

production of everyday products such as shaving cream and washing powder to complex technological processes. Many properties of foams are a direct consequence of their geometric structure.

Keywords: dry foam, liquid foam, dispersibility, surface tension, equation of state.

INFORMATION ABOUT THE AUTHOR

TUMANOV Grigory Alekseevich – Postgraduate Student, Tver State Technical University, 22, embankment of A. Nikitin, Tver, 170026, Russia. E-mail: grishatumanoff@yandex.ru

DOLUDA Valentin Yuryevich – Doctor of Science (Chemistry), Associate Professor, Department of Biotechnology, Chemistry and Standardization, Tver State Technical University, 22, embankment of A. Nikitin, Tver, 170026, Russia. E-mail: doludav@yandex.ru

MALKOV Alexander Anatolievich – Candidate of Technical Sciences, Associate Professor, Software Department, Tver State Technical University, 22, embankment of A. Nikitin, Tver, 170026, Russia. E-mail: kja227@list.ru

CITATION FOR AN ARTICLE

Tumanov G.A., Doluda V.Yu., Malkov A.A. Theoretical features of foam formation // Vestnik of Tver State Technical University. Series «Building. Electrical engineering and chemical technology». 2023. No. 4 (20), pp. 84–96.

УДК 004.42, 502.504

ТЕХНОЛОГИЯ ИЗВЛЕЧЕНИЯ МЕТЕОДАНЫХ С ИНТЕРНЕТ-САЙТА ДЛЯ ОЦЕНКИ ЭКОЛОГИЧЕСКОЙ БЕЗОПАСНОСТИ НАСЕЛЕННЫХ ПУНКТОВ

Е.Б. Сергеев, Э.С. Цховребов

*Всероссийский научно-исследовательский институт
по проблемам гражданской обороны
и чрезвычайных ситуаций МЧС России (г. Москва)*

© Сергеев Е.Б., Цховребов Э.С., 2024

Аннотация. В статье описана технология по извлечению из интернет-сайта метеоданных для большого числа населенных пунктов. Метод основан на анализе строк файлов, представляющих отдельные страницы сайта. Технология использует как стандартную методику парсинга, так и принципиально иной подход для извлечения данных, рассмотрению которого и посвящена данная статья. Планируемым результатом внедрения этой технологии, находящейся в стадии апробации, будет повышение уровня полноты и обоснованности экологической информации о состоянии окружающей среды в населенных пунктах и, как следствие, расширение возможности заблаговременного прогнозирования экологически неблагоприятных, природных и техногенных чрезвычайных ситуаций, планирование комплекса мер по защите населения, природной среды и объектов экономики от угроз возникновения чрезвычайных ситуаций и их негативных последствий.

Ключевые слова: прогнозирование, экологическая безопасность, информационная технология, метеорологические данные, населенные пункты.

DOI: 10.46573/2658-7459-2024-1-96-110

ВВЕДЕНИЕ

Для прогнозирования рисков чрезвычайных ситуаций природного характера (на уровне субъекта Российской Федерации и выше) необходимо обладать большим объемом подробных метеорологических данных для отдельных населенных пунктов (или метеостанций). Это необходимо для мониторинга и прогнозирования природных пожаров, наводнений, вызванных дождевыми паводками, и опасных метеорологических явлений. В интернете предлагается множество готовых программ [1], дающих метеорологическую информацию в отношении ограниченного перечня населенных пунктов. Получить данные для всей территории России или ее крупных образований (федеральных округов, субъектов) при помощи этих программ невозможно. Сайт <https://gp5.ru> позволяет получить метеоданные для всей Российской Федерации, включая новые территории. Проблема состоит в извлечении большого массива информации и записи его непосредственно в базу данных или в файл (например, в Excel-файл) для последующей обработки и анализа.

В рамках решения этой проблемы разработана технология извлечения комплекса необходимых для оперативного прогнозирования чрезвычайных ситуаций (ЧС) метеоданных с сайта <https://gp5.ru> для охвата большого числа населенных пунктов Российской Федерации. Указанная технология реализована в созданной автором для этой цели программе WebMeteo [2]. Настоящая работа в структурированном виде иллюстрирует результаты проведенных научных исследований авторов в части разработки нового методического подхода в информационных технологиях.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Материалами для проведения настоящего исследования послужили результаты в области информационных технологий поиска электронной информации, полученные отечественными [3, 5–9] и зарубежными авторами [4, 10–12, 17].

В статье использована как стандартная методика парсинга [13–16, 18, 19], так и принципиально новый подход для извлечения данных с сайта. Он заключается в последовательном (по блокам) извлечении каждого типа метеоданных. Предложенный подход апробируется в комплексных системах обеспечения экологической безопасности населенных пунктов [20, 21].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

На первом этапе исследования описана конфигурация искомого сайта, являющегося источником получения метеоданных для прогнозирования ЧС природного и техногенного характера, а также их последствий.

Сайт <https://gp5.ru> представляет собой иерархическую структуру. Загрузочная страница показана на рис. 1. Далее выбирается страна и переход на новую страницу. От нее следует переход на страницы отдельных регионов России. Затем осуществляется переход на страницы отдельных районов (городских округов), только затем переход на

страницы отдельных населенных пунктов, где расположены необходимые для прогнозирования подробные метеорологические данные.

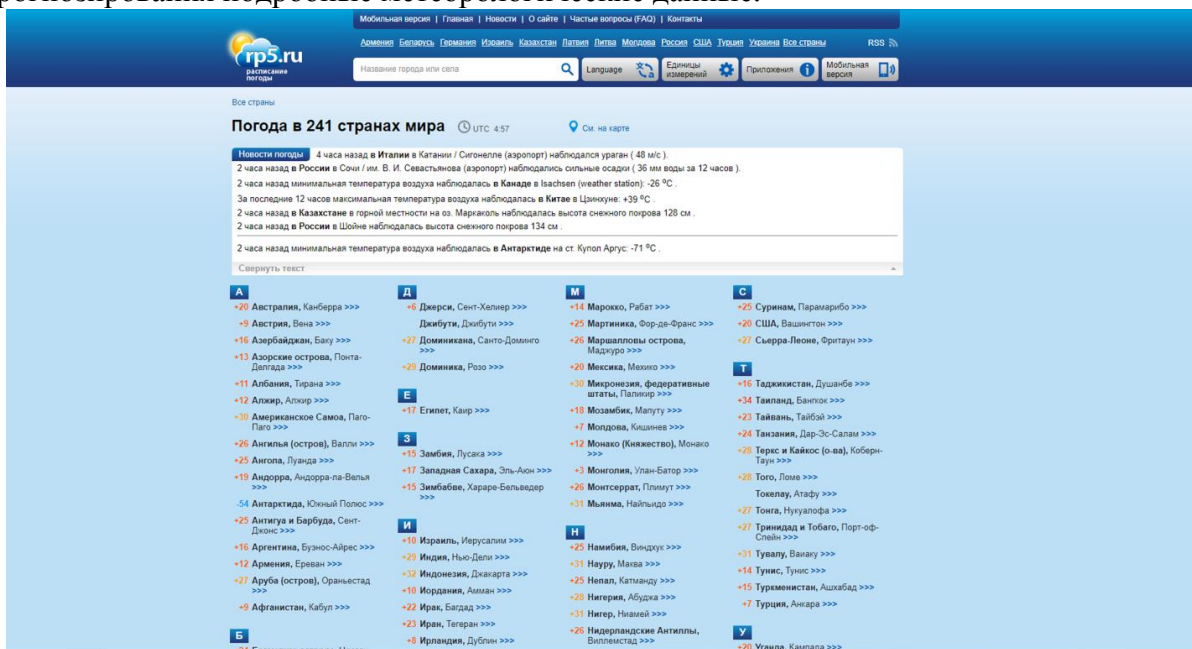


Рис. 1. Вид загрузочной страницы сайта <https://rp5.ru>

Таким образом, сайт <https://rp5.ru> представляет собой каталог, в котором содержатся страницы для стран, субъектов (провинций), районов и отдельных населенных пунктов. Общая формула обращения к этим страницам: <https://rp5.ru/+ссылка>. Во избежание значительных затрат времени в процессе оперативной работы органов МЧС России (и других заинтересованных ведомств) предлагается новый подход. Он состоит в следующем. Если известна ссылка на страницу какого-либо населенного пункта, то можно напрямую обратиться к данной странице (рассматриваемый сайт позволяет это делать) и в короткие сроки получить весь объем полноценной информации.

МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ

Первоначально программа **WebMeteo** обращалась к странице https://rp5.ru/Погода_в_России как к некоторому текстовому файлу и извлекала из нее данные о текущей температуре в главных городах субъектов и ссылки на страницы субъектов. Затем на основе этих ссылок обращались к файлам субъектов, где получали данные о текущей температуре районов и ссылки на страницы районов. Файлы страниц районов позволяли определять список отдельных населенных пунктов, текущую температуру и ссылки на страницы погоды в них. Наконец, на основе последних ссылок обращались именно к тем страницам, которые представляли для нас особый интерес.

На основе полученной информации в базе данных был сформирован справочник населенных пунктов Российской Федерации, каждая запись которого включала название населенного пункта, его код на сайте, географические координаты населенного пункта, код субъекта России, название района (городского округа) и ссылку на конкретный населенный пункт на сайте. В настоящее время на базе предлагаемого метода в рамках работы с новым справочником составляют список ссылок, после чего последовательно

просматривают файлы из каталога сайта, минуя ненужные и затратные по времени этапы просмотра файлов информации по стране в целом, субъектам и районам (рис. 2).

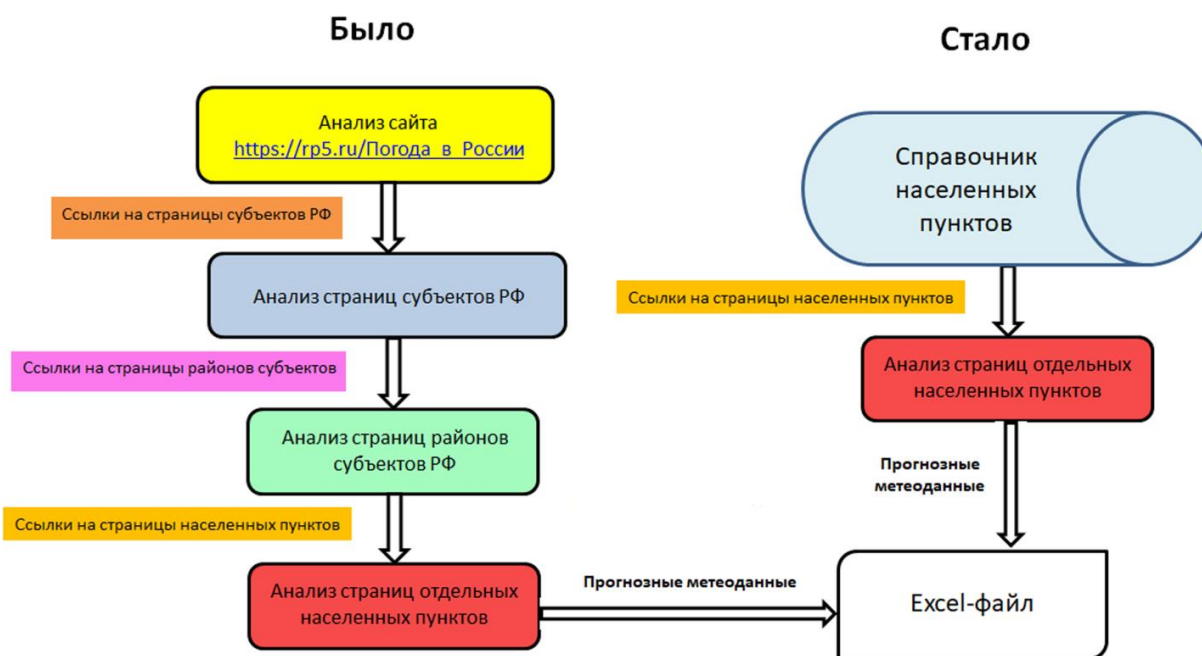


Рис. 2. Схемы извлечения метеоданных с сайта

В рамках предлагаемого автором нового подхода к получению требуемой для пользователя информации в короткие сроки была разработана инновационная технология извлечения специфического набора необходимых для оперативной работы данных с определенного сайта.

В программе **WebMeteo** использовалась единая технология загрузки файла из каталога сайта (код в Visual Basic):

```
Rep=getHTML(Url)
Private Function getHTML(ByVal Address As System.Uri) As String
'Загрузка файла сайта
Dim rt As String=""
Dim SR As StreamReader
Dim myHttpRequest As HttpRequest
Dim myHttpWebResponse As HttpWebResponse
myHttpRequest=HttpRequest.Create(requestUri:=Address)
myHttpRequest.Headers.Add("Accept-Language","ru")
myHttpWebResponse=myHttpRequest.GetResponse
SR=New StreamReader(myHttpWebResponse.GetResponseStream,True)
rt=SR.ReadToEnd
SR.Close()
getHTML=rt
End Function
```

Здесь Url – полный адрес страницы сайта (<https://rp5.ru/>+ссылка).

После загрузки файла страницы как текстового файла производились расщепление файла на отдельные строки и их анализ:

```

nLength=Rep.Length
If nLength>0 Then
'В качестве расщепления строк используется "<div>", а не Chr(10)
Dim masStr()=Rep.Split("<div>")
'Общее число строк
maxL=masStr.Length
For L=0 To (maxL - 1)
strF=masStr(L).Trim
'Вызов функции для анализа строки
    Ok = AnalisStr(strF,0)
.....
    Next L
.....
End If

```

Поиск данных о текущих температурах и ссылках из файлов, касающихся России в целом, субъектов и районов, осуществлялся с помощью методики парсинга [11, 12]. Не останавливаясь подробно на этой методике, сообщим необходимую информацию для осуществления поиска для каждого вида страницы. В табл. 1 и 2 приведены строковые переменные для поиска данных на страницах сайта, относящихся соответственно ко всей России и к ее субъектам. Для страниц, касающихся районов, значения строковых переменных аналогичны значениям таких переменных для субъектов. В табл. 3 приведены строковые переменные для поиска данных на страницах сайта о погоде в населенных пунктах.

Таблица 1

Строковые переменные для поиска на странице всей России

Переменная	Значение	Комментарий
aStr(0, 0)	"span class=" &Chr(34)&"Ajax-PointID"&Chr(34)	Начало цикла поиска температуры
aStr(1, 0)	"span class=" &Chr(34)&"t_0"&Chr(34)	Указание на использование шкалы Цельсия
aStr(2, 0)	"span class=" &Chr(34)&"red"&Chr(34)	Определение положительного значения температуры по Цельсию
aStr(3, 0)	"span class=" &Chr(34)&"blue"&Chr(34)	Определение отрицательного значения температуры по Цельсию
aStr(4, 0)	"style=" &Chr(34)&"color:#000"&Chr(34)&"href="	Определение ссылки на субъект
aStr(5, 0)	"title="	Определение названия субъекта
aStr(6, 0)	"class=" &Chr(34)&"windclose"&Chr(34)	Завершение процедуры поиска (выход)

Таблица 2

Строковые переменные для поиска на странице субъекта

Переменная	Значение	Комментарий
aStr(1, 1)	"span class="&Chr(34)&"Ajax-PointID"&Chr(34)	Начало цикла поиска температуры
aStr(2, 1)	"span class="&Chr(34)&"t_0"&Chr(34)	Указание на использование шкалы Цельсия
aStr(3, 1)	"span class="&Chr(34)&"red"&Chr(34)	Определение положительного значения температуры по Цельсию
aStr(4, 1)	"span class="& Chr(34)&"blue"&Chr(34)	Определение отрицательного значения температуры по Цельсию
aStr(5, 1)	"class="& Chr(34)&"href12"&Chr(34)&" href="	Определение ссылки на район
aStr(6, 1)	"title="	Определение названия района
aStr(7, 1)	"class="& Chr(34)&"windclose"&Chr(34)	Завершение процедуры поиска (выход)

Таблица 3

Строковые переменные для поиска на странице погоды в населенном пункте

Переменная	Значение	Комментарий
aStr(1, 3)	"onclick="& Chr(34)&"show_map"	Определение координат населенного пункта
aStr(2, 3)	"Время в данном населенном пункте. Учитывается летнее/ зимнее время"	Начало блока определения списка времен замера
aStr(3, 3)	"class="& Chr(34)&"d underlineRow"&Chr(34)	Определение дневного часа
aStr(4, 3)	"class="&Chr(34)&"d2 underlineRow"&Chr(34)	Определение дневного часа
aStr(5, 3)	"class="& Chr(34)&"n underlineRow"&Chr(34)	Определение ночного часа
aStr(6, 3)	"class="&Chr(34)&"n2 underlineRow"&Chr(34)	Определение ночного часа
aStr(7, 3)	"class="&Chr(34)&"t_cloud_cover"&Chr(34)	Начало блока определения облачности
aStr(8, 3)	"Осадки, мм"	Начало блока определения осадков
aStr(9, 3)	"Туман, %"	Начало блока определения тумана

Продолжение табл. 3

Переменная	Значение	Комментарий
aStr(10, 3)	"title="&Chr(34)&"Температура воздуха на высоте 1,5 метра (градусы Цельсия)"&Chr(34)&">Температура"	Начало блока определения температуры воздуха
aStr(11, 3)	"underlineRow toplineRow red"	Определение значения положительной температуры
aStr(12, 3)	"underlineRow toplineRow blue"	Определение значения отрицательной температуры
aStr(13, 3)	"div class="&Chr(34)&"t_0"&Chr(34)	Указание на использование шкалы Цельсия
aStr(14, 3)	""	Определение значения температуры в градусах Цельсия
aStr(15, 3)	"a class="&Chr(34)&"f_temperature"&Chr(34)	Начало блока определения температуры по ощущениям
aStr(16, 3)	"title="&Chr(34)&"Атмосферное давление у поверхности земли (миллиметры ртутного столба)"&Chr(34)	Начало блока определения давления воздуха
aStr(17, 3)	"div class="&Chr(34)&"p_0"&Chr(34)	Указание на использование шкалы давления в миллиметрах ртутного столба
aStr(18, 3)	"title="&Chr(34)&"Ветер на высоте 10 метров (метры в секунду)"&Chr(34)&">Ветер: скорость"	Начало блока определения скорости ветра
aStr(19, 3)	"div class="&Chr(34)&"wv_0 "&Chr(34)	Указание на использование шкалы скорости в метрах в секунду
aStr(20, 3)	"style="&Chr(34)&"padding-left: 5px;"&Chr(34)&">порывы,"	Начало блока определения порывов ветра
aStr(21, 3)	"Направление"	Начало блока определения направления ветра
aStr(22, 3)	"Ветер, дующий с "	Определение направления ветра
aStr(23, 3)	"Влажность"	Начало блока определения влажности
aStr(24, 3)	"class="&Chr(34)&"d underlineRow"&Chr(34)	Определение времени замера влажности (днем)

Окончание табл. 3

Переменная	Значение	Комментарий
aStr(25, 3)	"class="&Chr(34)&"d2 underlineRow"&Chr(34)	Определение времени замера влажности (днем)
aStr(26, 3)	"class="&Chr(34)&"n underlineRow"&Chr(34)	Определение времени замера влажности (ночью)
aStr(27, 3)	"class="&Chr(34)&"n2 underlineRow"&Chr(34)	Определение времени замера влажности (ночью)
aStr(28, 3)	"Солнце"	Завершение процедуры поиска (выход)

Перейдем к технологии поиска метеоданных для страниц поиска погоды в населенных пунктах. Как отмечено выше, расположение информации для этого типа страниц отличается от размещения информации на страницах других типов указанного сайта. Расположение данных в файле страницы блоковое для каждого вида данных, т.е. в каждом блоке находятся данные конкретного типа за весь период наблюдений. Таким образом, предлагаемая технология извлечения данных учитывает означенную выше специфику в целях осуществления поставленной задачи.

На первом этапе осуществляется поиск строки с содержанием aStr(1, 3) (см. табл. 3). В ней находятся географические координаты населенного пункта, которые вносятся в соответствующую переменную. Затем со строки, содержащей aStr(2, 3), начинается блоковое извлечение метеоданных. Обработывается в общей сложности одиннадцать блоков:

- 1) времен наблюдений;
- 2) облачности;
- 3) осадков;
- 4) тумана;
- 5) температуры воздуха;
- 6) ощущаемой температуры воздуха (с учетом влажности и ветра);
- 7) давления;
- 8) скорости ветра;
- 9) порывов ветра;
- 10) направления ветра;
- 11) влажности воздуха.

Факт обработки конкретного блока определяется соответствующей булевой переменной okData(i), где i изменяется от 0 до 10. Начало каждого из следующих блоков свидетельствует о завершении предыдущего блока. Вместо вызова функции для анализа строки выполняется блок обработки данных. Общий вид программного кода для каждого блока представлен ниже:

```
'Обработка блока времени наблюдения
If okData(0) Then
ok=DefTime(strF, 3)
If okData(0) Then
'переход на новую строку файла
GoTo Line20
```



```
Else  
'переход к обработке следующего блока  
GoTo Line...  
End If  
End If
```

Первый блок предоставляет список часов наблюдения метеоданных. Ниже показан программный код для извлечения этих часов (вместе с отдельной функцией поиска строки) и записи их в массив:

```
Private Function DefTime(ByVal Stroka As String, ByVal iL As Integer) As Boolean
```

```
'Анализ строки для определения часов наблюдения в населенных пунктах
```

```
DefTime=False
```

```
str1=""
```

```
npos=MyInStr(1, Stroka, aStr(7, iL))
```

```
If npos>0 Then
```

```
okData(0)=False
```

```
okData(1)=True
```

```
'Максимальное число наблюдений
```

```
iChasMax = iChas
```

```
GoTo Line1
```

```
End If
```

```
npos=MyInStr(1, Stroka, aStr(3, iL))
```

```
If npos>0 Then
```

```
GoTo Line2
```

```
End If
```

```
npos=MyInStr(1, Stroka, aStr(4, iL))
```

```
If npos>0 Then
```

```
GoTo Line2
```

```
End If
```

```
npos=MyInStr(1, Stroka, aStr(5, iL))
```

```
If npos>0 Then
```

```
GoTo Line2
```

```
End If
```

```
npos=InStr(1, Stroka, aStr(6, iL))
```

```
If npos>0 Then
```

```
GoTo Line2
```

```
End If
```

```
GoTo Line1
```

```
Line2:
```

```
npos1=InStr(1, Stroka, ">")
```

```
If npos1>0 Then
```

```
str1=Mid(Stroka, npos1+1).Trim
```

```
iChas=iChas+1
```

```
aTime(iChas)=str1
```

```
GoTo Line1
```

```
End If
```

```
Line1:
```

```
DefTime=True
```

```
End Function
```

```
.....
```

```
Public Function MyInStr(ByVal iBeg As Integer, ByVal Stroka As String, ByVal Stroka1 As
String) As Integer
'Поиск позиции вложения подстроки в строку
Dim iPos, iMax, iFirst, ni(10), ln2, ln2, npos, npos1 As Integer
Dim str1, nStr(10) As String
iPos=0
iMax=0
iFirst=0
If Stroka <> "" Then
Stroka=Stroka.Trim
ln2=Len(Stroka)
Else
GoTo Line4
End If
If Stroka1<>"" Then
Stroka1=Stroka1.Trim
ln1=Len(Stroka1)
Else
GoTo Line4
End If
If ln2<ln1 Then
GoTo Line4
End If
str1=Stroka1
Line1:
npos=InStr(1, str1, Chr(34))
If npos>0 Then
nStr(iMax)=Mid(str1, 1, npos - 1)
str1 = Mid(str1, npos + 1)
iMax = iMax+1
GoTo Line1
Else
If str1<>"" Then
nStr(iMax)=str1
iMax=iMax+1
End If
End If
If iMax>0 Then
str2=Stroka
For i=0 To (iMax-1)
npos1=0
If str2<>""Then
npos1=InStr(1, str2, nStr(i))
ln3=Len(nStr(i))
If i=0 Then
iFirst=npos1
End If
End If
If npos1>0 Then
str2=Mid(str2, npos1+ln3+1)
```

```
GoTo Line2
Else
GoTo Line4
End If
Line2:
Next i
Line3:
iPos=iFirst
Else
iPos=InStr(1, Stroka, Stroka1)
End If
Line4:
MyInStr=iPos
Exit Function
End Function
```

Массив `aTime(iChas)` сопоставлялся с датами наблюдений. Первая дата соответствует дате загрузки данных. В дальнейшем, если значение `aTime(iChas+1)` было меньше значения `aTime(iChas)`, дата для `aTime(iChas+1)` автоматически сдвигалась на один день вперед. Цикл повторялся всякий раз при `aTime(iChas+1) < aTime(iChas)`.

Следующий блок предоставлял данные об облачности в населенном пункте. Затем шла обработка блока осадков. Полученные данные об осадках записывались в соответствующий массив. Через номер `iChas` массив `aOsadki` ставился в соответствие массиву `aTime`, т.е. массив `aOsadki(0)` соответствовал времени `aTime(0)`, `aOsadki(1)` – времени `aTime(1)` и т.д. Программный код обработки этого блока выглядит следующим образом:

```
Private Function ObrabOsadki(ByVal Stroka As String, ByVal iL As Integer) As Boolean
'Анализ строки для обработки осадков в населенных пунктах
npos=MyInStr(1,Stroka,aStr(9,iL))
If npos>0 Then
okData(2)=False
okData(3)=True
iChas=0
Go To Line1
End If
npos=MyInStr(1,Stroka,aStr(10,iL))
If npos>0 Then
okData(2)=False
okData(4)=True
iChas=0
GoTo Line1
End If
strP="class=" & Chr(34) & "pr_0" & Chr(34) & " onmouseover"
npos=MyInStr(1,Stroka,strP)
If npos>0 Then
OsadkiTime = 0
npos1=InStr(1,Stroka,"Явления погоды отсутствуют")
If npos1>0 Then
aOsadki(iChas)=OsadkiTime
iChas=iChas+1
GoTo Line1
End If
```

```
npos2=InStr(1,Stroka,"Без осадков")
If npos2>0 Then
aOsadki(iChas)=OsadkiTime
iChas=iChas+1
GoTo Line1
End If
npos3=InStr(1, Stroka, "в среднем")
ln1=Len("в среднем")
If npos3>0 Then
npos4=InStr(npos3,Stroka,"мм воды за 1 час")
If npos4>0 Then
str1=Mid(Stroka,npos3+ln1,npos4-1-(npos3+ln1)).Trim
'Проверка строки на число
ok=NumData(str1)
If ok Then
'Замена точки на запятую в качестве разделителя дробной части
str1=Zamena_Point(str1)
OsadkiTime=CSng(str1)
aOsadki(iChas)=OsadkiTime*6
iChas=iChas+1
End If
End If
GoTo Line1
End If
End If
Line1:
ObrabOsadki=True
End Function
```

Полученный массив aPress(iChas) также ставился в соответствие массиву aTime(iChas). Все остальные блоки (тумана, температуры воздуха, ощущаемой температуры воздуха, давления, скорости ветра, порывов ветра, направлений ветра и влажности) имели программный код, подобный вышеприведенному. Поэтому в целях сокращения размеров статьи их тексты не приводятся. После обработки последнего блока (с данными по влажности воздуха), о чем извещает строка с содержанием aStr(28, 3), происходит завершение просмотра файла и переход к следующему файлу. Процесс повторяется до завершения просмотра всех необходимых файлов и получения полного объема требуемой информации.

Следует отметить, что программа **WebMeteo** загружает не более 30 % размещенной на сайте информации (иная информация не применяется в текущих прогнозах ЧС).

ЗАКЛЮЧЕНИЕ

Представленная в статье технология извлечения информации (в данном случае – метеорологических данных) принципиально отличается от применяемых в настоящее время технологий извлечения данных из Web-сайтов. Это обусловлено разработанным авторами сайта блоковым способом размещения данных. Предложенная инновационная технология основывается на последовательной обработке отдельных блоков сайта, соответствующих каждому виду данных.

При необходимости широкого использования специфической информации на конкретном сайте разработанная методика предоставляет возможность распространения предложенного авторами методического подхода на другие блоки данных на сайте. Рассмотренная технология может быть применена и на других сайтах, если структура размещения данных имеет схожий по соответствующим признакам блоковый характер.

Актуальность и практическая значимость предлагаемого метода и новой информационной технологии подтверждены их успешной апробацией и внедрением во Всероссийском научно-исследовательском институте по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России в процессе оперативного составления прогнозов неблагоприятных событий, опасных природных явлений и техногенных процессов по полученной в кратчайшие сроки полноценной адресной информации. В результате апробации нового подхода и информационной технологии при оперативном прогнозировании в системе с оптимизацией организационно-управленческих решений удалось добиться повышения показателя оправдываемости формируемых прогнозов возможных ЧС [20, 21]. Это позволило предупредить ряд негативных последствий опасных процессов и явлений в регионах РФ. Такой подход и алгоритм действий по его реализации предоставляет возможность заблаговременно предупредить или существенно смягчить риски, угрозы возникновения, а также опасные последствия чрезвычайных ситуаций, техногенных аварий для населения, территорий и окружающей среды.

СПИСОК ЛИТЕРАТУРЫ

1. Метео и связанные с ним приложения. URL: <https://www.softpedia.com/download/Tag/meteo> (дата обращения: 20.04.2023).
2. Патент РФ 2022613341. *Программа для ЭВМ WebMeteo* / Сергеев Е.Б.; Заявл. 02.03.2022. Оpubл. 14.03.2022.
3. Гуменюк Е. 10 лучших инструментов для сбора данных в Интернете // *Freelance.Today*. URL: <https://freelance.today/poleznoe/10-luchshih-instrumentov-dlya-sbora-dannyh-v-internete.html?ysclid=1h638g5cbp362477738> (дата обращения: 02.05.2023).
4. Да Х. Алгоритмы извлечения информации из текстов, парсинг веб-страниц с использованием языка программирования Python // *Актуальные исследования*. 2022. № 30 (109). С. 21–24.
5. Трушкова М. Как парсить сайт: 20+ инструментов на все случаи жизни. URL: <https://www.cossa.ru/imarketing/261951/?ysclid=1h644z5fvb507478044> (дата обращения: 02.05.2023).
6. Кармалев Д.А., Коршев Е.П., Сулейманова Е.А., Трофимов И.В. Технология извлечения текстов, основанная на знаниях // *Программные продукты и системы*. 2009. № 2. С. 62–66.
7. Брюхов Д.О., Скворцов Н.А. Извлечение информации из больших коллекций русскоязычных текстовых документов в среде Hadoop // *Электронные библиотеки: перспективные методы и технологии, электронные коллекции RCDL-2014: Труды 16-й Всероссийской научной конференции*. С. 391–398. URL: http://rcdl.ru/doc/2014/paper/RCDL2014_391-398.pdf (дата обращения: 26.04.2024).
8. Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. Автоматическая обработка текста на естественном языке и анализ данных. М.: НИУ ВШЭ, 2017. 269 с.
9. Федюшкин Н.А., Федосин С.А. Понятие, проблемы и разновидности интеллектуального анализа текста // *Проблемы и достижения в науке и технике: сборник научных трудов по итогам международной научно-практической конференции*. 2016. Вып. III. С. 39–41. URL: <https://izron.ru/articles/problemy-i-dostizheniya-v-nauke-i-tekhnikе-sbornik-nauchnykh-trudov->

po-itogam-mezhdunarodnoy-nauchno/sektsiya-2-informatika-vychislitel'naya-tehnika-i-upravlenie-spetsialnost-05-13-00/po-nyatiye-problemy-i-raznovidnosti-intellektual'nogo-analiza-teksta/ (дата обращения: 26.04.2024).

10. Gaikwad S.V., Chaugule A., Patil P. Text Mining Methods and Techniques // *International Journal of Computer Applications*. 2014. Vol. 85. № 17, pp. 42–45. DOI:10.5120/14937-3507

11. Использование HtmlAgilityPack и CSS Selectors. URL: <https://itvdm.com/ru/blog/article/using-html-agility-and-css-selectors> (дата обращения 25.04.2023).

12. Парсинг на C# с HtmlAgilityPack. URL: <https://vc.ru/dev/148017-parsing-na-c-s-htmlagilitypack> (дата обращения: 25.04.2020).

13. Парсинг сайтов, html-страниц и файлов. URL: https://vremya-ne-zhdet.ru/vba-excel/parsing-saytov/?ysclid=lh91dc86mv801207683#Parsing_html-stranic_msxml2xmlhttp (дата обращения: 02.05.2023).

14. Scraping a website HTML in VBA. URL: <https://www.wiseowl.co.uk/blog/s393/scrape-website-html> (дата обращения: 02.05.2023).

15. How to Import&Parse JSON Data with VBA. URL: <https://myexcelgenius.com/getting-data-from-a-website-in-json-format-using-vba> (дата обращения: 02.05.2023).

16. Парсинг текста с сайта. URL: https://www.cyberforum.ru/vb-net/thread_1063826.html?ysclid=lh923gqkp2314550261 (дата обращения: 02.05.2023).

17. Санчес Э. Скрапинг веб-сайтов с помощью Excel. URL: <https://baguzin.ru/wp/eduardo-sanches-skraping-veb-sajtov-s-pomoshhyu-excel> (дата обращения: 02.05.2023).

18. Импорт данных с web-страниц в Excel средствами VBA. URL: <https://ytikhonov.wordpress.com/2015/10/06/импорт-данных-с-web-страниц-в-excel-средствам> (дата обращения: 02.05.2023).

19. Анализ текстовых файлов с помощью объекта TextFieldParser (Visual Basic) // Learn. URL: <https://learn.microsoft.com/ru-ru/dotnet/visual-basic/developing-apps/programming/drives-directories-files/parsing-text-files-with-the-textfieldparser-object> (дата обращения: 03.05.2023).

20. Гаврилов Е.В., Исаков В.М., Цховребов Э.С. Проблемы обеспечения экологической безопасности на территории муниципального образования // *Экосинформ*. 2005. № 1. С. 17–21.

21. Исаков В.М., Цховребов Э.С. Правовые основы охраны окружающей среды. М.: МОФ МосУ МВД России, 2004. 100 с.

СВЕДЕНИЯ ОБ АВТОРАХ

СЕРГЕЕВ Евгений Борисович – научный сотрудник ФГБУ «Всероссийский научно-исследовательский институт по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России», 121352, Россия, г. Москва, ул. Давыдовская, д. 7. E-mail: alikh574@mail.ru

ЦХОВРЕБОВ Эдуард Станиславович – кандидат экономических наук, доцент, старший научный сотрудник ФГБУ «Всероссийский научно-исследовательский институт по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России», 121352, Россия, г. Москва, ул. Давыдовская, д. 7. E-mail: rebrovstanislav@rambler.ru

БИБЛИОГРАФИЧЕСКАЯ ССЫЛКА

Сергеев Е.Б., Цховребов Э.С. Технология извлечения метеоданных из интернет-сайта для оценки экологической безопасности населенных пунктов // Вестник Тверского государственного технического университета. Серия «Строительство. Электротехника и химические технологии». 2024. № 1 (21). С. 96–110.

TECHNOLOGY FOR EXTRACTING METEOROLOGICAL DATA FROM AN INTERNET SITE TO ASSESS THE ECOLOGICAL CONDITION OF SETTLEMENTS

E.B. Sergeev, E.S. Tshovrebov

*All-Russian Research Institute for Civil Defense and Emergency Situations
of the Ministry of Emergency Situations of Russia (Moscow)*

Abstract. The article describes a technology for extracting weather data from an Internet site for a large number of localities. The method is based on the analysis of strings of files representing individual pages of the site. The technology uses both a standard parsing technique and a fundamentally different approach to data extraction, which this article is devoted to. The planned result of the introduction of this technology, which is being tested, will be to increase the level of completeness and validity of environmental information on the state of the environment in human settlements and, as a result, expand the possibility of early forecasting of environmentally unfavorable, natural and man-made emergencies, planning a set of measures to protect the population, the natural environment and economic facilities from threats of emergency situations and their negative consequences.

Keywords: forecasting, environmental safety, information technology, meteorological data, settlements.

INFORMATION ABOUT THE AUTHORS

SERGEEV Evgeniy Borisovich – Researcher of All-Russian Research Institute for Civil Defense and Emergency Situations of the Ministry of Emergency Situations of Russia, 7, Davydkovskaya str., Moscow, 121352, Russia. E-mail: alik574@mail.ru

TSKHOVREBOV Edward Stanislavovich – Candidate of Economic Sciences in Economics, Associate Professor, Senior Researcher of All-Russian Research Institute for Civil Defense and Emergency Situations of the Ministry of Emergency Situations of Russia, 7, Davydkovskaya str., Moscow, 121352, Russia. E-mail: rebrovstanislav@rambler.ru

CITATION FOR AN ARTICLE

Sergeev E.B., Tskhovrebov E.S. Technology for extracting meteorological data from an internet site to assess the ecological condition of settlements // Vestnik of Tver State Technical University. Series «Building. Electrical engineering and chemical technology». 2024. No. 1 (21), pp. 96–110.