Crypto-World

Informační sešit GCUCMP

ISSN 1801-2140

Ročník 12, číslo 1/2010 **17. leden 2010**

1/2010

Připravil: Mgr. Pavel Vondruška

Sešit je přednostně distribuován registrovaným čtenářům.
Starší sešity jsou dostupné na adrese
http://crypto-world.info
(1345 registrovaných odběratelů)

Obsah :

A. Analýza Blue Midnight Wish – útok na vzor,

Vlastimil Klíma, kryptolog konzultant, Praha
(http://cryptography.hyperlink.cz, v.klima@volny.cz)
Prof. Danilo Gligoroski, Norwegian University of Science and Technogy, Norway (danilog@item.ntnu.no , http://www.item.ntnu.no/people/personalpages/fac/danilog/start)

Článek navazuje na předchozí příspěvek Crypto-Worldu 2009 v čísle 12, ale i v číslech 3 a 7-8. Je první z malé série, kterou chceme věnovat úvahám o bezpečnosti BMW a stimulovat útoky na něj nebo analýzy a prezentovat i otevřené problémy. Ty by se mohly stát předmětem studentských prací. Proč? Velkou výhodou oproti jiným tématům je, že tyto rozbory jsou nyní velmi žádané, ať s negativním nebo pozitivním výsledkem, takže když problém bude vyřešen, nebo naopak nebude vyřešen a ukáže se, že je složitý, je to potřebný, žádaný a velmi dobře publikovatelný výsledek.

Malá aktualizace – urychlení BMW

Reference platform: Intel Core 2 Duo, 2.4 GHz, Windows Vista Ultimate 64-bit edition
Compiler: Intel C++ 11.1.46

Mode: 32-bit (x86) Performance in cycles/byte for different message lengths (in bytes)

	1	8	64	576	1024	1536	4096	100000	eBash version
BMW224/256	1129	142.63	25.33	9.42	8.52	8.14	7.69	7.45	bmw256/optc02
BMW384/512	1321	165.13	22.33	6.29	5.73	5.28	4.73	4.4	bmw512/optx86sse2

Mode: 64-bit (x64) Performance in cycles/byte for different message lengths (in bytes)

	1	8	64	576	1024	1536	4096	100000	eBash version
BMW224/256	1081	135.13	24.77	9.02	8.18	7.76	7.29	7.02	bmw256/optc04
BMW384/512	1105	138.13	17.27	5.09	4.58	4.22	4.05	3.48	bmw512/optc04

Obr.1: Nejnovější urychlení BMW v SW, viz [1]

Na obrázku 1 jsou uvedeny nejnovější vynikající časy pro BMW, ještě lepší, než ty, uvedené v minulém čísle Crypto-Worldu. Tím se BMW stala ještě viditelněji nejrychlejším kandidátem v SW ve všech ukazatelích. Je to zásluha týmu BMW (vyjma mě).

Analýza celku a částí

BMW lze zkoumat jako celek, avšak to se velice brzo zastavíme, neboť dostáváme soustavu rovnic, jejíž řešení neumíme nalézt ani v té nej-nej-nejzjednodušenější podobě. Proto budeme chtít řešit alespoň dílčí úlohy, které by mohly k řešení vést. Ukážeme však, že ani dílčí úlohy, týkající se nejjednodušších (atomárních) stavebních bloků BMW, neumíme řešit. A teď doufám, že se najde někdo, kdo bude v opozici, kdo najde nějaké řešení, chybu, nedostatek, odchylku od očekávaného chování atomárních bloků nebo vyšších celků nebo BMW nebo její architektury. Jakákoliv opozice je vítaná, zejména zde, na stránkách Crypto-Worldu, e-mailem, v diskusní skupině sci.crypt, v poštovní konferenci NISTu k hašovacím funkcím apod.

Označení
Článek bude využívat označení zavedené v Crypto-Worldu 12/2009. Připomeňme jen šířku slova w = 32 nebo 64 bitů, délku bloku zprávy a průběžné haše n = 16*w (mají 16 slov) a výpočet haše:

1. Předzpracování
(a) Doplň zprávu M jednoznačným definovaným způsobem o délku zprávy v bitech a doplněk
(b) Rozděl zprávu na celistvý násobek (N) m-bitových bloků $M^{(1)}$, ..., $M^{(N)}$.
(c) Nastav počáteční hodnotu průběžné haše $H^{(0)}$ na konstantu ($CONST^0$).
2. Výpočet haše
For i = 1 to N: $H^{(i)} = f(M^{(i)}, H^{(i-1)})$.
3. Finalizace
$H^{final} = f(H^{(N)}, CONST^{final})$, kde $CONST^{final}$ je konstanta.
4. Závěr
H(M) = dolních n bitů z hodnoty H^{final}.

BMW vždy projde minimálně dvě iterace
Jak ukazuje následující schéma, BMW vždy projde minimálně dvě iterace - a to první a poslední. Kromě toho projde volitelně podle délky zpracovávané zprávy ještě určité množství tzv. vnitřních bloků mezi prvním a posledním, viz obr. 3 a 4. První a poslední blok mají pevně nastavenu hodnotu H. U prvního bloku je to konstanta $CONST^0$, u posledního bloku je to konstanta $CONST^{final}$. V první iteraci zpracovává kompresní funkce $f(M^{(1)}, CONST^0)$ první blok zprávy $M^{(1)}$ a konstantu $CONST^0$. Pokud tento blok zprávy není zároveň blokem posledním, následují ještě vnitřní iterace. Výsledkem je poslední průběžná haš H, která vstupuje v roli bloku zprávy do finalizace $f(H, CONST^{final})$.

Rychlost a složitost
U BMW jsme použili malý úskok nebo trik, chcete-li, stejně jako někteří další návrháři SHA-3. U funkcí SHA-2 postačí zkoumat jeden blok, který může být také blokem posledním. Tento blok lze zapsat rovnicemi, které udávají složitost problému, pokud se nenajde rychlejší řešení. Soustava rovnic u BMW je dvakrát větší, neboť k zápisu libovolného problému potřebujeme zapsat rovnice pro dva bloky. Trik je v tom, že na rychlost hašování toto má vliv jen u krátkých zpráv, zatímco u delších zpráv se zpracování přídavného posledního bloku „rozpustí“ v čase, potřebném na zpracování velkého počtu bloků předchozích. Rozdíl ve složitosti je ale ohromný, jak vidíme na obrázku 3. Na obrázku 4 vidíme realizaci BMW pro eventuelní libovolný počet vnitřních bloků a na obrázku 2 je jedna iterace zvětšená.

Úloha první - hledání vzoru
Pokud útočník bude znát vzor k dané haši (ať na něj přijde jak chce), bude zcela jistě znát vzor této haše v posledním bloku. K důkazu složitosti nalezení vzoru hašovací funkce postačí proto ukázat, že je příliš složité nalezení vzoru pro poslední iteraci. Situaci znázorňuje obrázek 5. Útočník zná výstupní haš o 8 slovech a hledá vstupní blok M o 16 slovech. Hodnota H, která vstupuje do posledního bloku je neměnná konstanta H^{final}. Útočník zná hnědou hodnotu a hledá žluté proměnné.

Hledání vzoru poslední iterace je ekvivalentní hledání řešení (M) soustavy rovnic na obrázku 6. Je to soustava, kterou můžeme zapsat atomárními operacemi jednoduše následovně:

$$\mathbf{Q_a = A_2(A_1(M \oplus H^{final})) + ROTL^1(H^{final}),} \quad \mathbf{(S1)}$$
$$\mathbf{Q_b = T^L(T^U(Q_a) + ((B(rotM) + K) \oplus ROTL^7(H^{final}))),} \quad \mathbf{(S2)}$$
$$\mathbf{G = (M \oplus L_a(Q_b)) + (Q_a \oplus L_b(Q_b)),} \quad \mathbf{(S3)}$$
$$\mathbf{hash = 8\_lswords\_of(\ f_6(G)\ ).} \quad \mathbf{(S4)}$$

Povšimněme si, že pokud bychom chtěli zkoušet vstup do posledního (kompresního) bloku f_2, hrubou silou nebo jakkoliv jinak, jedním ze vstupů je i M. Zpětný chod v bloku f_2 je tedy už přímo zpětným chodem celé kompresní funkce f.

Povšimněme si dále, že druhým vstupem bloku f_2 je Q_a, což je bijektivní obraz M, neboť $Q_a = A_2(A_1(M \oplus const1)) + const2$. Jakékoliv předvídání části nebo celé hodnoty Q_a je tedy ekvivalentní předvídání části nebo celé hodnoty M.

Třetí vstup do bloku f_2 je Q_b, což je pseudonáhodná funkce M. Zde můžeme provádět určité manipulace, protože jedné hodnotě Q_b může odpovídat žádný, jeden nebo mnoho vzorů M. Pokud bychom například pro více zpráv M docílili toho, že

$$\mathbf{Q_b = const,} \qquad \mathbf{(R0)}$$

máme vztahy:

$Q_a = A_2(A_1(M \oplus const)) + const,$
$const = T^L(T^U(Q_a) + ((B(rotM) + const) \oplus const)),$
$G = (M \oplus const) + (Q_a \oplus const),$
$hash = 8\_lswords\_of(f_6(G)),$

kde **const jsou nějaké konstanty (obecně různé)**. Druhou rovnici pro jednoduchost neuvažujme, tj. předpokládáme, že umíme nalézt několik (třeba velmi mnoho) zpráv M, které vedou na stejnou hodnotu Q_b a pracujeme pouze s těmito zprávami. Zbývají rovnice

$$\mathbf{Q_a = A_2(A_1(M \oplus const)) + const,} \qquad \mathbf{(R1)}$$
$$\mathbf{G = (M \oplus const) + (Q_a \oplus const),} \qquad \mathbf{(R2)}$$
$$\mathbf{hash = 8\_lswords\_of(f_6(G)),} \qquad \mathbf{(R3)}$$

tj.

$$G = (M \oplus const) + ((A_2(A_1(M \oplus const)) + const) \oplus const), \qquad (R4)$$
$$hash = 8\_lswords\_of(f_6(G)), \qquad (R5)$$

neboli

$$\mathbf{hash = 8\_lswords\_of(f_6((M \oplus const) + ((A_2(A_1(M \oplus const)) + const) \oplus const))).} \qquad \mathbf{(R6)}$$

Udělejme malou odbočku. Pokud uvažujeme nulové konstanty v (R6), máme tvar

$$\mathbf{hash = 8\_lswords\_of(f_6(\ M + A_2A_1(M) \ )).} \qquad \mathbf{(R6a)}$$

Ukazuje se, že by bylo dobré prozkoumat **vlastnosti funkce**

$$\mathbf{M + A_2A_1(M),} \qquad \mathbf{(R7)}$$

což nevypadá vůbec složitě (avšak složité je !!! a doufáme, že se najdou oponenti) nebo obecněji vlastnosti funkce

$$\mathbf{A_2(A_1(M \oplus const_1)) + const_2,} \qquad \mathbf{(R8)}$$

neboli vlastnosti Q_a jakožto funkce M. No ale teď se vraťme k hlavní rovnici.

Zdá se, že v soustavě (R6) máme velmi mnoho stupňů volnosti, neboť pevně je určeno pouze 8 slov (hash) bijektivního obrazu $f_6(G)$ hodnoty G, zatímco zpráva M má 16 volných slov. Bohužel, do soustavy (R6) nevstupuje M libovolné, ale pouze ty M, pro něž je Q_b = const pro zvolenou konstantu. Proto je velmi důležité zkoumat jak moc je Q_b náhodná a zda by nešlo najít hodně zpráv M se stejnou $Q_b(M)$. **Výzkum Q_b jako funkce M je klíčový.**

Zjednodušme úlohu (R0, R6):

- Předpokládejme, že výzkum Q_b přinesl velké ovoce a že každé řešení (R0) je už automaticky řešením (R6). Potom, pokud umíme řešit soustavu **(R0)**, umíme najít vzor finální iterace.
- Předpokládejme, že výzkum (R6) přinesl velké ovoce a že každé řešení (R6) je už automaticky řešením (R0). Potom, pokud umíme řešit soustavu **(R6)**, umíme najít vzor finální iterace.

Obě dvě zjednodušené úlohy mohou přispět k posouzení bezpečnosti BMW. Zejména úloha (R6) se zdá jednoduchá.

Pojďme ke složitější úloze. Pokud bychom neuměli nalézat zprávy M pro něž je Q_b = const, hodnota Q_b by byla obecně proměnná a řešili bychom soustavu (S1 - S4). Bez meziproměnných to dává sice jen jednu rovnici, ale nepříliš vábnou:

$$\mathbf{hash = 8\_lswords\_of(\ f_6((M \oplus L_a(T^L(T^U(A_2(A_1(M \oplus H^{final})) + ROTL^1(H^{final})) + ((B(rotM) + K) \oplus ROTL^7(H^{final}))))) + ((A_2(A_1(M \oplus H^{final})) + ROTL^1(H^{final})) \oplus L_b(T^L(T^U(A_2(A_1(M \oplus H^{final})) + ROTL^1(H^{final})) + ((B(rotM) + K) \oplus ROTL^7(H^{final})))))\ ).} \qquad \mathbf{(S5)}$$

Soustava (S5) je jednou rovnicí, ale pro osm známých slov na levé straně (hash) a 16 neznámých slov (M), čili je to soustava osmi rovnic o 16 neznámých. To je pro útočníka nadějné - je tu velmi mnoho stupňů volnosti pro nalezení řešení.

Jen pro zajímavost, pokud konstanty nahradíme symbolicky slovem const, máme

$$\mathbf{hash = 8\_lswords\_of(\ f_6((M \oplus L_a(T^L(T^U(A_2(A_1(M \oplus const)) + const) + ((B(rotM) + const) \oplus const)))) + ((A_2(A_1(M \oplus const)) + const) \oplus L_b(T^L(T^U(A_2(A_1(M \oplus const)) + const) + ((B(rotM) + const) \oplus const)))\ )),} \qquad \mathbf{(S5a)}$$

a pokud všechny konstanty uvažujeme nulové, máme soustavu

$$\mathbf{Q_a = A_2A_1(M),} \qquad \mathbf{(S1b)}$$
$$\mathbf{Q_b = T^L(T^U(Q_a) + B(rotM)),} \qquad \mathbf{(S2b)}$$
$$\mathbf{G = (M \oplus L_a(Q_b)) + (Q_a \oplus L_b(Q_b)),} \qquad \mathbf{(S3b)}$$
$$\mathbf{hash = 8\_lswords\_of(\ f_6(G)\ ).} \qquad \mathbf{(S4b)}$$

neboli

$$\mathbf{hash = 8\_lswords\_of(\ f_6((M \oplus L_a(T^L(T^UA_2A_1(M) + B(rotM)))) + (A_2A_1(M) \oplus L_b(T^L(T^UA_2A_1(M) + B(rotM))))),} \qquad \mathbf{(S5b)}$$

což je také hodné zkoumání. Nalezení vzoru pro finální iteraci je ekvivalentní úloze (S1-S4), resp. (S5), která je značně složitá. Nalezení jakékoliv zkratky v řešení by pomohlo lépe pochopit bezpečnost BMW.

Pojďme dále ke složitější úloze.

Předpokládejme, že útočník umí řešit předchozí úlohu nalezení vzoru pro finální iteraci. Umí tedy k zadané finální hodnotě haše H o n bitech nalézt blok zprávy X o 2n bitech, vedoucí společně s H^{final} k výsledné haši H = 8_lswords_of(f(X, H^{final})).

Tuto schopnost však útočník nemůže využít k dokončení útoku, neboť nyní má k dispozici výstup X a musí k němu nalézt skutečný blok zprávy M. Teď má ovšem zadanou výstupní hodnotu X o 2n bitech, zatímco v předchozí úloze měl jenom n bitů. Jinými slovy je teď neznámých stejně jako rovnic a nadbytečné stupně volnosti z minulé úlohy jsou pryč. Jedná se o úlohu

$X = f(M, CONST^{(0)})$.

Můžeme ji zapsat jako

$$Q_a = A_2(A_1(M \oplus CONST^{(0)})) + ROTL^1(CONST^{(0)}),$$
$$Q_b = T^L(T^U(Q_a) + ((B(rotM) + K) \oplus ROTL^7(CONST^{(0)})))$$
$$G = (M \oplus L_a(Q_b)) + (Q_a \oplus L_b(Q_b)),$$
$$X = (f_6(G)).$$

Protože hodnotu G můžeme dopočítat z X, je to pro nás známá hodnota. Takže úloha hledání vnitřního vzoru je ekvivalentní řešení soustavy (T1 - T3) pro zadané konstanty:

$$\mathbf{Q_a = A_2(A_1(M \oplus const_1)) + const_1,} \quad \mathbf{(T1)}$$
$$\mathbf{Q_b = T^L(T^U(Q_a) + ((B(rotM) + const_2) \oplus const_3)),} \quad \mathbf{(T2)}$$
$$\mathbf{const_4 = (M \oplus L_a(Q_b)) + (Q_a \oplus L_b(Q_b)).} \quad \mathbf{(T3)}$$

Poznámka. Dále je tu podmínka, že zpracovávaná zpráva by měla mít platný padding (tj. minimálně 65 bitů by mělo rezervovanou hodnotu), ale to bychom útočníkovi mohli prominout. K útoku na BMW by postačilo pouze najít nějakou negativní vlastnost, nemusí být zničena celá, protože v současné době je ve hře stále ještě 14 kandidátů. Pravda, BMW je k odstřelu na prvním místě, protože je nejrychlejší.

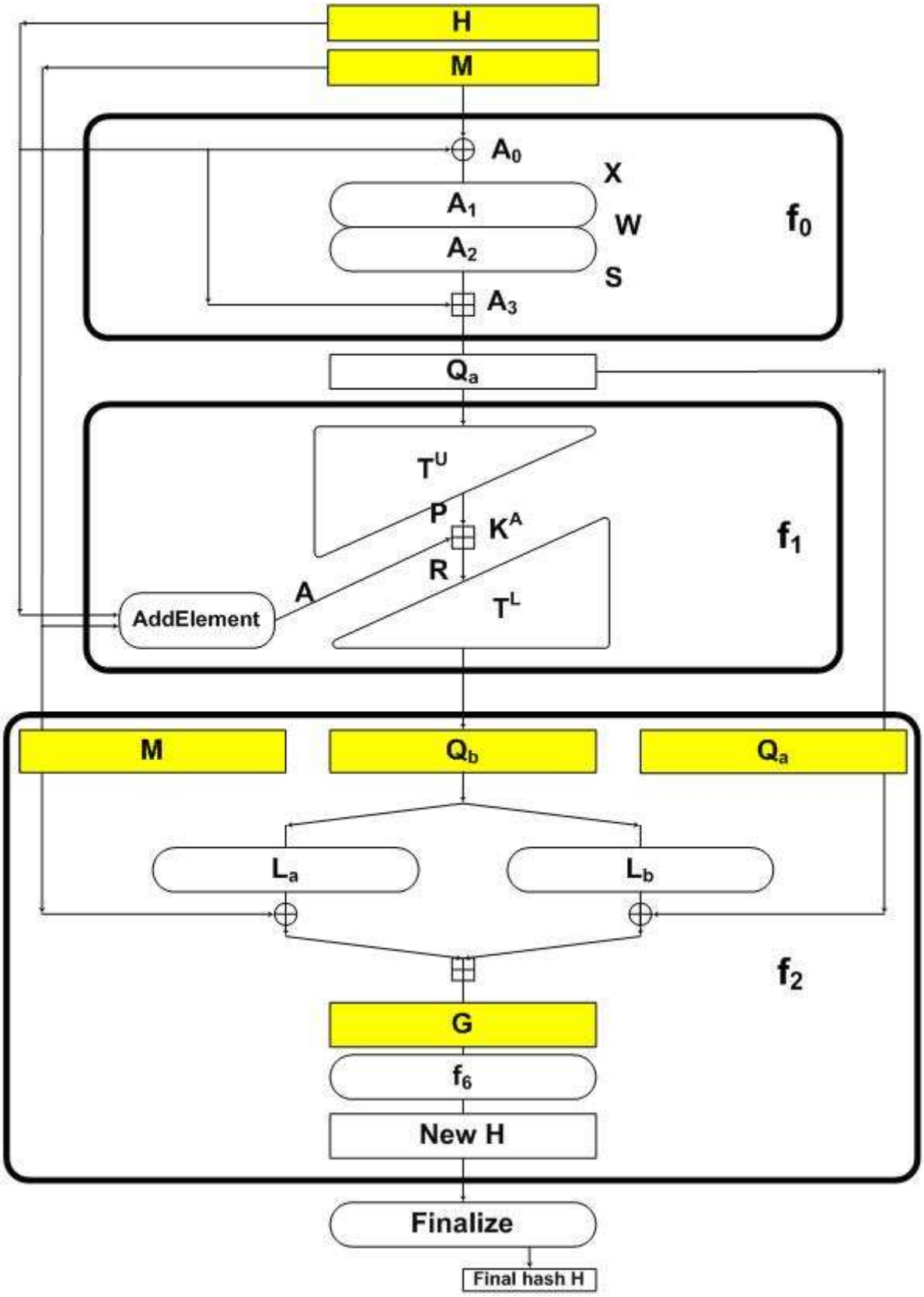
Závěr

V tomto článku jsme uvedli několik dílčích úloh, které se objevují při hledání vzoru k hašovací funkci BMW. Řešení všech těchto úloh je otevřené. Pro přehlednost jsou zvýrazněny tučně. Velice doporučujeme začít s analýzou těch nejjednodušších (R7) nebo (R8) a dále (R6a), (R0), (R6) a dále.

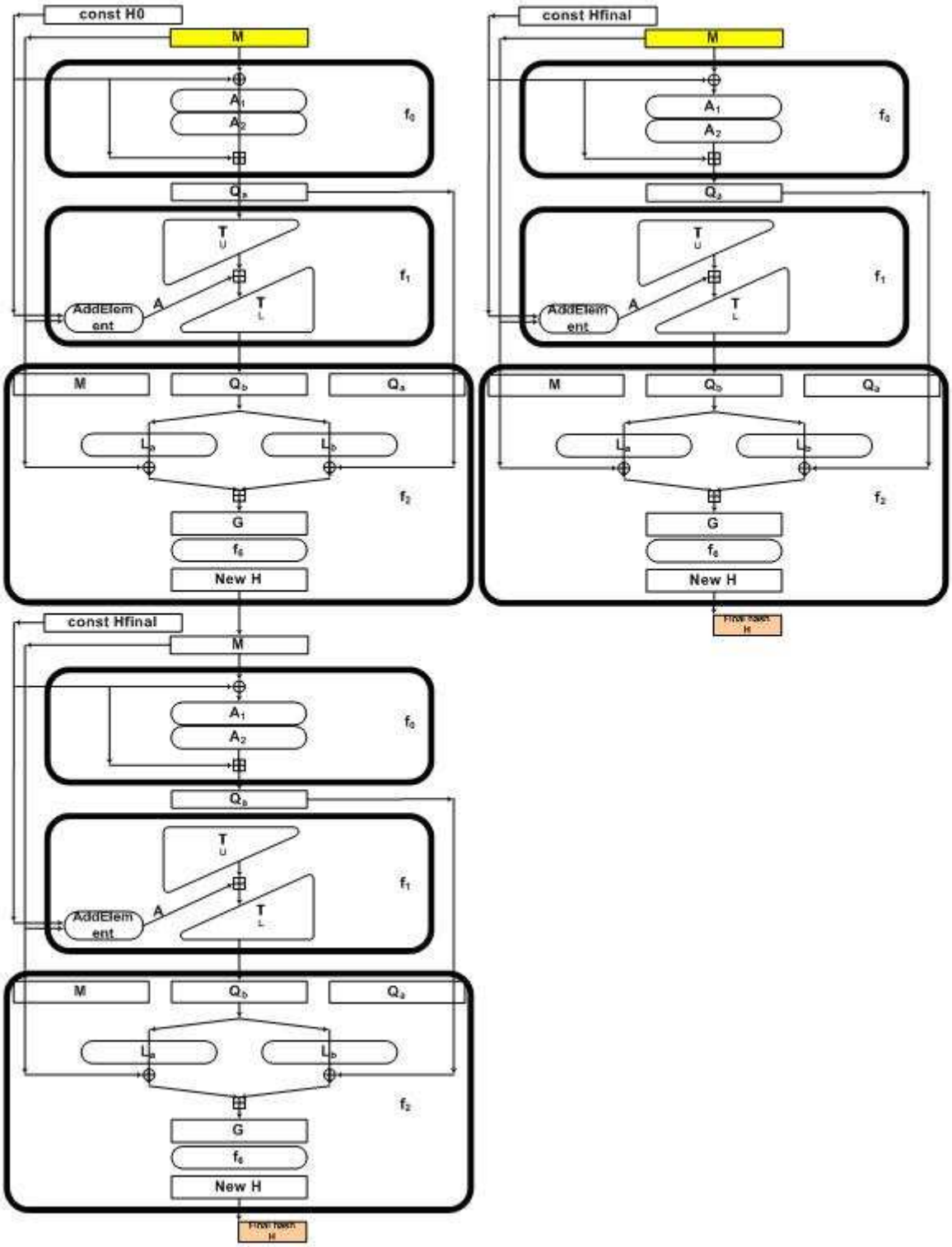
Literatura

[1] domácí stránka týmu BMW: http://www.q2s.ntnu.no/sha3_nist_competition/start
[2] stránka NIST k projektu SHA-3: http://csrc.nist.gov/groups/ST/hash/index.html
[3] dokumenty a analýzy BMW a průběžné novinky k projektu SHA-3: http://cryptography.hyperlink.cz/BMW/BMW_CZ.html
[4] Danilo Gligoroski, Vlastimil Klima, On Blue Midnight Wish Decomposition, SantaCrypt 2009, Dec. 3-4, 2009, Prague, Czech Republic, Proceedings of SantaCrypt 2009, ISBN 978-80-904257-0-5, pp. 41-51

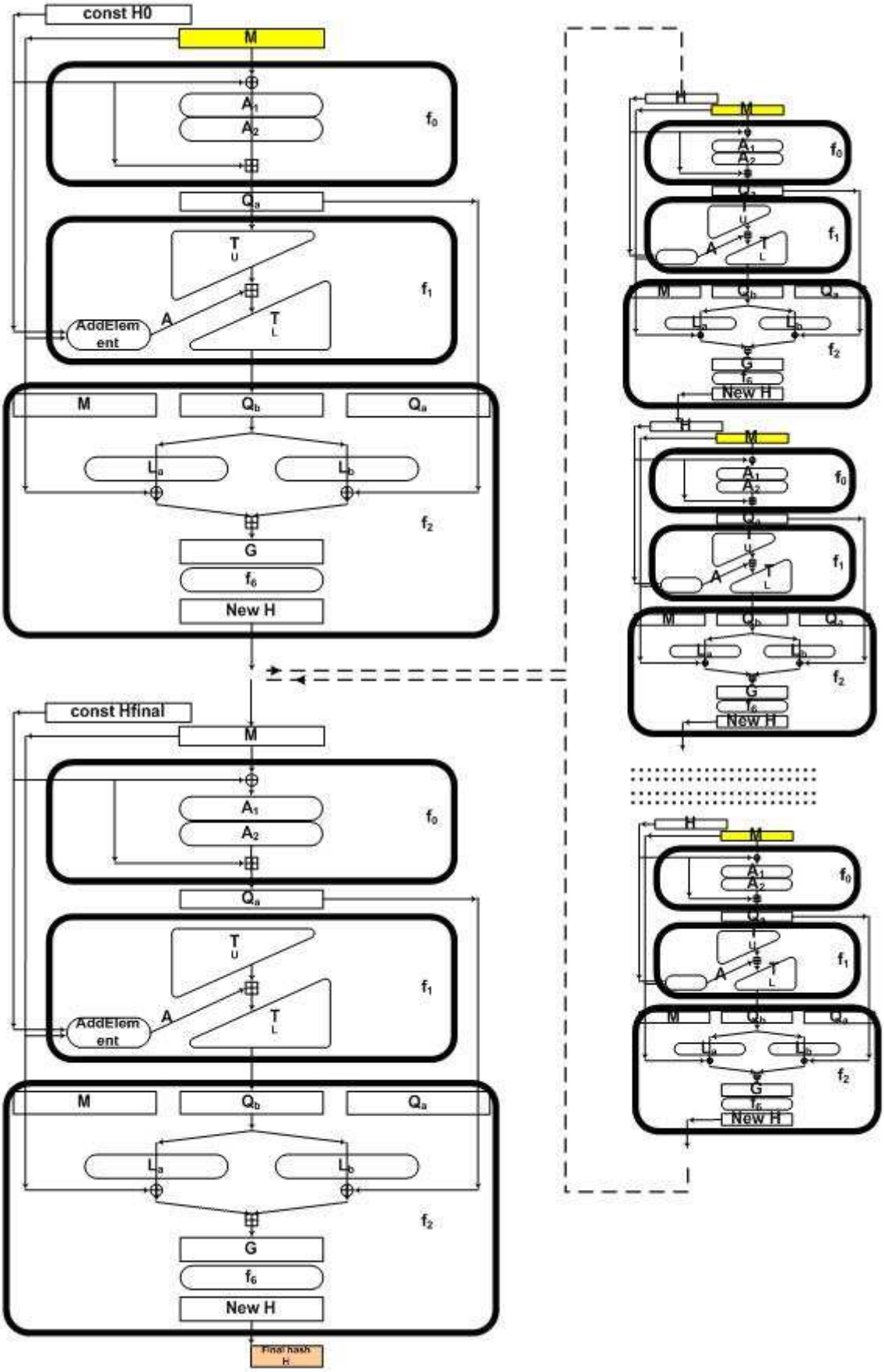
Příloha – Obrázky



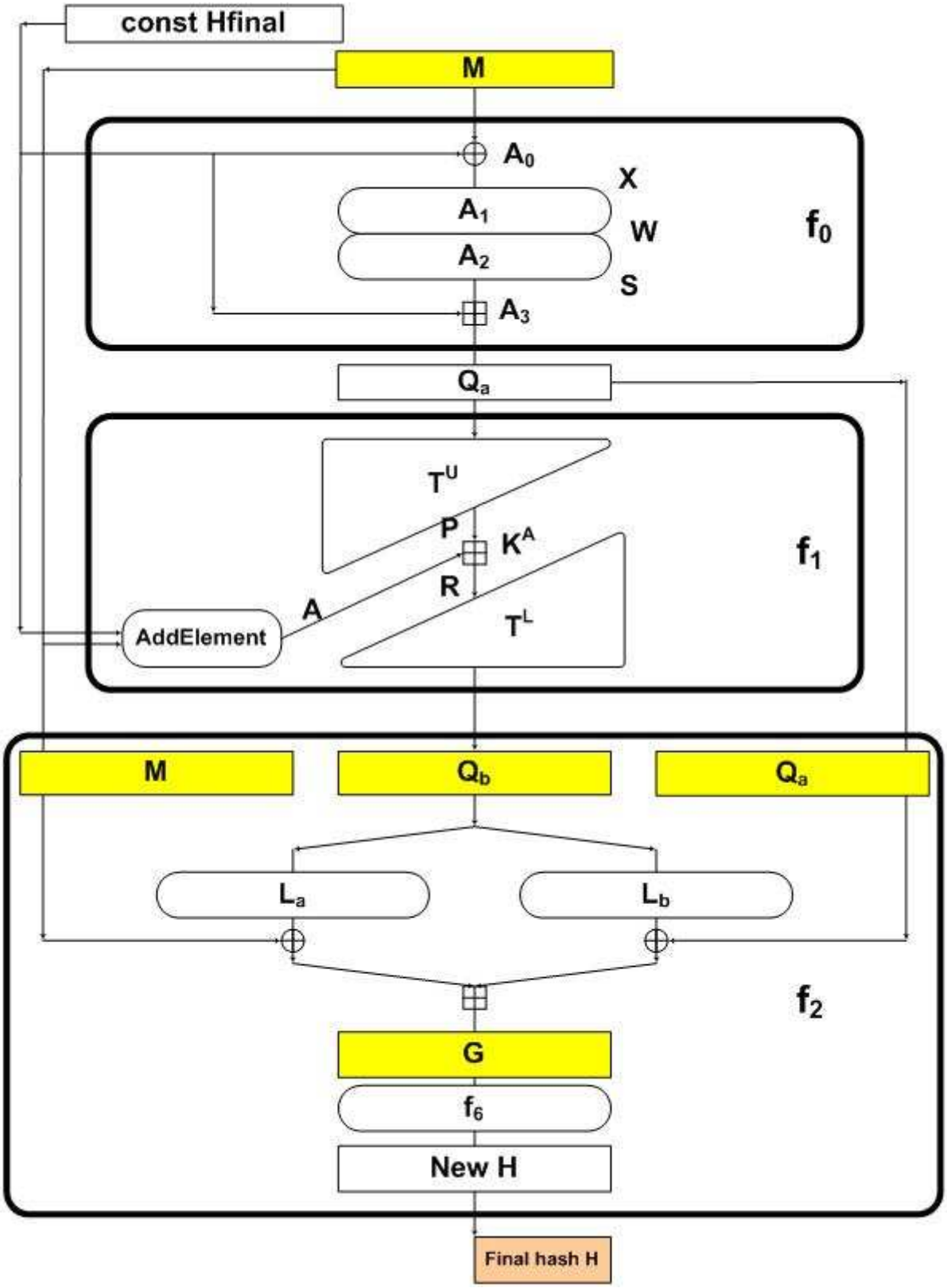
Obr.2: Kompresní funkce BMW



Obr.3: Rozdíl složitosti BMW bez a s přidanou finalizací (jediný vstup je jeden blok na počátku, výstup je na konci, vše ostatní je funkce zpracování jednoho bloku)



Obr.4: BMW s vnitřními bloky



Obr.5: Hledání vzoru u poslední iterace

Obr.6: Hledání vzoru poslední iterace je ekvivalentní řešení této soustavy rovnic, kde H je konstanta H^{final} a M je hledaný blok zprávy

B. Tajné písmo Martina Kukučína

Jozef Kollár, jmkollar@penguin.sk, Katedra matematiky a deskriptívnej geometrie, Stavebná fakulta STU v Bratislave

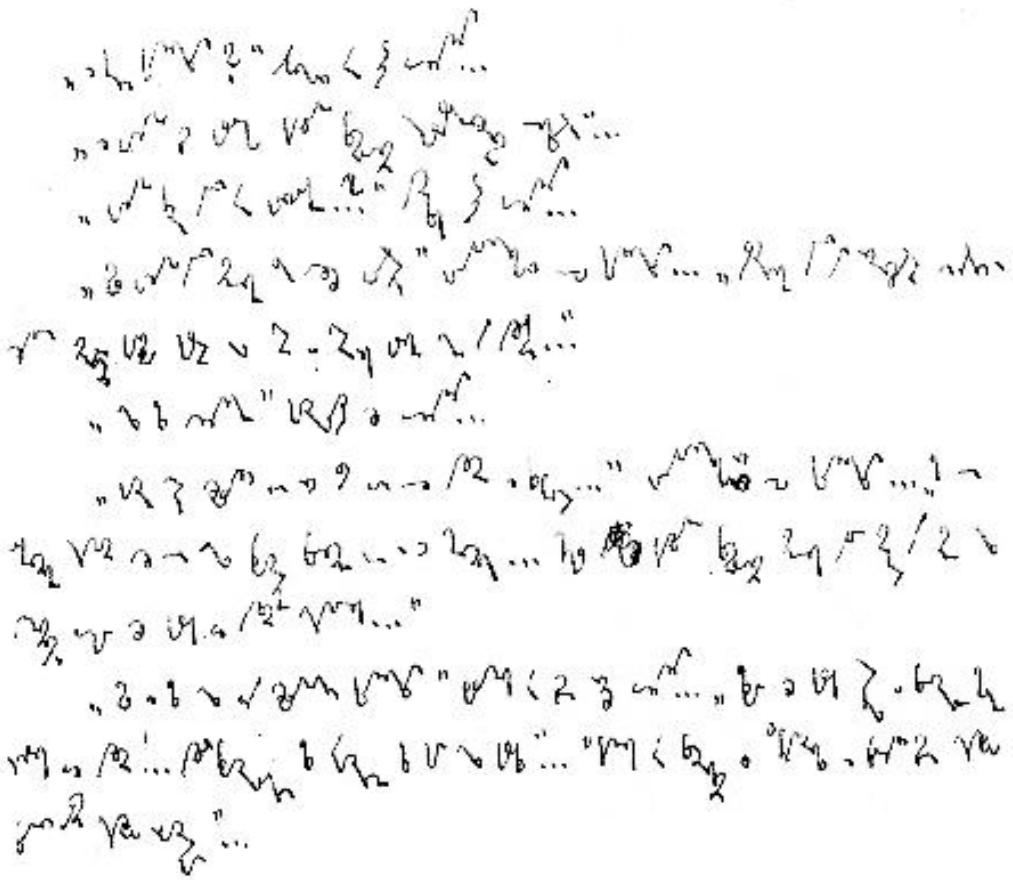
1. Curriculum Vitae

Martin Kukučín (1860-1928), vlastným menom Matej Bencúr, bol jedným z najvýznamnejších slovenských spisovateľov konca 19. a začiatku 20. storočia. Narodil sa v sedliackej rodine v obci Jasenová. Študoval na gymnáziách v Revúcej, Martine, Banskej Bystrici, Kežmarku a v Šoproni. Pôvodne zamýšľal študovať teológiu, ale v dôsledku maďarizačných tlakov sa napokon rozhodol pre štúdium medicíny v Prahe. Od roku 1893 pôsobil ako lekár na ostrove Brač na území dnešného Chorvátska. V roku 1908 odchádza do Punta Arenas v Chile, kde rovnako pôsobí ako lekár. V roku 1922 sa vracia na Slovensko a v rokoch 1924-1928 je opäť v Chorvátsku, na krátky čas v Chile a znova v Chorvátsku, kde v roku 1928 umiera v Pakraci.

Kukučín bol najvýznamnejším predstaviteľom slovenského literárneho realizmu a zakladateľ modernej slovenskej prózy. Medzi jeho diela patria napr.: *Rysavá jalovica*, *Neprebudený*, *Dom v stráni* a *Mať volá*.

2. Stenografia, tajné písmo, alebo šifra?

Kukučín bol autorom nielen mnohých významných slovenských literárnych diel, ale aj autorom zaujímavého písma, ktoré ako sa zdá, používal jedine on sám. Podľa [1], ako aj niektorých internetových zdrojov, toto písmo definitívne vzniklo až v roku 1910, počas jeho pobytu v Punta Arenas. Tam aj vznikali prvé literárne diela a zápisky písané týmto písmom. Na obr. 1 je ukážka z rukopisu románu *Bohumil Valizlosť Zábor* (str. 21 rukopisu). Kukučín pomocou tohto písma písal svoje poznámky, pôvodné rukopisy svojich diel a údajne aj svoje lekárske záznamy. Niektoré zdroje dokonca uvádzajú, že práve snaha o utajenie lekárskych záznamov bola dôvodom vytvorenia tohto tajného písma. Jeho prvé literárne diela zapísané týmto písmom boli *Mať volá* a *Prechádzky po Patagónii* [1].
Kukučín písal svoje zápisky, ako aj viaceré rukopisy svojich diel, pomocou tohto písma až do konca svojho života. Dôvod, prečo tak robil, zostáva dodnes záhadou a zrejme ňou zostane aj naďalej. V [1] sa stenografka Eva Ihnátová pokúšala rozoberať viaceré dôvody prečo tak robil, ale neprišla ku žiadnemu jednoznačnému záveru, pretože každý z uvedených dôvodov má zhruba rovnako veľa svojich argumentov pre a proti. Jeho písmo nebolo primárne určené na utajovanie informácii, rovnako ako neumožňovalo rýchly zápis, ako to umožňuje stenografia. Zrejme to ale nebol ani rozmar a chvíľková zábava, keďže Kukučín vydržal používať toto svoje písmo až do konca života, čo bolo takmer 20 rokov. Pravdepodobné je, že tento spôsob zápisu slúžil čiastočne na utajovanie, čiastočne na ozvláštnenie zápisu a bavenie sa a možno mal Kukučín aj nejaké iné pohnútky používať takéto písmo. Jediný dôvod, ktorý sa s vysokou pravdepodobnosťou dá vylúčiť je zrýchlenie zápisu, teda účel stenografie, pretože Kukučínove písmo zápis skôr spomaľuje než zrýchľuje.



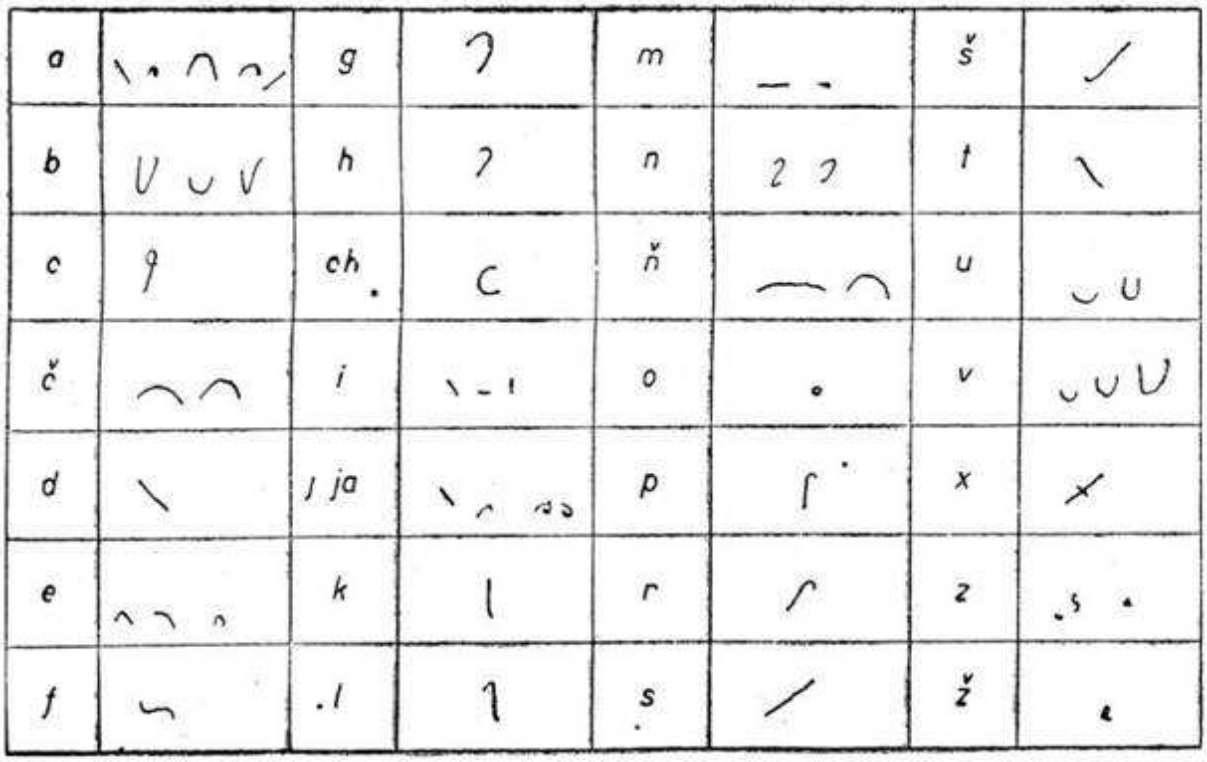
Obr. 1: Ukážka z rukopisu románu *Bohumil Valizlosť Zábor*

3. Princíp Kukučínovho písma a jeho lúštenie

Je veľmi pravdepodobné, že Martin Kukučín ako nesmierne vzdelaný človek ovládal stenografiu. Vzdelaním bol lekár, ale veľmi dobre sa orientoval aj v literatúre a umení a ovládal deväť cudzích jazykov. Jeho písmo je podobné stenografii, ale zároveň sa nejedná o stenografiu. Presnejšie povedané, je poskladané z viacerých stenografických systémov a v článku [1] autorka aj podrobne rozoberá z ktorých. V spomenutom článku autorka vymenúva znaky, ktoré Kukučín prevzal z Gabelsbergerovej, Pittmanovej a Greggovej stenografickej sústavy a niektoré ďalšie znaky, ktoré sú podobné znakom z iných stenografických sústav. Kukučín ale niektorým týmto znakov prisudzuje rozdielny význam než majú v spomenutých sústavách a podaktoré znaky sú jeho vlastným výtvorom.

Kukučínov spôsob písania sa ale od stenografie zásadne odlišuje tým, že Kukučín slová neskracoval, nepoužíval skratky za ustálené frázy a dokonca nepoužíval ani bežne používané a ustálené skratky (napr. jednotiek) ako sú *km=kilometer*, *kg=kilogram* a pod. Všetko vypisoval do posledného písmenka.
Po Kukučínovej smrti sa dlho nevedelo ako čítať jeho zápisky. Podľa [1] spomenul Kukučín vo svojej korešpondencii len jediný raz svoje tajné písmo. Bolo to v liste Škultétymu z roku 1921 a aj tam sa hovorilo len o existencii tohto písma a nie o jeho princípe. Ani vo svojich poznámkach Kukučín nezanechal popis svojho tajného písma.
Lúšteniu Kukučínovho písma sa dlhodobo a intenzívne venoval pán Ladislav Lorenc, ktorý

tomu s prestávkami venoval zhruba 12 rokov svojho života. Kukučínovým zápiskom sa začal venovať v roku 1943. Po štyroch rokoch neúspešných pokusov o rozlúštenie sa chcel vzdať, ale napokon s prestávkami predsa len pokračoval v práci [3]. Zlom údajne prišiel potom, keď v zápiskoch našiel citát opísaný z pomníka cisára Ferdinanda I., ktorého otvorený text poznal. Celú Kukučínovú abecedu definitívne dešifroval až v roku 1954. Do svojej smrti v roku 1964 prepísal veľkú časť Kukučínových zápiskov. Boli medzi nimi aj dovtedy nezverejnené a neznáme literárne diela. Pán Lorenc v [2] popísal princíp Kukučínovej šifry a uviedol kompletnú tabuľku znakov (obr. 2) ako aj zopár ukážok slov. Pokiaľ je známe, žiadny iný (kompletný) popis Kukučínovho písma nebol publikovaný.



Obr. 2: Tabuľka znakov z článku [2]

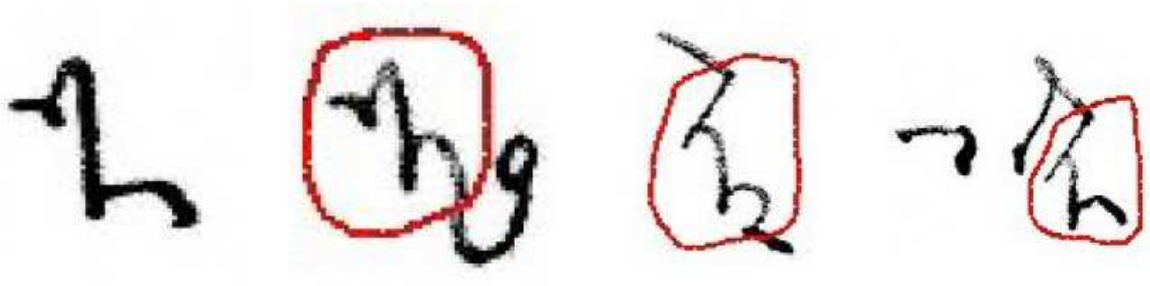
Žiaľ v článku [2] zo 16 uvedených ukážok slov je 11 chybných. Chyby sú však takého charakteru, že je zrejmé, že sa nejedná o chyby autora, ale chyby sadzača, ktorý zjavne nepoznal Kukučínove písmo a dodané obrázky navzájom povymieňal a rôznym spôsobom náhodne pootáčal.

Z kryptografického hľadiska je ale skôr zaujímavé to, že sa jednalo o homofónnu šifru. Homofónna šifra sa od jednoduchej zámeny (substitučnej šifry) líši tým, že niektoré znaky otvorenej abecedy sú nahradené nie jedným, ale hneď viacerými znakmi šifrovej abecedy. To, že sa v prípade Kukučínovho písma skutočne jednalo o homofónnu šifru, je na prvý pohľad zrejmé z tabuľky znakov uvedenej v článku [2], ktorá je aj na obr. 2. Zmysel homofónnej šifry spočíva v tom, že vyrovnáva percentuálny výskyt jednotlivých znakov v šifrovom texte, čím sa komplikuje štatistická analýza textu. Preto z kryptografického hľadiska má zmysel za homofóny (znaky, ktoré majú viac ekvivalentov v šifrovej abecede) voliť tie znaky, ktoré majú vysoký percentuálny výskyt v otvorenom texte. To sú predovšetkým samohlásky a niektoré častejšie sa vyskytujúce spoluhlásky. V Kukučínovom prípade je to splnené iba čiastočne. Napríklad často sa vyskytujúce spoluhlásky *s* a *t* majú v šifrovej abecede len jediný ekvivalent a naproti tomu niektoré menej často sa vyskytujúce znaky majú viacero ekvivalentov v šifrovej abecede. Už z toho je zjavné, že utajovanie zrejme nebolo

Kukučínovým primárnym cieľom. Z jeho rukopisov je vidno, že on používal rôzne verzie homofónnych znakov nie na vyrovnávanie frekvencie výskytu znakov, ale pre potreby väzby so susednými znakmi pri písaní textu. V rovnakom kontexte teda používal vždy rovnaké verzie znakov. Okrem toho homofónna šifra ako taká nie je, pri splnení istých podmienok, na lúštenie omnoho ťažšia než jednoduchá zámena (substitučná šifra). Potrebujeme iba dostatočné množstvo šifrového textu, musíme odhaliť homofóny pomocou digramových a trigramových väzieb (prípadne iným spôsobom) a potom je ďalší postup už triviálnou záležitosťou.
V prípade Kukučínovho písma bola situácia komplikovaná nie z dôvodu použitia homofónnej šifry, ktorej lúštenie bolo v tom čase už dobre známe, ale pre samotnú komplikovanosť zápisu textov. Kukučín totiž tvar a veľkosť jednotlivých znakov z tabuľky na obr. 2 menil veľmi rôzne a často sa stávalo, že niektoré znaky boli tak malé, že jednoducho „zanikli" medzi susednými znakmi. Preto bolo rozlúštenie Kukučínových rukopisov skutočne významným úspechom pána Ladislava Lorenca.

4. Ďalšie úlohy a ciele

Pán Lorenc vykonal veľmi veľký kus práce, rozlúštil Kukučínove písmo a prepísal väčšinu jeho zápiskov, avšak stále zostávajú neprepísané rukopisy. Okrem toho je ďalší veľký problém v tom, že v Národnej knižnici v Martine síce majú (zrejme) kompletné Kukučínove rukopisy, ako aj prepisy väčšiny z nich od pána Lorenca, avšak nie všetky rukopisy a ich prepisy sú jednoznačne spárované. Preto by bolo ideálne dostať všetky rukopisy do elektronickej podoby, ešte raz k nim spraviť prepisy a tieto dvojice spárovať. Samozrejme v súčasnosti túto prácu nikto nebude robiť ručne, ale bude ju treba zautomatizovať a spraviť to strojovo. Jedná sa totiž o tisíce strán rukopisov a už iba ich naskenovanie a prevedenie do elektronickej podoby bude nesmierne časovo náročná práca.
Preto bude potrebné vytvoriť aplikáciu, ktorá tieto rukopisy prevedie do otvorenej podoby, t.j. na čitateľný text. Úmyselne nehovoríme o OCR (Optical Character Recognition) aplikácii, pretože tento prístup v danom prípade nie je vhodný a navyše bežné OCR programy sú schopné pracovať len s tlačeným písmom a s rukopisom si buď nevedia poradiť vôbec, alebo majú pri ňom veľmi vysokú mieru chybovosti.

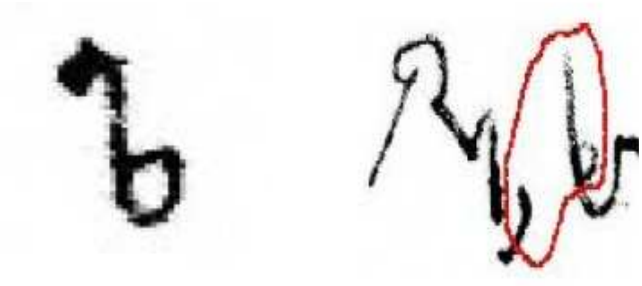


(a) **ale** (b) **alebo** (c) **ďalej** (d) **maštale**

Obr. 3: Rozdiely písania skupiny znakov **ale**

Ak si to zjednodušíme, tak môžeme povedať, že OCR programy pri prevádzaní obrázkov na text pracujú tak, že najskôr na obrázku identifikujú jednotlivé znaky (čo pri tlačenom písme a dobrej kvalite obrázku nie je až taký problém) a potom pomocou vhodnej vyhodnocovacej funkcie priradia jednotlivým znakom ich význam. Ak si ale pozrieme Kukučínove rukopisné texty, tak veľmi rýchlo zistíme, že rozbitie slov na jednotlivé znaky je prakticky nemožné. To bol aj najväčší problém pri lúštení tohto písma. Kukučín tak veľmi modifikoval tvar a

veľkosť jednotlivých znakov, že ak v texte nájdeme napr. oblúčik danej veľkosti a tvaru, tak len na základe tejto informácie nie je možné povedať, či sa jedná o písmeno *a*, *e*, *g*, *ň* alebo niektoré iné písmeno. Vždy to závisí od kontextu, v ktorom je tento oblúčik použitý. Okrem toho ako už bolo spomenuté skôr a ako píše aj pán Lorenc v článku [2], často sa stáva, že niektoré písmená sú tak nepatrne malé, že sa medzi inými znakmi v danom slove úplne stratia. Ďalším argumentom proti rozbíjaniu slov na znaky a následného rozpoznávania znakov je to, že Kukučínova stenografická abeceda je homofónna a v rôznych kontextoch píše ten istý znak rôznym spôsobom. Dva príklady takýchto skupín písmen sú uvedené na obr. 3 a obr. 4.
Preto vhodnejším spôsobom prevodu Kukučínovho rukopisu na otvorený text bude rozpoznávanie celých slov. Na to bude v prvom rade potrebné zostaviť si Kukučínove slovné bohatstvo, čo bude databáza niekoľko tisíc slov a nájsť k nim prislúchajúce vzory z Kukučínovho rukopisu. Potom v zoskenovaných rukopisných textoch bude nutné separovať jednotlivé slová v správnom poradí a tieto pomocou vhodnej vyhodnocovacej funkcie a zostavenej databázy slovného bohatstva identifikovať a prepísať do otvoreného textu.



(a) **ako** (b) **sedliakovi**

Obr. 4: Rozdiely písania skupiny znakov **ako**

Poďakovanie:

Tento článok vznikol s podporou grantu 1/0244/09.

Týmto by som sa chcel poďakovať Prof. RNDr. Otokarovi Grošekovi, CSc. za poskytnutie námetu ku zaujímavej práci, ako aj pracovníkom Národnej knižnice v Martine za poskytnutie potrebných materiálov.

Použité zdroje:

[1] Ihnátová, E.: Kukučínova stenografia. Literárny archív (1969), pp. 331-335.
[2] Lorenc, L.: Kukučínova ,,stenografia". Slovenská literatúra X, *4.* (1963), pp. 503-506.
[3] Sabová, A.: Ladislav Lorenc – slovenský Champollion, Prievidzský občasník č. 3, 2006
[4] http://sk.wikipedia.org/wiki/Martin_Kukučín
[5] http://www.osobnosti.sk/index.php?os=zivotopis&ID=59390

C. Chcete si zaluštit?

Martin Kolařík (marram.mail@gmail.com)

Geocaching *je turistická, navigační a trochu i internetová hra. Všechno najednou. Spočívá v tom, že někdo na neznámé místo ukryje schránku, které se v angličtině říká cache, čtěte a dnes už také většinou pište „keš“. Na internetu zveřejní její souřadnice a různé doplňující informace. Ostatní potom tuto schránku pomocí navigačních přístrojů hledají. Při nalezení se zapíšou do sešitku ve schránce, případně si vyberou něco z obsahu a výměnou do keše vloží něco svého. Po návratu svůj nález anebo i neúspěch svěří speciální internetové schránce.*

Úvodní citace je převzata z článku Lutonský, M.: Geocaching: hra pro mozek, nohy a vaši GPS, 12.8.2008 http://navigovat.mobilmania.cz/clanky/AR.asp?ARI=112930 .

V odkazovaném článku se dozvíte další užitečné informace a můžete shlédnout poutavě zpracované video o geocachingu.

Jedním z typů „keší“ jsou i tzv. Mystery cache. Jejich podstata spočívá v tom, že souřadnice se schránkou nejsou ihned k dispozici, ale až po zdolání nějakého úkolu. Forma úkolu je různá, vždy záleží na autorovi. Často se jedná o pouhé vyhledávání informací na internetu, ale některé využívají zajímavé, někdy dost netradiční, šifry. **Pokusím se nabídnout každý měsíc pár těchto šifrovacích mystery keší.** Luštit můžete samozřejmě i bez konečného odlovu, takže nezáleží, jak to máte k finální schránce daleko, nicméně radost z nálezu těžce vyluštěné keše umí potěšit.

Lednové GeoŠifry

1) Hvezdarenska http://coord.info/GC1J5H7
2) Cipher cache #2 http://coord.info/GC1ATJZ
3) Nostalgie - krizovatka (crossroads) http://coord.info/GC221DN

Užitečné je zjistit co je zašifrováno, někdy to jsou číslice, jindy slovní vyjádření čísel. Někdy přímo celý text, ve kterém jsou souřadnice zmíněny.

Přeji úspěšné luštění a šťastný lov.

Martin

D. Telefónica O2 poskytuje podklady pro stavební povolení elektronicky

Převzato - Tisková zpráva z 6.1.2010

Martin Žabka Tiskový mluvčí

S počátkem letošního roku spouští Telefónica O2 nový portál dokumentace sítě, který umožní všem zájemcům získat mapové podklady o poloze telekomunikačních sítí na jejich pozemcích elektronicky.

Tato aplikace je mimořádně užitečná pro všechny, kteří tyto podklady potřebují pro stavební nebo územní řízení. Díky elektronické komunikaci je budou schopni získat ve výrazně kratší době a bez nutnosti osobní návštěvy. Ročně přitom Telefónica O2 eviduje zhruba 150 000 takovýchto žádostí.

"Informace o našich sítích poptávají nejčastěji stavební firmy a experti, kteří plánují výstavbu nových objektů nebo přestavbu těch stávajících. Pro nás i všechny zájemce je pochopitelně nejvýhodnější, když dokážeme všechny úkony vyřešit elektronicky," říká Jiří Lagner, ředitel pro administraci a dokumentaci sítí ze společnosti Telefónica O2 Czech Republic. "Nová aplikace nejen výrazně zkrátí dobu potřebnou na vyřízení, ale také náklady spojené s jejím podáním."

Nyní probíhá celý proces elektronicky - od registrace zájemce a zadání požadavku prostřednictvím portálu, přes interní, centrální zpracování až po získání plnohodnotného vyjádření formou e-mailové zprávy a odkazu na elektronické úložiště, kde je příslušné vyjádření uloženo.

Automatické vyjádření s odpovědí obdrží žadatelé, na jejichž parcelu žádná síť nezasahuje. Po dokončení druhé etapy implementace umožní nová technologie automatické vydání vyjádření i zájemcům, kterých se současná infrastruktura přímo dotýká. Telefónica O2 bude touto cestou schopna obsloužit až 97,5 % všech podaných žádostí během zhruba deseti minut.

Postup při získávání informací o sítích je přitom velmi jednoduchý a intuitivní. Stačí, aby zákazník navštívil odkaz http://www.cz.o2.com/osobni/107010-dokumentace_site/ , kde najde přímý vstup do aplikace. V žádosti stačí vyplnit několik kontaktních údajů, vyznačit v mapě požadované území a po vygenerování dokumentů v elektronické podobě si je stáhnout. Na stránkách je rovněž videoprezentace, která zájemce celým procesem provede krok po kroku.

O kvalitě celého řešení svědčí i skutečnost, že na konci roku 2009 získal projekt prestižní mezinárodní ocenění Be Inspired Awards, které každoročně uděluje společnost Bentley Systems.

E. Science Café - Dobrodružství kryptologie
Pozvánka na únorovou akci Science Café Czech Republic
http://sciencecafe.cz/kalendar/

Co: Dobrodružství kryptologie

Kdy: úterý 9.února 2010, 19.00 hod.

Hosté: ***Vlastimil Klíma a Pavel Vondruška***

Kde: kavárna Potrvá (Srbská 2, Praha 6)

Poznámka: celý večer bude nahráván ČR Leonardo a vysílán v rámci víkendové univerzity.

Tajná písma, šifry a jejich luštění jsou inspirací pro řadu populárních knih a filmů, které baví čtenáře a diváky po celém světě. Vytváření šifer, respektive jejich luštění může být ale pro někoho jeho každodenní prací. O kryptologii jako nauce o metodách utajování informací a o tom, jaké to je být kryptologem jsme hovořili s předními českými kryptology-profesionály Vlastimilem Klímou a Pavlem Vondruškou.

Doprovodný rozhovor: Šifry mistrů kryptologů http://sciencecafe.cz/rozhovory/

Chcete-li se o kryptologii dozvědět více, přijďte na Science Café v úterý 9. února 2010 od 19 hodin do kavárny Potrvá (Srbská 2, Praha 6) a položte Vlastimilu Klímovi a Pavlovi Vondruškovi svou otázku.

F. O čem jsme psali v lednu 2000 – 2009

Crypto-World 1/2000

Crypto-World 1/2001

Příloha:
trustcert.pdf (upoutávka na služby Certifikační Autority TrustCert)

Crypto-World 1/2002

Crypto-World 1/2003

Příloha : Crypto_p1.pdf
CEN Workshop Agreements (dokumenty vztahující se k elektronickému podpisu)

Crypto-World 1/2004

Crypto-World 1/2005

Příloha :
Speciál 2004 - přehled článků a prezentací členů redakce Crypto-World za rok 2004 (http://crypto-world.info/casop6/prehled_2004.pdf)

Crypto-World 1/2006

Crypto-World 1/2007

Crypto-World 1/2008

Crypto-World 1/2009

G. Závěrečné informace

1. Sešit

Crypto-World je oficiální informační sešit "Kryptologické sekce Jednoty českých matematiků a fyziků" (GCUCMP). Obsahuje články podepsané autory. Případné chyby a nepřesnosti jsou dílem P.Vondrušky a autorů jednotlivých podepsaných článků, GCUCMP za ně nemá odbornou ani jinou zodpovědnost.
Adresa URL, na níž můžete najít tento sešit (zpravidla 3 týdny po jeho rozeslání) a předchozí sešity GCUCMP, denně aktualizované novinky z kryptologie a informační bezpečnosti, normy, standardy, stránky některých členů a další související materiály:
http://crypto-world.info

2. Registrace / zrušení registrace

Zájemci o e-zin se mohou zaregistrovat pomocí e-mailu na adrese pavel.vondruska@crypto-world.info (předmět: Crypto-World) nebo použít k odeslání žádosti o registraci elektronický formulář na http://crypto-world.info. Při registraci vyžadujeme pouze jméno a příjmení, titul, pracoviště (není podmínkou) a e-mail adresu určenou k zasílání kódů ke stažení sešitu.
Ke zrušení registrace stačí zaslat krátkou zprávu na e-mail pavel.vondruska@crypto-world.info (předmět: ruším odběr Crypto-Worldu!) nebo opět použít formulář na http://crypto-world.info. Ve zprávě prosím uveďte jméno a příjmení a e-mail adresu, na kterou byly kódy zasílány.

3. Redakce

E-zin Crypto-World

Redakční práce:	Pavel Vondruška
Stálí přispěvatelé:	Pavel Vondruška
	Jaroslav Pinkava
Jazyková úprava:	Jakub Vrána
Přehled autorů:	http://crypto-world.info/obsah/autori.pdf
NEWS	Vlastimil Klíma
(výběr příspěvků,	Jaroslav Pinkava
komentáře a	Tomáš Rosa
vkládání na web)	Pavel Vondruška
Webmaster	Pavel Vondruška, jr.

4. Spojení (abecedně)

redakce e-zinu	ezin@crypto-world.info ,	http://crypto-world.info
Vlastimil Klíma	v.klima@volny.cz ,	http://cryptography.hyperlink.cz/
Jaroslav Pinkava	Jaroslav.Pinkava@zoner.cz ,	http://crypto-world.info/pinkava/
Tomáš Rosa	t_rosa@volny.cz ,	http://crypto.hyperlink.cz/
Pavel Vondruška	pavel.vondruska@crypto-world.info ,	http://crypto-world.info/vondruska/index.php
Pavel Vondruška,jr.	pavel@crypto-world.info ,	http://webdesign.crypto-world.info
Jakub Vrána	jakub@vrana.cz ,	http://www.vrana.cz/