

新たな知見を用いた詰問題創作への取り組み

石飛太一^{†1} 飯田弘之^{†1}

探索中に得られた情報を探索以外の目的に利用することはときに有効であると考えられる。本研究では、証明数探索で求められた証明数・反証数および探索ノード数を詰め将棋コンクールにおける順位と比較し、これらの探索指標をパズル問題の感性評価のために用いる。これらの探索指標はパズル問題において難易度を表すことを確認し、コンクールの順位付けと難易度には密接な関係があることがわかった。

Mating Problem Composition with Insightful Knowledge

TAICHI ISHITOBI^{†1} HIROYUKI IIDA^{†1}

The information or indicators obtained during a search can be applied to other purpose than the search itself. In this study we compare the proof number, disproof number and the number of nodes visited in the proof number search framework, with the ranking of shogi mating problem competition. We draw a conclusion that 1) the indicators such as proof numbers obtained during the search show the difficulty of the problems, 2) the indicators have a close relationship with the ranking in the competition.

1. はじめに

ゲーム木探索において探索の方向性などをコントロールする何かしらの指標は、一般に探索の効率化を高める目的で導入される[8]。しかし、探索効率化以外にも活用できる可能性がある。例えば、探索した展開ノード総数は対象となる問題の複雑さを表す指標となりうる。パズル問題ではその指標は難易度を示すと考えてよい。このような考えに基づいて、先の研究[4]において詰将棋を題材とし、証明数探索で得られる証明数・反証数などの指標と感性評価（コンクールの順位）との関係性を調べた。その結果、証明数などの指標は人間の評価と密接な関係があることを確認した。ただし、例外となるケースや、探索指標が詰将棋のどの要素を示しているのかを説明することができなかった。

本研究では、短手数詰将棋コンテストを題材として、より多くのデータに基づく分析を実施し、これらの課題に挑戦した。また、ここで得られた知見を基に、先行研究において開発した Shape Keeping Technique [6]を利用し、詰将棋の自動創作を試みる。

詰将棋の感性評価についての代表的な研究として小山ら[2]による試みが挙げられる。詰将棋の様々な評価要素（玉の開放度、駒の位置、種類など）に着目し、対象となる問題群を統計的に分析することで、人間による感性評価の傾向を分析した。詰将棋の評価で重要とされる様々な要素を考慮することは、感性評価というあいまいな要素を扱う際に大切である。しかし、これら全要素の関係性を理解し、コンテストの順位付けを明快に説明することは極めて困難

である。

そこで本研究では、関係性を説明し得る可能性がある探索の指標を利用する。具体的には、証明数探索において使用される証明数・反証数等の指標に着目し、1) 人間の評価を最もよく説明する指標は何か、2) その指標が人間の評価をどの程度説明でき、どの評価要素を包含するのか探求する。

先行研究[4]では、1992年刊行の将棋世界誌において実施された5手詰め・7手詰めコンテストの上位3位問題のデータと、練習問題集から得た一般問題30題の平均データを比較した。その結果、コンテストの上位問題は多くの場合、練習問題の平均よりも大きな証明数・反証数を持っていることを確認した。しかし例外もあり、コンテストの順位を定める指標の解明までには至らなかった。そこで本研究では、さらに多くのコンテストから情報を収集し、不足していたと考えられるデータ数の補強を行うことで評価の優劣と探索指標の関係解明に取り組む。また、コンテストの結果から得られた探索指標データという一方通行の情報だけでなく、得られたデータから問題創作に挑戦することで双方向の結びつきについて確認する。

2. 探索の指標と問題評価

詰将棋を解くにあたり、本研究では Proof-Number Search[1] (PNS) を用いる。PNS は AND/OR 木を構築するゲームにおいてルートノードの値を効率的に決定できる手法であり、詰将棋の探索によく用いられる。PNS は証明数・反証数という二つの指標を用いて探索を展開する。証明数

^{†1} 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

とは、あるノードの値を True に決定するために調べなければならない末端ノードの数であり、反証数とはあるノードの値を False に決定するため調べなければならない末端ノードの数である。True と False は詰将棋において、詰みと不詰みに相当する。証明数は詰みを証明するために調べなければならない末端ノード数となり、証明数が大きいと詰みを証明することが困難になる。変化が複雑で詰み手順を特定するのが難しくなるからである。また、反証数は不詰みを証明するために調べなければならない末端ノード数となり、反証数が大きいと不詰みを証明することが困難になる。よって、証明数・反証数はコンピュータにとっては詰将棋を解く上での難易度を表す指標と言える。

詰将棋のようなパズル問題の感性評価において、重要だと考えられる要素は難易度の他にも存在する。これら複数の評価要素をどのように計算するか考察したものとして、Iqbal[5]が挙げられる。ここでは、次のような計算式が提案されている。

$$P = \sum_{k=1}^m C_k \Rightarrow A = \sum_{i=1}^n P_i$$

P はある評価要素、C は要素の成分、A は全体評価を表している。すなわち、評価とはそれを構成する要素の合計であり、各要素はその要素を構成する成分の合計とする。

本研究で扱う証明数・反証数は難易度の要素、すなわち 1 つの P に当たる。重要なのは、難易度という要素が全体の評価においてどれほどの割合を占めているかという点である。本研究では、コンクールの順位と証明数・反証数を比較した結果から難易度と評価の関係を明らかにし、評価手法に探索指標がどの程度有効であるかを調査する。

3. 実験

詰め将棋パラダイス誌[3]等で実施された詰将棋短編コンクールの全作品を対象として、証明数などのデータ収集を行い、コンクール順位（人間による感性評価）と照らし合わせる。コンクールの実施概要は以下の通りである。最初に、同誌 12 月号で、詰将棋エキスパートの選考を経て入選した全問題が発表される。次に、全入選問題を対象として、解答および評価（A、B、C の 3 段階評価）を読者に募集する。その結果から、同誌翌年 4 月号においてコンテスト順位が発表される。順位は評価スコア A、B、C にそれぞれ 3、2、1 点を与える。ただし、誤答は 3 点とする。各問題の合計点を回答者数で割った点数を問題の評価点として順位付けされている。

次に、PNS を実装したソルバを用いて探索を実行し、証明数、反証数、展開ノード数などのデータ収集を行った。証明数、反証数は難易度と関係があることから、コンクールの順位（人間による感性評価）にも大きく関係すると予想した。また探索中の展開ノード総数も多いほど探索が困難になるため、同じく難易度と深く関係し、詰将棋の感性

評価と深い関わりがあると予想した。なお今回は通常の詰将棋についてのみ取り扱い、各コンクールに 1、2 問ほどある双玉問題についてはデータから除いた。以下実験結果を示す。

まず、実施されたコンクールの内、9 手詰めを対象とした 2007、2012、2014 年度及び、7 手詰めを対象とした 2013 年度結果発表のデータを収集した。得たデータの例を表 1 に示す。

表 1 2014 年度詰将棋コンクール実験結果の一例

問題 番号	証明数 最大	反証数 最大	ノード数 最大	順位
6	13	13520	534847	1
17	18	16907	1186086	2
15	120	14755	5986066	3

まず、問題番号とはコンクールにおいて各問題に与えられる番号である。順位は、コンクールの結果与えられた優劣である。また証明数、反証数、ノード数とは PNS の探索中に与えられる値であり、その最大とは探索中にルートノードにおいて記録した内、最も高い値を示す（ただし探索終了時に与えられる無限の値は除く）。

まず、各探索指標と順位との相関を調べた結果を示す。初めに例として 2014 年度コンクールの順位と証明数、反証数、ノード数最大の分布を図 1～図 3 に示す。

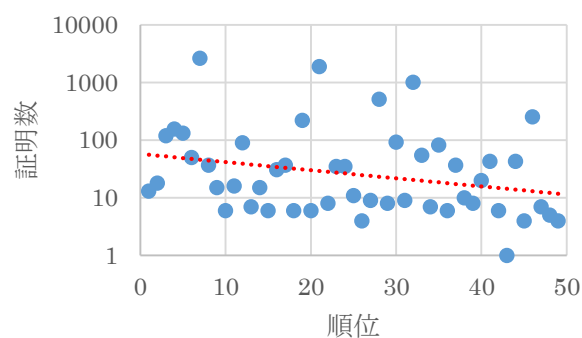


図 1 2014 年度コンクールの順位と証明数最大

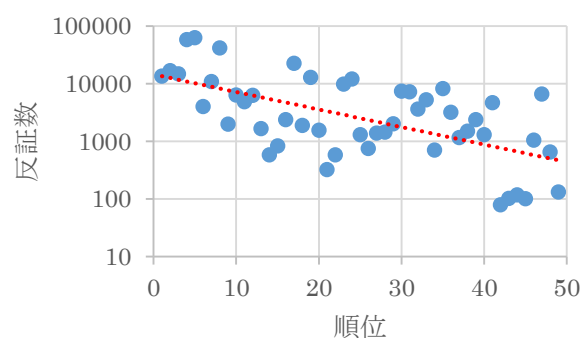


図 2 2014 年度コンクールの順位と反証数最大

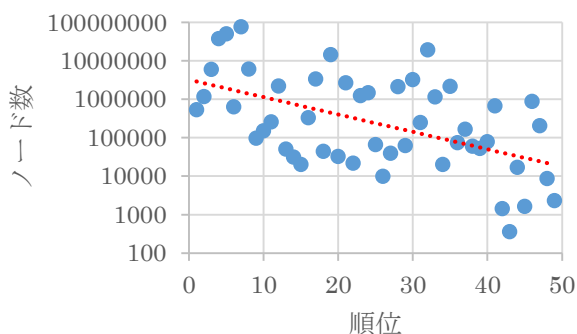


図3 2014年度コンクールの順位とノード数最大

各図の点線は近似線である。便宜上、縦軸が対数軸になっていることに注意する。図を見ると分かるように、それぞれの値は分散しているが、近似線から分かるように順位に従い、それぞれの値は下降傾向にあることが分かる。表2に各年度の順位と各探索指標の相関について調べた結果を示す。

表2 順位と各指標の相関係数 R

年度	証明数最大	反証数最大	ノード数最大
2013(7手)	0.0633	-0.1703	-0.1004
2007	-0.0445	-0.3452	-0.3056
2012	-0.1665	-0.3586	-0.2854
2014	-0.1570	-0.4992	-0.3589

2013年度が7手詰め、それ以外は9手詰めのコンクールである。それぞれの値は順位に対する各指標の相関係数であり、それぞれ負の相関を持っていることが分かる。すなわち、評価の高い問題は大きな探索指標の値を持っていることになる。それぞれを比較してみると、証明数最大が最も相関が低く、3つの指標の中では特に順位との関係を示せていない。反証数最大、及びノード数最大は小さいながらも一定の相関係数を持っており、多少の順位との関連を示せている。特に反証数は全年度において最も強い相関を持っていた。7手詰めと9手詰めの違いとしては、7手詰めコンクールの方が、全体を通して相関が少ない。ただし、どの相関も弱く、順位との関係性を示すことができたとは言いがたい。これは、難易度以外の要素が大きく評価され高評価に繋がった問題がある場合、今回の探索指標には数値として現れ難いためである。難易度についてのみ見たい場合、細かく順位との相関を見るとノイズが多く、正確に見ることは難しい。そこで、各順位を10位ごとにグループ化し、全問題をランクとしてクラス分けした後、それぞれの中央値について比較する。これにより、コンクールの評価と難易度の関係についての全体傾向を分析する。

それぞれの年度毎のランクにおける探索指標の中央値、及びランクと指標の相関係数を表3～表6に示す。

表3 2013年度(7手詰め)コンクールのランクと中央値

ランク	証明数	反証数	ノード数
1	65.5	2975.5	923727.5
2	10.5	1689.0	88262.5
3	12.0	1256.0	179185.0
4	21.5	1689.0	211931.0
5	3.0	226.0	2921.0
相関	-0.723	-0.880	-0.737

表4 2007年度コンクールのランクと中央値

ランク	証明数	反証数	ノード数
1	26.0	9921.0	1024955.0
2	28.0	4776.5	814401.0
3	13.0	2660.0	148461.5
4	12.0	3921.0	197406.0
5	7.0	1618.0	58143.0
6	15.0	827.0	180782.0
相関係数	-0.764	-0.880	-0.844

表5 2010年度コンクールのランクと中央値

ランク	証明数	反証数	ノード数
1	25.5	6615.0	1199794.5
2	21.0	2668.0	233510.0
3	16.5	4428.5	267442.0
4	8.0	926.0	45660.0
5	10.5	771.0	20886.0
相関係数	-0.940	-0.859	-0.829

表6 2014年度コンクールのランクと中央値

ランク	証明数	反証数	ノード数
1	43.5	14137.5	3586076.0
2	16.0	2388.0	257515.0
3	10.0	1423.5	63915.0
4	23.5	3426.5	206623.0
5	6.5	393.0	12874.5
相関係数	-0.716	-0.749	-0.736

表のランクとは、ランク1ならば1位から10位の問題、ランク2なら11位から20位といった用に全問題をクラス分けしたものである。基本的にクラス5までであり、2007年度は全60題だったため、ランク6まで存在している。証明数、反証数、ノード数はそれぞれ証明数最大、反証数最大、ノード数最大の各ランクにおける中央値を表す。相関係数はこれら中央値とランクについての数値である。まず全体を通して相関係数がどれも高い値を示していることが分かる。問題ごとに調べた場合に比べ、ランクを使用して調べ

た場合では、コンクールにおける評価の傾向を見ることができる。従って、コンクールでは探索指標の大きさが、評価において重要視される傾向にあると分かる。

次にそれぞれの指標について見る。まず証明数は、2010年度以外では最も低い相関を示しており、ノード数よりも低い相関である。ノード数という全体の難易度から、証明数という攻め手側の難易度を分けて考えることは、あまり良い結果にならない。そこで反証数について注目すると、全てのコンクールにおいてノード数よりも良い相関を示していることが分かる。反証数とは、受け手側の変化を解くための難易度である。従って短編詰め将棋コンクールの評価傾向として、変化の難しい問題が評価される傾向にあることが分かった。

最後に表のデータを取りまとめたグラフを示す。ここでは、例として2014年度のグラフを図4～図6に示す。

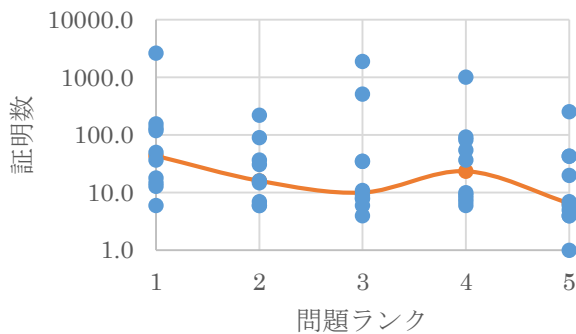


図4 ランク毎の最大証明数分布と中央値

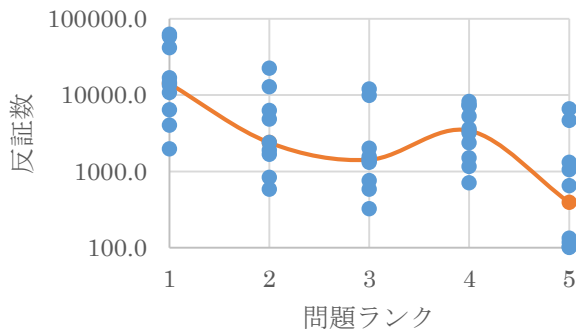


図5 ランク毎の最大反証数分布と中央値

グラフを見ると表の相関係数と同じく、各指標はランク1から5にかけて下降傾向にあることが分かる。他年度のグラフもおおよそ同じでありここでは省略する。しかし中央のランク2から4にかけては相関に不安定な部分があることがグラフからはっきり分かる。他年度においても同様の結果が得られている。評価の高い問題及び、低い問題については探索指標の大小が強く表れるが、中間のランクは

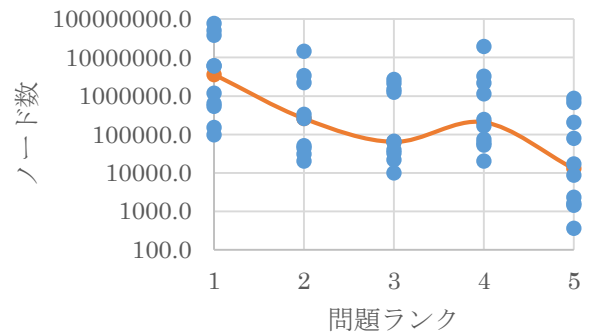


図6 ランク毎の最大ノード数分布と中央値

不安定である。このことから、難易度が非常に難しい、または簡単な問題は、評価の善し悪しに繋がりがやすい一方、中間のランクに位置するような同程度の難易度を持つ問題においては、そのわずかな差をもった評価は難しく、難易度以外の要素が評価されていると推察できる。

4. 詰め問題自動生成への応用

これまでに得られた知見を基に Shape Keeping Technique [6]を用いた詰問題の自動創作に取り組む。Shape Keeping Technique はあらかじめ用意された詰将棋問題の初形を利用し、その駒配置を利用し、駒の向きや種類だけを変更することで詰め問題を自動創作する。この時、創作上の基準を与えることで様々な詰め問題が生成できる。上述した実験結果より、反証数の大きい問題は良問題となりやすいという結論が得られたため、反証数が大きくなるように駒の向きや種類を変更させる。ここでは詰め将棋パラダイス誌において2013年度に実施された7手詰めコンテストの1位受賞作品を利用し、新たな詰め問題の生成に取り組んだ。図7に利用した問題を示す。

次に得られた詰め問題を図8に示す。こちらは、余詰めなどの存在により、詰将棋のルールを完全に満たしていないため、単なる詰め問題である。オリジナル問題の最大反証数が3192に対して、こちらは57660もあるという特徴を有する。

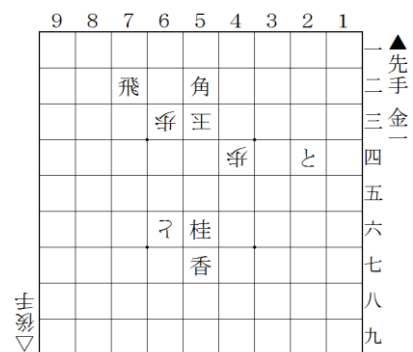


図7 オリジナルの詰め将棋問題 (2013年度 詰め将棋パラダイス誌 7手詰めコンテスト1位 武島宏明氏作)

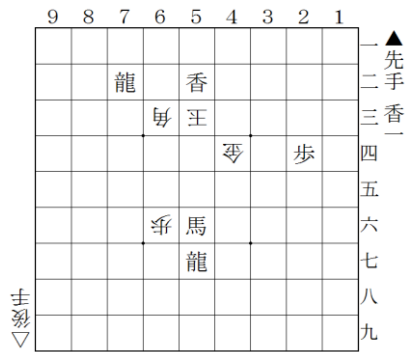
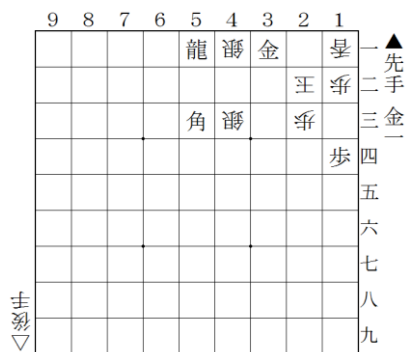


図8 自動生成された詰め問題

反証数の高い問題はその複雑さにより、詰将棋のルールを満たすように局面を変更することが難しい。上記の例では、反証数を可能な限り高くした問題を用意し、その後、人手によってルールを満たすよう手を加えようとしたが、うまくいかなかった。そこで、もう少し簡単な局面を用意し、詰め問題を生成した後、ルールを満たすよう盤面に手を加える実験を行った。図9に用意した局面を示す。

図9 オリジナルの詰将棋 (明解5手詰 39題
北村昌男 将棋世界5月号付録 問題番号05)

Shape Keeping Technique を利用した後、人の手を加え、詰将棋のルールを満たした局面を図10に示す。オリジナルでは反証数最大が398だったのに対し、自動創作された問題では1527となった。詰め手数が変わっているため、簡単に比較することはできないが、今後このように創作された問題が元の問題よりも高い評価を得られるのかどうかについて調査することが今後の課題である。

5. まとめ

本研究を通して得られた知見を以下にまとめる。

- 詰め問題における証明数・反証数は、攻め方と受け方における難易度に相当する事を再確認した。

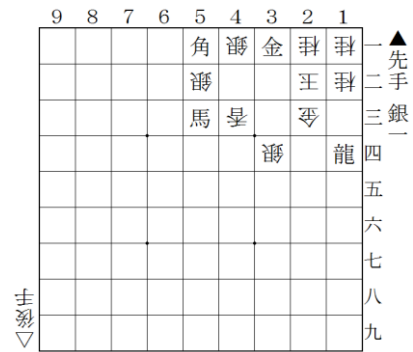


図10 自動創作に手を加えた詰将棋

- コンクールの順位と探索指標は密接に関係する。
- コンクールにおいて難易度という評価要素は重要視される傾向にある。
- 短手数の詰将棋の評価において、受け方の難易度が重要視されている可能性がある。
- 中間の順位において、難易度以外の評価要素が重要視されている可能性がある。
- 最大反証数によって、評価の非常に高い問題と低い問題とある程度見分けることが可能である。

先行研究として挙げた小山ら[2]の成果では評価の高い問題は探索局面数が多いという結果が得られていた。また先行研究[4]では証明数の大きい問題が高い評価となる結論を得ていた。今回、反証数の重要性を発見したことで既存研究をより発展させることができたと考える。

これらの知見から、コンピュータの探索によって得られる反証数などの指標は探索効率化以外の利用価値があり、詰将棋の感性評価に応用できる可能性があるという感触を得た。また今後作成した問題を詰将棋へと直し、評価してもらうことも検討している。

参考文献

- 1) L. V. Allis, M. van der Meulen, and H. J. van den Herik. Proof-number search. *Artificial Intelligence*, 66(1):91-124, 1994.
- 2) 小山 謙二, 河野 泰人. 名作詰め将棋における感性の定量的評価, 情報処理学会論文誌, 35(11):2338-2346, 1994
- 3) 詰め将棋パラダイス, 2014年, 3月号, pp. 47-68
- 4) 石飛 太一, 飯田弘之. 詰将棋問題の感性評価と証明数に関する考察, *Game Programming Workshop in Japan 2012*, pages 163-166, Hakone, Japan, 2012.
- 5) Azlan Iqbal. Is Aesthetics Computable? *ICGA Journal* 29(1): 32-39 (2006)
- 6) T. Ishitobi, A. Cincotti, H. Iida, Shape-Keeping Technique and Its Application to Checkmate Problem Composition, *Artificial Intelligence and Game Aesthetics: the 2013 AIIDE Workshop (WS-13-19)*, Boston, MA, 2013/10/14
- 7) 将棋世界5月号付録, 明解5手詰め39題 北村昌男, 1999
- 8) Mark Winnands, *Informed Search in Complex Games*. Ph.D. thesis, Universiteit Maastricht, Maastricht, The Netherlands, 2004.

●参考資料

データ例として 2014 年度の詰め将棋短編コンクール（9 手詰め）を添付する.

問題番号	証明数最大	反証数最大	最大ノード数	ランク	順位
3	103	133046	59584675	1	1
24	7	1099	28127	1	2
11	26	31426	2859731	1	3
34	91	7159	2423672	1	3
27	20	6605	509818	1	5
5	25	1899	182112	1	6
8	8	5368	184902	1	7
41	50	16986	3437889	1	8
46	82	6625	1889771	1	9
36	2	283	2370	1	10
25	25	22304	1807901	2	11
22	35	2530	362912	2	12
44	8	2806	89334	2	13
23	3	611	7381	2	14
17	5	273	4550	2	15
7	10	863	38415	2	16
39	42	1429	207099	2	16
31	17	3979	259921	2	18
29	108	20148	7841034	2	19
12	240	11277	8890669	2	20
32	10	4326	157264	3	21
26	11	27354	951660	3	22
45	2	299	2281	3	23
21	90	11187	3504103	3	24
15	105	23358	10029857	3	25
14	16	4531	301552	3	26
1	17	3967	233332	3	27
33	5	1834	27396	3	28
47	163	195	67242	3	29
9	143	5763	3474269	3	30
20	6	2457	49488	4	31
18	8	1943	62442	4	32
37	9	6787	270451	4	33
42	3	784	8475	4	34
30	9	330	12380	4	35
48	10	926	45660	4	36
16	5	471	7015	4	38
38	114	1333	843163	4	39
19	6	306	4623	4	40
50	5	666	15289	5	42
4	5	1221	26483	5	43
43	21	156	9777	5	45
2	19	201	11572	5	46
49	11	2493	118088	5	47
10	11	5389	258143	5	48
40	10	876	32751	5	49
13	4	523	6749	5	50