

UOT 004.056

V.Ə. QASIMOV, E.Ə. MUSTAFAYEVA

## WORD SƏNƏDLƏRİNDƏ İNFORMASIYANIN GİZLƏDİLMƏSİNİN SİMVOLLARARASI İNTEVALLAR ÜSULU

*Məqalədə mətn tipli fayllarda informasiyanın steqanoqrafik gizlədilməsi üsullarına baxılmış, Word sənədlərində simvollar arası intervalların dəyişdirilməsi yolu ilə informasiyanın gizlədilməsi üsulu işlənilib hazırlanmışdır. Etibarlılığının yüksəldilməsi məqsədilə MS Word sisteminin sənədlərin şifrələnməsi və qorunması funksiyalarının imkanlarından istifadə təklif olunmuşdur.*

**Açar sözlər:** kompüter steqanoqrafiyası, mətn steqanoqrafiyası, linqvistik steqanoqrafiya, informasiyanın gizlədilməsi, simvollar arası intervallar üsulu

**1. Giriş.** İnformasiyanın ötürülməsi faktının gizlədilməsi ilə məşğul olan steqanoqrafiya elminin müasir inkişaf mərhələsi kimi kompüter və rəqəmli steqanoqrafiya bir sıra üsul və texnologiyalara əsaslanır. Bu üsul və texnologiyalar, bir qayda olaraq, informasiyanın gizlədilməsi üçün istifadə olunan konteynerlərin xüsusiyyətlərindən asılı olur. Belə ki, konteynerlərin strukturunda kiçik və ya böyük həcmli əlavə informasiyanın yerləşdirilməsinə imkan verən ehtiyat sahələr (boşluqlar və izafiliklər) mövcud olmalıdır. Praktikada informasiya daşıyıcılarında, fayllarda və fayl strukturlarında, şəbəkə protokollarında, rəqəmsal obyektlərdə (rəqəmsal təsvirlərdə, qrafiklərdə, audio-video fayllarda) informasiyanın gizlədilməsi üçün müəyyən üsullar reallaşdırılır [1, s.260-269; 2].

Kompüter steqanoqrafiyasında informasiyanın gizlədilməsi üçün mətn tipli konteynerlərdən də geniş istifadə olunur, başqa sözlə, mətnlərdə informasiyanın gizlədilməsi üsullarında geniş yayılmışdır [2; 3, s.231-245; 4-7].

Steqanoqrafiyanın informasiyanın mətnlərdə gizlədilməsi üsulları ilə məşğul olan istiqaməti mətn steqanoqrafiyası və ya linqvistik steqanoqrafiya adlanır. Mətnlərdə informasiyanın gizlədilməsi üsulları mütəxəssislər tərəfindən üç əsas qrupa bölünür: sintaksis üsul, semantik üsul və intervallar üsulu.

Linqvistik steqanoqrafiyanın sintaksis üsullarına mətnlərdə istifadə olunan durğu işarələrinin, mətnlərin strukturunun və stilinin dəyişdirilməsi üsulları aid edilir. Aydın ki, mətnlərdə durğu işarələri müxtəlif qrammatik qaydalardan istifadə edilməklə tətbiq olunur. Bu üsul vasitəsilə durğu işarələrinin dəyişdirilməsi nəticəsində mətnə məna və məzmun baxımından dəyişiklik yaratmır, yəni mətn ekvivalent şəkildə qalır. Belə dəyişikliklər kənar şəxslərin diqqətini çəkmir və şübhə yaratmır. Ona görə də, durğu işarələrinin və sintaksis qaydalarının müxtəlif formalarda istifadəsi yolu ilə verilmiş informasiyanı mətnə gizlətmək olar. Bu məqsədlə, həmçinin, durğu işarələri ilə yanaşı bağlayıcılardan, qısaltmalardan, abreviaturadan, ədədi ifadələrin rəqəmlər və ya sözlə yazılışından, sözlərin, cümlələrin, abzasların birinci hərflərindən, müxtəlif dillərdə olan eyni şəkili hərflərdən, cümlələrin sintaksis stillərindən və s. istifadə oluna bilər [4,8].

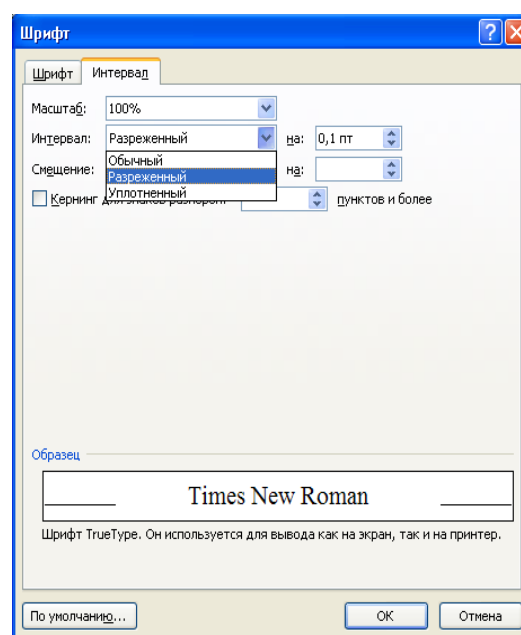
Semantik üsullar mətnə sözlərin və ya onların sinonimlərinin istifadəsi yolu ilə informasiyanın gizlədilməsinə əsaslanır. İnformasiyanın mətnə gizlədilməsi zamanı istifadə olunan bütün mümkün sözlər və onların sinonimlərini özündə əks etdirən cədvəl əvvəlcədən qurulur. Sözlərin nə qədər çox sinonimi olarsa, bu zaman daha çox informasiya kodlaşdırmaq və mətnə gizlətmək mümkün olur [4, 8].

İntervallar üsuluna mətnin formatlaşdırılması üsulları aid edilir. Burada mətn fayllarında informasiyanın gizlədilməsi üçün adi yazı üslubunun ehtiyat sahələrindən, mətn formatlarından, hərflərin registrindən, sətirlər, sözlər və abzaslar arasındakı intervallardan, kerninqdən, tabulyasiya, sətir sonu və abzas sonu işarələrdən və s. istifadə olunur [1, s.260-269; 2; 3, s.231-245; 4].

**2. Məsələnin qoyuluşu.** Qeyd olunan üsulların əksəriyyəti konteynerlərdə insan gözü ilə görünəcək dərəcədə dəyişikliklərin yaradılması, ilk baxışda aşkar olunması, konteyner üzərində həyata keçirilən müəyyən əməliyyatlara həssas olması, kiçik informasiya yerləşdirmək üçün böyük həcmli konteynerin tələb olunması və s. kimi çatışmazlıqlara malikdir. Bu çatışmazlıqların aradan qaldırılması üçün məqalədə Microsoft Word sənədlərində ayrı-ayrı sözlərin simvolları arasındakı intervalların dəyişdirilməsinə əsaslanan üsul təklif olunur. İnformasiya gizlədilmiş konteynerin etibarlılığını təmin etmək üçün Word redaktorunun sənədlərin şifrələnməsi və qorunması funksiyalarının imkanlarından istifadə edilir. Beləliklə, bu üsul vasitəsilə informasiya yeridilmiş konteyneri İnternet və digər informasiya daşıyıcıları vasitəsilə daha etibarlı şəkildə ünvana ötürmək mümkündür.

### 3. Sənəddə informasiyanın gizlədilməsi üçün simvollar arası intervallardan istifadə.

Məlumdur ki, bu gün praktikada istifadə olunan mətn tipli fayllar arasında MS Word redaktorunda hazırlanmış sənədlər əksəriyyət təşkil edir. Məhz belə sənədlərdə informasiyanın gizlədilməsi üçün daha çox imkanlar mövcuddur. Məhz belə imkanlardan biri də mətndə sözlərin simvolları arasındakı intervallardan istifadə olunmasıdır. Bunun üçün verilmiş gizli informasiya əvvəlcə ikilik koda (məsələn, ASCII standartına uyğun olaraq) çevrilir. Bundan sonra, MS Word redaktorunun **Шрифт** (Şrift) pəncərəsinin **Интервал** (Interval) hissəsində olan **Интервал** (Interval) açılan siyahısının **Разреженный** (Seyrək) və **Уплотненный** (Sıx) parametrlərinin qiymətlərini dəyişməklə həmin informasiya sənədlərin mətnlərində gizlədilir (şək.1).



Şək.1. Şrift pəncərəsi

Qeyd edilməlidir ki, Word redaktorunda hazırlanmış sənəddə sözlərin simvolları arasındakı intervallar üçün **Разреженный** (Seyrək) və **Уплотненный** (Sıx) parametrlərinin qiyməti 0–0,3 pt intervalında təyin edildikdə mətndə baş verən dəyişikliklər insan gözü ilə görünəcək həddi aşmır. Ona görə də simvollar arasındakı intervalı 0,3-ə qədər dəyişməklə ikilik formatda verilmiş informasiyanı kodlaşdırmaq və sənədə yerləşdirmək mümkündür.

Məlum olduğu kimi, Word redaktorunda bu parametrlər üçün başlanğıc qiymət 0, artırıb-azaltma addımı isə 0,05-dir. Beləliklə, simvollar arası intervalı 0-dan 0,25-dək 0,05 addımı ilə dəyişdikdə 10 kodlaşdırma variantı yaranır. Bu isə sənəddə daha çox informasiya gizlətmək imkanı yaradır (cədv.1).

Cədvəldən görüldüyü kimi, simvollar arası intervalın ilk səkkiz qiyməti (hər iki parametr üçün 0,05-dən 0,2-dək dörd qiymət) ilə “000”-dan “111”-dək ikilik bitlərdən ibarət ardıcılıqların kodlaşdırılması üçün istifadə olunur. Son iki qiymət (hər iki parametr üçün 0,25 qiyməti) isə “0” və “1” simvollarının kodlaşdırılması üçün nəzərdə tutulur. Belə ki, bəzi hallarda ikilik ardıcılıq hər biri 3 bit olmaqla qruplara bölündükdən sonra yekunda bir və ya iki bit qalır.

**Cədvəl 1**

İkilik bitlər ardıcılıqları və onların kodlaşdırılması üçün təklif olunan simvollar arası intervallar

Kodlaşdırılan bitlər	Simvollar arası intervallar	
	Seyrək (Разреженный)	Sıx (Уплотненный)
000	0,05	-
001	-	0,05
010	0,1	-
011	-	0,1
100	0,15	-
101	-	0,15
110	0,2	-
111	-	0,2
0	0,25	-
1	-	0,25

Məsələn, “LG” sözünü sənəddə gizlədilməsi üçün əvvəlcə bu söz ikilik koda çevrilir: “01001100010001111” (2 bayt). Alınmış ikilik ardıcılıq üç-üç qruplara ayrılır və nəticədə altı qrup alınır: “010”, “011”, “000”, “100”, “011”, “1”. Bu bitlərin kodlaşdırılması üçün sənəddə ilk altı sözün simvolları arasındakı intervallar 1 sayılı cədvələ əsasən aşağıdakı kimi təyin edilməlidir:

- birinci sözdə “010” bitləri gizlədilir – seyrək 0,1 pt;
- ikinci sözdə “011” bitləri gizlədilir – sıx 0,1 pt;
- üçüncü sözdə “000” bitləri gizlədilir – seyrək 0,05 pt;
- dördüncü sözdə “100” bitləri gizlədilir – seyrək 0,15 pt;
- beşinci sözdə “011” bitləri gizlədilir – sıx 0,1 pt.
- altıncı sözdə “1” biti gizlədilir – sıx 0,25pt.

Qeyd edilməlidir ki, ikilik kodda verilmiş informasiyanın sənəddə gizlədilməsi üçün tələb olunan konteynerin həcmi mətn steqanoqrafiyasının digər mövcud üsullarında istifadə edilən konteynerlərin əhəmiyyətli dərəcədə kiçik olur. Belə ki, sözlər arası probelin istifadəsi alqoritmində hər sözdə bir bit, sətir sonu probellər üsulunda bir sətirdə 1-3 bit gizlədildiyi halda, təklif olunan üsulda hər sözdə 3 bit gizlətmək mümkündür. Başqa sözlə, 300 simvoldan ibarət informasiyanı gizlətmək üçün birinci üsulda 2400 sözdən, ikinci üsulda təxminən 8000 sözdən (~800 sətir), təklif olunan üsulda isə 800 sözdən ibarət mətn tələb olunur.

Təklif olunan üsulun iş prinsipini daha əyani nümayiş etdirmək üçün aşağıdakı nümunəyə baxaq. Tutaq ki, MS Word sənədində “*Seminar sabah keçiriləcək*” informasiyasının gizlədilməsi tələb olunur. Bunun üçün informasiya əvvəlcə ASCII standartına uyğun olaraq ikilik koda çevrilir:

```
11010001 11100101 11101100 11101000 11101101 1110000011110000
1111000111100000 11100001 11100000 1111100111101010 11100101
11110111 1110100011110000 1110100011101011 11111111 11100110
11111111 11101010
```

Göründüyü kimi, 23 hərfdən ibarət olan informasiya ikilik koda çevrildikdən sonra 184 ikilik simvoldan ibarət ardıcılıq alınır. Alınmış ardıcılıq hər biri 3 bit olmaqla qruplara bölünür və cəmi 62 qrup alınır. Hər bir bitlər qrupu üçün cədvəl 1-dən uyğun qiymətlər seçilir (cədv.2). Ən azı 62 sözdən ibarət sənəd-konteyner seçilir. Verilən informasiyanın gizlədilməsi üçün sənəd-konteynerin sözlərində simvollar arasındakı intervallar müvafiq bit qruplarının qiymətlərinə uyğun təyin edilir.

Burada nümunə qismində seçilmiş sənəd-konteynerin informasiya gizlədilməzdən əvvəlki və gizlədildikdən sonrakı görüntüləri uyğun olaraq 2 və 3 sayılı şəkillərdə verilmişdir.

**Cədvəl 2**

Mətnə yeridilən bitlər və onlara uyğun simvollararası intervallar

№	Kodlaşdırılan bitlər	Simvollararası interval	№	Kodlaşdırılan bitlər	Simvollararası interval	№	Kodlaşdırılan bitlər	Simvollararası interval
1	110	Seyrək -0,2	22	111	Sıx -0,2	43	001	Sıx -0,05
2	100	Seyrək -0,15	23	100	Seyrək -0,15	44	111	Sıx -0,2
3	011	Sıx -0,1	24	000	Seyrək -0,05	45	000	Seyrək -0,05
4	110	Seyrək -0,2	25	111	Sıx -0,2	46	011	Sıx -0,1
5	010	Seyrək -0,1	26	000	Seyrək -0,05	47	101	Sıx -0,15
6	111	Sıx -0,2	27	011	Sıx -0,1	48	000	Seyrək -0,05
7	101	Sıx -0,15	28	110	Seyrək -0,2	49	111	Sıx -0,2
8	100	Seyrək -0,15	29	000	Seyrək -0,05	50	010	Seyrək -0,1
9	111	Sıx -0,2	30	011	Sıx -0,1	51	111	Sıx -0,2
10	010	Seyrək -0,1	31	111	Sıx -0,2	52	111	Sıx -0,2
11	001	Sıx -0,05	32	001	Sıx -0,05	53	111	Sıx -0,2
12	110	Seyrək -0,2	33	111	Sıx -0,2	54	111	Sıx -0,2
13	110	Seyrək -0,2	34	010	Seyrək -0,1	55	100	Seyrək -0,15
14	111	Sıx -0,2	35	101	Sıx -0,15	56	110	Seyrək -0,2
15	100	Seyrək -0,15	36	110	Seyrək -0,2	57	111	Sıx -0,2
16	000	Seyrək -0,05	37	010	Seyrək -0,1	58	111	Sıx -0,2
17	111	Sıx -0,2	38	111	Sıx -0,2	59	111	Sıx -0,2
18	100	Seyrək -0,15	39	110	Seyrək -0,2	60	110	Seyrək -0,2
19	001	Sıx -0,05	40	111	Sıx -0,2	61	101	Sıx -0,15
20	111	Sıx -0,2	41	111	Sıx -0,2	62	0	Seyrək -0,25
21	000	Seyrək -0,05	42	010	Seyrək -0,1			

Fiziki proseslərdə bir neçə enerji forması mövcuddur. Onları bir neçə qrup daxilində birləşdirmək mümkündür. Enerji forması onun daxil olduğu qrupdan asılı olmayaraq sistemin halını göstərən xarakterik kəmiyyətdir. Burada enerjinin saxlanması qanunu hökm sürür, yəni qapalı sistemin ümumi enerjisi həmişə sabitdir. Yalnız sistemə kənardan təsir etdikdə (əlavə enerji verdikdə) onun ümumi enerjisi dəyişir. Mexaniki sistemin enerjisini kibernetik və potensial enerjinin cəmi kimi təsvir etmək olar.

**Şək.2.** Orijinal konteyner

Fiziki proseslərdə bir neçə enerji forması mövcuddur. Onları bir neçə qrup daxilində birləşdirmək mümkündür. Enerji forması onun daxil olduğu qrupdan asılı olmayaraq sistemin halını göstərən xarakterik kəmiyyətdir. Burada enerjinin saxlanması qanunu hökm sürür, yəni qapalı sistemin ümumi enerjisi həmişə sabitdir. Yalnız sistemə kənardan təsir etdikdə (əlavə enerji verdikdə) onun ümumi enerjisi dəyişir. Mexaniki sistemin enerjisini kibernetik və potensial enerjinin cəmi kimi təsvir etmək olar.

**Şək.3.** İnformasiya daxil edilmiş konteyner

Bu üsulun çatışmazlığı qismində sənəd-konteyner üzərində həyata keçirilən formatlaşdırma (stil) dəyişikliklərinə həssas olması və Word redaktorunda şrift parametrlərinə baxılan zaman simvollararası intervalların fərqli olmasının görünməsinə göstərmək olar. Bu çatışmazlığın aradan qaldırılması üçün sənədin açılmasına, redaktə olunmasına, stil və şrifflərlə bağlı informasiyanın baxılmasına Word redaktorunun imkanlarından istifadə etməklə məhdudiyətlər qoymaq mümkündür.

Belə ki, bu məqsədlə MS Word redaktorunda **Рецензирование** tab-vərəqində olan **Защитить документ** (Sənədi qorumaq) əmrindən istifadə etmək olar. **Защитить документ** (Sənədi qorumaq) əmrini seçdikdə **Ограничить форматирование и редактирование** (Formatlaşdırma və redaktəyə məhdudiyətlərin qoyulması) dialoq pəncərəsi açılır (şək. 4). Burada **Ограничение на редактирование** (Redaktəyə məhdudiyətin qoyulması) hissəsində  işarəsini qoymaq, açılan siyahıdan **Только чтение** rejimini seçmək, **Включить защиту** (Qorumanı qoşmaq) bölməsində **Да, включить защиту** (Bəli, qorumanı qoşmalı) əmrini yerinə yetirmək və açılan dialoq pəncərəsində qoruma parolunu daxil edərək **OK** düyməsini basmaq lazımdır (şək.5).

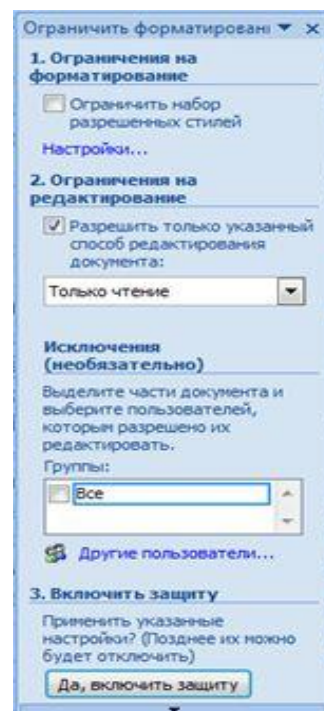
Parol vasitəsilə sənədin qorunması sənəd üzərində formatlaşdırma və sitilin dəyişdirilməsi kimi əməliyyatlara məhdudiyətlər qoymaqla yanaşı sənəddə stil və formatlaşdırma ilə bağlı parametrlərin qiymətlərini də görünməz edir.

Bununla yanaşı, Word redaktorunda yekun sənədin açılmasına paraol qoymaq (**Сохранить как** → **Сервис** → **Общие параметры**) və архиватор (məsələn, WinZip) vasitəsilə parol qoymaqla sənədi sıxmaq olar. Belə ki, sıxılma zamanı sənədin istifadə olunan parametrləri dəyişikliyə uğramır.

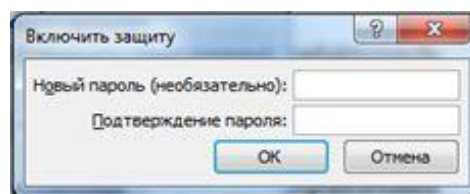
Beləliklə, əgər parollar tərəflər arasında əvvəlcədən razılaşıdırlarsa və ya digər etibarlı kanalla ötürülsə, onda konteynerdə saxlanılan informasiyanın daha etibarlı qorunması təmin edilmiş olar.

**4. Nəticə.** Təklif olunan üsul Word sənədindən ibarət konteynerdə digər mətn steqanoqrafiyası üsullarına nisbətən daha böyük həcmdə (mətnin hər sözündə 3 bit) informasiyanı gizlətməyə imkan verir. Burada mətnin strukturundan (adi mətn, şərh, siyahı, cədvəl və s.) asılı olmayaraq istənilən sənəd-konteynerdən istifadə oluna bilər. İnformasiyanın gizlədilməsi üçün sənəddə aparılan dəyişikliklər (sözlərin simvolları arasındakı intervalların artırılıb-azaldılması) insan gözü ilə görünəcək həddən aşağı olur ki, bu da konteynerdə informasiyanın gizlədilməsi faktının aşkarlanmasını çətinləşdirir.

Eyni zamanda, informasiya gizlədilmiş sənəd üzərində aparılan format və stil dəyişiklikləri zamanı informasiya itkisinin və mətnin şriffləri ilə bağlı parametrlərin baxılmasının qarşısını almaq üçün MS Word redaktorunun sənədin şifrlənməsi və qorunması funksiyasının imkanlarından istifadə olunur.



Şək.4. Sənədin qorunması



Şək.5. Parolun qoyulması

### Ədəbiyyat

1. Qasimov V.Ə. İnformasiya təhlükəsizliyinin əsasları. Bakı. MTN-in nəşriyyat-poliqrafiya mərkəzi. 2009. 340 s.
2. Bender W., Gruhl D., Morimoto N., Lu A. "Techniques for data hiding", IBM Systems Journal, vol. 35, Issues 3&4, 1996, pp. 313-336.
3. Конахович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика// МК-Пресс, 2006. – 288 с.
4. Текин В. Текстовая стеганография // Мир ПК. - 2004. - №11. с. 63. <http://www.osp.ru/pcworld/2004/11/169154/>
5. Бондарчук С.С., Давыдова Е.М., Костюченко Е.Ю. Встраивание цифровых знаков для обеспечения защиты информации. // Доклады ТУСУРа, № 2(24), часть 3, декабрь 2011.
6. Урбанович Н., Пласковицкий В. Метод текстовой стеганографии на основе модификации цветковых координат символов. // 21я Международная Конференция по Компьютерной графике и зрительное восприятие. Москва. 25-30 сентября 2011.
7. Барильник С.С., Минин И.В., Минин О.В., Щетинин Ю.В. Текстовая стеганография в HTML: реализация скрытых каналов передачи данных. // Ползуновский альманах. № 3, 2007, с.28-29.
8. Алиев А.Т., Щербаков А.Н. Стеганографический метод синонимичных преобразований открытого текста с учетом контекста. // III Международная Научно-практическая конференция «Актуальные проблемы информационных технологий». Красноярск. -2009.

УДК 004.056

**В.А. Касумов, Э.А. Мустафаева**

#### **Метод межсимвольных интервалов сокрытия информации в Word документах**

*Рассмотрены методы стеганографического сокрытия информации в текстовых файлах, разработан метод сокрытия информации в Word документах путем изменения межсимвольных расстояний слов. Для повышения надежности предложено использовать функции шифрования и защиты документов MS Word.*

**Ключевые слова:** компьютерная стеганография, текстовая стеганография, лингвистическая стеганография, сокрытие информации, метод межсимвольных интервалов

**V.A. Gasimov, E.A. Mustafaeva**

#### **Method of information concealment by intercharacter spacing in Word documents**

*The authors consider steganographic methods of information concealment in text file, a method of information concealment in Word documents by changing intercharacter spacing is developed. The use of MS Word encryption and protection feature is propose to improve reliability.*

**Keywords:** computer steganography, textual steganography, linguistic steganography, information concealment, intercharacter spacing method

Milli Təhlükəsizlik Nazirliyinin H.Əliyev adına Akademiyası  
AMEA İdarəetmə Sistemləri İnstitutu

Təqdim olunub 24.06.14