

Further developing a word mark similarity measurement framework

by David Barnett

Part 1 – Testing the algorithm

Introduction

In my introductory article on mark similarity measurement¹, I presented an initial version of a suggested algorithm for objectively quantifying the degree of similarity of a pair of word marks. Similarity assessment is a key element of the decision-making process in many trademark disputes, relating to the likelihood of competing marks generating confusion with each other.

The methodology proposed in the previous study makes use of four calculated components: two of these² generate scores (between 0 and 100) indicating the degree of *visual* (i.e. spelling) similarity; the other two³ quantify the degree of *aural* (i.e. pronunciation) similarity. These two similarity types can therefore be quantified *separately*, or can be combined to create an *overall similarity metric (S)* for the two marks⁴. The basic metrics take no account of the *meaning* of the terms (i.e. *conceptual* similarity) or their level of distinctiveness (which is more relevant to the determination of *likelihood of confusion*), and also ignore any associated classes of goods and services, any logos, imagery or fonts (and the *case* (i.e. upper vs lower) of the marks).

In this first part of the follow-up, I compare the outputs of the similarity metric with those from other contexts in which similarity is assessed. Throughout the article, I use the term ‘inconsistency’ to refer to any deviation from the scenario in which a *fixed* ‘decision tree’ (i.e. the situation where, *all other factors being equal*, an objectively equivalent input will produce the same output) can be seen to have been applied. It is important to note, however, that in real dispute cases, the (manual) analysis will be highly nuanced and following established practice, so different decisions may (correctly) be reached in similar cases. Specific details of individual cases are not, however, considered further in this analysis.

Analysis and testing – Case study 1: Recent UK trademark dispute case decisions

As an initial test of the meaningfulness of the similarity metrics, it is informative to compare their output against the point-of-law decisions from a series of recent trademark disputes. The dataset utilises only word mark vs word mark cases from UK courts (in part, because some of the algorithms are ‘tuned’ to be most appropriate for English-language content) in the last year, comprising 243 cases. The data⁵ gives, for each case, the decision on whether the pairs of marks were deemed to be similar or different, according to each of four similarity types (aural, visual, conceptual, overall) (see Appendix A), but does not (in the summary overview) give more granular information (e.g. on the assessed *degree* of similarity in each case).

¹ <https://circleid.com/posts/towards-a-quantitative-approach-for-objectively-measuring-the-similarity-of-marks>

² The Python-based *fuzz.ratio* metric (F_{lev}), and Jaro-Winkler similarity (sim_j)

³ The *fuzz.ratio* metric applied to the Soundex (F_{sou}) and NYSIIS (F_{NYSIIS}) phonetic representations of the pair of marks

⁴ The overall *visual* similarity (S_{vis}) can simply be quantified as $S_{vis} = (F_{lev} + sim_j) / 2$; the overall *aural* similarity (S_{aur}) as $S_{aur} = (F_{sou} + F_{NYSIIS}) / 2$; and the overall (total) similarity (S) as $S = (S_{vis} + S_{aur}) / 2$ (or, equivalently, as the mean of the four individual components)

⁵ Taken from the Darts-ip tool (<https://app.darts-ip.com/darts-web/login.jsf>)

The first point to note is that – looking purely at the data as presented, with no deeper analysis of case background or context – the decisions include significant degrees of subjectivity, and the decision-making process is not consistent throughout. Arguably, this is why a more objective approach could add value to the process of decision-making, but it does seem unavoidable that some subjective elements will always be involved.

Examples of the inconsistencies in the decisions across the dataset (beyond just cases where different decisions may have been reached on the *same case* at different times) are outlined below.

- Some pairs of marks which are clearly different if viewed *in their entirety* have been assessed to be similar, suggesting that the main consideration in these cases has been on (only) the primary distinctive element of the mark. Examples include: *Dose & Co. vs Dose Labs* ('Dose'); *Catapult Vision / Catapult One vs Catapult Consulting* ('Catapult'); *Alpha Boxing vs Alpha Force* ('Alpha'), etc. – though this is not always the case. Furthermore, in some cases where one mark is entirely contained within the other, the pair have been deemed to be similar (e.g. *Amazon vs Amazon Food Trader Ltd*; *Very vs Veryco*; *Life vs ChopLife / Chop Life*; *Carbon vs MyCarbon*; *PXG Pharma vs PXG*; *Bad Boy vs BBCC / Bad Boy Chiller Crew*) and, in some cases, different (in at least either an aural or visual sense) (e.g. *Yoga vs Yoga Man*; *Ghost vs Ghostdancer CBD*; *Purple Computing / Purple vs Purplecube*). Part of the rationale in some of these cases seems to have been the disregarding of non-distinctive terms (e.g. 'Co', 'My', etc.), but this also is not consistent, and requires *a priori* assumptions.
- It might be reasonable to expect it not to be possible for a pair of marks to be similar (overall) if they differ according to at least one of the similarity types, but there are a number of exceptions to this (e.g. *Boss vs Bossvel*, *Goddess vs Godless*, *VFH vs VFHOnline*, *Folium Science vs Folium*, etc. – all deemed conceptually different, but similar overall; *Sky / Sky X vs Ysky* – deemed aurally and visually different, but similar overall). There is also no consistent application of 'rules' here – e.g. *Matters vs M4tter* and *MFlor vs HFlor / H Flor* were both considered to be aurally and visually similar but conceptually different, but the first pair was deemed similar overall and the second pair different.
- There are also a number of other case-specific inconsistencies – for example, in two cases where the contested marks were almost identical between the cases (*X12 / X14 / X16 / X15 / X13 vs vivo X15 / vivo X12 / vivo X14 / vivo X16 / vivo X13* and *6X / 7X / 9X / 8X / 5X vs vivo X6 / vivo X5 / vivo X7 / vivo X9 / vivo X8*), they were found aurally and visually similar in one case, and different in the other.

With these inconsistencies in mind, it is difficult to formulate a test of the similarity algorithm(s) to determine whether it produces an output consistent with the case law. Furthermore, this would arguably not really be the desired behaviour for any such algorithm designed to produce a *consistent, objective* output.

The simplest test is to consider only those cases where both of the contested marks are a single word (to avoid having to making any judgement as to which terms within a multi-word mark are primarily being assessed). There are 90 such cases within the database.

Figures 1 and 2 show the visual and aural similarity scores⁶ (respectively) for the pairs of single-word marks assessed in the case decisions as being similar (shown in green) or different (shown in red) according to the same criterion / similarity type.

⁶ Noting that, in the calculation, any accented characters have been replaced with their non-accented equivalents

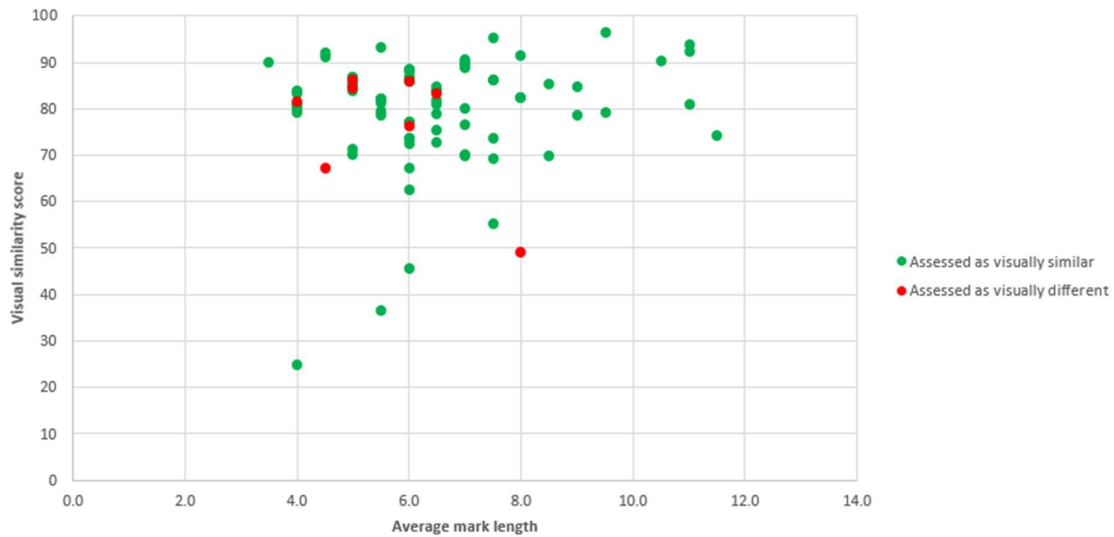


Figure 1: Visual similarity scores for the pairs of single-word marks assessed in the case decisions as being visually similar (green) or different (red), as a function of the average of the lengths of the two marks in each case

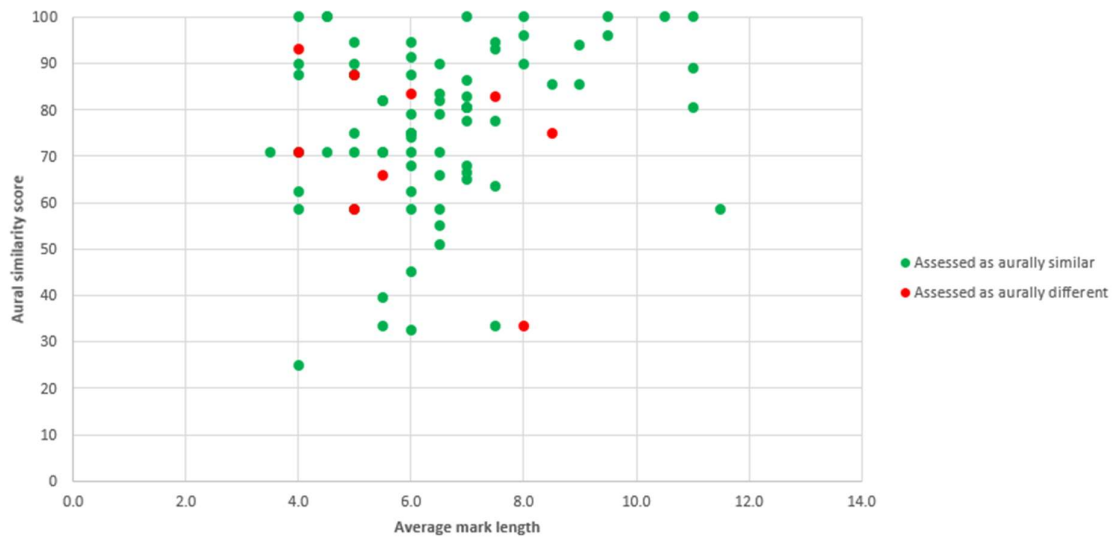


Figure 2: Aural similarity scores for the pairs of single-word marks assessed in the case decisions as being aurally similar (green) or different (red), as a function of the average of the lengths of the two marks in each case

The analysis shows that there is little meaningful correlation between the assessments given in the case decisions and the scores assigned by the algorithm (and this comment is true regardless of the *length* of the marks) – although all pairs assigned a visual similarity score greater than 86, and an aural similarity score greater than 93, were manually deemed to be similar according to the same criterion.

Part of this lack of overall correlation is, however, probably due to the significant subjectivity (and apparent ‘inconsistency’) in the manual decisions. Table 1 shows the scores assigned by the

algorithms in each of the cases (with the pairs ranked by overall similarity score), and the results to appear – by manual inspection – to be meaningful (noting that there are some shortcomings, such as relatively poor reflection of distinct vowel sounds).

Mark 1	Mark 2	Avg. mark length	Visual sim. score	Aural sim. score	Overall sim. score
consiglieri	consigliera	11.0	94	100	97
fashiongo	fashionego	9.5	97	96	96
demiegod	demigods	8.0	92	100	96
1link	link	4.5	91	100	96
intellicare	intelecure	10.5	90	100	95
lovello	lovelle	7.0	90	100	95
configon	configo	7.5	95	93	94
chooey	chooee	6.0	88	100	94
asos	asas	4.0	81	100	90
prinker	prink	6.0	89	92	90
nutravita	nootrovita	9.5	79	100	90
resolution	resolute	9.0	85	94	89
naturli'	natureal	8.0	83	96	89
gobox	g-box	5.0	84	95	89
prinz	prinse	5.5	81	95	88
realme	realmz	6.0	88	88	88
cintra	citra	5.5	93	82	88
testex	test-x	6.0	88	87	87
curve	crv	4.0	82	93	87
energeo	enerjo	6.5	84	90	87
pyra	prya	4.0	84	90	87
kramer	cramer	6.0	86	88	87
billionaire	zillionaire	11.0	92	81	86
vidas	vidya	5.0	85	88	86
yorxs	yorks	5.0	85	88	86
burgerme	burgerly	8.0	83	90	86
mbet	m-bets	5.0	85	88	86
pikdare	pi-kare	7.0	89	83	86
geneverse	genv3rse	8.5	85	86	85
retaron	retlron	7.0	90	81	85
goddess	godless	7.0	90	81	85
snuggledown	snugglemore	11.0	81	89	85
philips	philzops	7.5	86	83	85
noughty	naughtea	7.5	74	95	84
tygrys	tigris	6.0	74	95	84
george	georgine	7.0	91	78	84
mbfw	mvfw	4.0	80	88	84
hanson	hansol	6.0	88	79	84
zemo	zoomo	4.5	67	100	84
ellesse	elliss	6.5	83	84	83
scaffeze	scaffx	7.0	80	87	83
maplab	maplab.world	9.0	79	86	82
matters	m4tter	6.5	82	82	82
createme	create.	7.5	86	78	82

patter	yatter	6.0	86	78	82
treca	trea	4.5	92	71	82
foltene	foltex	6.5	84	79	81
lucite	luci	5.0	87	75	81
dcsl	dcs	3.5	90	71	81
fransa	fanza	5.5	79	82	80
kelio	kleeo	5.0	70	90	80
zara	zareus	5.0	71	88	79
very	veryco	5.0	87	71	79
saypha	shaype	6.0	74	84	79
carbon	mycarbon	7.0	89	68	78
z-biome	biome	6.0	87	68	77
chef	chefchy	5.5	82	71	77
atma	atmaspa	5.5	82	71	77
rabe	rase	4.0	81	71	76
azure	azurity	6.0	77	74	76
noughty	nouti	6.0	76	75	76
suntech	suntank	7.0	70	81	75
live	vive	4.0	79	71	75
repevax	epvax	6.0	87	63	75
sherco	charco	6.0	72	75	74
cazoo	carkoo	5.5	79	66	73
coversyl	covixyl-v	8.5	70	75	72
zara	zarzar	5.0	86	59	72
hypr	hypernft	6.5	73	71	72
fido	fiio	4.0	81	63	72
pockit	mypocket	7.0	76	67	71
axis	traxis	5.0	84	59	71
ulma	luma	4.0	83	59	71
idee	idee-home	6.5	75	66	71
live	life's	5.0	70	71	71
salio	saliogen	6.5	85	55	70
e-bulli	bullit	6.5	81	59	70
waken	wakeful	6.0	77	59	68
airbnb	airbrick	7.0	70	65	68
resolva	consolva	7.5	69	64	66
glenfiddich	inverfiddich	11.5	74	59	66
hotpatch	patch	6.5	79	51	65
boss	bossvel	5.5	82	40	61
bloo	bluuwash	6.0	46	71	58
vfh	vfhonline	6.0	67	45	56
boss	kissboss	6.0	63	33	48
swift	microswift	7.5	55	34	44
rolex	dermarollex	8.0	49	34	41
sane	cbdsane	5.5	37	34	35
md	intimd	4.0	25	25	25

Table 1: Visual, aural and overall similarity scores (as assigned by the algorithm(s)) for the pairs of single-word marks

Analysis and testing – Case study 2: Trademark similarity searches

As a second validation, it is informative to compare the similarity score(s) as described above (hereafter described as the '**similarity score(s)**') against those given as outputs from a trademark watching service. As an example, I consider a set of around 1,800 findings from watches being run for around 30 distinct marks. The trademark watching service reports on identified third-party trademark applications deemed in each case to be similar to the mark being watched, and returns a similarity measurement (hereafter described as the '**TM-watch similarity measurement**') (expressed as a percentage) for the pair of marks, also generated via a proprietary algorithm. This algorithm takes into account a number of factors, including numbers of letters in common and a series of alphabetical, phonetic and grammatical rules⁷.

For simplicity:

- I consider only trademark watches where the watched mark is a single-word word mark.
- I focus only on watches monitoring *all* product and service classes, to ensure that the similarity score output by the service reflects *only* the similarity of the word marks, and *not* the degree of overlap of the classes.
- I exclude any results where the identified trademark contains non-English characters (approx. 300 cases).
- I convert all marks (watched *and* identified) to lower-case text.

In theory (if both the similarity score and the TM-watch similarity measurement algorithms work well), we would expect their values (for the same pairs of marks) to correlate well with each other. However (as for the case decisions considered in Case Study 1), there do seem to be some inconsistencies in the TM-watch similarity measurement algorithm. For example:

- In some cases, *identical* matches, or cases where the identified mark is identical to the watched mark plus just an additional word, are given low TM-watch similarity measurement values (~5%).
- In many cases, identified marks *containing* the watched mark are allocated very high values, regardless of the length of the identified mark and the degree to which the watched mark may feature just as an unrelated sub-string within it. This may be problematic / non-meaningful if (for example) the identified mark is a very long / multi-word string, and the corresponding watched mark is only (say) a two-character string contained within one of the words.

However, other apparent inconsistencies (such as instances of the same type of *visual* variation – e.g. the replacement of one character in a two-character mark – not always being allocated the same TM-watch similarity measurement value) *might* be intended behaviour (e.g. being intended to reflect the effect on *pronunciation* – i.e. *aural* similarity / difference). Admittedly, the full details of the algorithm used by the trademark watching service are not publicly available, and may include additional sophisticated elements which impact on the outputs.

Overall, it will be instructive to assess whether the similarity score is able to provide a better means of sorting (by potential relevance) the findings from the watch service, meaning that it could therefore be used to 'post-process' the watch-service data and build efficiencies into the review process.

⁷ Trademark watch service provider, *pers. comm.*, 09-Aug-2024

The figures below show the relationship between the TM-watch similarity measurement value for each pair of marks and the overall similarity score (Figure 3) and the score reflecting just visual (i.e. spelling only) similarity, for the same pairs.

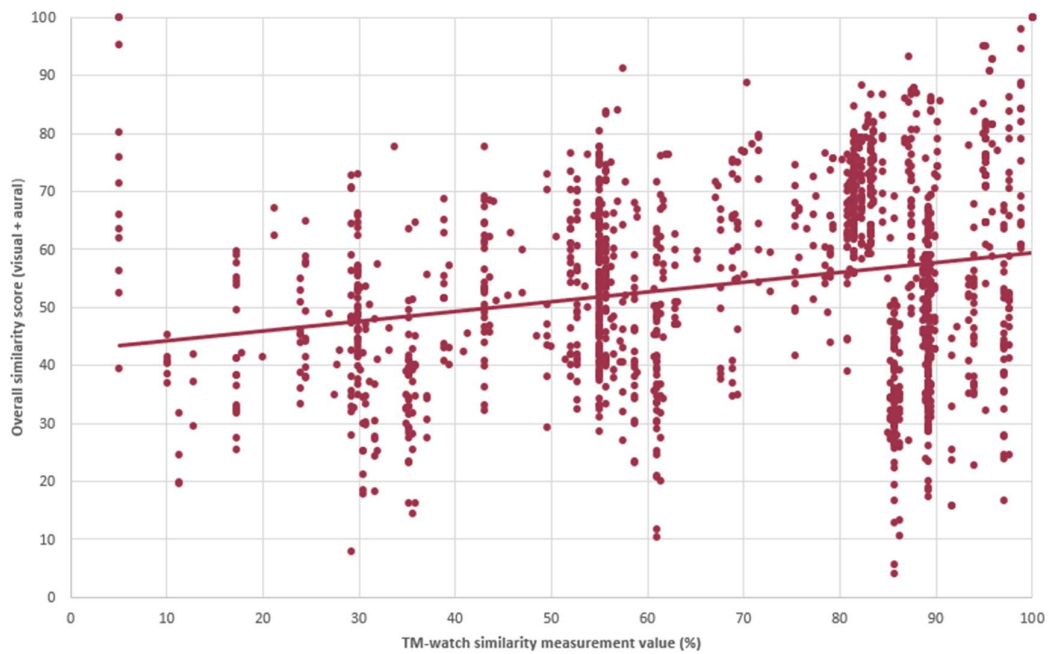


Figure 3: The relationship between TM-watch similarity measurement value and overall similarity score (i.e. considering both spelling and pronunciation) for the (~1,500) pairs of marks in the dataset

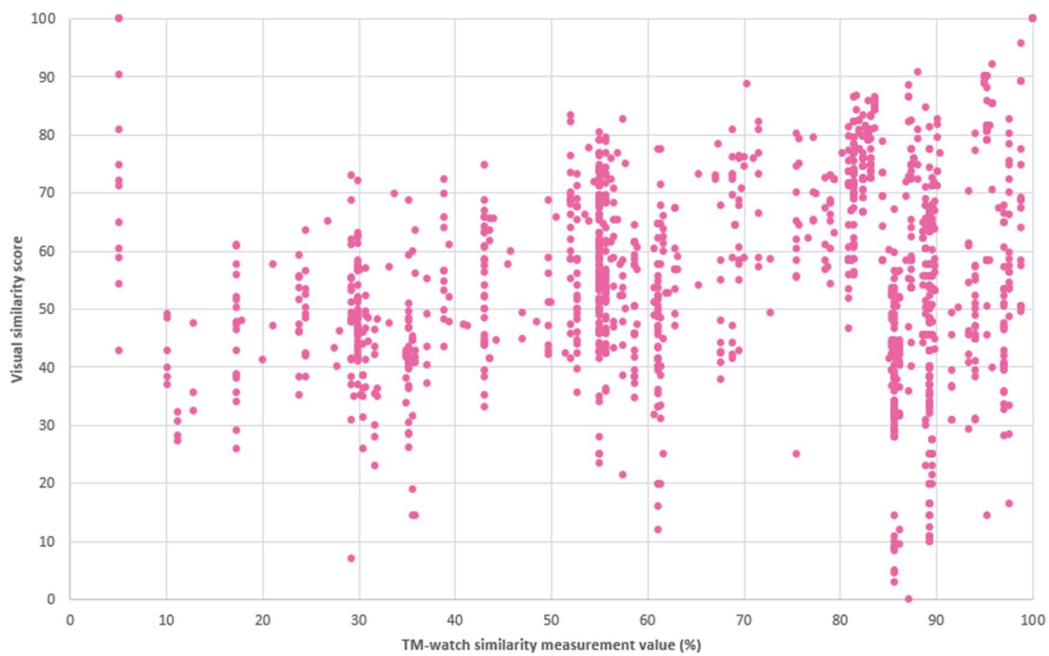


Figure 4: The relationship between TM-watch similarity measurement value and visual similarity score (i.e. considering spelling only) for the (~1,500) pairs of marks in the dataset

The data shows little meaningful correlation between the two metrics (with a weak positive correlation (corr. coefficient = +0.233) for the overall similarity score). In part, this seems to be because there is a lot of 'noise' / inconsistency in the TM-watch similarity measurement value (noting, for example, the anomalous cluster of pairs of marks assigned a very low value of ~5%).

The behaviour of the algorithms, and the relationship between them, may be easier to understand if we consider only the simpler cases where the identified mark (as well as the watched mark) is also just a single word (rather than a multi-word phrase). This relationship is shown in Figure 5, where the results are also categorised by the length (in characters) of the watched mark (to determine whether one or both algorithms perform better with marks of different lengths).

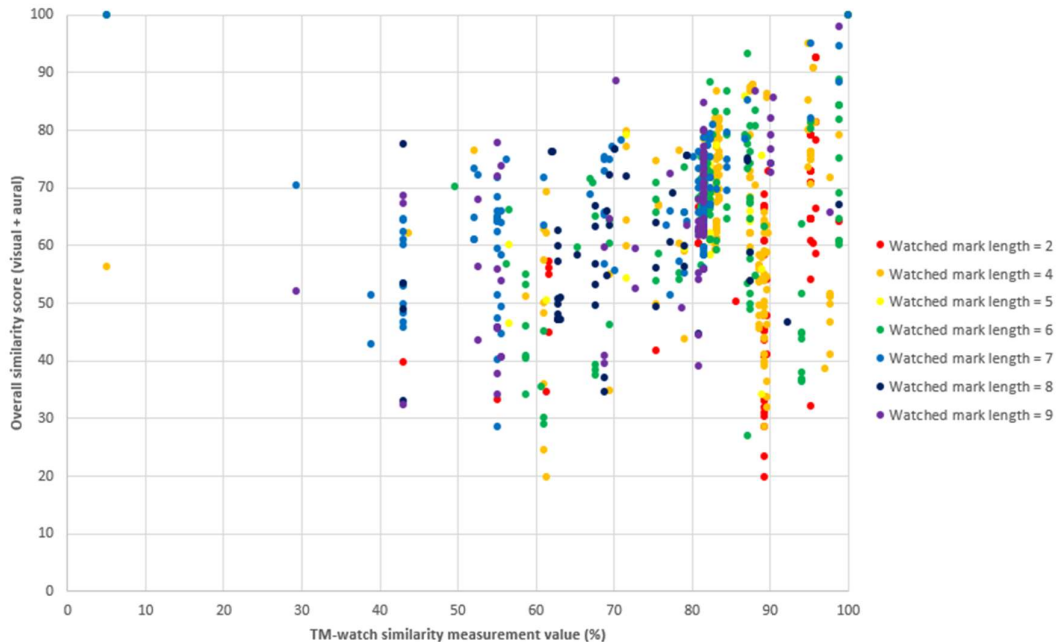


Figure 5: The relationship between TM-watch similarity measurement value and overall similarity score for the pairs of marks in which both the watched and the identified mark are single-word strings

Once again, there is no strong correlation (corr. coefficient = +0.231) and no obvious pattern in the data.

The lack of correlation actually seems to be largely due (as discussed) to the relatively poor performance of the TM-watch similarity measurement algorithm. The similarity score calculation seems to provide a much better measurement of (a subjective manual judgement of) the 'actual' similarity between the pairs of marks. This is illustrated by Tables 2 and 3, showing (in encoded form⁸) (for the full dataset, including multi-word identified marks) the top- and bottom-ranked pairs of marks by overall similarity score – which seems to provide a (subjectively) reasonable assessment of similarity – and the fact that there is a wide range of TM-watch similarity measurement values present within the groups of most-similar and least-similar pairs. It does also seem, therefore, that

⁸ In these tables, all marks are shown in an encoded form, to obfuscate the names of the actual marks being watched, so as to maintain confidentiality. The encoding is carried out by replacing every instance of each letter with the same alternative letter. This thereby generates pseudo-random strings, but in which a visual assessment of similarity across the individual pairs can still be carried out.

the application of the similarity score could be an effective tool in post-processing the results from the trademark watch, and allowing for a quicker review process and identification of genuinely high-risk marks.

Mark 1 (watched)	Mark 2 (identified)	Visual sim. score	Aural sim. score	Overall sim. score	TM-watch sim. meas. value
jddi	jddi	100	100	100	100
dgwf	dgwf	100	100	100	100
rjoct	rjoct	100	100	100	100
fggwoc	fggwoc	100	100	100	5
fccjfg	fccjfg	100	100	100	100
qgffiwo	qgffiwo	100	100	100	5
dwbgcwi	dwbgcwi	100	100	100	100
bfigegiri	bfigegir	96	100	98	99
ajwbgeioj	ajwbg & eioj	91	100	95	5
dwbgcwi	dwbgcii	90	100	95	95
dgwf	dgw	90	100	95	95
dwbgcwi	dwjgcwi	89	100	95	99
dgjaaj	dgja	87	100	93	87
ft	ftj	86	100	93	96
ft	fti	86	100	93	96
qgffiwo	qgff fwof	83	100	91	57
dgwf	dgwc	82	100	91	95
dgwf	dgww	82	100	91	95
rjbjci	rjb	78	100	89	99
bfigegiri	bfigegirft	89	89	89	70
fccjfg	fccjii	77	100	88	82
dwbgcwi	dwpgcwi	89	88	88	99
dgwf	dgwci	76	100	88	88
dgwf	dgwcg	76	100	88	88
dwaf	dwwac	75	100	87	87
bfigegiri	jbfigegir	91	83	87	88
icge	ifgef	74	100	87	83
dwaf	dfafg	74	100	87	87
rjbjci	rfjffi	74	100	87	84
qgffiwo	qgffiwo ffjc	83	90	86	98
dgwf	4dgwc	73	100	86	89
rjoct	rcojtc	72	100	86	87
ajwbgeioj	ajwbgefwb	77	95	86	90
dgwf	dgwc21	71	100	86	89
dgwf	dgwc 3	71	100	86	89

Table 2: Top-ranked pairs of marks (shown in encoded form) by overall similarity score

Mark 1 (watched)	Mark 2 (identified)	Visual sim. score	Aural sim. score	Overall sim. score	TM-watch sim. meas. value
rjbjci	igt tjdc tjggfc	41	6	23	35
qp	op ocawfwjid pwfjww jdtce	29	18	23	86
fccjfg	tcaiw dfqf fcofg 77	40	7	23	35
rjbjci	ifiwg tjdc	37	9	23	59
fccjfg	fdwdbjfo fwjf	31	15	23	94
qp	qf ejfdbjfg jdiwofdbc bifafdw	29	16	22	86
rjbjci	wdtco wgc tjdc	35	8	21	30
rjbjci	iwffco tjdc	33	9	21	61
rjbjci	iaojdq tjdc	33	9	21	61
qp	iwfo-qf	11	29	20	89
dgwf	ridtcdgwc	20	20	20	61
rjbjci	qgidfgtji qowai tjbjff	31	9	20	11
gcggidcc	wfffd twf ichigj	28	11	20	11
qp	jd tj qi	10	29	20	86
dgwf	rjcdco dgww	20	18	19	89
qp	foiff qw	10	27	19	89
rjbjci	cfdfbc cwoiacfd tjdc	26	11	18	30
icge	fwigci wcf icdifij jcdjffwfd & jcawfifd wcg ficgj jdtidcijf	30	7	18	32
rjbjci	biotjdf fe fjbjci & bctjbjci ffojpwcojf dfo ffojt-fe	31	5	18	30
qp	iifii fp	10	25	18	89
qp	gjec wi qi	9	25	17	86
ft	aoie c. & cggjc'i fjttgc ibgiig ie ibjcdbc ft fiwof aco fiacof	28	5	17	97
rjbjci	oiwg afgcwwcd wdt tjcgci fcgo	29	4	16	35
ft	gif bid fttw	15	18	16	36
jddi	wcoofj wcoofj igt tjdc	31	1	16	92
jddi	wcoofj wcoofj igt tjdc	31	1	16	92
dgwf	fcejbi bjww dgwci	15	15	15	35
ft	wgc iwaocfc tft	12	15	13	86
qp	ffjc wiwo qf gfaaw	5	21	13	86
rjbjci	gfofidw tjdc	16	8	12	61
ft	wgc iwaocfc tft.bif	10	12	11	86
rjbjci	dofdt tjdcq	12	9	11	61
ft	idc-bgjbj fa	7	9	8	29
qp	fa rrr.fwwi-afowi.qo	5	7	6	86
qp	gcw jw jd	3	5	4	86

Table 3: Bottom-ranked pairs of marks (shown in encoded form) by overall similarity score

Conclusions

The analysis does not show any strong correlation between the overall similarity score and the other two measures of similarity considered in this article – namely, (a) the subjective case decisions from recent UK trademark disputes, and (b) the more deterministic TM-watch similarity measurement algorithm. However, a significant factor in this lack of correlation seems to be the apparent inconsistencies (i.e. non-determinism) in these other types of assessment, rather than reflecting significant shortcomings in the similarity score algorithm in itself.

The similarity score does provide what appears (subjectively) to provide a meaningful measure of ‘actual’ similarity, meaning its application could lend it to uses in post-processing data drawn from a range of sources (such as trademark watching services) and potentially in helping to allow greater consistency and a more scalable, reproduceable, quantitative and objective basis for dispute resolution and case-law formulation.

There are, however, a number of enhancements still to be made to the algorithm. The current version takes no account of the meaning of terms (i.e. *conceptual* similarity) or of other factors such as the similarity of associated goods and services classes, or the level of distinctiveness of the marks (which is perhaps more relevant specifically to a determination of likelihood of confusion, rather than simply similarity).

Finally, it would certainly not be reasonable to suggest that this type of formulaic approach is preferable to a full manual assessment, taking into account the wide range of additional nuanced and subjective factors which are relevant in a dispute case. Rather, I am suggesting that these types of quantitative algorithm can be used as (consistent, reproduceable) ‘tools’ to be utilised as part of the overall similarity assessment frameworks relevant to the resolution of intellectual property disputes.

Part 2 – Subsequences and substrings

Introduction

Following on from the analysis presented in Part 1, it is useful to consider some additional characteristics of word marks which can further be used in a determination of similarity. In this follow-up, I again consider the dataset of recent UK trademark disputes (Case study 1 of Part 1), though now considering *all* 243 cases, rather than just the single-word marks⁹.

In Part 1, I noted that some of the case decisions appeared to have been influenced, at least in part, by a subjective assessment of the elements of the marks which were deemed to be the ‘distinctive’ parts (e.g. ‘Dose’ in *Dose & Co. vs Dose Labs*).

In order to try and build a formulation to try and address this point, it is useful to consider characteristics of the subsequences and substrings (i.e. groups of characters) contained within the marks.

First steps towards formulating a methodology

One characteristic often considered in analyses of these types is the concept of the **longest common subsequence** (hereafter referred to as the ‘**LCSSQ**’) – as also referenced in my original study on mark similarity – between a pair of strings. This is defined as the longest set of characters which appears in the same *order* (though not necessarily in *consecutive positions*) in both strings (i.e. generally (through not necessarily) **non-contiguous characters**). This parameter is also the basis of the Ratcliff-Obershelp similarity metric^{10,11}, defined as twice the length of the LCSSQ divided by the sum of the lengths of the two strings (giving a value between 0 and 1) – essentially, the length of the LCSSQ expressed as a proportion of the average of the length of the two strings.

Whilst the LCSSQ is a useful characteristic – particularly in cases of alternative spellings, such as *Sole Sister vs Soulsistar*, where it is equal to ‘solsistr’ – it can lead to some misleading results. Consider, for example, the case of *Dreams vs Dream Big Make Dua Move Mountains*; the LCSSQ in this case is (the apparent full word) ‘dreams’, although the ‘s’ identified in the latter case is actually the character at the end of ‘mountains’. Accordingly, it can be instructive also to consider the **longest common substring** (hereafter referred to as the ‘**LCSST**’), in which the characters must appear in both strings in the same order *and* in consecutive positions (i.e. **contiguous characters**) (which, in the case of the above marks is ‘dream’).

Having identified the LCSSQ and LCSST, useful insights can be obtained by determining the ‘remainders’ of each string, i.e. the parts which remain after the (first instance of) the common sub-elements are removed. The analysis of the full set of pairs of disputed marks is shown in Appendix B. Arguably a better metric than (just) Ratcliff-Obershelp similarity is to also calculate an equivalent score, using the LCSSTs (rather than the LCSSQs), and then calculate the *average* of these two scores

⁹ In creating a ‘clean’ dataset, the following modifications to the ‘raw’ data have been made:

- In cases where multiple distinct marks have been cited in a case, just one has been considered. The selected mark is generally the ‘simplest’ of the set and/or the one (subjectively) most similar to the other disputed mark or, all other factors being equal, the first mark in the list
- All marks have been converted to lower-case characters (i.e. case is disregarded)
- All accented characters have been replaced by their non-accented equivalents
- All ampersands (&) appearing as space-separated words have been replaced by the word ‘and’

¹⁰ <https://www.drdoobs.com/database/pattern-matching-the-gestalt-approach/184407970?pgno=5>

¹¹ <https://xlinux.nist.gov/dads/HTML/ratcliffObershelp.html>

(to produce a metric which is here termed the ‘modified Ratcliff-Obershelp similarity score’); the data in Appendix B is sorted by this score.

Invariably, the data produced by this purely mathematical approach may require some ‘cleaning up’ prior to any further analysis, as it will generate some apparent anomalies. For example, in *Holiday People vs The Holiday People*, the LCSSQ is ‘holiday people’, which actually generates a remainder for the second mark of ‘te h’ – since the extraction algorithm removes the LCSSQ’s initial ‘h’ from the word ‘the’ (the first place it appears in the overall string) rather than the word ‘holiday’ – and so the remainder should instead probably be ‘corrected’ to ‘the’. Similarly, for *PT Powerpod vs Powerpod*, the remainder for the first mark is extracted as ‘t p’, rather than the ‘pt’ it arguably ‘should’ be.

Taking the above caveat on board, the analysis does provide some useful insights. The most obvious (mathematically trivial) case is where the two marks are identical, in which case both the remainders are non-existent (blank / null / empty), and the modified Ratcliff-Obershelp similarity score is 1 (as for the top four examples in the Table in Appendix B). Cases where *one* remainder is blank imply that one mark is *contained* within the other.

Beyond this, the LCSSQ / LCSST analysis does provide a framework making it possible to conveniently review the elements which the marks have in common. For example, in the *Dose & Co. vs Dose Labs* case, the remainders after extracting the LCSST are, respectively ‘and co.’ and ‘labs’. When assessing overall similarity, it might be reasonable to disregard remainders which are extremely non-distinctive (such as ‘and co.’); arguably, however, ‘labs’ might be deemed to be a more distinctive / descriptive term, which might mean that the overall assessed level of similarity should be lower than if the remainders for *both* marks were less distinctive terms. Other non-distinctive terms in the dataset include ‘the’, ‘my’, and so on.

Attempting to quantify the degree of distinctiveness (or non-distinctiveness) of the terms in the mark remainders is a more difficult prospect. In the original study, the numbers of results returned by search engines were explored as a proxy for this parameter. However, in these types of analysis, this approach may not be very productive, as even more distinctive terms such as ‘labs’ would generate large numbers of results. Perhaps a better question is whether the remainder keywords overlap significantly with the *goods and services* of the other mark (e.g. if *Dose & Co.* offered laboratory services, the assessed degree of similarity (by virtue of the use of the term ‘labs’ in the other mark) should potentially be assessed as being much higher).

Similar comments might also apply to other pairs which differ only in a keyword relating to business area. Examples in the dataset include *Stones vs Stone Brewing*, *Whitehorse Liquidity Partners vs Whitehorse*, *PXG vs PXG Pharma*, and *Unity vs Unity Real Estate Ltd*. It might be reasonable to expect that similarity should be assessed to be lower if the remainders (i.e. the differences between the marks) are both distinctive and relate to *different* business areas or brand descriptors – such as *Catapult Vision vs Catapult Consulting* or *Spear & Jackson Predator vs Predator Gutter Vacuum*.

It is essential that consideration of goods and services – and the degree to which these are likely to be familiar to general consumers – should always be incorporated into any overall similarity assessment framework. For example, in *Honda vs Honbike* – objectively not highly similar as *words* – relevant points to consider might be the extent to which the Honda brand is known to be associated with motorbikes, and whether it is well-known enough that even an abbreviated variant (‘Hon’) would evoke a brand association in the mind of the average consumer.

In other cases, similarity between the word types in the remainders might imply an overall greater similarity between the marks – such as *Lemon Perfect vs Peach Perfect*, where the remainders after

the removal of the LCSST are 'lemon' and 'peach' (both fruits), *Glenfiddich vs Inverfiddich*, where the remainders are 'Glen' and 'Inver' (both common terms in Scottish place-names), or *Karmacoin vs Karmacash*.

Another consideration is that if the remainders are very short and/or meaningless (say, single letters), this might also imply a greater degree of similarity between the marks. However, this assertion may also be dependent on the positions of the distinct elements within the strings – for example, 'Consiglieri' and 'Consigliera' might be deemed to be more similar to each other than are 'Lotus' and 'Motus' (both cases differing by a single character) – particularly when factors such as local-language significance (e.g. just a potential difference in gender) are taken into account.

Conclusions

The discussion presented in this article has stopped a long way short of attempting to define a full framework for mark comparison using analysis of LCSSQs and LCSSTs. Nevertheless, an analysis of pairs of marks (as presented in Appendix B) does provide a useful framework for conveniently reviewing the elements that pairs of marks have in common, and those which are distinct from each other – which is key to an overall assessment of similarity. In this analysis, it is useful to consider both LCSSQs and LCSSTs together, as marks which are distinct in differing ways may be more suitable for analysis using one characteristic than the other.

The modified Ratcliff-Obershelp similarity score – which quantifies the size of the common sub-elements as a proportion of the length of the original strings – is also a useful metric in its own right, which could easily be incorporated into an enhanced version of the similarity score presented in the previous studies.

Building on these ideas, it should be possible to build the beginnings of a frameworks of mark similarity comparison utilising sub-element analysis, to augment the ideas presented previously. Such a framework would likely need to incorporate a number of key ideas, such as analysis of the 'remainders' for each of the marks (i.e. the portions which are left when the common elements are removed) and how descriptive they are, how closely they overlap with the goods and services of the other mark, how thematically similar they are to each other, and the positions within the original marks at which they appear.

Part 3 – Phonemizing

Introduction

The initial study on mark similarity measurement utilised two distinct models to generate the *phonetic* representation of word marks¹², from which the degree of (aural) similarity in the pronunciation of pairs of marks could be quantified (by measuring the similarity of the phonetic representations using the Python-based *fuzz.ratio* algorithm¹³). The two models were Soundex, which represents each mark as a four-character string, and NYSIIS (New York State Identification and Intelligence System). However, both of these encodings have certain shortcomings, not least the fact that they poorly handle (or disregard entirely) the vowel sounds within the strings and – in the case of Soundex – do not encode anything beyond the fourth consonant.

In this article, I consider the application of a (Python-based) model (named ‘*Phonemizer*’) for encoding the marks in International Phonetic Alphabet¹⁴ (IPA) format, which better handles the full string in each case and – unlike some other text-to-IPA encoders – can handle *arbitrary* strings, rather than just dictionary terms.

Methodology

The IPA is a means of representing strings phonetically using (mainly) Latin or Greek script¹⁵, but also utilising other special characters, such as a colon-like character denoting that the preceding sound is long and, in some versions, high or low vertical lines denoting the primary and secondary stressed syllables¹⁶.

The *Phonemizer* package is a Python-based tool for converting text strings to IPA format using a back-end text-to-speech software application named *espeak-ng*^{17,18,19,20,21}.

In this article, I again consider the same pairs of marks, taken from previous dispute cases, as utilised in the initial study. The marks are converted to their IPA representations using the *Phonemizer* package, and the phonetic representations again compared using the *fuzz.ratio* algorithm.

Analysis

Table 4 shows the IPA representations of each of the marks, as given by *Phonemizer*, and the corresponding (phonetic) similarity scores for the pairs of marks (compared with the scores obtained using the Soundex and NYSIIS representations, from the initial study).

¹² In addition to other models for quantifying the *spelling-based* (visual) similarity

¹³ <https://pypi.org/project/fuzzywuzzy/>

¹⁴ <https://www.internationalphoneticassociation.org/content/ipa-chart>

¹⁵ https://en.wikipedia.org/wiki/International_Phonetic_Alphabet_chart; a pronunciation guide can be found at: <https://www.vocabulary.com/resources/ipa-pronunciation/>

¹⁶ [https://en.wikipedia.org/wiki/Stress_\(linguistics\)](https://en.wikipedia.org/wiki/Stress_(linguistics))

¹⁷ M. Bernard and H. Titeux (2021). ‘Phonemizer: Text to Phones Transcription for Multiple Languages in Python’, *J. Open Source Software*, 6(68), p.3958.

¹⁸ <https://pypi.org/project/phonemizer/>

¹⁹ <https://github.com/bootphon/phonemizer>

²⁰ <https://bootphon.github.io/phonemizer/install.html>

²¹ <https://github.com/espeak-ng/espeak-ng#espeak-ng-text-to-speech>

Mark 1	Mark 2	Mark 1 (IPA)	Mark 2 (IPA)	Similarity scores		
				IPA	Soundex	NYSIIS
kresco	cresco	kɹɛskɔʊ	kɹɛskɔʊ	100	75	100
casoria	castoria	kæso:ɹiə	kæsto:ɹiə	95	75	91
seiko	seycos	seɪkɔʊ	seɪkɔʊz	93	100	67
starbucks	charbucks	stɑ:ɹbʌks	tʃɑ:ɹbʌks	90	50	77
mahendra	mahindra	mæhɛndɹə	mæhɪndɹə	89	100	100
bacchus	cacchus	bækəs	kækəs	83	75	67
trucool	turcool	tɹu:kɔ:l	tɜ:kɔ:l	82	100	83
lucozade	glucos-aid	lu:kəzeɪd	glu:kɔʊzeɪd	82	50	93
louis vuitton	chewy vuiton	lu:i vju:ɪʔŋ	tʃu:i vju:ɪtən	76	50	71
mdh	mhs	ɛmɹi:ɛɪtʃ	ɛmɛɪtʃɛs	74	75	80
intelec	entelec	ɪntɛlɛkt	ɛntɛlɛk	71	75	80
starbucks	sardarbuksh	stɑ:ɹbʌks	sɑ:ɹdɑ:ɹbʌkʃ	70	75	71
cana	canya	kɑ:nə	kænjə	67	100	100
simoniz	permanize	sɪmənɪz	pɜ:mənəɪz	67	50	62
zirco	cozirc	zɜ:kɔʊ	kɑ:zɜ:k	67	50	60
bisleri	bilseri	bɑ:ɪsɹɪ:ɹi	bɪɪsɹɪ:ɹi	67	75	83
magnavox	multivox	mægnevɑ:ks	mʌltɪvɑ:ks	64	50	75
nike	nuke	nɑ:k	nu:k	60	100	100
lakme	likeme	lækmi	lɑ:kɪ:m	57	100	89
puma	coma	pʊ:mə	kɔʊmə	50	75	67
hpnotiɔ	hopnotic	ɛɪtʃpi:nɔʊɹɪk	həpno:ɹɪk	50	100	80
mcdonalds	mcsweet	mækda:nɛldz	mækswi:t	48	75	43
louboutin	lubov	lɑʊbɑʊtɪn	lu:bɑ:v	33	50	67

Table 4: IPA representations (from *Phonemizer*) and similarity scores (using the *fuzz.ratio* algorithm) for the pairs of marks, compared with the scores using the Soundex and NYSIIS representations

Overall (subjectively!), the IPA-based similarity score seems to perform well in ranking the pairs of marks by aural similarity, and provides a more satisfactory analysis than either of the two other phonetic models explored previously (which is perhaps unsurprising, in view of the shortcomings discussed above).

There are also a number of other specific observations of note:

- The algorithm correctly (in my opinion!) rates the ‘kresco’ and ‘cresco’ marks as phonetically identical, and with the same IPA representation.
- The IPA representation provides a convenient way of comparing the degree of similarity (according to *Phonemize*) between sub-elements of the strings when the primary difference is disregarded; for example, with *Lucozade* vs *Glucos-Aid*, if the initial ‘g’ is removed from the IPA representation of the latter, the remaining strings are *lu:kəzeɪd* and *lu:kɔʊzeɪd*, i.e. differing only in the middle vowel sound.
- In portions of the words which the algorithm deems ‘unreadable’, the phonetic representation conveys a series of letter *names*. For example, ‘mdh’ is expressed as ‘em-dee-aitch’, and ‘hpnotiɔ’ as ‘aitch-pee-notic’. Whilst this may be the desired behaviour in some cases, it may not always be appropriate.

- This particular implementation of IPA is built around American, rather than English, pronunciation – for example, syllables such as ‘vox’ and ‘not’ are encoded using a long ‘ah’ sound (ɑ:), and ‘nuke’ is represented as ‘nooke’ rather than ‘nyooke’. Again, this may not always be appropriate for marks targeting an English audience (although perhaps less of an issue if comparing like-with-like).

One option for handling the above issues wherever they arise is to manually modify the strings, to ensure that they are encoded ‘correctly’ (or simply to modify the IPA representation before calculating the similarity score) – though of course this removes the objectivity of the approach. Nevertheless, there may be cases where this is unavoidable, if it is relatively indisputable that the algorithm has got the encoding ‘wrong’ (based on the – admittedly subjective – intended pronunciation). Examples in the dataset include:

- ‘hypnotic’ – encoded as ‘aitch-pee-notic’, where it would be preferable to modify the mark or directly edit the IPA representation to ensure that it is represented as *həpna:ɾɪk* (in which case the similarity score will be 100)
- ‘likeme’ – this has been encoded as it would be pronounced if intended to be a single readable word (*laɪki:m* – ‘lyekeem’), whereas it is presumably intended to be read as ‘like me’. This can be addressed by re-writing the mark as ‘like-me’, in which case it is encoded as *laɪkmi:*, giving an increased similarity score (compared with ‘lakme’) of 71.

Conclusion

Converting marks to their full IPA representations, as a means of comparing aural (pronunciation) similarity, appears to provide improved performance than using the Soundex or NYSIIS algorithms described previously, and would probably be preferable for inclusion in an improved metric for quantifying the overall similarity of marks.

The *Phonemizer* package offers a convenient method for generating the required phonetic encoding, although there remain some potential issues to be addressed, such as the emphasis on American pronunciation, which may not be appropriate in all cases, and the handling of lower-readability strings, or those marks which (subjectively) appear intended to be read in particular ways. These issues can be addressed by employing manual edits, but this runs the risk of breaching the wholly objective nature of the approach.

Appendix A: Assessments of word mark* similarity in recent UK trademark dispute cases

*Neglecting certain variants which differ only in the case of the characters

s = similar

d = different

? = inconsistent – i.e. different decisions reached at different times

Ref #	Mark 1	Mark 2	Word trademark similarity			
			Aural (pronun.)	Visual (spelling)	Conceptl. (m'ning)	Overall
1	GIZEBRA	DEBRA THE GIZEBRA				s
2	SUNTECH	SunTank	s	s	d	d
3	IBACCY	Biccy Baccy	d	d	d	d
4	JOLLY	JOLLY PECKISH	s	s	s	s
5	DREAM COACH / DREAM BIGGER / DREAMS	Dream Rite	s	s	s	s
6	DOSE & CO.	DOSE LABS	s	s	s	s
7	AMAZON	AMAZON FOOD TRADER LTD	s	s	s	s
8	SKULLCANDY / SKULL-IQ	SKULL GAMING	s	s	d	d
9	PRINKER	Prink	s	s		s
10	3MONKEYS	3 Monkeys Communications	s	s	s	s
11	BOSS	Bossvel	s	s	d	s
12	CAZOO	CARKOO	d	s		d
13	LOTUS	Motus Group UK (<i>and variants</i>)	d	d		d
14	CARTILS	CARTEL DESIGN	s	s	d	d
15	PRINZ	Prinse		s		s
16	SHERCO	CHARCO	s	s		s
17	WATERFORD / WATERFORD TEIREOIR	LADY LOUISA WATERFORD			d	d
18	LIVE	LIFE'S			d	d
19	FLOWERS	FLOWER CAFE / FLOWER DRINKS	s	s	?	?
20	STR8 GO FOR GREAT / STR8	ST8	s	d	d	d
21	myGeneCare / myGeneWisdom / myGeneDiary / myGenePredict / myGeneHelp	MYGENES	s	s	s	s
22	NRJ / ENERGY NRj	NRG	d	d	d	d
23	OYSTER	Oyster and Pop	s	s	s	s
24	SKY	Sky Force	s	s	s	s
25	THE GOOD SCHOOLS GUIDE	The Good School	s	s	s	s
26	NOUGHTY	NAUGHTEA	s	s	s	s
27	GLOW UP / Glow Up: Britain's Next Make-Up Star	glow up: britain's next make-up star	s	s	s	s
28	DEMIEGOD	DEMIGODS	s	s	s	s
29	Retaron	RetIron	s	s		s
30	THIS GIRL CAN	This Girl Came	s	s	d	d
31	GODDESS	GODLESS	s	s	d	s
32	PARIS-MATCH	PARIMATCH / PARIMATCH TECH / PARI MATCH	s	s	s	s
33	ANYTIME FITNESS / ANYTIME HEALTH	ANYTIME PRO	s	s	s	s
34	PIKDARE	PI-KARE	s	s		s
35	Kramer	CRAMER		s	s	s
36	FIDO	FIIO	s	s		s
37	EVOLUTIQ	ESSENTIAL EVOLUTION	d	d	d	d
38	TRECA	TREA	s	s		s
39	EASYJET / EASYGYM / EASYHOTEL / EASYBUS /	Easycosmetic	d	d	d	d

	EASYCAR / easyProperty / EASYCOFFEE					
40	X12 / X14 / X16 / X15 / X13	vivo X15 / vivo X12 / vivo X14 / vivo X16 / vivo X13	s	s	d	d
41	Victoria / victoria Dear World	Victoro	s	s	s	s
42	VIZRT	Vizst TECHNOLOGY / Vizst	s	s	s	s
43	ASOS	ASAS	s	s		s
44	MAKEUP WARDROBE	MAKEUP WARDROBING	s	s	s	s
45	LUSH	Lush Lights	s	s	s	s
46	SIX DAYS	DAY6	s	s	s	s
47	GENEVERSE	GENV3RSE	s	s		s
48	DIAMOND MIST	VAPES BARS DIAMOND / DIAMOND BAR 600 / MAX DIAMOND / DIAMOND MAX / DIAMOND PRO	d	d	d	d
49	EATALIANO	EATalia / EAT-alia	s	s	s	s
50	VFH	VFHOnline	s	s	d	s
51	MBFW	MVFW	s	s		s
52	TYGRYS	TIGRIS	s	s		s
53	Meta Technology / META META / Meta	META PORTAL / META PLATFORMS / META / META QUEST / META HORIZON / META VIEW	s	s	s	s
54	Burgerme	BURGERLY	s	s	d	d
55	SIR / SIR	SIRO	s	s	d	d
56	OYSTER PERPETUAL / PERPETUAL	PERPÉTUEL / PERPETUEL	s	s	s	s
57	Spear & Jackson Predator	Predator Gutter Vacuum	s	s	s	s
58	ULMA	LUMA	s	s		s
59	AZURE	Azurity	s	s	s	s
60	1LINK	LINK	s	s	s	s
61	CATAPULT VISION / CATAPULT ONE	Catapult Consulting	s	s	s	s
62	6X / 7X / 9X / 8X / 5X	vivo X6 / vivo X5 / vivo X7 / vivo X9 / vivo X8	d	d		d
63	ASPREY / Asprey LONDON	DAVE ASPREY	s	s	d	d
64	ENERGEO	ENERJO	s	s	s	s
65	MUTANT	MEGA MUTANT	s	s	s	s
66	PHILIPS	PhilzOps	d	s		d
67	Last Shelter:Survival	Doomsday: Last Survivors	d	d	s	d
68	ACTIVIST	Activist Ingredients / Davines Activist Ingredients	?	?	?	?
69	FiTTIPALDi / FiTTIPALDI	EMERSON FiTTIPALDI / eFittipaldi / FiTTIPALDI AUTOMOBILI	s	s	s	s
70	SWIFT	MicroSwift	s	s	s	s
71	ArmaLight / ArmaGel	ARMATHERM	s	s	d	s
72	FlexoLid	kp FlexiLid	s	s		s
73	VERY	VERYCO	s	s	s	s
74	ZARA	ZARZAR	s	d		d
75	VAULT IP / VAULT INTELLECTUAL PROPERTY	BRANDVAULT	d	d	d	d
76	Alpha Boxing	ALPHA FORCE	s	s	s	s
77	LEMON PERFECT	PEACH PERFECT	s	s	s	s
78	life	ChopLife / Chop Life	s	s	s	s
79	RYZEN / AMD RYZEN	RYZEUP / RyzeUp	s	s	d	s
80	IPING 2.0 / PING	pingNpay	d	s	s	s
81	NUTRAVITA	Nootrovita	s	s	s	s
82	PUSHER	Pushers Only	s	s	s	s
83	Zemo	ZOOMO	s	d	d	d
84	WEAR THE CHANGE	WEAR THE FUTURE	s	s	s	s

85	IntelliCare	Intelecare	s	s	s	s
86	PT Powerpod / PT Powerpods	POWERPOD	s	s		s
87	AGRHO S-ROX / AGRHO	agro S	s	s		s
88	SIZZLING FORTUNES / SIZZLING MOON / SIZZLING REELS / SIZZLING HOT	SIZZLING BELLS / SIZZLING MOON / SIZZLING REELS / SIZZLING KINGDOM	s	s	s	s
89	FOLTENE	FOLTEX	s	s		s
90	du Feu	DU FEU DESIGN	s	s		s
91	VIDAS	VIDYA	s	s		s
92	VOLVO	VOLTA TRUCKS / VOLTA ZERO / VOLTA / VTRUCKS / V TRUCKS / V-TRUCKS	d	?	d	d
93	CROC odor WC / CROC odor / Croc'Odor / Croc Odor the kitchen expert	cocod'or	s	s	d	d
94	RABE	RASE	s	s		s
95	CINTRA	CITRA	s	s	d	d
96	MATTERS	M4TTER	s	s	d	s
97	MFLOR	HFLOR / H FLOR	s	s	d	d
98	MBET	M-Bets	s	s	s	s
99	next	NXTWEAR S	s	s	s	s
100	STONES	STONE BREWING	s	s	s	s
101	CHEF	CHEFCHY	s	s	s	s
102	Bones	Bones Of Barbados	s	s	d	d
103	Satisfyer	SIMPLY SATISFY	d	d	s	d
104	FRANSA	FANZA	s	s	s	s
105	GAP	GAL London	s	d	d	d
106	CARBON	MyCarbon	s	s	s	s
107	COMPAL	COPALLI / COPAL TREE	?	?	d	?
108	DENNIS / DENNIS AND GNASHER	Dennis G	s	s	s	s
109	SAVANT	SAVANT POWER	s	s	s	s
110	HYPRR	HYPERNFT	s	s	s	s
111	POCKIT	MyPocket	s	s	s	s
112	SHEPHERD	WOLF & SHEPHERD	s	s	s	s
113	HONWAVE / HONDA / Honda e	Honbike	s	s	d	s
114	skin ² / NITRILE SKIN ²	SKINS	s	s	s	s
115	Hanson	HANSOL	s	s		s
116	CULT BEAUTY / CULT CONCIERGE	PERFUME CULT	s	s	s	s
117	BOSS / HUGO BOSS / BOSS HUGO	BOSSEUR	s	d		d
118	Kelio	KLEEO	s	s		s
119	e-BULLI	BULLIT	s	s		s
120	realme	REALMZ	s	s	d	d
121	MUSTANG	MUSTANG / FORD MUSTANG / MUSTANG MACH-E	s	s		s
122	EUREKA!	EUREKA EDUCATION	s	s	s	s
123	Rebelle Copenhagen	reBELLE BEAUTY	s	s	s	s
124	PI DATABOOK / PI	PIANYWHERE	d	d	s	d
125	CONSIGLIERI	CONSIGLIERA	s	s		s
126	Saypha	SHAYPE	d	s	d	d
127	FOLIUM SCIENCE	FOLIUM	s	s	d	s
128	MD	IntiMD	s	s	s	s
129	Higicol – AMMA / amma / AMMA COLORS	Amma Wellness	s	s		s
130	LIP INJECTION	LIPJECTION GLOSS	s	s		s
131	RESOLVA	CONSOLVA	s	s		s
132	CHOOEY	CHOOEE		s		s
133	Bloo	bluuwash	s	s	s	s
134	COVERSYL	COVIXYL-V	d	s		d

135	CLEAN V / CLEAN W / CLEANCO / CLEAN G / CLEAN R / CLEAN T	Drink Clean.	d	d	s	d
136	ZARA	ZAREUS	d	s	d	d
137	PXG Pharma	PXG	s	s	s	s
138	Click-EAT / CLICK EAT	SUBWAY CLICK & EAT	s	d	s	s
139	REPEVAX	EPVAX	s	s		s
140	CURVE	CRV	d	d	d	d
141	GEORGE	GEORGINE	s	s	s	s
142	One4All Favourites / One4all	OneFor / ONE FOR		s	s	s
143	PYRA	PRYA	s	s		s
144	HALLOUMI	GRILLOUMAKI / GRILLOUMI	s	s	d	d
145	CANE AND GRAIN	CANE & GRAIN INTERNATIONAL	s	s		s
146	BAD BOY	BBCC / BAD BOY CHILLER CREW	s	s	s	s
147	XACTIMATE / XACTANALYSIS / XACTWARE	BUILDXACT	d	d	d	d
148	SALIO	SALIOGEN	s	s	d	s
149	HotPatch	Patch	s	s	s	s
150	JUST	Just The Ticket	s	s	d	d
151	THINKSMART / THINKPAD / THINKBOOK / THINKSHIELD	XHINKCAR	d	d	d	d
152	AESULAP	AESKUCARE / AESKUCARE Allergy / AESKUCARE Food Intolerance	d	d		d
153	RESOLUTION	RESOLUTE	s	s	s	s
154	PUMA	Huma / Huma London			d	d
155	YOGA	Yoga Man		d		d
156	LIVE	VIVE	d	s	s	d
157	PLANET BOTTLE	One Planet	s	s	d	d
158	MySugardaddy	Sugar Daddy / Sugardaddy	s	s	s	s
159	PATTER	Yatter		d		d
160	HOLIDAY PEOPLE	The Holiday People	s	s	s	s
161	PRADA Invites / RE-PRADA / PRADA TIMECAPSULE SERIAL N	PADRA	s	s		s
162	THE IVY LEAGUE	The New Ivy League	s	s	s	s
163	UNITY TRUST BANK / UNITY	UNITY REAL ESTATE LTD	s	s	s	s
164	DREAMS / DREAM BIGGER Dreams / DREAM BIGGER / DREAM COACH	Dream Big Make Dua Move Mountains	s	s	s	s
165	MOON PRINCESS	Time Princess		d		d
166	FASHIONGO	FASHIONEGO	s	s	s	s
167	THE SECRET GARDEN PARTY	The Secret Garden Glamping / The Secret Garden				s
168	ELLESSE	ELLISS	s	d	d	d
169	SOLE TRADER / SOLE / SOLE SOLE / SOLE SISTER	SoulSistar	s	s	s	s
170	SWOOP / FLY SWOOP	SWOOP TAXIS / swooptaxis	s	s	s	s
171	WASHTOWER	Washing Tower / Washer Tower	s	s	s	s
172	POTETTE PLUS	Pote Plus	s	s	s	s
173	IDÉE	idee-home	s	s	d	s
174	JOY	BODY IN JOY	s	s	s	s
175	ROLEX	DERMAROLLEX	d	d	d	d
176	BARRIER	Barrier Coat	?	?	?	?
177	FLEX	FLEXX BY BBOXX	s	s	s	s
178	LOVELLO	LOVELLE	s	s		s
179	easyLand / EASYNETWORKS / EASYHUB	EasyMap	s	d	d	d
180	VARSITY / VARSITY SPIRIT FASHIONS	VARSITY HEADWEAR	s	s		s
181	WAKEN	Wakeful	s	s	s	s
182	AK DAMM	BLACK ADAM	d	d	d	d

183	YALU	[HYALU] BIOTIC				d
184	Z-BIOME	BIOME	s	s		s
185	Sleep Doctor / Sleep Dr	Dr.sleep	s	s	s	s
186	sane	CBDSANE	s	s	s	s
187	Closet. LONDON / CLOSET	THE LUXURY CLOSET	s	s	s	s
188	Lucky Strike	LUCKY BAR	s	s	s	s
189	ACTIVON MANUKA / ACTIVON	ARTIVION	s	s	d	s
190	YORXS	Yorks		s		s
191	AGROS	agro S	s	s	s	s
192	IVALUA / IVALUA VALUE BEYOND SAVINGS / IVALUA BUYER	iValue Solutions	s	s	s	s
193	Maplab	maplab.world	s	s	s	s
194	EVERY BODY	Everybodies	s	s	s	s
195	MATCH.COM / match / MG MatchGroup	MatchMate	s	s	s	s
196	NOUGHTY	NOUTI	s	d		d
197	Red Bull	ELFBULL	s	s	s	s
198	SMART	SMARTCARE / SMART CARE / SMARTBUSINESS / SMART BUSINESS / SMARTCLASS / SMART CLASS / SMARTSPACE	s	s	s	s
199	SNAP	Snap Nurse / SnapNurse	s	s	s	s
200	IV BOOST	IV PATCH	s	s	s	s
201	UNITY	INVIVO X UNITY	s	s	s	s
202	Eye of Horus	EYE OF ATUM	s	s	d	d
203	THUNDER PRODUCTIONS	SUN AND THUNDER	d	d	d	d
204	VISTAINTRA / VISTAVOX / VISTAPANO	VISTO	s	s	d	d
205	X WAY / XWAY	Exway	s	s		s
206	XCODE / CORE ANIMATION / CORE HAPTICS	XCORE	s	s	d	s
207	PETROL	PETROL REVOLT	s	s	s	s
208	SCAFFEZE	SCAFFX	s	s		s
209	BUBBLE / BUBL	BUBBLE ROCKET	s	s	s	s
210	THE LEONARDO COLLECTION	LEONARDO	s	s	s	s
211	LUCITE	Luci	s	s	d	d
212	Sense / Essence	SENSE	s	s	d	d
213	TESTEX	TEST-X		s	d	s
214	Lister's Brewery / Lister's	Listers		s		s
215	Nisha	Misha Cosmetics	s	s	d	s
216	AIRBNB	AIRBRICK	s	s		s
217	MOAT	Moat Systems / MOATSYSTEMS	s	s	s	s
218	ME YOU	YOU ME	s	s		s
219	KARMACOIN	KARMACASH / KARMAGIVES / KARMAPAY / KARMASHOPPING	s	s	s	s
220	GOBOX	G-Box	s	s	s	s
221	ATMA	AtmaSPA	s	s	s	s
222	BOSS	Kissboss	s	s	s	s
223	CHAINLINK / CHAINLINK LABS	LINK	s	s	d	d
224	JESSICA LONDON	JESSICA JOY LONDON	s	s	s	s
225	SNUGGLEDOWN	Snugglemore	s	s	s	s
226	AXIS	TRAXIS	d	d	d	d
227	Gadget Centre	Gadget Centre UK Ltd	s	s	s	s
228	WHITEHORSE LIQUIDITY PARTNERS	WHITEHORSE / H.I.G. WHITEHORSE	s	s	s	s
229	BILLIONAIRE	ZILLIONAIRE	s	s	s	s
230	GOLFER	The Golphers	s	s	s	s
231	PRED	PRD TECHNOLOGY	d	d		d
232	GHOST	GHOSTDANCER CBD	s	d	s	d

233	NATURLI'	NATUREAL	s	s	s	s
234	GLENFIDDICH	Inverfiddich	s	s	s	s
235	CONFIGON	CONFIGO	s	s		s
236	PURPLE COMPUTING / PURPLE	PURPLECUBE	d	d	d	d
237	CREATEME	Create.	s	s	s	s
238	sky / SKY X	YSKY	d	d	s	s
239	RING	Home Ring	s	s	s	s
240	CHLOE / Chloé	chloédigital	s	s	s	s
241	YOU BEAUTY DISCOVERY / YOU BEAUTY / YOU	YOU·OLOGY	d	d	d	d
242	SIO / SIO BEAUTY	RIO COSMETICS	d	d	d	d
243	DCSL	DCS	s	s		s

Appendix B: LCSSQs and LCSSTs, and 'remainders', for the 243 pairs of marks

Mark 1	Mark 2	LCSSQ	Rem. 1	Rem. 2	LCSST	Rem. 1	Rem. 2	Mod. Rat.-Ob. similarity
glow up: britain's next make-up star	glow up: britain's next make-up star	glow up: britain's next make-up star			glow up: britain's next make-up star			1.000
meta	meta	meta			meta			1.000
mustang	mustang	mustang			mustang			1.000
sense	sense	sense			sense			1.000
fittipaldi	efittipaldi	fittipaldi		e	fittipaldi		e	0.957
configon	configo	configo	n		configo	n		0.941
consiglieri	consigliera	consigliere	i	a	consigliere	i	a	0.917
billionaire	zillionaire	illionaire	b	z	illionaire	b	z	0.917
mysugardaddy	sugardaddy	sugardaddy	my		sugardaddy	my		0.917
1link	link	link	1		link	1		0.909
xway	exway	xway		e	xway		e	0.909
this girl can	this girl came	this girl ca	n	me	this girl ca	n	me	0.897
dcs1	dcs	dcs	l		dcs	l		0.889
eataliano	eatalia	eatalia	no		eatalia	no		0.889
sir	siro	sir		o	sir		o	0.889
sky	ysky	sky		y	sky		y	0.889
lister's	listers	listers	'		lister	's	s	0.882
makeup wardrobe	makeup wardrobing	makeup wardrob	e	ing	makeup wardrob	e	ing	0.882
holiday people	the holiday people	holiday people		te h	holiday people		the	0.882
lovello	lovelle	lovell	o	e	lovell	o	e	0.875
carbon	mycarbon	carbon		my	carbon		my	0.875
dennis	dennis g	dennis		g	dennis		g	0.875
fashiongo	fashionego	fashiongo		e	fashion	go	ego	0.857
chooey	chooee	chooe	y	e	chooe	y	e	0.857
prinker	prink	prink	er		prink	er		0.857
realme	realmz	realm	e	z	realm	e	z	0.857
kramer	cramer	ramer	k	c	ramer	k	c	0.857
hanson	hansol	hanso	n	l	hanso	n	l	0.857
patter	yatter	atter	p	y	atter	p	y	0.857
z-biome	biome	biome	z-		biome	z-		0.857
pt powerpod	powerpod	powerpod	t p		powerpod	pt		0.857
the secret garden party	the secret garden	the secret garden	party		the secret garden	party		0.857
perpetual	perpetuel	perpetul	a	e	perpetu	al	el	0.850
agros	agro s	agros			agro	s	s	0.846
lucite	luci	luci	te		luci	te		0.833
very	veryco	very		co	very		co	0.833
axis	traxis	axis		tr	axis		tr	0.833
lotus	motus	otus	l	m	otus	l	m	0.833

mflor	hflor	flor	m	h	flor	m	h	0.833
skin ²	skins	skin	²	s	skin	²	s	0.833
createme	create.	create	me	.	create	me	.	0.824
victoria	victoro	victor	ia	o	victor	ia	o	0.824
the good schools guide	the good school	the good school	s guide		the good school	s guide		0.821
treca	trea	trea	c		tre	ca	a	0.818
george	georgine	george		in	georg	e	ine	0.813
resolution	resolute	resolut	ion	e	resolut	ion	e	0.800
foltene	foltex	folte	ne	x	folte	ne	x	0.800
live	vive	ive	l	v	ive	l	v	0.800
salio	saliogen	salio		gen	salio		gen	0.800
e-bulli	bullit	bulli	e-	t	bulli	e-	t	0.800
hotpatch	patch	patch	hot		patch	hot		0.800
diamond mist	diamond max	diamond m	ist	ax	diamond m	ist	ax	0.800
puma	huma	uma	p	h	uma	p	h	0.800
gadget centre	gadget centre uk ltd	gadget centre		uk ltd	gadget centre		uk ltd	0.800
the ivy league	the new ivy league	the ivy league		new	ivy league	the	the new	0.794
testex	test-x	testx	e	-	test	ex	-x	0.786
potette plus	pote plus	pote plus	tte		te plus	potet	po	0.783
burgerme	burgerly	burger	me	ly	burger	me	ly	0.778
purple	purplecube	purple		cube	purple		cube	0.778
str8	st8	st8	r		st	r8	8	0.778
cintra	citra	citra	n		tra	cin	ci	0.769
prinz	prinse	prin	z	se	prin	z	se	0.769
chef	chefchy	chef		chy	chef		chy	0.769
atma	atmaspa	atma		spa	atma		spa	0.769
boss	bossvel	boss		vel	boss		vel	0.769
sane	cbdsane	sane		cbd	sane		cbd	0.769
ryzen	ryzeup	ryze	n	up	ryze	n	up	0.769
boss	bosseur	boss		eur	boss		eur	0.769
barrier	barrier coat	barrier		coat	barrier		coat	0.762
gobox	g-box	gbox	o	-	box	go	g-	0.750
vidas	vidya	vida	s	y	vid	as	ya	0.750
yorxs	yorks	yors	x	k	yor	xs	ks	0.750
mbet	m-bets	mbet		-s	bet	m	m-s	0.750
zara	zarzar	zara		zr	zar	a	zar	0.750
vizrt	vizst	vizt	r	s	viz	rt	st	0.750
xcode	xcore	xcoe	d	r	xco	de	re	0.750
moon princess	time princess	m princess	oon	tie	princess	moon	time	0.750
scaffeze	scaffx	scaff	eze	x	scaff	eze	x	0.750
nrj	nrg	nr	j	g	nr	j	g	0.750
match	matchmate	match		mate	match		mate	0.750
mygenecare	mygenes	mygene	care	s	mygene	care	s	0.737
asprey	dave asprey	asprey		dve a	asprey		dave	0.737
mutant	mega mutant	mutant		ega m	mutant		mega	0.737

halloumi	grilloumi	lloumi	ha	gri	lloumi	ha	gri	0.737
energeo	enerjo	enero	ge	j	ener	geo	jo	0.733
matters	m4tter	mtter	as	4	tter	mas	m4	0.733
demiegod	demigods	demigod	e	s	demi	egod	gods	0.722
naturli'	natureal	naturl	i'	ea	natur	li'	eal	0.722
azure	azurity	azur	e	ity	azur	e	ity	0.714
waken	wakeful	wake	n	ful	wake	n	ful	0.714
boss	kissboss	boss		kiss	boss		kiss	0.714
ping	pingnpay	ping		npay	ping		npay	0.714
yoga	yoga man	yoga		man	yoga		man	0.714
repevax	epvax	epvax	re		vax	repe	ep	0.714
snuggledown	snugglemore	snuggleo	dwn	mre	snuggle	down	more	0.708
resolva	consolva	solva	re	con	solva	re	con	0.706
swift	microswift	swift		micro	swift		micro	0.706
smart	smart care	smart		care	smart		care	0.706
jessica london	jessica joy london	jessica london		joy	jessica	london	joy london	0.706
philips	philzops	philps	i	zo	phil	ips	zops	0.706
asos	asas	ass	o	a	as	os	as	0.700
curve	crv	crv	ue		rv	cue	c	0.700
mbfw	mvfw	mfw	b	v	fw	mb	mv	0.700
rabe	rase	rae	b	s	ra	be	se	0.700
fido	fiio	fio	d	i	fi	do	io	0.700
ulma	luma	uma	l	l	ma	ul	lu	0.700
maplab	maplab.world	maplab		.world	maplab		.world	0.700
flowers	flower cafe	flower	s	cafe	flower	s	cafe	0.700
pusher	pushers only	pusher		s only	pusher		s only	0.700
savant	savant power	savant		power	savant		power	0.700
karmacoin	karmacash	karmac	oin	ash	karmac	oin	ash	0.700
intellicare	intelecure	intelcare	li	e	intel	licare	ecare	0.696
paris-match	pari match	parimatch	s-		match	paris-	pari	0.696
washtower	washer tower	washtower		er	tower	wash	washer	0.696
agrho	agro s	agro	h	s	agr	ho	o s	0.692
pikdare	pi-kare	pikare	d	-	are	pikd	pi-k	0.688
retaron	rettron	retron	a	l	ret	aron	Iron	0.688
goddess	godless	godess	d	l	god	dess	less	0.688
pockit	mypocket	pockt	i	mye	pock	it	myet	0.688
ibaccy	biccy baccy	ibaccy		bccy	baccy	i	biccy	0.684
cane and grain	cane and grain international	cane and grain		international	cane and grain		international	0.682
glenfiddich	inverfiddich	efiddich	gln	invr	fiddich	glen	inver	0.680
eye of horus	eye of atum	eye of u	hors	atm	eye of	horus	atum	0.680
lemon perfect	peach perfect	e perfect	lmon	pach	perfect	lemon	peach	0.679
ellesse	elliss	ellss	ee	i	ell	esse	iss	0.667
compal	copalli	copal	m	li	pal	com	coli	0.667
sizzling fortunes	sizzling bells	sizzling es	fortun	bll	sizzling	fortunes	bells	0.667

shepherd	wolf and shepherd	shepherd		wolf and	shepherd		wolf and	0.667
zara	zareus	zar	a	eus	zar	a	eus	0.667
dreams	dream rite	dream	s	rite	dream	s	rite	0.667
life	chop life	life		chop	life		chop	0.667
du feu	du feu design	du feu		design	du feu		design	0.667
volvo	volta	vol	vo	ta	vol	vo	ta	0.667
swoop	swoop taxis	swoop		taxis	swoop		taxis	0.667
sleep dr	dr.sleep	sleep	dr	dr.	sleep	dr	dr.	0.667
petrol	petrol revolt	petrol		revolt	petrol		revolt	0.667
bubble	bubble rocket	bubble		rocket	bubble		rocket	0.667
chainlink	link	link	chain		link	chain		0.667
ring	home ring	ring		home	ring		home	0.667
idee	idee-home	idee		-home	idee		-home	0.667
wear the change	wear the future	wear the e	chang	futur	wear the	change	future	0.656
every body	everybodies	everybod	y	ies	every	body	bodies	0.652
lucky strike	lucky bar	lucky r	stike	ba	lucky	strike	bar	0.652
noughty	naughtea	nught	oy	aea	ught	noy	naea	0.647
easyland	easymap	easya	lnd	mp	easy	land	map	0.647
red bull	elbull	ebull	rd	lf	bull	red	elf	0.647
activon	artivion	ativon	c	ri	tiv	acon	arion	0.647
anytime fitness	anytime pro	anytime	fitness	pro	anytime	fitness	pro	0.643
saypha	shaype	sayp	ha	he	ayp	sha	she	0.643
noughty	nouti	nout	ghy	i	nou	ghty	ti	0.643
sherco	charco	hrco	se	ca	rco	she	cha	0.643
satisfyer	simply satisfy	satisfy	er	imply s	satisfy	er	simply	0.640
varsity	varsity headwear	varsity		headwear	varsity		headwear	0.640
catapult vision	catapult consulting	catapult sin	vio	conultg	catapult	vision	consulting	0.639
zemo	zoomo	zmo	e	oo	mo	ze	zoo	0.636
oyster	oyster and pop	oyster		and pop	oyster		and pop	0.636
folium science	folium	folium	science		folium	science		0.636
geneverse	genv3rse	genvrse	ee	3	rse	geneve	genv3	0.632
chloe	chloedigital	chloe		digital	chloe		digital	0.632
suntech	suntank	sunt	ech	ank	sunt	ech	ank	0.625
airbnb	airbrick	airb	nb	rick	airb	nb	rick	0.625
waterford	lady louisa waterford	waterford		lady louisa	waterford		lady louisa	0.625
snap	snap nurse	snap		nurse	snap		nurse	0.625
nutravita	nootrovita	ntrvita	ua	ooo	vita	nutra	nootro	0.619
flexolid	kp flexilid	flexlid	o	kp i	flex	olid	kp ilid	0.619
fransa	fanza	fana	rs	z	an	frsa	fza	0.615
cazoo	carkoo	caoo	z	rk	ca	zoo	rkoo	0.615
gizebra	debra the gizebra	gizebra		debra the	gizebra		debra the	0.615
x15	vivo x15	x15		vivo	x15		vivo	0.615
lip injection	lipjection gloss	lipjection	in	gloss	jection	lip in	lip gloss	0.613

sole sister	soulsistar	solsistr	e e	ua	sist	sole er	soular	0.609
pyra	prya	pya	r	r	a	pyr	pry	0.600
alpha boxing	alpha force	alpha o	bxing	frce	alpha	boxing	force	0.600
thinksmart	xhinkcar	hinkar	tsmt	xc	hink	tsmart	xcar	0.600
hypr	hypernft	hypr	r	enft	hyp	rr	ernft	0.600
md	intimd	md		inti	md		inti	0.600
jolly	jolly peckish	jolly		peckish	jolly		peckish	0.600
activist	activist ingredients	activist		ingredients	activist		ingredients	0.600
vault ip	brandvault	vault	ip	brand	vault	ip	brand	0.600
rebelle copenhagen	rebelle beauty	rebelle ea	copnhgen	buty	rebelle	copenhagen	beauty	0.588
lush	lush lights	lush		lights	lush		lights	0.588
vistaintra	visto	vist	aintra	o	vist	aintra	o	0.588
live	life's	lie	v	f's	li	ve	fe's	0.583
skullcandy	skull gaming	skullan	cdy	gmig	skull	candy	gaming	0.583
croc'odor	cocod'or	cocodor	r'	'	od	croc'or	coc'or	0.579
ak damm	black adam	ak dam	m	blca	dam	ak m	black a	0.579
tygrys	tigris	tgrs	yy	ii	gr	tyys	tiis	0.571
easyjet	easycosmetic	easyet	j	cosmic	easy	jet	cosmetic	0.571
vfh	vfonline	vfh		online	vfh		online	0.571
sky	sky force	sky		force	sky		force	0.571
six days	day6	day	six s	6	day	six s	6	0.571
stones	stone brewing	stone	s	brewing	stone	s	brewing	0.571
honda	honbike	hon	da	bike	hon	da	bike	0.571
clean v	drink clean.	clean	v	drink .	clean	v	drink .	0.571
unity	invivo x unity	unity		invivo x	unity		invivo x	0.571
me you	you me	you	me	me	you	me	me	0.571
you	you-ology	you		-ology	you		-ology	0.571
dose and co.	dose labs	dose a	nd co.	lbs	dose	and co.	labs	0.565
eureka!	eureka education	eureka	!	education	eureka	!	education	0.560
planet bottle	one planet	planet	bottle	one	planet	bottle	one	0.560
closet	the luxury closet	closet		the luxury	closet		the luxury	0.560
rolex	dermarollex	rolex		demarl	lex	ro	dermarol	0.556
moat	moat systems	moat		systems	moat		systems	0.556
evolutiq	essential evolution	evoluti	q	ssential eon	evoluti	q	essential on	0.552
bad boy	bad boy chiller crew	bad boy		chiller crew	bad boy		chiller crew	0.552
armalight	armatherm	armah	ligt	term	arma	light	therm	0.550
click eat	subway click and eat	click eat		subway and	click	eat	subway and eat	0.548
cartils	cartel design	cartls	i	e deign	cart	ils	el design	0.545
the leonardo collection	leonardo	leonardo	the collection		leonardo	the collection		0.545
ghost	ghostdancer cbd	ghost		dancer cbd	ghost		dancer cbd	0.545

whitehorse liquidity partners	whitehorse	whitehorse	liquidity partners		whitehorse	liquidity partners		0.537
re-prada	padra	pada	re-r	r	ra	re-pda	pad	0.533
pxg pharma	pxg	pxg	pharma		pxg	pharma		0.533
one4all	onefor	one	4all	for	one	4all	for	0.533
coversyl	covixyl-v	covyl	ers	ix-v	cov	ersyl	ixyl-v	0.526
aesculap	aeskucare	aesca	ulp	kure	aes	culap	kucare	0.526
amma	amma wellness	amma		wellness	amma		wellness	0.526
golfer	the golphers	golher	f	the ps	her	golf	the golps	0.524
kelio	kleeo	keo	li	le	k	elio	leeo	0.500
iv boost	iv patch	iv t	boos	pach	iv	boost	patch	0.500
3monkeys	3 monkeys communicatio ns	3monkeys		communicat ions	monkeys	3	3 communicat ions	0.500
bones	bones of barbados	bones		of barbados	bones		of barbados	0.500
xactimate	buildxact	xact	imate	build	xact	imate	build	0.500
joy	body in joy	joy		body in	joy		body in	0.500
flex	flexx by bboxx	flex		x by bboxx	flex		x by bboxx	0.500
yalu	[hyalu] biotic	yalu		[h] biotic	yalu		[h] biotic	0.500
ivalua	ivalue solutions	ivalu	a	e solutions	ivalu	a	e solutions	0.500
just	just the ticket	just		the ticket	just		the ticket	0.476
next	nxtwear s	nxt	e	wear s	xt	ne	nwear s	0.467
amazon	amazon food trader ltd	amazon		food trader ltd	amazon		food trader ltd	0.467
nisha	misha cosmetics	isha	n	m cosmetics	isha	n	m cosmetics	0.455
thunder productions	sun and thunder	thunder	productions	sun and	thunder	productions	sun and	0.444
bloo	bluuwash	bl	oo	uuwash	bl	oo	uuwash	0.429
pi	pianywhere	pi		anywhere	pi		anywhere	0.429
unity	unity real estate ltd	unity		real estate ltd	unity		real estate ltd	0.429
last shelter:survival	doomsday: last survivors	last surviv	helter:sal	doomsday: ors	last s	helter:survi val	doomsday: urvivors	0.404
gap	gal london	ga	p	l london	ga	p	l london	0.400
cult beauty	perfume cult	cult	beauty	perfume	cult	beauty	perfume	0.400
spear and jackson predator	predator gutter vacuum	pear ac	sandjksn predator	rdtogutterv uum	predator	spear and jackson	gutter vacuum	0.360
pred	prd technology	pre	d	d tchnology	pr	ed	d technology	0.350
sio	rio cosmetics	si	o	rio cometcs	io	s	r cosmetics	0.333
dreams	dream big make dua move mountains	dreams		big make dua move mountain	dream	s	big make dua move mountains	0.317
6x	vivo x6	6	x	vivo x	6	x	vivo x	0.182