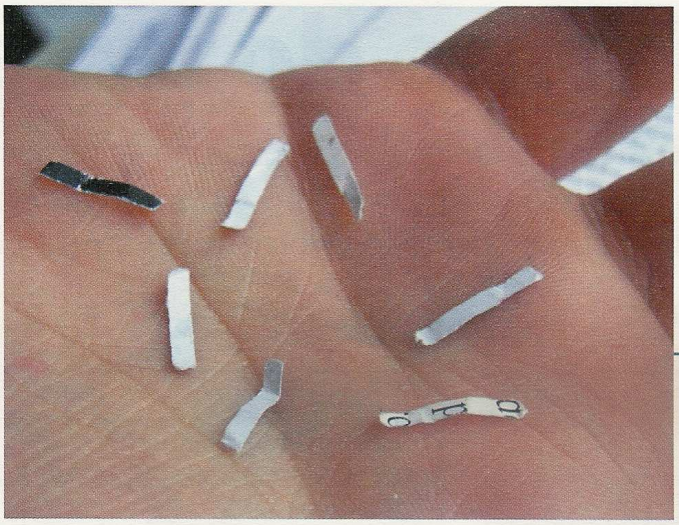
małą ilością tekstu na stronie zajmuje od 8 do 20 minut. Natomiast dokument, w którym znaki wypełniają prawie całą kartkę można odzyskać średnio w sześć minut. Wszystkie podane przedziały czasowe dotyczą procesu rekonstrukcji dokumentu, którego treść nie była wcześniej znana badaczowi.

Wydawać by się mogło, że podczas procesu niszczenia ścinki poszczególnych kartek mieszają się ze sobą w koszu, przez co odtworzenie wielu stron jest bardzo skomplikowane. Problem tasowania się poszczególnych kartek wydawał się na tyle istotny, że postanowiono przeprowadzić odpowiednie badania. Rezultaty były zaskakujące. Okazało się, że ścinki kartek, będące produktem niszczarek klasy 1–3 prawie w ogóle ze sobą się nie mieszają. Warto wspomnieć, że jeżeli z jakiegoś powodu fragmenty pisma zostałyby przetasowane, to i tak będzie można je rozróżnić po gramaturze, barwie papieru lub po kącie wkładania dokumentu do niszczarki. Ponadto na wymieszanie się elementów pociętego papieru nie ma większego wpływu to, czy kosz jest pusty, czy też wypełniony do tego stopnia, że nowe ścinki są mocno upychane przez mechanizm niszczarki.

(Nie)tajna czwórka

Pamiętajmy, że podczas badań PAN zajmowano się tylko i wyłącznie urządzeniami o klasie tajności 1–3. Niszczarki sklasyfikowane wyżej pozwalają ciąć kartki na małe prostokąty. Oczywiście zrekonstruowanie tak zniszczonego dokumentu jest dużo trudniejsze. Wynika to z większej liczby ścinków, które dodatkowo mieszają się z fragmentami innych kartek. Nie oznacza to jednak, że nie można odtworzyć takiego pisma. Jak widać na **zdjęciu**, na ścinku o wymiarach 2 × 15 mm mieści się cała litera, podczas gdy do odtworzenia całego tekstu wystarczy, że na fragmentach pisma widoczne są tylko połówki znaków.

Jest to jednak tylko teoria, gdyż doświadczenia przeprowadzone przez PAN pokazują, że czas składania kartki „posiekanej" na prostokąty jest tak długi, iż ręczna rekonstrukcja wymieszanych dokumentów jest pozbawiona sensu. Czy zatem posiadając niszczarkę o czwartej klasie tajności, możemy czuć się bezpiecznie? Niestety, urządzenia z tego segmentu nie dają nam 100% pewności, że zniszczonych danych nikt nie odczyta. O ile odtworzenie dokumentu z pasków jest dość łatwą czynnością, o tyle rekonstrukcja strony z małych równoległoboków, ze względu na dużą liczbę możliwych połączeń, jest zadaniem dużo bardziej skomplikowanym. Co prawda człowiek nie jest w stanie ręcznie wypróbować wszystkich możliwości. Jednak z punktu widzenia matematyka tworzącego algorytm rekonstrukcji składanie pasków różni się od składania prostokątów tylko jedną cechą. Otóż paski układane są w jednym wymiarze, natomiast prostokąty w dwóch. Dla człowieka wielowymiarowość jest sporym problemem, lecz dla komputera nie stanowi ona dużej przeszkody – wymagana jest jedynie większa moc obliczeniowa.



Fragmenty dokumentu zniszczonego przez **niszczarkę czwartej klasy tajności**. Urządzenia tego typu pracują np. w MSZ.

Kłopoty złodzieja

Spójrzmy na zagadnienie odtwarzania dokumentów z punktu widzenia osoby, która chciałaby przechwycić nasze dokumenty. Jeżeli w biurze nie ma niszczarki, to cała trudność polegałaby tylko na przechwyceniu urzędowych odpadów oraz wyselekcjonowaniu z nich odpowiednich dokumentów. Natomiast jeżeli informacje byłyby „przemielone", to atakujący, oprócz grzebania w śmieciach, musiałby dokonać ich rekonstrukcji. Dlatego też niszczenie danych – w porównaniu z czynnościami, które złodziej musiałby wykonać, by przechwycić odpowiednie dane – jest opłacalne tylko wtedy, gdy używamy niszczarek o wysokich klasach tajności. Tak więc niszcząc papierowe materiały, róbmy to solidnie i starannie, nie dając potencjalnemu złodziejowi pola do popisu. Niezależnie od pocięcia dokumentów, warto rozważyć zniszczenie samych ścinków. Dobrą i skuteczną metodą jest ich spalenie.

Więcej o projekcie badawczym

Opisywane w artykule badanie przeprowadzone zostały w ramach letniej praktyki badawczej (**www.praktyki.ibspan.waw.pl**), prowadzonej przez Polską Akademię Nauk. Projekty, w których uczestniczą studenci, zorganizowane są na wzór brytyjskich *study groups*. Tematyka badań dotyczy głównie bieżących problemów przemysłowych i biznesowych, których rozwiązaniem może być podejście badawczo-naukowe. Praktyki są prowadzone w PAN od kilku lat, obecnie są organizowane wspólnie przez Instytut Badań Systemowych PAN, Instytut Matematyczne PAN i Wyższą Szkołę Informatyki Stosowanej i Zarządzenia pod auspicjami PAN.

FOT. SZ. PIŁAT

Szymon Piłat jest studentem Wydziału Fizyki Uniwersytetu Warszawskiego. Od kilku lat zajmuje się m.in. analizą sygnałów oraz problematyką bezpieczeństwa i odzyskiwania danych. W trakcie zaprezentowanych badań pełnił funkcję lidera projektu.

Kamil Kulesza jest adiunktem w Instytucie Badań Systemowych PAN oraz pracownikiem naukowym University of Cambridge. W swojej pracy badawczej zajmuje się zagadnieniami związanymi z informatyką i zastosowaniami matematyki, ze szczególnym uwzględnieniem bezpieczeństwa informacji. Jest też zaangażowany w Letnie Praktyki Badawcze PAN od momentu ich powstania.