И. В. БАХМУТОВА, В. Д. ГУСЕВ, Т. Н. ТИТКОВА

Институт математики им. С. Л. Соболева Сибирского Отделения РАН

УДК 001:002;002.53:004.89;78

КОМПЬЮТЕРНЫЙ ПОИСК ИНВАРИАНТНЫХ СТРУКТУРНЫХ ЕДИНИЦ ЗНАМЕННОГО РАСПЕВА

роблема нотолинейной реконструкции знаменных песнопений носит дешифровочный характер и в общем случае (беспометные тексты) остаётся нерешённой. Основные трудности дешифровки связаны с неоднозначностью соответствия «знамя - нота»: одна и та же цепочка знамён может иметь разные интерпретации в зависимости от гласа, типа структурной единицы, в состав которой она входит, её позиции в тексте и ряда других факторов. Начиная с XVII века, для уточнения певческого значения знамён используется система степенных и указательных помет. Это облегчает дешифровку, но не гарантирует однозначного результата из-за наличия в пометных текстах значительного числа беспометных знамён с не всегда очевидной интерпретацией, а также многочисленных разночтений и исключений в трактовке самих помет1.

Общеязыковым приёмом, используемым для устранения неоднозначности, является привлечение контекста, что эквивалентно переходу от отдельных знамен к структурным единицам - попевкам, лицам и фитам. Однако для построения формальных алгоритмов выделения из текста структурных единиц знаменного распева явно недостаточно описаний тех их свойств, которые содержатся в теоретических руководствах. Использование общеязыковых принципов² даёт хорошие результаты лишь при выделении массовых попевок гласа. Немаловажно и то, что известные подборки структурных единиц представлены в форме, малопригодной для дешифровки беспометной нотации. Например, у В. М. Металлова³ попевки приведены только в нотолинейной форме, у М. В. Бражникова⁴ лица и фиты снабжены степенными пометами и лишь половине из них сопутствует нотолинейная интерпретация.

В вышеуказанных и других подборках слабо отражён фактор вариативности, что затрудняет процесс выделения структурных единиц в реальных песнопениях. Так, покрытие нотных текстов двознаменников попевками из подборки Металлова составляет в среднем (по разным гласам) лишь порядка 25–30% объёма текста при поиске на точное соответствие. При поиске на приближённое соответствие (допускалось не более двух искажений типа замены, вставки или устранения символа в нотной записи попевки) коэффициент покрытия возрастал до 60%.

Целью предлагаемой статьи является определение и исследование специфического класса структурных единиц, условно названных нами внутригласовыми инвариантами (ВИ). Их можно охарактеризовать как цепочки беспометных знамён с минимальной (в идеале нулевой) степенью неоднозначности в плане интерпретации. Судить об интерпретации можно по текстам, представленным в формате «знамя — нота» (двознаменники). Условность понятия ВИ связана с его зависимостью от исходных данных. Если подборка достаточно представительна, то построенное по ней множество ВИ можно применять с известной осторожностью для дешифровки беспометных текстов.

Система обозначений

В работе использованы следующие обозначения. Обиходный звукоряд представлен нотами: G, A, H (малой октавы), c, d, e, f, g, a, b (первой октавы), С, D (второй октавы). Длительности звуков мая). Для обозначения высоты и длительности звука используем комбинацию буквы и цифры (например, H4 – это четвертная нота «си» малой октавы). Знак (*) используется в качестве разделителя между нотолинейными интерпретациями разных знамён, а (~) – как символ эквивалентности, отделяющий знаменную цепочку от её нотолинейного представления. Например, запись н ~ d4c4d2*c4e4 означает, что цепочка из двух знамён, стоящая слева от знака (~), интерпретируется в тексте двознаменника, соответственно, цепочками из трёх и двух нот, разделёнными знаком (*).

Формирование множества ВИ

Как уже отмечалось выше, метод решения предполагает наличие достаточно представительных подборок двознаменников. Число известных двознаменников невелико⁵, но достаточно для наших целей. Нами были переведены в электронную форму три двознаменных Октоиха конца XVII — начала XVIII века (РНБ, С.— Петербург, Соловецкое собрание, шифры 619/647, 618/644 и QI 188). Два первых представлены подборками $T\Gamma_i$ ($1 \le i \le 8$, i—номер гласа) из 25—27 песнопений в каждом гласе; в QI188 этот показатель на треть ниже.

Под внутригласовым инвариантом (ВИ) относительно заданной подборки мы понимаем любую повторяющуюся цепочку беспометных знамён произвольной длины L (L=1,2,3,K), имеющую одинаковую интерпретацию во всех своих вхождениях в песнопения данного (i-го) гласа.

Множество *ВИ* формировалось для каждого из трёх Октоихов отдельно и в каждом разбивалось по гласам на подмножества $BU(T\Gamma_{I})$, К, $BU(T\Gamma_{g})$. Ранее по этим Октоихам нами была построена электронная азбука знаменного распева⁶, содержащая частоты встречаемости отдельных знамён и всевозможных их интерпретаций во всех гласах. Уже тогда было отмечено, что существуют знамёна, однозначно интерпретируемые во всех песнопениях одного гласа инварианты длины 1. В данной работе исследуется вопрос о существовании инвариантов длины 2 и выше. Основная посылка - это отсутствие формальных ограничений на возникновение однозначно интерпретируемой цепочки путём конкатенации многозначных знамён. Процедура выявления ВИ произвольной длины состоит из трёх этапов.

Объединяем песнопения i—го гласа в один текст $T\Gamma_i = T_{il} @ T_{i2} @ ... T_{ik(l)}$, где k(i) — число песнопений в i—ом гласе, T_{il} — песнопение с номером l ($l \le l \le k(i)$), записанное в виде цепочки знамён, «@» — символ — разделитель, не принадлежащий алфавиту знамён. Каждое знамя из текста T_{il} сопровождается нотолинейной интерпретацией.

Этап второй. Для каждого текста $T\Gamma_i$ ($l \le i \le 8$) подсчитываем внутригласовую встречаемость всевозможных цепочек знамён длины L, не содержащих

разделитель $(L=1,2,3,\mathrm{K},L_{imax},\,\,\mathrm{где}\,\,L_{imax}-\,\,\mathrm{длина}\,\,\mathrm{мак-}$ симальной повторяющейся цепочки в $T\Gamma_i$). Процесс ведётся итеративно по L. Сравнение и подсчёт цепочек (L–грамм) осуществляется за один просмотр текста без непосредственного сопоставления каждой L–граммы с каждой (процедура хеширования). По завершении L—ой итерации получаем полный спектр L—грамм, представленных в песнопениях i—го гласа, с указанием частоты встречаемости каждой цепочки и списка её возможных интерпретаций (тоже с частотами).

Этап третий. Для каждого гласа выделяем из списка найденных L–грамм (L=1,2,3,... , L_{imax}) лишь те, что имеют частоту встречаемости $F \geq F_{nop} > 1$ и единственную интерпретацию. Они и образуют списки внутригласовых инвариантов ВИ ($T\Gamma_i$) ($1 \leq i \leq 8$). Например, биграмма I=1 встретилась 15 раз в 1-м гласе двознаменника 619/647 с единственной интерпретацией: I=1 встремент в список I=1 другая же биграмма I=1 имеет в этом гласе две интерпретации: I=1 имеет в этом гласе две интерпр

Описание экспериментальных результатов

Сформированные для каждого из трёх двознаменников *словари ВИ* проиллюстрируем на Октоихе 619/647, привлекая другие двознаменники для оценки чувствительности методики к объёму исходных данных и проявлениям вариативности в разных Октоихах.

Основные выводы таковы:

1) Многообразие ВИ разной длины во всех гласах довольно велико и существенно зависит от порога отбора по частоте встречаемости ВИ в гласе (F_{nop}) . Мы выбрали $F_{nop}=3$. ВИ с меньшей частотой обладают невысокой предсказательной способностью. В Таблице 1 указано число ВИ разной длины с частотой встречаемости $F \geq 3$ в каждом из 8 гласов Октоихов 619/647 и QI188 (его объём примерно на треть меньше).

Таблица 1 Размеры словарей ВИ разной длины ($F \ge 3$)

	Октоих 619/647									Октоих QI 188							
L	Гласы								L	L Гласы							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
1	10	12	9	12	11	8	13	9	1	6	5	9	5	6	3	9	4
2	66	49	49	77	55	66	82	73	2	45	<u>38</u>	33	46	38	31	<u>35</u>	31
3	<u>79</u>	<u>58</u>	<u>67</u>	<u>96</u>	<u>81</u>	<u>73</u>	<u>109</u>	<u>95</u>	3	<u>46</u>	<u>38</u>	<u>47</u>	<u>58</u>	<u>52</u>	<u>36</u>	<u>35</u>	<u>40</u>
4	72	52	<u>67</u>	82	<u>81</u>	52	98	84	4	32	15	45	37	36	30	26	27
5	57	32	48	77	65	41	76	68	5	18	5	30	22	22	21	15	15
6	43	18	31	50	44	27	44	46	6	6	2	15	11	9	9	12	7
7	25	3	15	27	28	15	27	30	7	2	1	5	6	5	1	6	1
8	12	-	6	14	16	5	12	18	8	-	-	-	4	2	-	2	-

Примечание: BU с длиной L > 8 не представлены (их мало).

Длины BU в Октоихе 619/647 меняются от 1 до 16, а в QI188 — от 1 до 9 знамён. Сравнение размеров словарей BU для двух Октоихов (при фиксированных i и L) показывает, что меньшему объёму исходных данных (QI188) соответствуют меньшие размеры словарей BU. Этот вывод заранее не очевиден, так как с увеличением объёма исходных данных растут проявления вариативности, ограничивающие рост числа BU.

2) С увеличением L число BU в каждом гласе сначала растёт, при $L=3\div 4$ достигает максимума (эти значения выделены и подчёркнуты), а затем монотонно падает. Такая зависимость числа BU от L объясняется влиянием двух факторов. Первый фактор связан с увеличением многообразия L—грамм при переходе от L=1 к L=2, L=2 к L=3 и т. д., что способствует увеличению числа BU с ростом L. Особенно показателен переход от L=1 к L=2. Например, в гласе 4 Октоиха 619/647 выделено всего 12 BU длины 1, но уже 77 BU длины 2. При этом лишь 27 из них «наследственные», то есть получены всевозможными расширениями BU длины 1, а 50 BU (L=2) образованы конкатенацией многозначных знамён.

Второй фактор способствует уменьшению числа BH с ростом L. Это связано с уменьшением частоты встречаемости L—грамм по мере увеличения их длины. Если частота снижается до $F < F_{nop}$, то L—грамма устраняется из рассмотрения. При малых L превалирует первый фактор и число BH растёт, при больших L—второй и число BH падает.

3) Частоты встречаемости BU колеблются в широком диапазоне и зависят от их длины, гласовой принадлежности и типа структурной единицы, в которую входит BU (см. ниже Таблицу 2).

Из Таблицы 2 видно, что пересечений между множествами BH одинаковой длины мало, даже для параллельных гласов — таких, как 4 и 8. При L=I в пересечение попало лишь знамя . Гласоспецифичность этого BH проявляется в секвентном переносе в 8-м гласе его распева на кварту вверх и уменьшении длительностей всех звуков вдвое. При L=2 в пересечение попадают цепочки . \square и \square с одинаковой интерпретацией в обоих гласах, но существенно отличающимися частотами, поэтому не все из них включены в топ-списки Таблицы 2. Эти примеры показывают, что словари BH содержат полезную информацию о проявлениях гласоспецифичности знаменного распева.

4) Важен вопрос о соотношении *ВИ* с традиционными структурами знаменного распева, в первую очередь, с попевками. В классификации А. Н. Кручининой⁷ попевки разбиты на 24 семейства в соответствии с тремя кадансовыми знамёнами, образующими «архетип». Ему предшествует более вариативный «подвод».

Короткие ВИ $(L=1\div 3)$, как и следовало ожидать, часто реализуют ритмический останов и относятся к категории кадансов (см., например, в табл. 2, глас 4, знамена (x, y), цепочки (x, y), (x,

Таблица 2 Примеры наиболее частых ВИ для двух гласов Октоиха 619/647

P1		ГЛА	C 4			ГЛА	C.8
L	ВИ	F	Интерпретация — — — — — — — — — — — — — — — — — — —	L	ВИ	F	Интерпретация
1	سننه	23	e4f4g2	1	220	21	f4.g8a4g4f4a2.g4
1	منبر	14	f4g4a2	1		18	f4g2f4
1		12	e2g2	1	2 <	12	a2.g4
1		9	c4d4e1	1	منتذ	4	g4a4b2
1	^ ∩	8	H4c4d4e4	1	7:	3	d2
1	**	6	e4d4c2d1	1	ຶ່	3	e2d4c4
1	**	4	f4e2	1	**	3	a8g8f4g2
2))	37	g2*f1	2	ب ب	38	d4e4*d4c4
2	75	27	d4e4*f1	2	~ ~ ~	35	f4.e8d4f4*c4d2e4
2	ا ناز	18	f4e4*d4e4f2	2	٥٠	24	f2*d4e4
2	֓֞֝֝֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֓֓֡֓֓֓֓֓֡֡	17	e2*f4e4	2	L- N	24	f4e4*d2
2	5 \†	14	e1*d1	2	إخا	22	g2*f4e4
2	١٠ ١٦	12	d2e4f4*g1	2	SAND	22	f4g4f4.e8*d4c4d2
2	λ÷	11	e4f4*g4f4	2	٦٤٤	21	f4e4*f4.g8a4g4f4a2.g4
2	一	10	e2g2*f1	2	# Z	20	d4f4* e4d4e4f4
2	nt ñ	10	d2*e4f4	2	~~ ñ	16	f4g2f4* e4f4

Иногда короткие BU возникают на стыке разных структурных единиц, как в случае с BU = $\mathbb{L} \sim d1*f2*f4e4$ (F=5, $2\pi.4$). Здесь «статья» = завершает предшествующую попевку, а «параклит» \mathbb{L} начинает другой раздел песнопения. С параклитом связаны и BU в началах песнопений. В частности, с помощью таких BU в 1-м гласе идентифицируется \mathbb{A} :

BИ средней длины $(4 \le L \le 8)$ обычно представлены либо законченными попевками, либо достаточно крупными их фрагментами, зацепленными друг с другом. Примером первых являются приведённые выше попевки семейств «грунка» (глас 4), «кулизма» (глас 8), а также: 1000 10

Длинные BU(L ≥ 9) представляют комбинации разных структурных единиц (двух попевок, тандемного повтора с попевкой и т. п.), например:

В случае б) имеем комбинацию *тандемного пов-тора*, образуемого нотными цепочками «тандем 1» и «тандем 2», отличающимися лишь секвентным переносом, и двух попевок — «подъём» и «срединка» (?).

"срединка"(?)

"подъём"

Оценки полноты множества ВИ

Интересно оценить степень покрываемости выделенными BH (порог отбора по частоте $F_{nop}{=}3$) как исходного текста, так и контрольных текстов, не использовавшихся при формировании множества BH.

Коэффициент покрытия k конкретного песнопения подсчитывается следующим образом. Выделяем в песнопении все цепочки знамён, являющиеся BH, и подсчитываем суммарное число покрываемых ими позиций n. Каждая позиция учитывается лишь odun pas, чтобы нивелировать эффект перекрываемости отдельных BH. Тогда k = n/N, где N – число знамён в песнопении (его длина).

Коэффициенты покрытия песнопений меняются в широких пределах: в первом гласе от 0.23 до 0.85, во втором – от 0.30 до 0.73, в третьем – от 0.29 до 0.84 и т.д. Это, скорее всего, обусловлено ограниченностью материала (отдельные жанры в гласах представлены 2-3 песнопениями, поэтому некоторые жанроспецифичные цепочки знамён отсеивались по частоте при F_{nop} =3). Ниже приведены средние значения коэффициента покрытия \bar{k} по каждому гласу двознаменника 619/647:

Глас	1	2	3	4	5	6	7	8
\bar{k}	0.64	0.49	0.61	0.68	0.62	0.51	0.79	0.62

Нетрудно видеть, что в среднем покрываемость песнопений каждого гласа на материале обучения достаточно высока – порядка 60% и выше.

Для получения более реалистичной оценки был проведён эксперимент в режиме скользящего контроля, когда из обучающей подборки песнопений гласа одно («контрольное») удалялось, а по оставшимся строилось множество ВИ и подсчитывался коэффициент покрытия этими ВИ контрольного песнопения. Затем песнопение возвращалось в обучающую подборку, а в качестве «контрольного» рассматривалось следующее песнопение и т. д. Ниже показаны средние значения коэффициента покрытия в режиме скользящего контроля:

Глас	1	2	3	4	5	6	7	8
\bar{k}	0.55	0.38	0.51	0.51	0.57	0.38	0.69	0.54

Можно видеть, что в этом режиме наблюдается ощутимое снижение показателей покрываемости по сравнению с предыдущим случаем.

Ещё один эксперимент был проведён с привлечением в качестве контрольного материала песнопений Октоиха 618/644. Ниже приведены среднегласовые коэффициенты покрытия песнопений этого Октоиха инвариантами, построенными по Октоиху 619/647:

Глас	1	2	3	4	5	6	7	8
\bar{k}	0.49	0.33	0.41	0.37	0.43	0.22	0.54	0.46

Наблюдается снижение показателей покрываемости по сравнению с вариантом скользящего контроля, то есть степень различия между Октоихами носит значимый характер, что уже отмечалось⁸.

 $(2\pi.4, F=3)$

Обсуждение результатов

Может показаться, что методика выделения BM очень чувствительна к проявлениям вариативности знаменного распева. Например, высокочастотный BM из первого гласа Октоиха 618/644 — $L_{-5} \sim e4f4d4$ (F=24) не является BM в Октоихе 619/647 из-за наличия двух интерпретаций: e4f4d4 (F=18) и d4f4d4 (F=3). Однако, расширения знамени L_{-5} (см.гл. 1, 619/647) уже образуют BM длины 3: $L_{-5} \approx e4f4d4*e1*d1$ (F=9) и $L_{-5} \approx 1 \sim e4f4d4*e1*d1$ (F=6). Эти же цепочки фигурируют среди BM в Октоихе 618/644 с частотами (F=13) и (F=6), соответственно, но там они носят наследственный характер. Итак, отсутствие знамени L_{-5} среди BM в Октоихе 619/647 не мешает однозначной идентификации архетипа попевки «колесо» ($L_{-5} \approx 2$) в обоих Октоихах.

ВИ можно трактовать как фрагменты, наилучшим образом сохранившиеся в ходе эволюции. Ниже приведена половина стихиры «Врата адовы....» (глас 4, 619/647), где эти фрагменты (режим скользящего контроля) выделены подчёркиванием:

Коэффициент покрытия $k=20/34\sim0.59$, что несколько выше среднего значения для гласа 4 в режиме скользящего контроля ($\bar{k}=0.51$). Результаты выглядят достаточно логично: в качестве BU выделены отде-

льные попевки (см. фрагменты 2, 3, 5), структура 4 (расширенный подвод к архетипу «грунка» (см. 5) и устойчиво повторяющаяся «внепопевочная» структура 1. Непокрытыми остаются начальный фрагмент и межпопевочные структуры, отличающиеся повышенной вариативностью.

Значения коэффициентов покрытия были получены при пороге отбора F_{nop} = 3. Фактически могут быть использованы и BU с частотой 2. Использование низкочастотных BU повышает покрываемость, но уменьшает надёжность предлагаемого решения.

Заключение

Исследуется специфическая система структурных единиц знаменного распева, названных нами внутригласовыми инвариантами. В её основе лежит выделение цепочек знамён различной длины, однозначно интерпретируемых (в пределах исходной обучающей подборки) в конкретном гласе. На материале двознаменников конца XVII - начала XVIII века построены словари ВИ разной длины для всех гласов, проведена их количественная и качественная характеризация. Получены оценки покрываемости внутригласовыми инвариантами беспометных песнопений, не входивших в исходную подборку. Места вхождения ВИ в текст можно трактовать как своего рода «островки надёжности», существенно снижающие уровень неоднозначности в интерпретации знамён. Ввиду этого ВИ могут быть использованы в любых методиках дешифровки беспометной нотации.

ПРИМЕЧАНИЯ

- ¹ Бражников М. В. Древнерусская теория музыки. Л.: Музыка, 1972. С. 422; Бахмутова И. В., Гусев В. Д., Титкова Т. Н. О функциях указательных помет // Сибирский музыкальный альманах. Новосибирск: НГК, 2002. С. 81− 92; Бахмутова И. В., Гусев В. Д., Титкова Т. Н. Факторы, влияющие на точность нотолинейной реконструкции пометных знаменных песнопений // Сибирский музыкальный альманах. Новосибирск: НГК, 2004. С. 51−59.
- 2 Бахмутова И. В., Гусев В. Д., Титкова Т. Н. L-граммные азбуки для дешифровки знаменных песнопений // Сибирский журнал индустриальной математики. Т. 1, № 2. Новосибирск, 1998. С. 51–66.
- 3 Металлов В. М. Осмогласие знаменного распева (сборник нотолинейных попевок). М., 1899. С. 1–50.

- 4 Бражников М. В. Лица и фиты знаменного распева. Л.: Музыка, 1984. С. 302.
- 5 Бражников М. В. Древнерусская теория музыки. Л.: Музыка , 1972. Глава 12.
- ⁶ Бахмутова И. В., Гусев В. Д., Титкова Т. Н. Электронная азбука знаменного распева: предварительная версия. Вычислительные системы. Новосибирск: ИМ СО РАН, 2005. Вып. 174. С. 29–53.
- ⁷ Кручинина А. Н. Попевка в русской музыкальной теории XVII века: дис. ... канд. искусствоведени. Л., 1979.
- 8 Бахмутова И. В., Гусев В. Д., Титкова Т. Н. Электронная азбука знаменного распева...

Бахмутова Ирина Владимировна

научный сотрудник Института математики Сибирского отделения Российской Академии Наук

Гусев Владимир Дмитриевич

кандидат технических наук, старший научный сотрудник Института математики Сибирского отделения Российской Академии Наук

Титкова Татьяна Николаевна

кандидат технических наук, научный сотрудник Института математики Сибирского отделения Российской Академии Наук

