Проблемы статистического оценивания данных мониторинга в задачах безопасности компьютерных сетей.

А.А.Макаров, Г.И.Симонова, Н.Л. Ковба, А.Полищук

НИИ механики МГУ

Основные параметры мониторинга трафика на входе канала

- 1. IP - число пакетов на входе канала
- 2. IB - объем переданной информации в байтах
- 3. IF - число соединений
- 4. IT - совокупное время соединений

Различные срезы трафика

- 1. По типу сетевых протоколов (http, ftp и т.д.)
- 2. По подсетям
- 3. По отдельным IP адресам
- 4. И т. д.

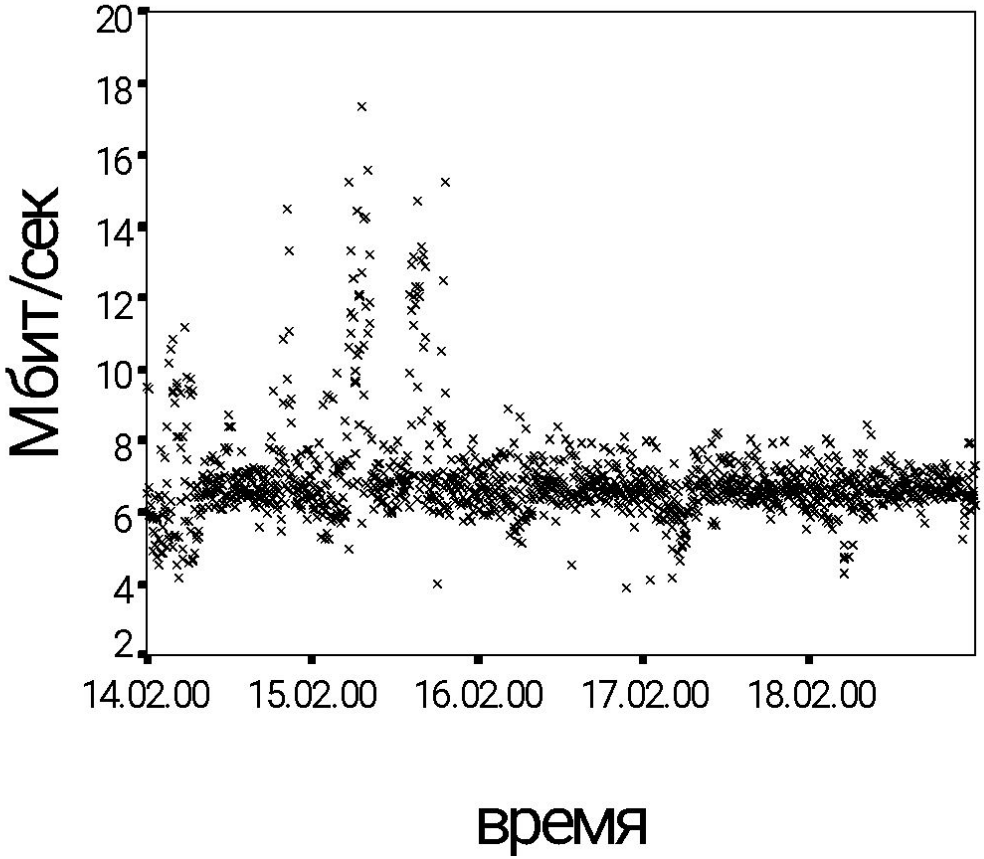
Традиционные характеристики описательной статистики

- Среднее значение
- Дисперсия
- Корреляция и автокорреляция
- Спектр
- Асимметрия, эксцесс
- Моменты старших порядков

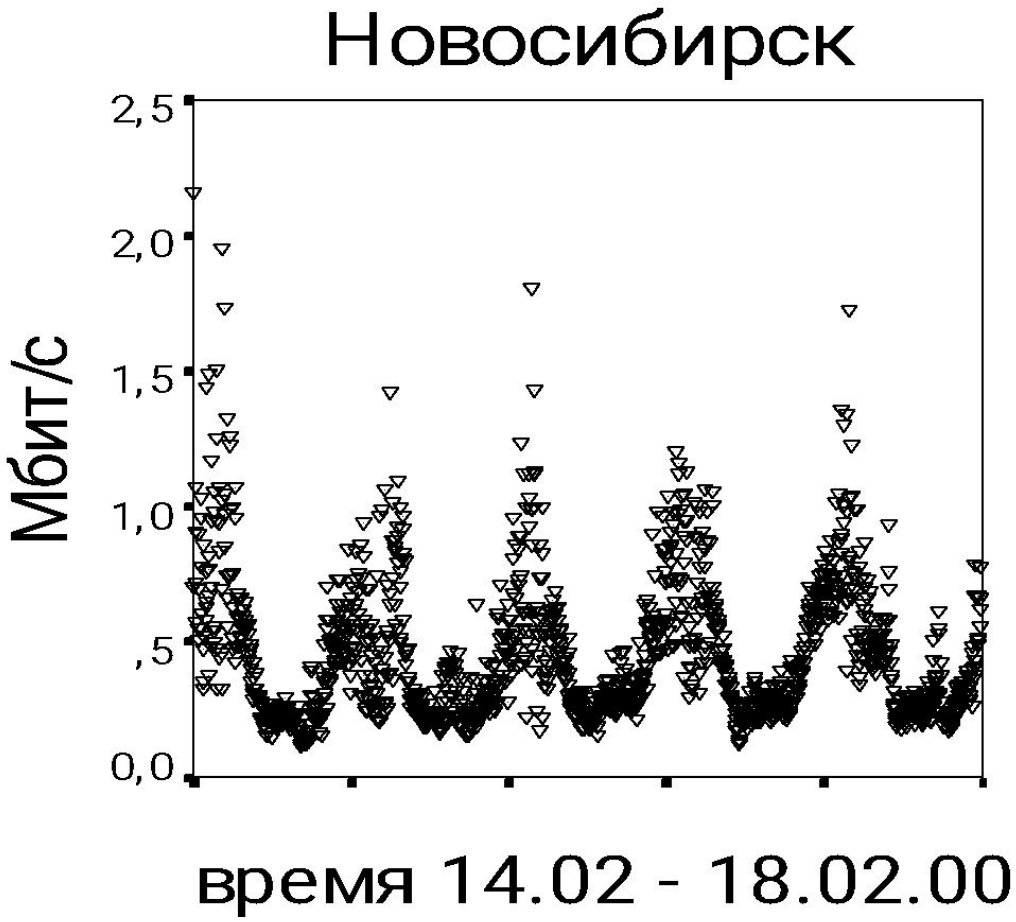
Ведущие научно-образовательные сети на первое полугодие 2000г. (по объему трафика извне)

Наименование сети	Число хостов	Средняя доля загрузки входа канала
МГУ им. М.В.Ломоносова	4200	9.7%
Фед. зел RUNnet	4100	9.4%
РАН у в НГУ	2000	7.2%
ОИЯИ (ДУБНА)	2800	6.3%
РАДИО-МГУ	5800	6.2%
Томская гор. образ. науч. сеть	2900	5.5%
Сеть Екатеринбург. ун-тов	3100	4.4%
RELARN-MSK	10000	3.8%
Сеть FREEnet	7000	3.3%
ИППИ РАН (Москва)	1650	3.0%

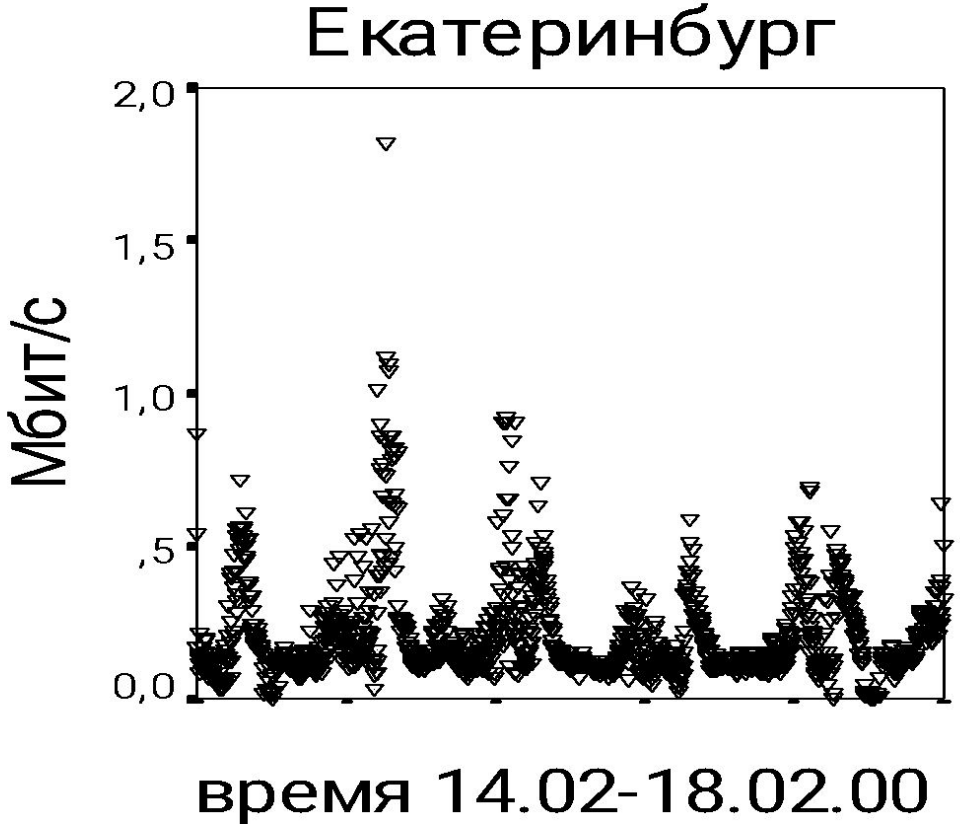
Характер загрузки входа канала (за 5 мин.)



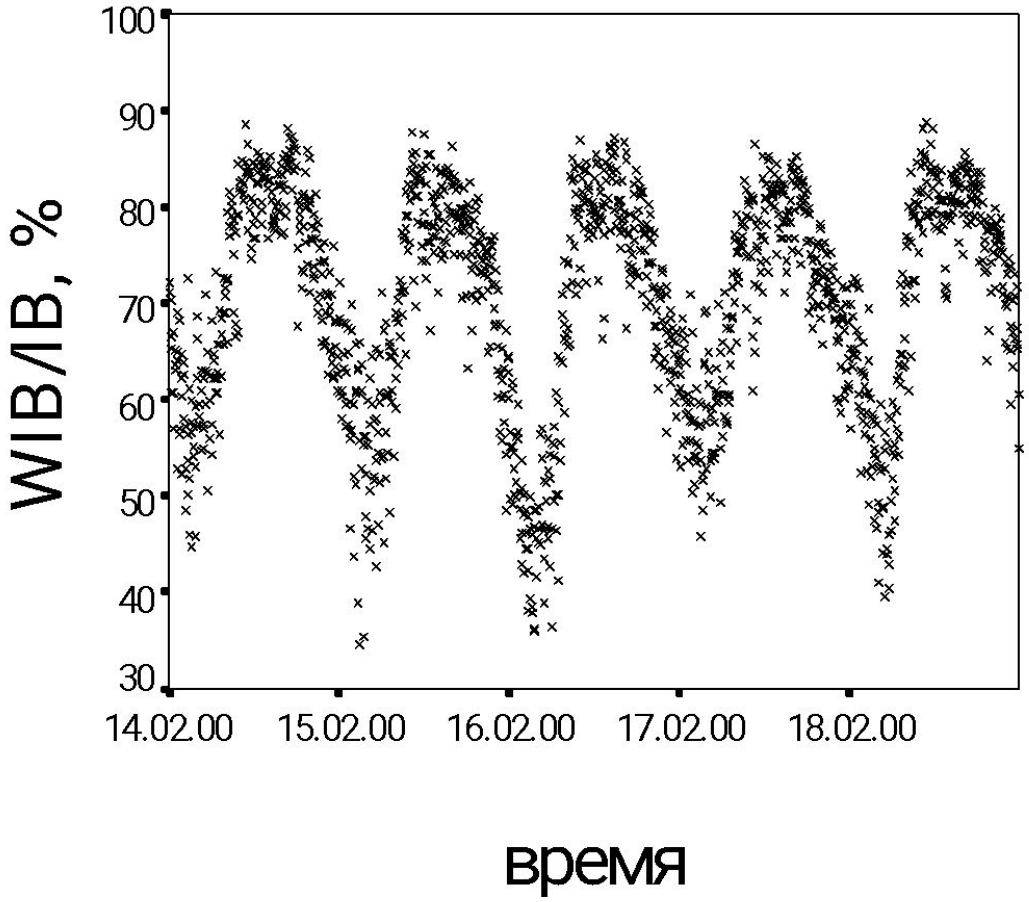
Характер загрузки входа канала (за 5 мин.)



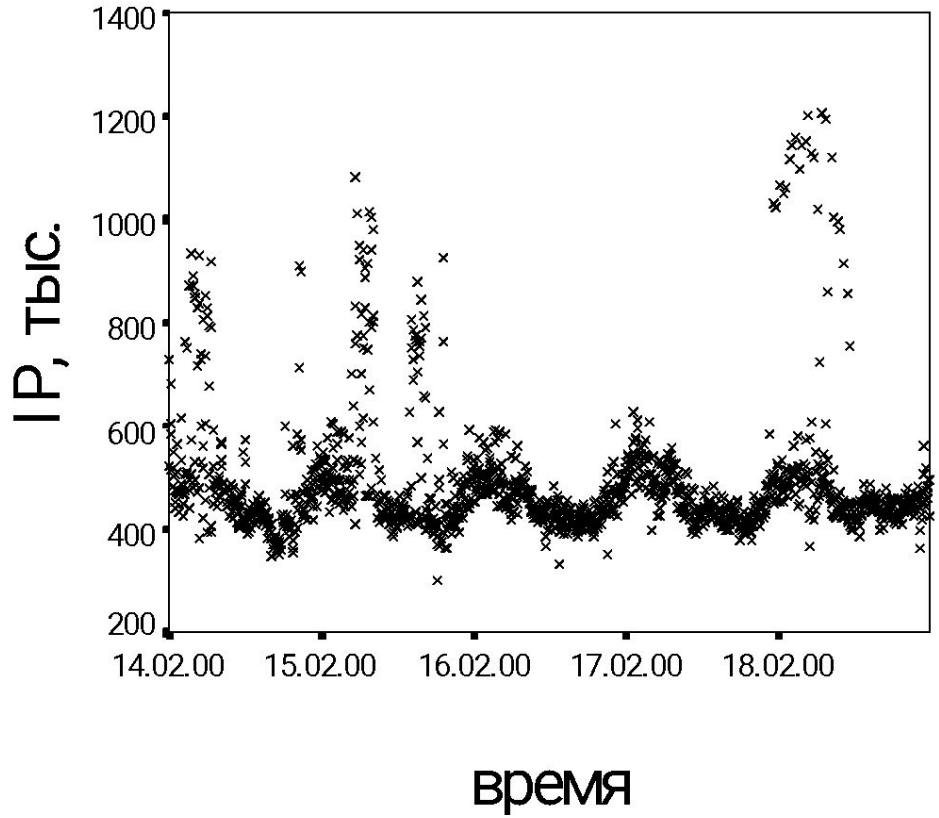
Характер загрузки входа канала (за 5 мин.)



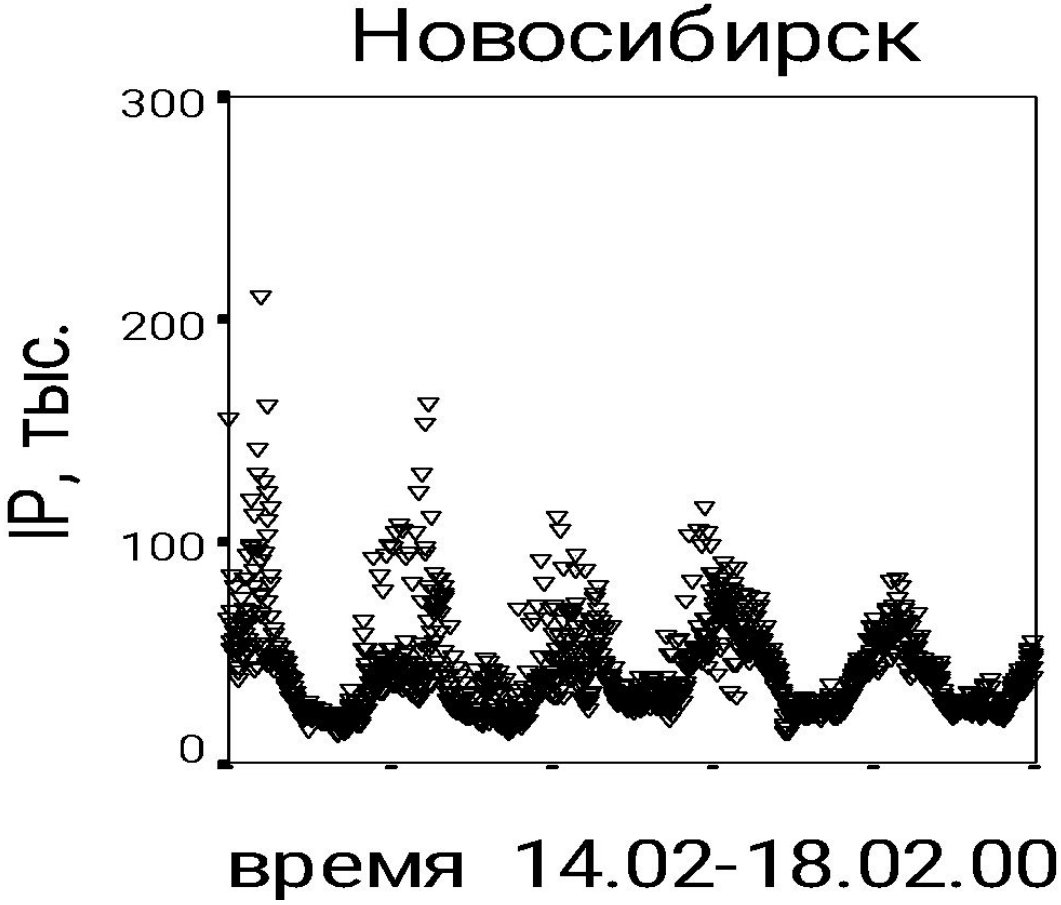
Доля протокола Http на входе канала (за 5 мин.)



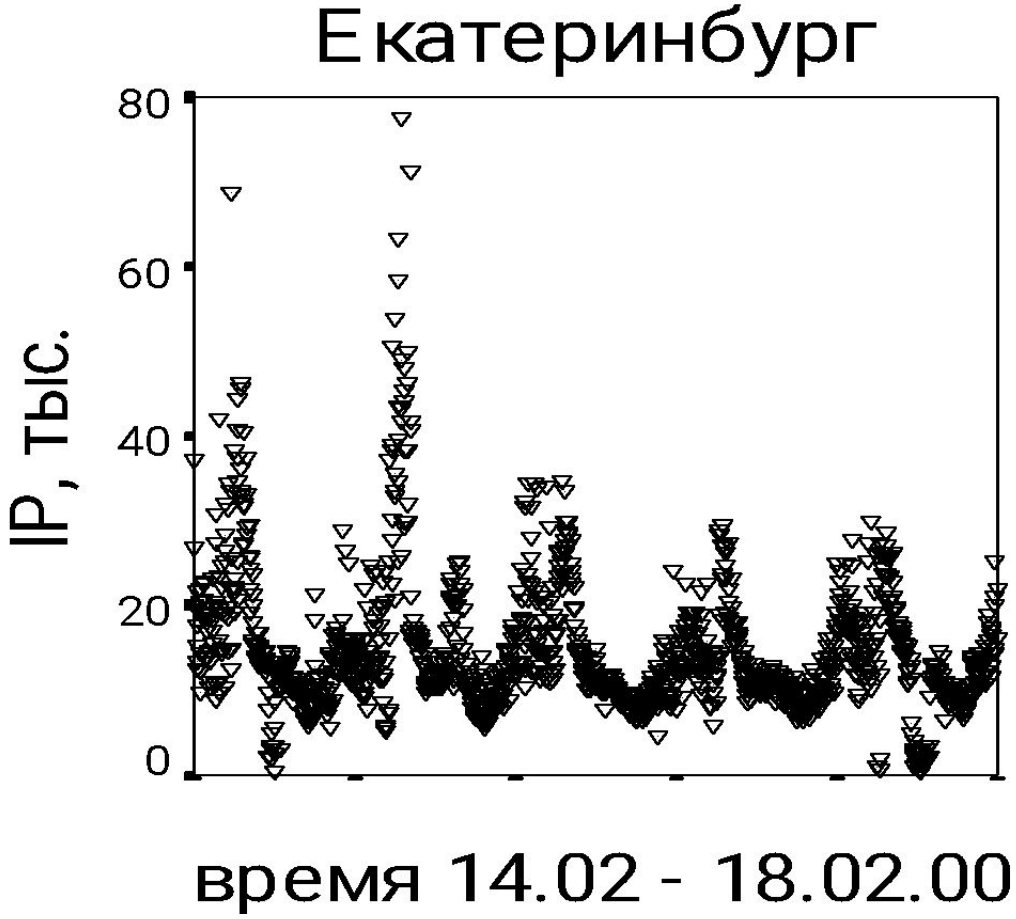
Число пакетов на входе канала (за 5 мин.)



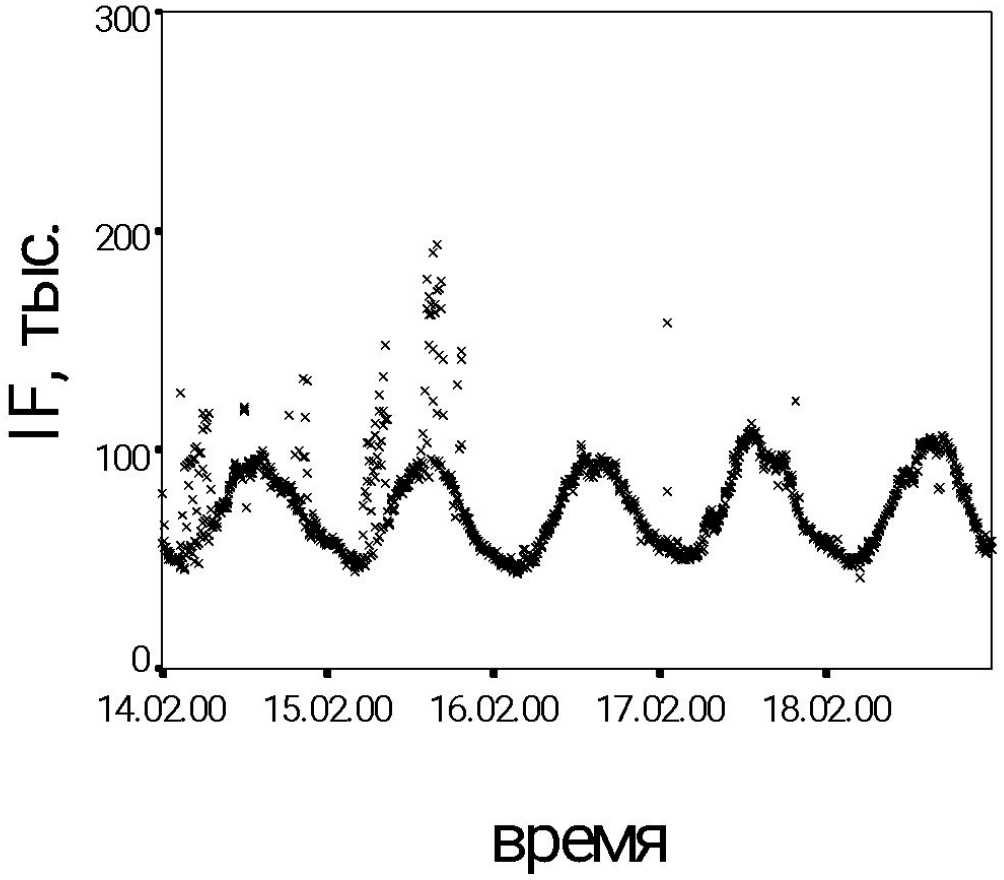
Число пакетов на входе канала (за 5 мин.)



Число пакетов на входе канала (за 5 мин.)

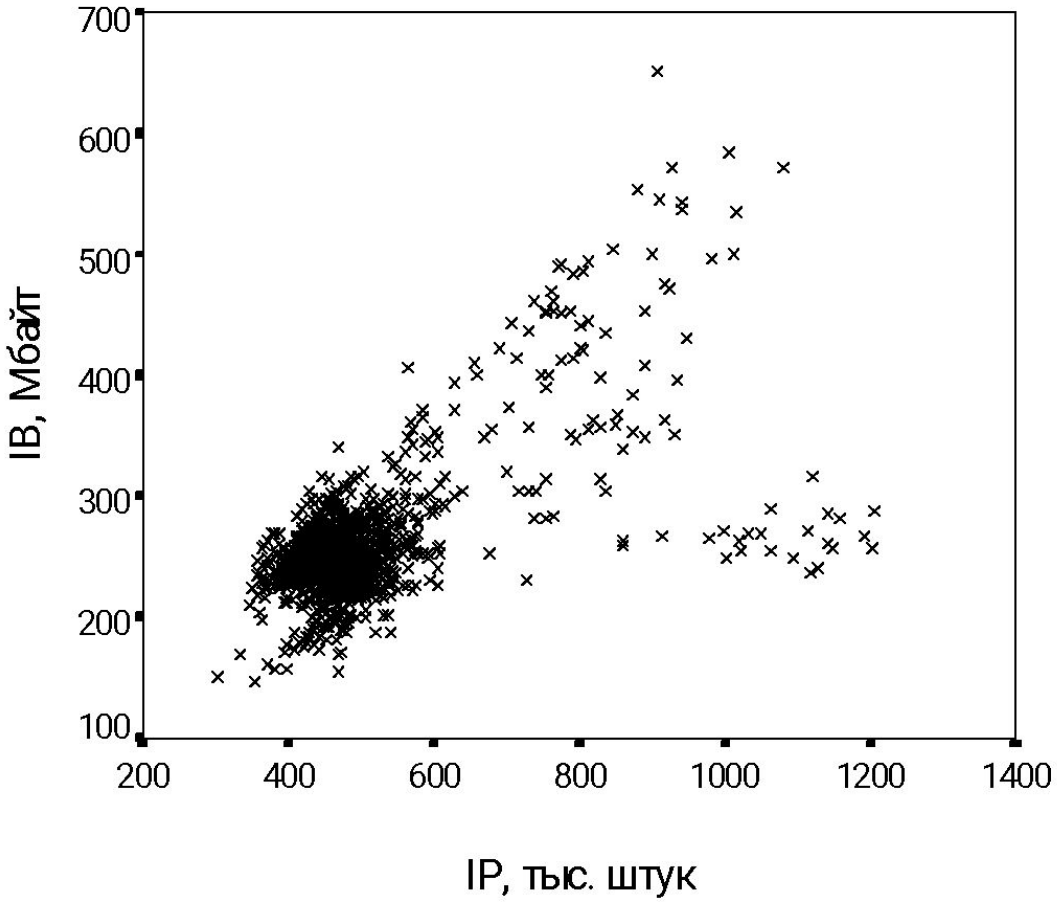


Число соединений на входе канала (за 5 мин.)

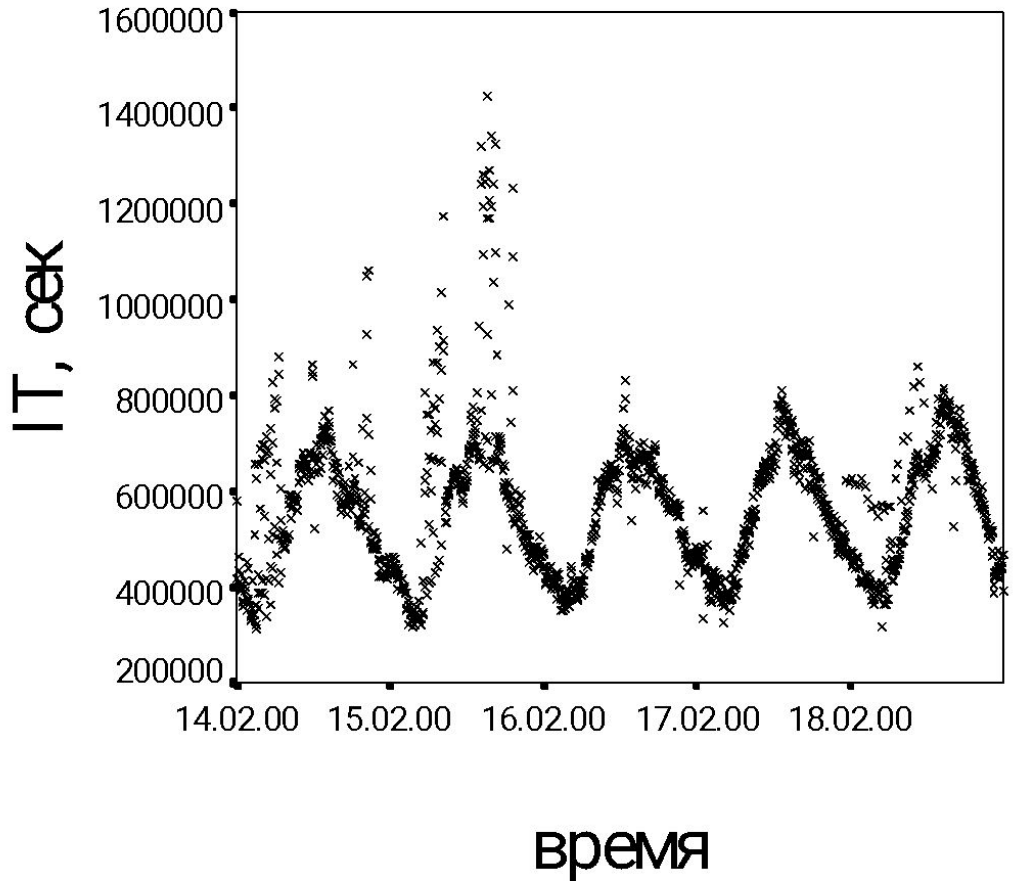


Зависимость объема трафика от числа переданных пакетов

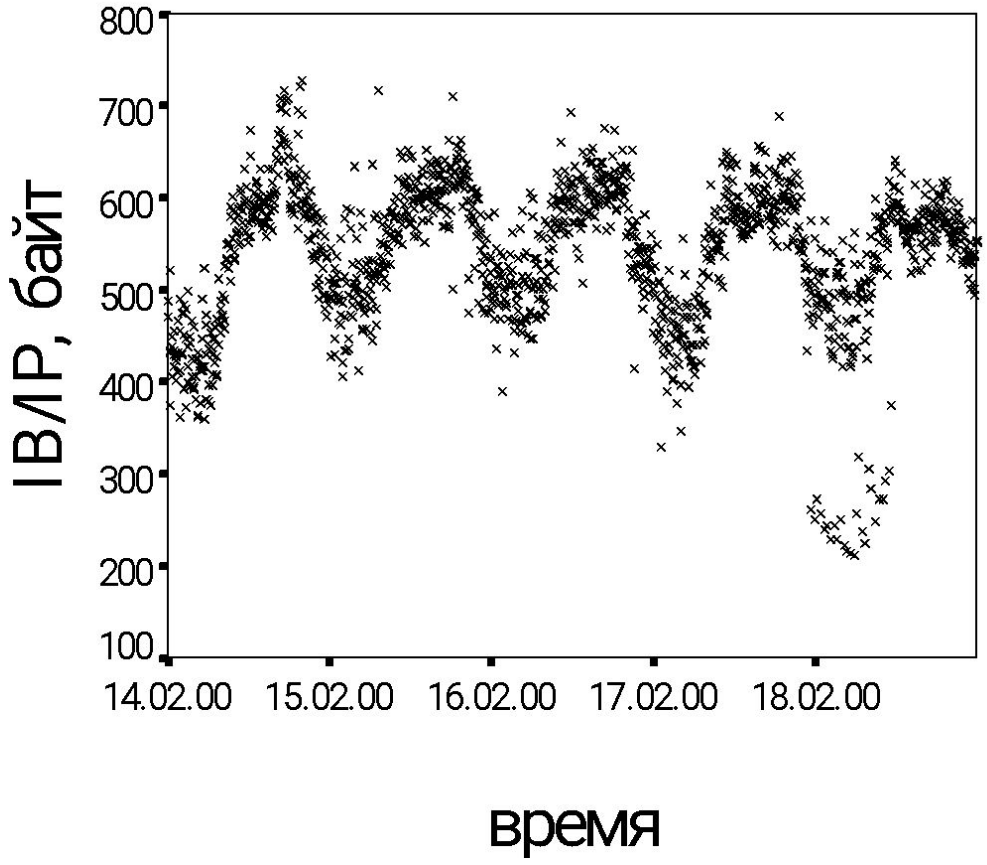
(средние за 5 мин. в течение 5 рабочих дней)



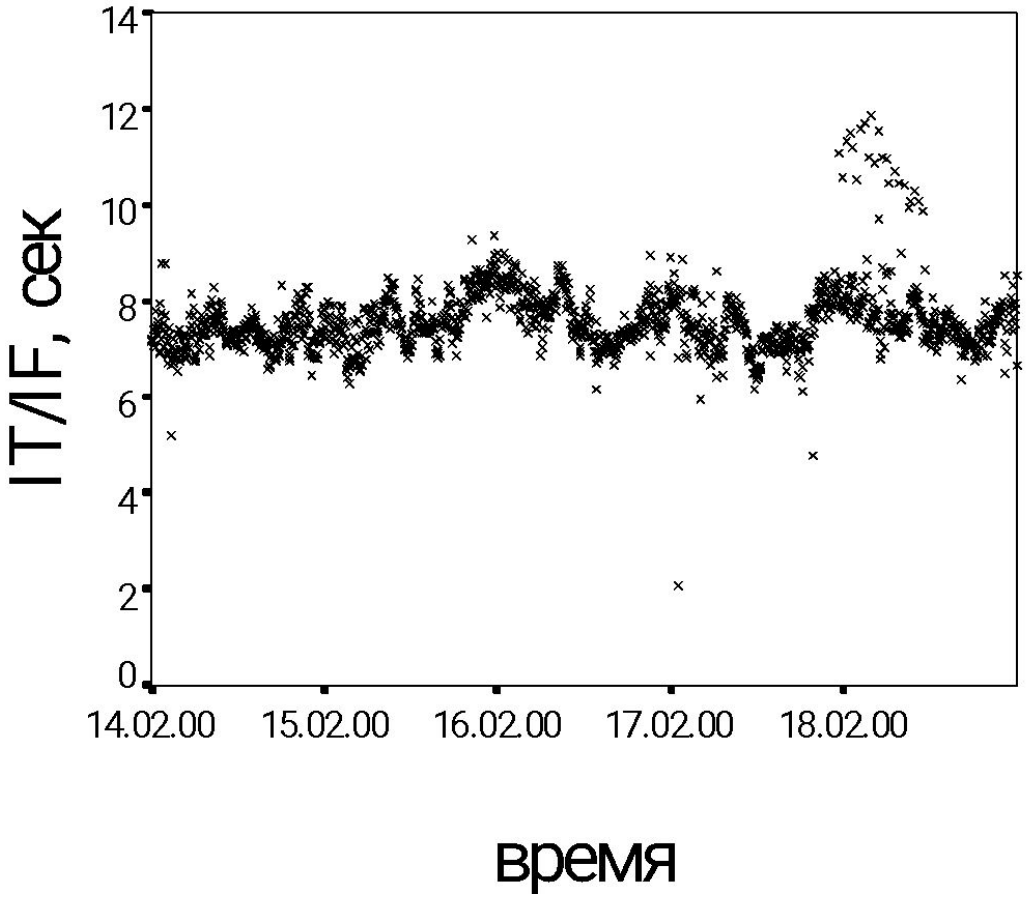
Общее время соединений на входе канала (за 5 мин.)



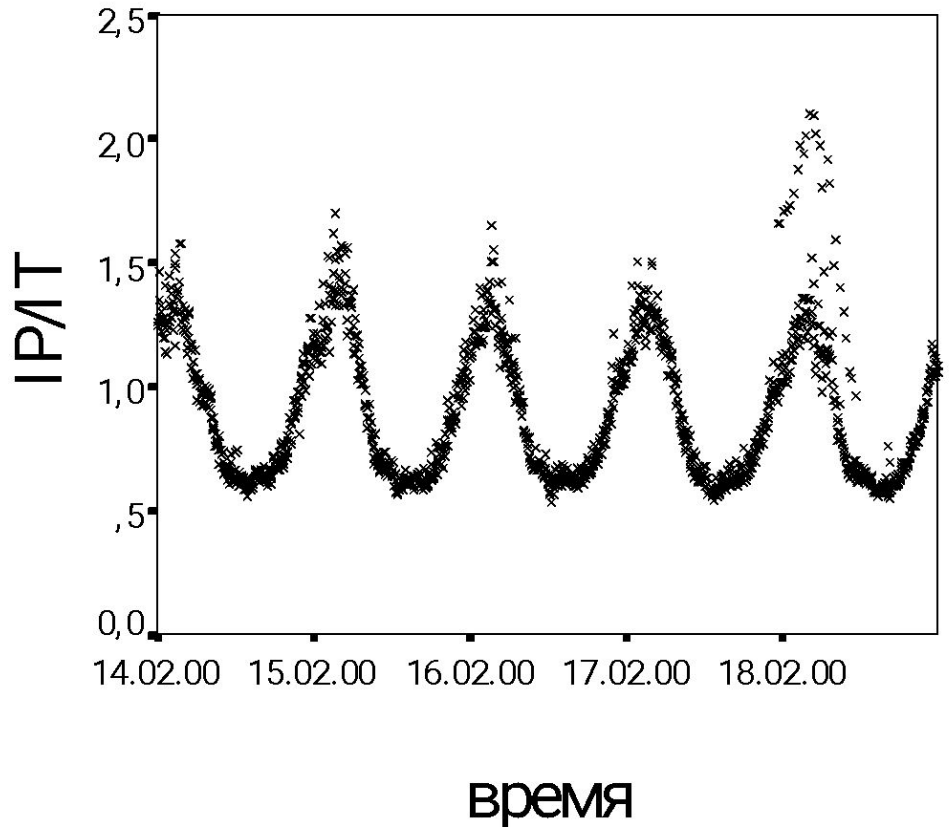
Средний размер пакета на входе канала (за 5 мин.)



Среднее время соединения на входе канала (за 5 мин.)



Число пакетов в единицу времени (за 5 мин.)



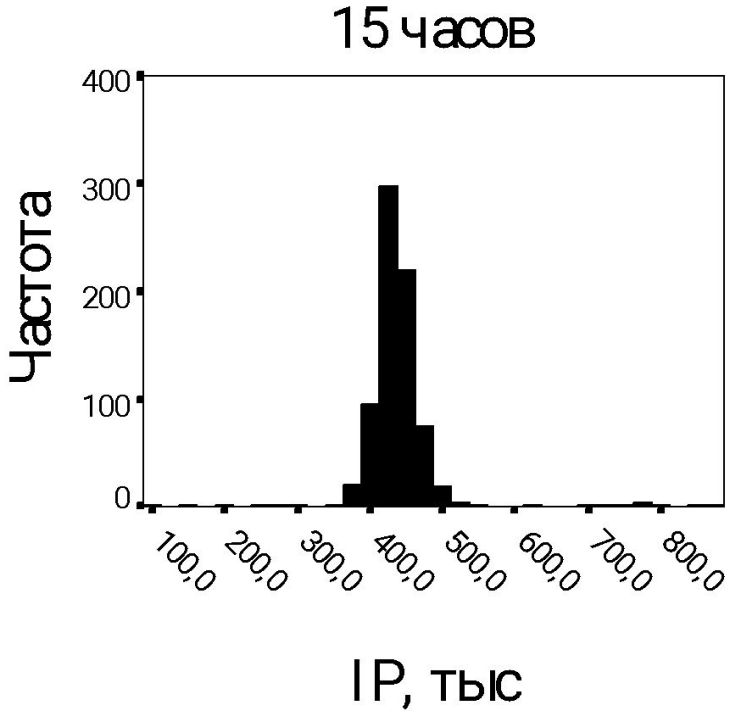
Сравнение средних значений и разбросов числа пакетов на входе в 14-15 и 15-16 часов

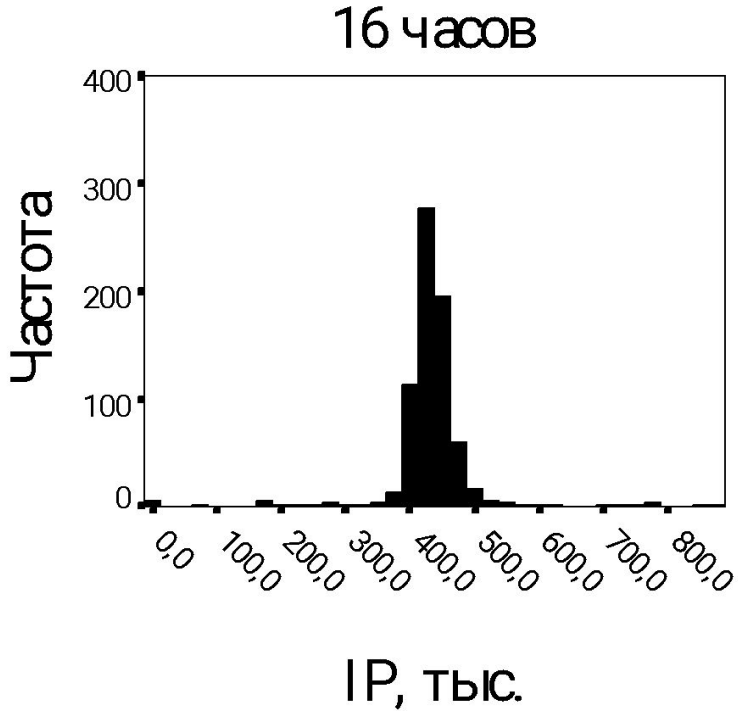
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
IP	Equal variances assumed	0,205	,010	2,137	174	,033
	Equal variances not assumed			2,158	1350,935	,031

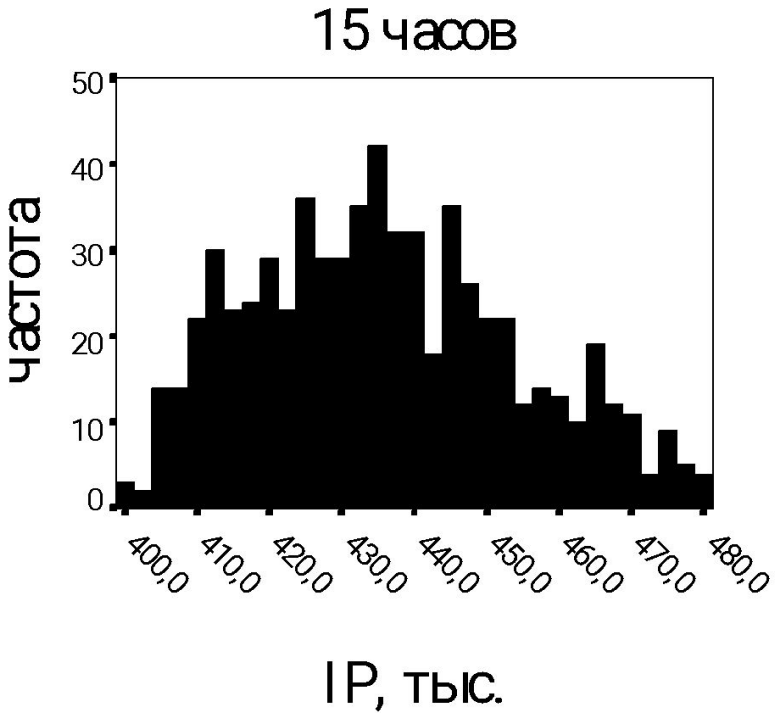
Гистограммы не усеченных данных

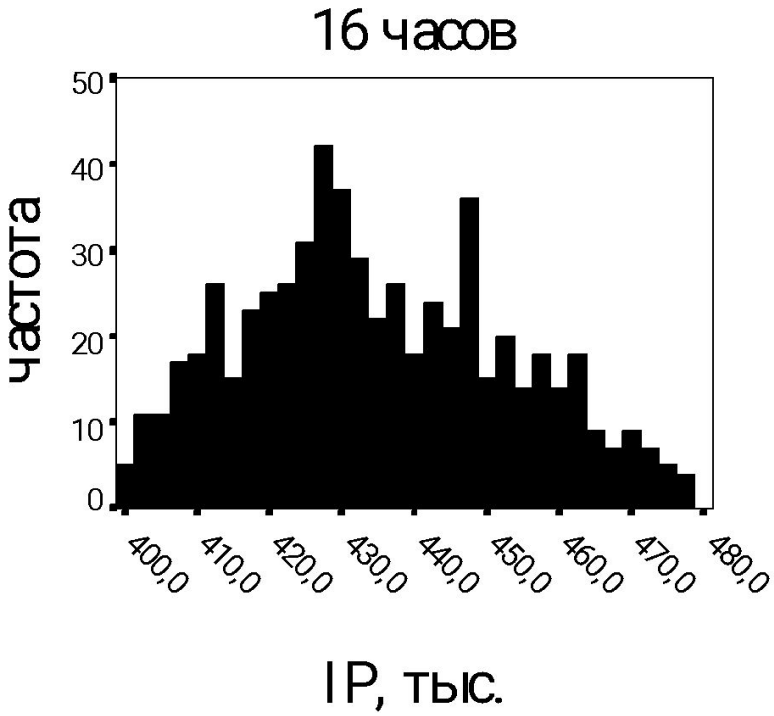
(число пакетов на входе за 5 мин.)





Гистограммы после усечения
(число пакетов на входе за 5 мин.)





Сравнение средних значений и разбросов числа пакетов на входе в 14-15 и 15-16 часов после усечения

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
IP	Equal variances assumed	,178	,670	,363	1520	,336
	Equal variances not assumed			,363	1517,675	,336

Алгоритм проверки согласия условных распределений

- 1. Вычисление базовых описательных статистик: min, max, медиана, межквартильный размах (range).
- 2. Выбор предварительных границ для наиболее вероятных значений
- медиана-1.5 range медиана+1.5 range
- 3. Определение числа шагов усечения и размера шага снизу и сверху.
- Расчет двухвыборочной статистики Колмогорова-Смирнова для условных распределений.

Протокол расчета двухвыборочного критерия согласия Колмогорова-Смирнова для условных распределений усеченных выборок

час	час	границы		статистика	уровень	процент отсечения		
		мин	мах		значимости	в целом	снизу	сверху
2	6	1004	49797	1,421724	0,035	0	0	100
2	6	1103	49067	1,480169	0,025	0,6	0,2	99,6
2	6	1203	48337	1,459563	0,028	0,6	0,2	99,6
2	6	1302	47607	1,459563	0,028	0,6	0,2	99,6
2	6	1401	46878	1,459563	0,028	0,6	0,2	99,6
2	6	1501	46148	1,438906	0,032	0,6	0,2	99,6
2	6	1600	45418	1,438906	0,032	0,6	0,2	99,6
2	6	1699	44688	1,438399	0,032	0,7	0,2	99,5
2	6	1798	43958	1,396875	0,04	0,7	0,2	99,5
2	6	1898	43228	1,374317	0,046	0,7	0,2	99,5
2	6	1997	42498	1,374851	0,046	1	0,2	99,3
2	6	2096	41768	1,311171	0,064	1,1	0,2	99,1
2	6	2196	41039	1,290036	0,072	1,1	0,2	99,1
2	6	2295	40309	1,246839	0,089	1,2	0,2	99
2	6	2394	39579	1,225466	0,099	1,5	0,2	98,8
2	6	2494	38849	1,203031	0,111	1,8	0,2	98,4
2	6	2593	38119	1,136525	0,151	1,8	0,2	98,4
2	6	2692	37389	1,202037	0,111	2,2	0,2	98
2	6	2791	36659	1,158159	0,137	2,4	0,2	97,8
2	6	2891	35930	1,135787	0,151	2,7	0,2	97,6
2	6	2990	35200	1,135349	0,152	2,9	0,2	97,3
2	6	3089	34470	1,202026	0,111	3,3	0,2	96,9
2	6	3189	33740	1,065094	0,207	3,7	0,2	96,6
2	6	3288	33010	1,052745	0,218	4	0,2	96,2

Группы часов с согласованными распределениями вероятностей числа пакетов на входе для различных сетей

Сеть/час	1	2	3	4	5	6	7	8	9	10	11
Общий	1	1	1		2	2	2				
Новос.	1	1	2	2	2	2	3	3			
Екатер.		1	1	1	1	1					

Сеть/час	12	13	14	15	16	17	18	19	20	21	22	23	24
Общий		3	3	4	4	4	4	5	5				
Новос.	4	4	4	5	5	5	5		5				
Екатер.	2	2	2	2				3	3				