

УДК 681.3

А.А. Стенин, В.П. Пасько, А.Н. Губский, М.А. Каурковская

ПОСТРОЕНИЕ И АНАЛИЗ СИСТЕМ ИНТЕРНЕТ ГОЛОСОВАНИЯ И РЕГИСТРАЦИИ

Аннотация: в статье предложена математическая модель систем интернет-выборы и интернет-регистрация, в основу которой положена теория систем массового обслуживания.

Ключевые слова: Ключевые слова: система массового обслуживания, интернет-голосование.

Введение

В настоящее время голосование и регистрация являются одним из обязательных процессов при проведении различного рода собраний, телевизионных ток-шоу, работе законодательных органов, проведении референдумов, регистрации документов в различных государственных инстанциях и т. д. В частности, одними из наиболее важных для каждой страны являются процессы голосования на выборах, начиная от формирования местных органов власти и кончая выборами главы государства, и ежегодная регистрация абитуриентов при поступлении в высшие учебные заведения.

В связи с очевидными недостатками “ручного” голосования и регистрации, особенно в Украине [1], внедрение современных информационных технологий имеет несомненную актуальность.

Наиболее перспективным является внедрение электронного голосования и регистрации. Системы электронного голосования и регистрации предполагают различные виды технического исполнения на основе современных информационных систем и технологий [2], [6], среди которых наиболее распространенными являются телефонная связь, перфокарты, системы оптического сканирования и специализированные терминалы. Особую популярность во многих странах мира приобретают интернет-системы, так как в этом случае можно проголосовать или зарегистрироваться из любой точки земного шара (рис. 1).

Наряду с очевидными преимуществами, системы электронного голосования и регистрации имеют и некоторые проблемы реализации их эффективной работы.

Постановка задачи

Основными проблемами работы систем СИВ и СИР являются проблемы связанные как с соблюдением правовых норм – это аутентификация, недопустимость неоднократного доступа, “прозрачность” и др., так и с техническими требованиями к нормальной работе систем в различных режимах работы информационно-коммуникационной сети.

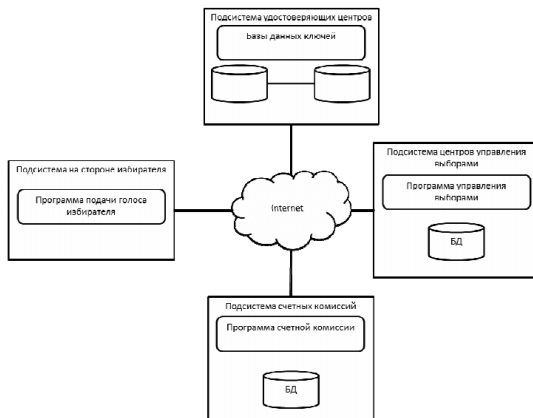


Рис. 1 – Обобщенная схема системы электронного голосования и регистрации с использованием Интернет-ресурсов

Несмотря на то, что для решения указанных проблем уже предложены некоторые конкретные пути, в частности инфраструктура управления открытыми ключами РКІ, которая позволяет проводить безопасную идентификацию и аутентификацию граждан с использованием цифровых подписей и электронных ID-карт, многие вопросы построения эффективных систем интернет-выборы и интернет-регистрации являются открытыми. Как показывает практика, одной из основных проблем, является проблема обеспечения эффективной работы в пиковых режимах, когда в короткий промежуток времени голосует или регистрируется большое число респондентов.

Построение и анализ СИВ и СИР

Поскольку СИВ и СИР являются по сути системами социального опроса, то наиболее подходящим математическим аппаратом организации их работы является теория массового обслуживания.

Будем рассматривать работу СИВ и СИР как работу системы массового обслуживания (СМО) с ожиданием, ограниченным временем пребывания заявки в системе и упорядоченном обслуживанием [3]. В качестве каналов обслуживания здесь выступают специально выделенные серверы, количество которых равно n .

Работу СИВ и СИР в этом случае можно описать следующим образом. Если к моменту запроса респондента (заявки) свободен хотя бы один из каналов (серверов), то этот запрос принимается к обслуживанию только одним (любым) из свободных каналов. Запрос, заставший все каналы занятыми может быть в дальнейшем обслужен, если за время пребывания его в очереди освобождается хотя

бы один из каналов. Максимальное число мест в очереди ограничивается числом m , то есть в системе может на данный момент находится $(n + m)$ заявок респондентов. При превышении этого числа приходит отказ и предложение повторить запрос через некоторое время.

Находясь в очереди респондент может проявлять “нетерпение”. Обозначим интенсивность потока уходов респондентов из очереди величиной, входного потока запросов λ и потока обслуживания μ . В общем случае параметры λ, ν и являются функциями времени $\lambda(t), \nu(t), \mu(t)$. Граф состояния представлен на рис. 2.

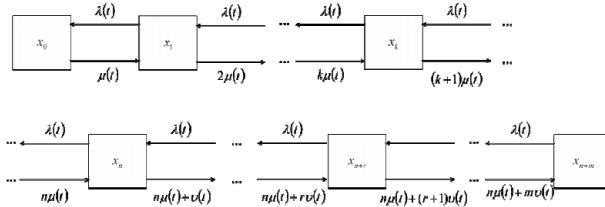


Рис. 2 – Граф состояний

Здесь $x_k (k = \overline{1, n})$ и $x_{n+r} (r = \overline{1, m})$ – характеризуют состояние, в котором одновременно находятся $k + r$ запросов. Используя известное мнемоническое правило [4], запишем систему дифференциальных уравнений для вероятностей состояния системы.

$$\left\{ \begin{array}{l} \dot{\rho}_0(t) = -\lambda(t)\rho_0(t) + \mu(t)\rho_1(t) \\ \dots \\ \dot{\rho}_k(t) = -[\lambda(t) + k\mu(t)]\rho_k(t) + \lambda(t)\rho_{k-1}(t) + (k + 1)\mu(t)\rho_{k+1}(t) \\ \dots \\ \dot{\rho}_n(t) = -[\lambda(t) + n\mu(t)]\rho_n(t) + \lambda(t)\rho_{n-1}(t) + [n\mu(t) + \nu(t)]\rho_{n+1}(t) \\ \dots \\ \dot{\rho}_{n+r}(t) = -[\lambda(t) + n\mu(t) + rv(t)]\rho_{n+r}(t) + \lambda(t)\rho_{n+r-1}(t) + \\ + [n\mu(t) + (r + 1)v(t)]\rho_{n+r+1}(t) \\ \dots \\ \dot{\rho}_{n+m}(t) = -[n\mu(t) + mv(t)]\rho_{n+m}(t) + \lambda(t)\rho_{n+m-1}(t) \end{array} \right. \quad (1)$$

Данная система решается на интервале $[0, T]$, где T – период голосования(регистрации) при известных, или полученных статистическим путем $\lambda(t), \nu(t)$, заложенным в СИВ(СИР) $\mu(t)$ и начальными условиями $p_0(0) = 1, p_i(0) = 0$, где $i = \overline{1, n}$ причем для вероятностей $\rho_i(t)$ выполняется условие связи

$$\sum_{i=0}^n p_i(t) = 1, t \in [0, T] \quad (2)$$

С помощью найденных $p_i(t)$ можно определить основные показатели качества работы СИВ(СИР), такие как среднее время обслуживания, среднее время нахождения респондента в системе, вероятность обслуживания, вероятность отказа и другие.

Как уже указывалось ранее, особый интерес представляет работа СИВ и СИР в пиковых режимах. Предположим, что на некотором периоде времени $(t_1, t_2) \subset [0, T]$ установился стационарный пиковый режим. Это возможно если предположить, что параметры λ , ν и μ на протяжении данного периода $(t_1, t_2) \in [0, T]$ голосования являются постоянными, тогда вместо системы (1) можно рассматривать систему алгебраических уравнений, из которых в аналитическом виде можно получить интересующие нас показатели работы в пиковом режиме. Данное предположение не обязательно, оно лишь упрощает процедуру анализа СИВ и СИР.

В частности, при определенных допущениях [3], можно показать, что вероятность обслуживания респондента в пиковом режиме будет равна

$$P_{обсл} = \frac{\mu \bar{k}}{\lambda} \quad (3)$$

где $\bar{k} = \sum_{i=0}^n i p_i + \sum_{r=1}^m n p_{n+r}$ – математическое ожидание числа занятых каналов. Фактически формула (3) определяет вероятность наличия в этот период свободного канала.

Из формулы (3) следует, что при наличии пикового режима, когда $\lambda \gg 0$ вероятность обслуживания значительно снижается вплоть до “зависания” всей системы. Во избежание такого случая, очевидно, что необходимо либо существенно увеличить величину μ , то есть максимально минимизировать время обработки одного запроса, либо увеличить длину очереди при неизменном числе каналов ($\bar{k} \leq r$), то есть снизить вероятность отказа. Эта задача решается разработкой специального программного обеспечения, о чем будет сказано ниже (первый путь решения). Второй путь заключается в увеличении числа каналов (серверов), то есть числа n . Второй связан с созданием необходимого запаса ресурса. При наличии достоверной информации второй путь достаточно тривиален, не во всегда выполним (количество серверов всегда ограничено). Кроме того второй путь органически связан с реализацией первого пути.

Техническая реализация

Рассмотрим техническую реализацию СИВ и СИР. Учитывая выше изложенное, предлагается строить СИВ и СИР как многоканальную СМО с ожиданием, ограниченным пребыванием заявок в системе и упорядоченным обслуживанием. Респонденты со своих специально организованных терминалов формируют поток “заявок”, который поступает на n -каналов (серверов) обслу-

живання. Число каналов обычно определяется административно-территориальным устройством той или иной страны.

Как видно из рис. 3, информационно-коммуникационную сеть СИВ и СИР предлагается строить по радиальному принципу, а схему взаимодействия как систему “клиент-сервер”. Для крупной страны эта блок-схема может быть частью более общей схемы, построенной по тому же принципу. Вся информация стекается в центральный информационный комплекс (ЦИК).

Как уже говорилось выше, анализ статистических данных предыдущих избирательных компаний дает возможность определить количественные оценки основных показателей работы СИВ и СИР.

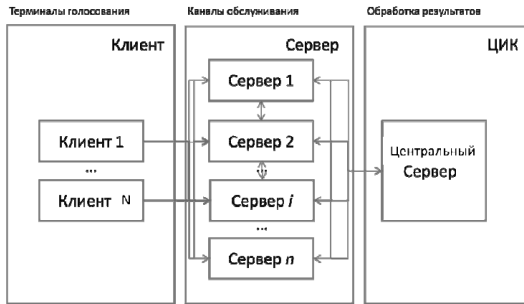


Рис. 3 – Блок-схема СИВ(СИР)

Как было показано выше, наиболее проблемными для СИВ и СИР является рассмотренный выше пиковый режим, когда практически одновременно в системах находится большое число респондентов.

В частности для СИВ, на рис. 4, по данным СМИ приведена диаграмма активности избирателей на выборах в Верховную Раду Украины в 2012 году.

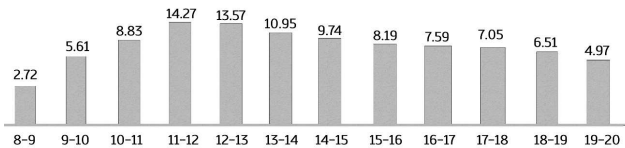


Рис. 4 – Диаграмма активности избирателей. Выборы в Верховную Раду Украины 2012 г.

Из диаграммы (рис. 4) видно, что во время пиковой активности избирателей с 11 до 12 часов, количество проголосовавших состав-

ляло 14,27% за час от их общего числа. По данным Центральной избирательной комиссии Украины количество зарегистрированных избирателей в 2012 году на выборах в Верховную Раду составляло ≈ 36 млн. человек. Соответственно в часы пиковой нагрузки при стопроцентной явке избирателей количество голосующих могло достигать $5 \cdot 4 \cdot 10^6$ человек в час (т.е. $\lambda_{\max} \approx 1.5 \cdot 10^3$ в секунду. Аналогично можно оценить и $\mu(t)$).

Согласно формуле (3) при пиковой нагрузке в СИВ и СИР резко снижается вероятность обслуживания респондентов. Для ее увеличения необходимо, как указывалось выше, либо значительно повысить интенсивность обслуживания (сделать $\mu \gg 0$), либо увеличить число каналов обслуживания. Поскольку число реальных каналов n ограничено, то и первая и вторая задачи решаются программным образом. В частности используя метод виртуальных каналов, предлагается для каждого i -го сервера организовать k_i дополнительных динамических каналов обслуживания. Такой канал устанавливается при вызове, по нему передается информация, после чего канал закрывается (уничтожается). В этом случае общее число текущих каналов обслуживания будет определяться как

$$n_{\text{обсл}}(t) = n + \sum_{i=1}^{n_i} k_i(t) \quad (4)$$

Такой подход помогает существенно снизить жесткие требования к величине интенсивности обслуживания в пиковых режимах.

В качестве примера практической реализации СИР можно рассматривать автоматизированную информационную систему “Конкурс”, которая в 2012-2013 годах стала одним из основных механизмов обеспечения открытости и прозрачности вступительной кампании абитуриентов, поступавших в ВУЗы Украины.

Работа системы “Конкурс” выявила проблемы, присущие многоканальным системам массового обслуживания. В первую очередь, это проблема колебания нагрузки, т. е. проблема определения порога пиковой нагрузки, которая была рассмотрена выше и были предложены пути ее решения.

Кроме того, наиболее оптимальным для решения проблемы колебаний нагрузки было бы наличие в каналах обслуживания модуля самодиагностики [5], который отслеживает:

- нагрузки на процессор;
- потребляемую память;
- количество подключений;
- количество запросов к базе данных.

В случае перегрузки этот модуль связывается с подсистемой самодиагностики и передает информацию о проблеме.

Система самодіагностики на основі даних, получаемых от разных модулей, принимает решение либо о:

- перераспределении нагрузки между всеми каналами системы;
- переводе отдельных каналов системы в автономный режим (когда им выделится только ограниченный объем ресурсов);
- отключение отдельных каналов системы и перераспределение их ресурсов между остальными каналами.

При наличии подобной подсистемы самодіагностики можно обеспечить частичную работу системы и локализовать проблему, не останавливая всю систему.

Заключение

В статье предложена математическая модель систем интернет-выборы и интернет-регистрация, в основу которой положена теория систем массового обслуживания. Рассмотрен и проанализирован наиболее проблемный режим их работы - пиковый режим. Предложены пути реализации нормального функционирования систем в этом режиме.

Список использованных источников

1. Ключковский Ю. Стабілізація виборчого законодавства – одне із основних завдань сучасної України / Вісник центральної виборчої комісії – 2008 – №3(13) – С. 19–23.
2. Глобальні інформаційні системи та технології і моделі ефективного аналізу, опрацювання та захисту даних/ Под ред. В.В. Пасічника – Львів. Видавництво Львівської політехніки 2006, – 348 с.
3. Овчаров Л.А. Прикладные задачи теории массового обслуживания. – М. Машиностроение – 1969. – 324 с.
4. Вентцель Е.С. Теория вероятностей. – М.:Наука 1964. – 402 с.
5. Риордан Дж. Вероятностные системы обслуживания – М.: Связь, 1966 – 184 с.
6. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы технологии, протоколы – СПб: Питер. – 2006 – 958 с.

Отримано 22.03.2014 р.