Tutoriál: Čipové karty v informačních systémech – Rozšířený abstrakt

Pert HANÁČEK [1], Vašek MATYÁŠ [2]

[1] *Fakulta informačních technologií, Vysoké učení technické v Brně*
Božetěchova 2, 612 66 Brno
hanacek@fit.vutbr.cz

[2] *Fakulta informatiky, Masarykova univerzita v Brně*
Botanická 68a, 602 00 Brno
matyas@fi.muni.cz

1 Úvod

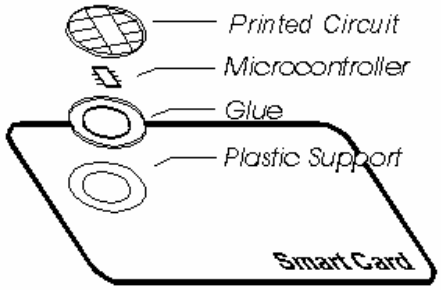
Čipová karta (angl. *smart card*) je plastická karta o velikosti běžné platební karty, která obsahuje zabudovaný čip, zpravidla vybavený jednočipovým procesorem. Skutečnost, že tento procesor je umístěn v jediném čipu, zajišťuje zvýšenou odolnost čipové karty proti útoku a předurčuje čipovou karty především pro aplikace, kde je třeba zajištění vyššího stupně bezpečnosti. V minulých letech byla odolnost čipových karet proti útokům nezpochybnitelným dogmatem a mnoho návrhářů zabudovávalo čipové karty do svých systémů s důvěrou, že platí „čipová karta=bezpečnost“. Dnes tomu tak už není – v následujícím příspěvku se pokusíme ukázat proč a v jakých aplikacích je nasazení čipových karet vhodné, potažmo jakým způsobem se nasazují.

2 Typy a vlastnosti čipových karet

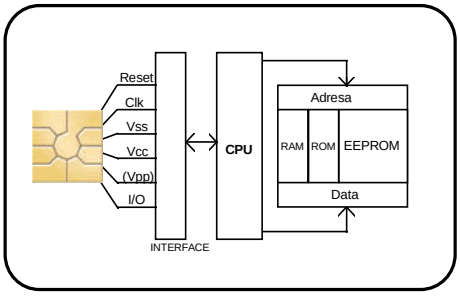
Čipové karty jsou dvojího druhu – kontaktní a bezkontaktní. Kontaktní čipové karty se dají snadno rozeznat podle typického tvaru osmice kontaktů, jak je známe například z telefonních karet. Kontaktní karta musí být v okamžiku její činnosti zasunuta v konektoru nebo v takzvané čtečce. Bezkontaktní kartu může její majitel držet v ruce, nosit v kapse nebo ji může mít připnutou na svém oděvu. Bezkontaktní karty totiž používají pro styk s vnějším světem a pro napájení elektromagnetickou indukční smyčku. Karta se skládá ze tří základních částí: z plastového nosiče o rozměrech 85.60 mm x 53.98 mm x 0.80 mm, z plošného spoje, který obsahuje osm kontaktů konektoru a z čipu integrovaného obvodu. Schopnosti čipové karty jsou dány schopnostmi jejího čipu. Podle toho dělíme čipové karty na tři skupiny: na paměťové čipové karty, na paměťové čipové karty se speciální logikou a na procesorové čipové karty. Posledně zmíněné karty mají na svém čipu jednočipový mikroprocesor, který je schopen sám provádět více či méně složité výpočty. Mezi tyto výpočty velmi často patří schopnost zašifrovat nebo dešifrovat data pomocí klíče, který je uložen na kartě, což dává tomuto typu karet značné možnosti pro kryptografickou autentizaci nebo pro realizaci platebních transakcí. Jelikož tyto

Lubos Popelínský (ed.), DATAKON 2003, Brno, 18.-21. 10. 2003, pp. 1-7.

čipové karty patří k nejuniverzálnějším a v oblasti zabezpečení k nejpoužívanějším, budeme se dále v tomto článku zabývat pouze kontaktními procesorovými čipovými kartami.



Obr. 1: Fyzická konstrukce čipové karty.



Obr. 2: Struktura kontaktní procesorové čipové karty.

Kontaktní procesorová čipová karta je spojena s vnějším světem konektorem. Ačkoli ISO standard 7816-2 definuje osm kontaktů tohoto konektoru, v praxi se z nich využívá pouze pět nebo šest. karty. Mimo interface pro styk s vnějším světem obsahuje procesorová čipová karta mikroprocesor a paměti. Pokud má karta provádět kryptografické operace s veřejným klíčem, které jsou značně výpočetně náročné (např. pomocí algoritmů RSA nebo DSS), je vybavena speciálním hardwarovým koprocesorem pro operace modulární aritmetiky s velkými čísly. Paměť bývá v čipové kartě rozdělena na několik oblastí, které jsou realizovány různou technologií. Karta obsahuje paměť RAM (Random Access Memory), která při vyjmutí karty ze čtečky ztrácí svůj obsah a slouží pouze pro uložení dočasných výsledků při výpočtech. Kapacita této paměti u většiny karet nepřevyšuje jednotky kilobajtů. Další oblast paměti bývá realizována pamětí ROM (Read Only Memory), která je nepřepisovatelná a obsahuje operační systém karty a někdy také samotnou aplikaci karty. Kapacita této paměti bývá několik desítek kilobajtů, výjimečně i několik málo stovek kilobajtů. Poslední část paměti je provedena technologií EEPROM (Electrically Erasable Programmable Read Only Memory), která je elektricky programovatelná a mazatelná. Na rozdíl od paměti RAM si však pamatuje informace i po vytažení karty ze čtečky a obsahuje všechny vitální data aplikace. Kapacita této paměti do značné míry ovlivňuje použitelnost karty a na zvyšování kapacity této paměti bývá kladen největší tlak. U běžných současných karet bohužel kapacita paměti EEPROM obvykle nepřevyšuje desítky KB.

Paměť EEPROM, která obsahuje data aplikace, je nejdůležitější paměťovou oblastí čipové karty. Proto je strukturování těchto dat věnována velká pozornost. V současnosti je tato paměť u většiny moderních karet rozdělena na soubory. Tyto soubory mohou být organizovány do stromové struktury, kořenový adresář je nazýván u čipové karty master file (MF), podadresář je nazýván dedicated file (DF) a datový soubor je v terminologii čipových karet elementary file (EF). Přístup k jednotlivým souborů je chráněn pomocí přístupových práv. Tato přístupová práva jsou uživateli přidělena buď na základě toho, že zadá správně některé z definovaných hesel (která

jsou nazývána PIN – Personal Identification Number) nebo na základě úspěšného provedení některé kryptografické operace. Pro manipulaci se soubory čipové karty slouží sada příkazů, které jsou definovány ve standardu ISO 7816. Tyto příkazy nepokrývají všechny potřebné činnosti a často jsou jednotlivými výrobci rozšiřovány o proprietární příkazy.

3 Aplikace čipových karet

Aplikace čipových karet jsou velmi rozmanité a je obtížně vyjmenovat všechny možné aplikace. Pokusíme se naznačit alespoň základní oblasti nasazení čipových karet.

Telefonní karty jsou nejrozšířenější a každému známá aplikace čipových karet. V současnosti jsou jako telefonní karty používány paměťové karty se speciální logikou, které neodpovídají standardům ISO 7816. Vzhledem k nevyhovujícím bezpečnostním vlastnostem těchto čipových karet jsou již v některých zemích už postupně nahrazovány modernějšími kartami s kryptografickým zabezpečením.

Čipová karta je také důležitou součástí systému digitálních mobilních telefonů GSM, kde je nazývána *SIM modul* (zkratka SIM znamená Subscriber Identity Module), ve kterém hraje klíčovou úlohu v kryptografickém zabezpečení telefonního hovoru.

Zdravotní čipová karta slouží jako identifikační průkaz (potenciálního) pacienta a nosič informací o pacientovi.

Čipové karty jsou také klíčovou součástí mnoha systémů *placené televize*, šířené například kabelovým rozvodem nebo satelitním vysíláním, a slouží jako hlídač", zda si divák sledovaný kanál nebo pořad skutečně zaplatil.

Čipové karty hrají také důležitou úlohu při zabezpečení *přístupu k počítačovým systémům*, kde provádějí kryptografickou autentizaci uživatele a mohou sloužit i pro zabezpečení samotného přenosu dat.

Identifikační čipová karta může mít spoustu podob nebo názvů. Podle způsobu použití může fungovat jako elektronický služební průkaz, elektronický občanský průkaz, elektronický řidičský průkaz, studentská karta, průkaz ke vstupu do budovy a podobně.

Elektronické platební systémy s čipovými kartami (obvykle nazývané *elektronické peněženky*) mají sloužit jako náhrada hotovosti pro drobné platby. Jejich funkce je na první pohled podobná funkci magnetických platebních karet, avšak je zde několik významných rozdílů. Hlavní rozdíl spočívá v tom, že elektronická peněženka v sobě obsahuje elektronické peníze a je s ní tedy možno provádět transakce i bez on-line spojení s bankou. Dalším významným rozdílem je zvýšená bezpečnost oproti systémům s magnetickými kartami. PIN zákazníka je ověřován lokálně čipovou kartou a po určitém počtu pokusů o zadání nesprávného PINu se karta zablokuje.

Jednodušší variantou elektronických platebních karet jsou předplatní karty (např. parkovací karty) a věrnostní (anglicky loyalty) čipové karty. Tyto karty neobsahují elektronické peníze, ale tzv. žetony nebo body, které je možno směnit pouze za velmi omezený okruh zboží nebo služeb.

4 Bezpečnost čipových karet

Proč jsou čipové karty tak těsně svázány s bezpečností informačních systémů? Čipové karty totiž poskytují velmi levnou implementaci jednoho z bezpečnostních konceptů, který se anglicky nazývá „tamper resistant hardware". Tento pojem je možno přibližně přeložit do češtiny jako „hardware odolný proti fyzickému útoku", zjednodušeně „bezpečný hardware". Bezpečný hardware je hardwarový modul, obvykle vybavený mikroprocesorem, který obsahuje nějaká chráněná data a algoritmy, které na základě příkazů z vnějšího světa s těmito daty manipulují.

Tato vlastnost se obvykle využívá dvojím způsobem:

1. Bezpečný hardware v sobě obsahuje data, se kterými je možno manipulovat pouze jistým způsobem. Příkladem může být předplatní (telefonní) telefonní čipová karta, která v sobě obsahuje čítač impulsů (chráněná data), který je možno pouze snižovat a nikdy ne zvyšovat.
2. Bezpečný hardware má v sobě tajný kryptografický klíč, který nikdy nevypustí ven a je pouze ochoten s tímto klíčem provést jistou kryptografickou operaci (například zašifrovat data zaslaná z vnějšího světa). Příkladem může být autentizační čipová karta, která prokazuje svou totožnost pomocí zašifrování vložených dat uloženým tajným klíčem. Tomuto typu bezpečného hardware se někdy také říká kryptografický bezpečný hardware.

Velmi rozšířeným jevem, se kterým je možno se setkat u čipových karet, je utajování algoritmů. Algoritmy použité v aplikacích s čipovými kartami (kryptografické i nekryptografické) bývají poměrně často utajovány nebo aspoň „nezveřejňovány". Děje se tak v daleko větší míře než u softwarových aplikací. Proč tomu tak je? Důvodem je to, že vlastnost bezpečného hardware (tj. ochránit „chráněná data" před neoprávněným přístupem) se dá velmi snadno využít i pro ochranu použitých algoritmů před prozrazením. Zatímco u čistě softwarového systému nelze efektivně utajit žádný algoritmus, protože nakonec to vždycky někdo „zreverzuje" a zveřejní (o čemž svědčí příklad kdysi utajovaných kryptografických algoritmů RC2 a RC4), u hardwarového systému tato možnost existuje. A vývojáři aplikací s čipovými kartami ji také zhusta využívají. Zvlášť v minulých letech panovaly v této oblasti až paranoidní názory. Nejen že se utajovaly algoritmy, kryptografické protokoly a datové struktury, ale „nezveřejňovaly" se i samotné příkazy čipových karet. Zvláště komické bylo to, že čipové karty se dodávaly pouze „spolehlivým" a „prověřeným" odběratelům, po podepsání různých závazků a prohlášení doprovázených vysokými smluvními pokutami.

Tento způsob zabezpečení, zvaný „security through obscurity", což znamená přibližně „zabezpečení pomocí zatemnění", je svou účinností asi tak bezpečný, jako ukládání klíče pod rohožku. Samozřejmě, pokud se celá domácnost domluví, že klíč bude pod rohožkou a nikdo o této dohodě neví, je tento způsob bezpečný. Ale stačí, aby jediný člen domácnosti byl spatřen, jak ukládá klíč pod rohožku a bezpečnost je okamžitě zkompromitována. Je pak třeba změnit algoritmus (např. ukládat klíč pod květináč). V současné době, kdy se snažíme o maximální interoperabilitu, o standardizaci a o veřejný audit algoritmů, je způsob zabezpečení „security through obscurity" považován za nevhodný a v běžných softwarových aplikacích se vyskytuje stále méně. Dokonce i v oblasti čipových karet je vidět jistý posun od utajovaných algoritmů a protokolů k veřejným. Svědčí o tom jak silná standardizace (kde standard samozřejmě nemůže být tajný), tak také například technologie JavaCard (která znamená čipovou kartu, programovanou v jazyce Java), která je zcela veřejná.

4.1 Levné útoky na kryptografické moduly

Asi nejjednodušším způsobem, jak "obelstít" čipovou kartu, je využít některé chyby v software čipové karty. Přestože se může zdát, že pravděpodobnost chyby v software je velmi malá opak je pravdou. V software čipových karet se vyskytují často jak funkční chyby (např. špatná kontrola mezních hodnot parametrů některých příkazů a nesprávná reakce na chybné příkazy), tak i chyby v kryptografických algoritmech. Existence funkčních chyb je způsobena především tím, že se jedná o chyby, které nenarušují funkčnost karty v bezpečném prostředí, a proto zůstávají při běžném testování systému neodhaleny. Vzhledem k poměrně důslednému dodržování principu "security through obscurity" tyto chyby často v systému přetrvávají značnou dobu. Nedostatky kryptografických algoritmů bývají často způsobeny poměrně malou paměťovou kapacitou čipové karty a problémy vývojáře s paměťovým prostorem. Je jasné, že první věc, na které se začne šetřit, je kryptografie.

Nenalezne-li útočník v software karty chyby, má další možnosti. Jednou z nich je diferenciální chybová analýza (zkráceně DFA, Differential Fault Analysis). DFA vychází z následující myšlenky: software čipové karty (nebo jiného hardwarového kryptografického zařízení) byl navržen tak, aby se choval správně za předpokladu, že při činnosti procesoru karty nedojde k hardwarové chybě. Pokud ale při provádění programu čipové karty dojde k hardwarové chybě (například k nesprávnému provedení některé instrukce procesoru) je pravděpodobné, že karta se zachová nestandardně způsobem, který napomůže útočníkovi při útoku na kartu. Je jasné, že útočník nebude čekat, až při provádění programu dojde k chybě, ale pokusí se chybu sám vyvolat. Jaké má proto možnosti? Velmi široké. Jednou z prvních diskutovaných možností je ozařování čipu rentgenovými paprsky nebo jiným podobným zářením. Tento útok při dostatečné intenzitě záření vede k úspěchu, ale je poměrně nesnadno proveditelný v amatérských podmínkách. Útočník však má mnoho jiných možností. Může vyvolat chybu pomocí drastického zvýšení frekvence hodinového signálu (uvědomme si, že hodinový signál je téměř u všech čipových karet generován externě) nebo pomocí velmi krátkých impulsů na některém signálovém vodiči (tzv. „glitch attack"). Útočník také může dostat čip do nestandardního stavu zvýšením nebo

snížením teploty (tzv. teplotní útok) a zvýšením nebo snížením napájecího napětí (tzv. napěťový útok). Z těchto typů útoku je nejvíce používaný „glitch attack“, neboť dovoluje útočníkovi zaměřit se na chybné provedení konkrétní instrukce procesoru.

Časový útok (timing attack) objevil poměrně nedávno (v roce 1996) Paul Kocher. Tento útok vychází z předpokladu, že některé algoritmy potřebují různý čas pro zpracování různých vstupních dat.

Příčiny této časové závislosti jsou rozmanité. Mezi nejčastější příčiny časové závislosti programů na datech patří:

- optimalizace programů (například neprovedení některých matematických operací při nulové nebo jedničkové hodnotě operandů)
- podmíněné skoky
- nestejná lokalita odkazů do paměti v případě použití vyrovnávací paměti (cache hits)
- existence instrukcí s různou dobou provádění (např. násobení, dělení, rotace, posuvy)

Výsledkem je, že čas provádění programu závisí na zpracovávaných datech i na klíči, což oboje útočníka zajímá. Vzniká tedy časový skrytý kanál, kterým ”vytéká” část informace (někdy jen několik bitů, ale i to je pro útočníka cenná informace) z bezpečného hardware ven.

Principu "vytékání" informace z bezpečného hardware využívá i jiný zajímavý útok - výkonová analýza (PA, Power Analysis) a diferenciální výkonová analýza (DPA, Differential Power Analysis). Tyto útoky měří během činnosti kryptografického mosulu jeho proudový odběr a podle něj se znaží zjistit, jaké instrukce modul právě provádí a s jakými daty je provádí.

4.2 Náročnější útoky

Výše uvedené útoky patřily mezi útoky logické, neboť nevyžadovaly fyzickou manipulaci s čipem karty. Pokud tyto útoky nejsou pro danou čipovou kartu použitelné, nastupují útoky, které přímo manipulují s čipem karty, a které se nazývají útoky fyzické. Přestože pro tyto útoky je již třeba speciální technické vybavení, tyto útoky nemusí být mimo možnosti amatérů, protože často se toto vybavení nachází na univerzitách nebo bývá možné si toto vybavení na několik hodin pronajmout v laboratoři. A pro průmysl hackerských satelitních karet, který disponuje poměrně značným kapitálem, je většina těchto útoků cenově dostupná.

Jako příklad uveďme alespoň příklady některých fyzických útoků:

Pasivní útoky

- Mikroskopické sondy (microprobes), které často obsahují až 9 jehel, umožňují elektrické sledování signálů na čipu.

- Elektronový mikroskop umožní sledování signálů na sběrnici.
- Elektrooptické vzorkování sleduje krystal niobátu lithia laserovým paprskem a tím zjišťuje přítomnost elektrostatického pole pod krystalem.
- Spodní rentgenování umožňuje pozorování tranzistorů zespod čipu na vlnové délce, pro kterou je křemíkový substrát průhledný.

Aktivní útoky

- Laserový nůž umožňuje přerušení spojů a odstranění pasivační vrstvy
- Iontový paprsek umožní vytvoření nových spojů
- Selektivní suché leptání umožňuje oklamat senzory, testující přítomnost pasivační vrstvy

Výše uvedeným výčtem samozřejmě nekončí možnost útoků na čipové karty. Na druhé straně čipová karta má možnost těmto útokům odolat a skutečně dobré čipové karty jim odolají. Vše zůstává na vývojářích systémů, jaké čipové karty si z poměrně široké nabídky vyberou a jakým způsobem je použijí.

5 Co na závěr

Čipové karty jsou jako každý jiný produkt: jsou čipové karty kvalitní a tedy bezpečné, jsou čipové karty méně kvalitní a tedy i méně bezpečné. Kvalitu čipových karet však nemůžeme nikdy posuzovat podle marketingových tvrzení výrobce nebo vývojáře aplikace, ale pouze podle výsledků nezávislého auditu. V tutoriálu uvedené příklady nejsou selháním čipové karty jako takové, ale selháním člověka, který ve své nafoukanosti a domýšlivosti nerespektoval základní bezpečnostní zásady a použil tyto karty způsobem, který nebyl správný.

Literatura

1. Ross J. Anderson, Markus G. Kuhn: Low Cost Attacks on Tamper Resistant Devices. *Security Protocols, 5th International Workshop* (M. Lomas, ed.), Lecture Notes in Computer Science no. 1361, Springer Verlag, 1997.
2. Ross J. Anderson: Security Engineering: A Guide to Building Dependable Distributed Systems. John Wiley & Sons, UK, 2001. ISBN: 0-471-38922-6.
3. U.S. Federal Information Processing Standard 140-2: Security Requirements for Cryptographic Modules - http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf.
4. Oliver Kommerling, Markus G. Kuhn: Design Principles for Tamper-Resistant Smartcard Processors. *USENIX Workshop on Smartcard Technology (Smartcard '99)*, USENIX Association, 1999.